From root tips to phylogenetic branches : a functional and evolutionary study of peroxidases

Passardi, Filippo

UNIVERSITÉ DE GENÈVE
Département de botanique
et de biologie végétale

FACULTÉ DES SCIENCES
Professeur C. Penel
Dr. C. Dunand

From root tips to phylogenetic branches : a functional and evolutionary study of peroxidases

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences, mention biologique

par
Filippo PASSARDI
de
Torricella-Taverne (TI)

Thèse N°3822

GENÈVE
Atelier de reproduction d'Uni-Mail
2006

UNIVERSITÉ DE GENÈVE
Département de botanique
et de biologie végétale

FACULTÉ DES SCIENCES
Professeur C. Penel
Dr. C. Dunand

From root tips to phylogenetic branches : a functional and evolutionary study of peroxidases

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences, mention biologique

par
Filippo PASSARDI
de
Torricella-Taverne (TI)

Thèse N°3822

GENÈVE
Atelier de reproduction d'Uni-Mail
2006

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
mention biologique

Thèse de *Monsieur Filippo PASSARDI*

intitulée :

"From Root Tips to Phylogenetic Branches : a Functional and Evolutionary Study of Peroxidases "

La Faculté des sciences, sur le préavis de Messieurs C. PENEL, professeur titulaire et directeur de thèse (Département de botanique et biologie végétale), Ch. DUNAND, docteur et co-directeur de thèse (Département de botanique et biologie végétale), L. FALQUET, docteur (Institut Suisse de Bioinformatique – Epalinges, Suisse), et P. BOLWELL, professeur (University of London – Royal Holloway – School of Biological Sciences – Lonson, United Kingdom), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 14 décembre 2006

Thèse - 3822 -

Le Doyen, Pierre SPIERER

N.B.- La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Nombre d'exemplaires à livrer par colis séparé à la Faculté : - 7 -

Remerciements

Mon travail de thèse n'aurait pu se faire sans l'aide de nombreux membres du laboratoire. A ce titre, je tiens particulièrement à remercier **Claude** et **Christophe**, qui m'ont accueilli et m'ont suivi tout au long de mon doctorat. Ils m'ont appris à persévérer, à réfléchir, à me « débrouiller » et à me faire découvrir le monde fascinant des peroxydases ; ils m'ont aussi permis de concilier une vie de famille chargée avec le bon déroulement de ma thèse. Je souhaite à Claude de profiter pleinement du temps libre qui l'attend. J'espère que Christophe pourra continuer dans le domaine de la recherche, pour lequel il s'est beaucoup investi, tant par son travail que par sa capacité à encadrer et à motiver un grand nombre de gens dans ce laboratoire ; j'espère que ses grandes qualités autant scientifiques qu'humaines seront reconnues à leur juste valeur, afin qu'il puisse continuer à l'amélioration des connaissances sur les peroxydases.

Merci aussi à tous les collègues que j'ai eu la chance de croiser, et qui ont contribué à rendre une joyeuse atmosphère au quotidien, sans oublier leur aide et leurs conseils durant mon travail. Merci à **Luisa** (pour son sourire… et pour améliorer mon portugais), **Maria** (pour sa recette des « frijoles »), **Jan** (pour ses discours décalés), **Yann** (pour son aide à rechercher l'ancêtre des peroxydases), **Dale** (pour me faire découvrir le métal), **Claudia** (pour parler de peroxydases et… de bébés, et pour une journée de ski inoubliable), **Mireille** (pour son aide lors de nombreuses expériences, et pour sa philosophie « Rock&Roll »), **Jennifer** (pour sa bonne humeur communicative et…pour son multicrêpes), **Nenad** (pour toutes les glaces, chocolats et autres délices), **Virginie** (pour me faire découvrir l'art gothique), **Marion** (pour son immense gentillesse… et pour améliorer mon allemand), **Michele** (pour améliorer mon italien), **Sonia** (pour parler des « stars » californiennes… et pour améliorer mon anglais), **Ruben&Marta** (pour parler d'autres sujets que les peroxydases… et pour améliorer mon espagnol), **Lucette** (pour son aide précieuse lors des chromatographies et pour bien d'autres conseils), **Christophe dit « Junior »** (pour apprendre de nouvelles expressions régionales), **Dan** (pour m'aider à cultiver les algues, et pour ses nombreuses cartes postales), **Sergiu** (pour ses chansons poétiques).

En-dehors du laboratoire, je remercie aussi **Alain**, **Bernard** et **Yves**, les jardiniers du département, qui ont eu beaucoup de travail lors des nombreux changements de phytotrons ; **Michèle**, pour ses conseils en microscopie ; **Jocelyne**, pour sa collaboration lors de mon travail sur les catalases-peroxydases chez les bactéries ; **Patricia**, pour son aide à mettre au point mon protocole de marquage au X-Gluc, et bien sûr pour sa grande gentillesse au quotidien ; **Vassilios** et **Grégory**, pour avoir permis à la PeroxiBase de devenir un outil performant ; enfin, **Antonio Satorre**, de l'imprimerie, auquel j'ai demandé des délais d'impression très serrés.

Merci aussi au professeur **Paul Bolwell** et à **Laurent Falquet**, qui ont accepté de faire partie de mon jury de thèse, et qui ont eu la patience de lire ce travail dont je ne m'attendais pas qu'il atteindrait plus de 300 pages.

Finalement, merci à mes parents et mon frère, qui m'ont toujours encouragé à continuer durant ma thèse ; à mes beaux-parents, qui m'ont aidé à pouvoir travailler au laboratoire en gardant mes enfants ; à mes enfants, **Rafael** et **Lara**, ainsi qu'à ma femme **Maria**, tous les trois sources inépuisables de motivation et de bonheur.

RESUME EN FRANÇAIS

1. Introduction

La superfamille dite des « peroxydases de plantes, champignons et bactéries » (Welinder, 1992) est composée de protéines qui présentent toutes une structure tridimensionnelle similaire (Edwards et al., 1993; Smulevich et al., 2006). Toutes ces peroxydases réduisent l'eau oxygénée en eau, mais oxydent des substrats très divers. La superfamille est séparée en trois classes distinctes, qui diffèrent principalement de par leur séquence peptidique primaire. La classe I comprend trois types de peroxydases : les ascorbate peroxydases (APX), les cytochrome c peroxydases (CCP) et les catalases-peroxidases (CP). Les APX se trouvent dans les cellules chloroplastiques de plantes, et utilisent principalement l'ascorbate pour éliminer l'excès de H_2O_2 produit lors de la photosynthèse (Shigeoka et al., 2002). Les CCP sont liées aux mitochondries et servent aussi à réduire l'eau oxygénée (produite par la respiration) en oxydant le cytochrome c (Skulachev, 1998; Erman et Vitello, 2002). Quant aux CP, elles ont principalement le rôle de catalases (Obinger et al., 1999), et se retrouvent chez de nombreuses bactéries et quelques champignons.

Les peroxydases de classe II jouent un rôle essentiel dans la dégradation des déchets organiques du sol : elles sont capables d'oxyder des substrats à très haut potentiel redox, en particulier la lignine. Les lignine peroxydases (LiP) oxydent directement leur substrat, alors que les manganèse peroxydases (MnP) relâchent des ions Mn^{3+} (sous forme de chélates organiques) qui, à leur tour, oxydent des molécules à haut potentiel redox (Gold et Alic, 1993). Enfin, les peroxydases dites « versatiles » (VP) peuvent effectuer les deux types de mécanismes (Ruiz-Dueñas et al., 2001).

Les peroxydases de classe III sont typiques des plantes vertes : toutes les plantes, des hépatiques aux plantes supérieures les plus évoluées, possèdent des peroxydases de classe III. Au contraire des deux classes précédentes, les peroxydases de classe III forment des familles multigéniques de plusieurs dizaines (voire plus de 100) de gènes. Une des conséquences majeures du nombre important d'isoformes est la difficulté de pouvoir déterminer un rôle précis pour chacune de ces peroxydases dans les plantes. En plus du nombre élevé d'isoformes, les peroxydases de classe III ne présentent pas d'affinité particulière pour une molécule : *in vitro*, elles peuvent oxyder efficacement un nombre incalculable de substrats. *In vivo*, il est presque impossible ainsi de savoir si elles ont un substrat particulier, et si oui,

lequel. Malgré ces difficultés, on sait maintenant que les peroxydases participent (mais on ne sait pas quelle peroxydase, et comment) à de très nombreuses fonctions, telles que la construction de la paroi, la croissance par élongation, la défense contre les pathogènes ou la sénescence. Suite à de nombreuses publications récentes sur les nombreuses facettes de peroxydases de classe III, et à la caractérisation d'un deuxième cycle réactionnel des peroxydases, le cycle dit « hydroxylique » (Berglund et al., 2002; Liszkay et al., 2003), j'ai entrepris, avec plusieurs co-auteurs, d'écrire deux revues : la première traite de l'implication des espèces actives de l'oxygène (EAO) dans le remodelage de la paroi ; la deuxième décrit le rôle des peroxydases dans la vie d'une plante, de la germination à la sénescence. Ces deux revues sont publiées et incluses dans cette thèse. Leurs références sont :

Passardi F, Penel C, Dunand C.
Performing the paradoxical: how plant peroxidases modify the cell wall.
Trends Plant Sci. 2004 Nov; 9(11):534-40.

Passardi F, Cosio C, Penel C, Dunand C.
Peroxidases have more functions than a Swiss army knife.
Plant Cell Rep. 2005 Jul; 24(5):255-65.

2. Etude fonctionnelle des peroxydases AtPrx33 et AtPrx34

Comme énoncé dans l'introduction, les peroxydases de classe III forment des grandes familles multigéniques, et n'ont pas de spécificité connue particulière pour un substrat. Ainsi, une des difficultés majeures que rencontrent les chercheurs est de trouver un rôle précis pour une peroxydase. Un des buts de mon travail de thèse a été d'identifier le profil d'expression et la fonction de la peroxydase AtPrx33 chez *Arabidopsis thaliana*. En parallèle, sa plus proche orthologue (92% d'identité), AtPrx34, a aussi été comparée afin de mieux comprendre l'intérêt de la présence de peroxydases très semblables chez *Arabidopsis*.

Les premières expériences, effectuées avec le mutant d'insertion *atprx33*, m'ont permis de constater que la perte d'AtPrx33 a pour conséquence une croissance réduite de la racine principale. Des plantes transgéniques AtPrx33 ::*AtPrx33* ont par la suite été obtenues dans le mutant *atprx33*, et le phénotype « racine courte », comme prévu, a disparu. De plus,

des plantes transgéniques surexprimant AtPrx33 (35S ::*AtPrx33*) ont donné des plantes avec des racines plus longues que les plantes sauvages. Des expériences ultérieures avec des plantes AtPrx33 ::*GFP* (promoteur-GFP) ont confirmé une forte expression d'*AtPrx33* dans la zone d'élongation cellulaire, et les lignées 35S ::*AtPrx33* ::*GFP* ont affiné cette localisation aux parois des cellules. Enfin, la mesure de la longueur des cellules dans les différentes plantes décrites ci-dessus a confirmé que le rôle d'AtPrx33 chez *Arabidopsis* est de participer à l'élongation des cellules de la racine, et ceci au niveau de la paroi. Les résultats de cette étude sont détaillés dans l'article suivant :

Passardi F, Tognolli M, De Meyer M, Penel C, Dunand C.
Two cell wall associated peroxidases from *Arabidopsis* influence root elongation.
Planta. 2006 Apr; 223(5):965-74.

Dans une seconde partie, j'ai essayé d'élucider le mécanisme par lequel la peroxydase AtPrx33 pouvait provoquer l'élongation cellulaire. Un ancien doctorant (Michael Tognolli) avait remarqué, à une seule reprise toutefois, que le mutant *atprx33* pouvait spontanément former un cal et que ce cal pouvait par la suite former des structures différenciées, sans l'adjonction d'hormones de croissance. Or, la formation de cals et leur différenciation dépendent de la présence et de l'équilibre de deux hormones : les auxines et les cytokinines (Skoog et Miller, 1957). L'auxine étant un substrat bien connu, *in vitro*, des peroxydases (Ricard et Nari, 1966; Yamazaki et Yamazaki, 1973), j'ai entrepris d'étudier la relation qu'il pouvait exister entre l'auxine (AIA) et la peroxydase AtPrx33 *in vivo*. En parallèle, j'ai mis au point un protocole permettant de comparer la formation de cals et leur différenciation de manière standardisée chez les différentes plantes transgéniques citées précédemment (*atprx33*, *atprx33* x AtPrx33 ::*AtPrx33*, AtPrx33 ::*GFP*).

Afin d'étudier la répartition de l'auxine chez des plantes sauvages et le mutant *atprx33*, j'ai obtenu des plantes transgéniques exprimant le gène marqueur de la glucuronidase (GUS) sous influence du promoteur GH3 inductible par l'auxine. Aucune différence n'est apparue entre le mutant *atprx33* et les lignées contrôle. En revanche, l'auxine était majoritairement localisée dans la zone d'élongation, comme AtPrx33. Différents traitements chimiques indiquent que l'auxine et AtPrx33 ne varient pas de la même façon, et qu'ainsi l'auxine ne contrôle pas directement l'expression de cette peroxydase. En revanche, la co-localisation de l'auxine et de la peroxydase suggère qu'elles pourraient interagir *in vivo*.

L'expérience suivante a été d'effectuer un dosage d'auxine par chromatographie dans les racines et les parties aériennes de différents mutants (*atprx33*, *atprx34*, double-mutant *atprx33*/*atprx34*, surexpresseur 35S ::*AtPrx34*). Les résultats démontrent que soit la perte, soit la surexpression des peroxydases AtPrx33 et AtPrx34 conduisent à une diminution marquée des niveaux d'auxine dans les parties aériennes. Dans les racines en revanche, seule une dérégulation de l'expression de la peroxydase AtPrx34 conduit à une variation des niveaux d'auxine : contrairement aux parties aériennes, l'auxine augmente dans les racines. Les résultats des expériences avec GH3 ::*GUS* sont difficiles à interpréter de façon globale, mais indiquent qu'AtPrx33 et AtPrx34 interagissent probablement avec l'auxine et que cette interaction pourrait conduire à l'élongation cellulaire par le biais d'espèces actives de l'oxygène. Afin de confirmer ou non cette relation, j'ai par la suite essayé de produire les deux peroxydases et de tester leur capacité à oxyder l'auxine. Les peroxydases sont intégrées à un ADN viral (baculovirus) et produites dans des cellules d'insecte (*Spodoptera frugiperda*), un système déjà utilisé par ce laboratoire précédemment (Dunand et al., 2002). Les premiers résultats, vu la petite quantité produite, on permis de confirmer que les deux peroxydases s'attachent à la pectine, et donc à la paroi cellulaire (cf. lignées 35S ::*AtPrx33* ::*GFP*). Cependant, avec Mireille de Meyer, nous n'avons pas encore réussi à obtenir des quantités suffisantes pour effectuer des tests d'activité auxine oxydase, et la production a été relancée récemment.

En même temps que ma recherche sur les plantes GH3 ::*GUS*, j'ai étudié un effet dérivé de l'auxine : la formation des cals et leur différentiation. J'ai ainsi mis au point un protocole précis pour générer des cals à partir de racines de plantules d'*Arabidopsis*, en m'inspirant de la littérature (Cary et al., 2001). Une fois ce protocole établi, j'ai pu démontrer, grâce à l'étude des plantes transgéniques *atprx33*, *atprx33* x AtPrx33 ::*AtPrx33*, 35S ::AtPrx33, *atprx34* et *atprx33*/*atprx34* que la peroxydase AtPrx33 participe à la différenciation des cals en organes aériens. La peroxydase AtPrx34, en revanche, ne semble pas jouer de rôle particulier. De plus, le retard de différenciation observé avec *atprx33* se retrouve aussi avec les cals qui surexpriment *AtPrx33*. Ce résultat démontre qu'un équilibre précis régit le processus de différenciation, et que la perturbation de cet équilibre, dans un sens ou un autre, conduit aux mêmes effets. Il est tentant, à ce point, de suggérer que le facteur contrôlant cet équilibre est la concentration d'auxine. D'ailleurs, des cals transgéniques GH3 ::*GUS* montrent, comme dans les pointes des racines, une co-localisation de l'auxine et de la peroxydase AtPrx33. Enfin, des plantes adultes surexprimant *Atprx33* ont montré un défaut marqué de différenciation dans leurs parties aériennes : dans certaines

conditions de croissance (partiellement définies), les plantes adultes avaient des tiges très courtes, voire inexistantes, un taux de mortalité accru ainsi qu'une fasciation des tiges (Fig. 1). Etant donné que la peroxydase *Atprx33* est peu ou pas exprimée dans les parties aériennes, ce résultat confirme le rôle important de cette enzyme dans la différenciation d'organes.

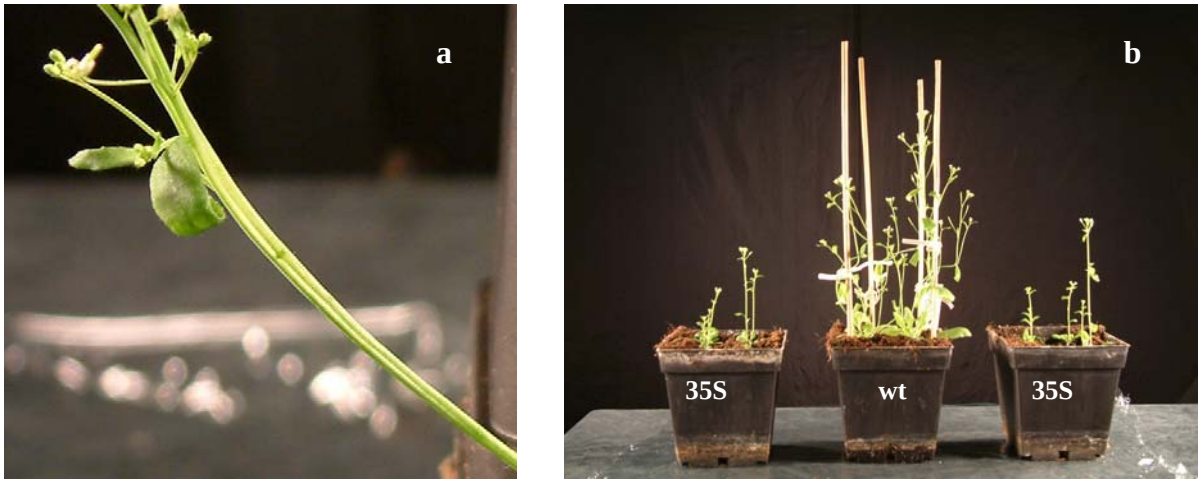


Figure 1

Plantes contrôles et 35S::*AtPrx33* agées d'un mois.

a Exemple de tiges fasciées chez une plante adulte 35S::*AtPrx33*

b Comparaison de plantes 35S::*AtPrx33* (35S) et sauvages (wt)

Malgré une observation répétée de ce phénotype dans plusieurs lignées indépendantes, ces mêmes plantes ont aussi poussé normalement lors d'autres expériences. Des recherches sont encore en cours pour déterminer dans quelles conditions ce phénotype apparaît.

Lors des expériences citées dans ce chapitre, une difficulté majeure s'est présentée dès le début : les mutants *atprx33* et *atprx34* ont été isolés à partir de deux écotypes différents, respectivement Wassilewskija (Ws) et Columbia (Col). Malgré leur forte ressemblance, le profil d'expression des peroxydases ainsi que de nombreuses variations de traits phénotypiques particuliers se sont vite révélés d'une grande importance dans l'interprétation des résultats et l'établissement de stratégies expérimentales. Avec plusieurs collègues du laboratoire, nous avons ainsi entrepris une étude détaillée des caractéristiques de trois écotypes couramment utilisés au laboratoire: Ws, Col et La*er* (Landsberg *erecta*). Mon travail a principalement consisté dans l'étude de la différenciation de cals chez ces écotypes, et à la rédaction de l'article scientifique suivant :

Passardi F, Dobias J, Valério L, Guimil S, Penel C, Dunand C.
Morphological and physiological traits of three major *Arabidopsis thaliana* accessions.
J Plant Physiol. 2006 *In press.*

3. Peroxydases de classe III chez d'autres plantes

Suite au séquençage du génome d'*Arabidopsis thaliana*, Michael Tognolli avait, dans notre laboratoire, répertorié tous les gènes de peroxydase présent chez cette plante (Tognolli et al., 2002). Au début de ma thèse, le projet de séquençage du génome du riz (*Oryza sativa ssp. japonica* cv. Nipponbare) touchait à sa fin. Avec Christophe Dunand, nous avons ainsi entrepris de chercher tous les gènes de peroxydase du riz. Cette étude nous a permis de comparer les génomes des deux plantes, de mieux comprendre les mécanismes ayant conduit à la formation de familles multigéniques et de déterminer l'évolution des peroxydases chez les plantes terrestres. Concernant ce dernier point, nous avons émis l'hypothèse, en accord avec une autre étude (Duroux et Welinder, 2003), que les peroxydases de classe III seraient apparues lors de la colonisation de la terre ferme par les plantes. Ce travail a fait l'objet de la publication suivante :

Passardi F, Longet D, Penel C, Dunand C.
The class III peroxidase multigenic family in rice and its evolution in land plants.
Phytochemistry. 2004 Jul; 65(13):1879-93.

Du point de vue taxonomique, il est reconnu que les plus proches représentants des ancêtres des plantes terrestres font partie de l'ordre des Charales (McCourt et al., 2004). A ce jour, aucune donnée n'existe cependant sur des séquences génomiques ou d'ADNc chez les Charales, ni chez la plupart des algues vertes de l'embranchement des Streptophytes (Fig. 2). Des tests d'oxydation du gaïacol (activité typique des peroxydases de classe III), m'ont confirmé que les peroxydases de classe III, ou du moins toute trace de leur activité, étaient absentes des algues vertes (Chlorophytes et Streptophytes). Deux exceptions sont apparues : une très faible activité a été détectée chez une Coleochaetale (*Coleochaete scutata*) et une activité significative chez une Charale (*Chara zeylanica*). Suite aux essais non concluants d'une approche génomique, j'ai essayé d'isoler une protéine de peroxydase chez *Chara*

contraria, une espèce relativement commune vivant dans le lac Léman (Auderset Joye et al., 2002). Les premiers essais ont montré que cette algue comprend en fait plusieurs peroxydases, et des purifications ultérieures m'ont permis d'isoler un isoforme acide. Malgré cette découverte intéressante, très récemment une petite banque d'ADNc d'une algue encore plus primitive, *Closterium peracerosum-strigosum-littorale* (Zygnematophycées), a été rendue publique : une recherche BLAST m'a permis d'identifier avec certitude un fragment de séquence de peroxydase. Les peroxydases sont donc apparues plus tôt que prévu et, même si on ne peut pas exclure un rôle dans la colonisation de la terre ferme, leur apparition n'a pas coïncidé avec la sortie des algues hors de l'eau.

Parallèlement aux recherches effectuées sur le génome du riz et sur différents organismes clés du règne végétal, la quantité de données sur les séquences de peroxydases a rapidement augmenté. J'ai ainsi (avec Christophe Dunand) créé une base de données visant à regrouper les séquences de peroxydases et de pouvoir ainsi les comparer de manière rapide et efficace. Afin de partager nos informations, nous avons établi une collaboration avec l'Institut Suisse de Bioinformatique (Vassilios Ioannidis, Laurent Falquet) et rendu notre base de données, baptisée « PeroxiBase », publique sur internet (http://peroxidase.isb-sib.ch). Suite à l'engagement d'une personne à plein temps (Nenad Bakalovic) pour entrer des données, la PeroxiBase a pris une taille considérable (plus de 2000 séquences), et nous avons ainsi validé notre travail par la publication suivante :

Bakalovic N, **Passardi F (premier co-auteur)**, Ioannidis V, Cosio C, Penel C, Falquet L, Dunand C.
PeroxiBase: a class III plant peroxidase database.
Phytochemistry. 2006 Mar; 67(6):534-9.

La PeroxiBase n'a depuis lors jamais cessé de s'agrandir et de s'améliorer avec, en particulier :

- l'introduction des peroxydases de la superfamille
- l'établissement d'une nomenclature unique
- l'addition d'outils bioinformatiques (BLAST, Finger Print Scan)
- la création de liens (« cross-references ») entre SwissProt et la PeroxiBase
- l'extension de la PeroxiBase à tous les êtres vivants et toutes les peroxydases (hème et non-hème)

J'ai officiellement présenté la PeroxiBase au Japon (Fukuoka) en 2005, lors du VII^{e} symposium international des peroxydases. De nombreux scientifiques se sont depuis joints à ce projet, et nous comptons actuellement trois collaborations principales avec des groupes en Slovaquie, Brésil et en France. Le potentiel de cette base de données est très prometteur, et permettra, nous l'espérons, de faciliter les travaux de recherche en bioinformatique, phylogénétique, biologie moléculaire, cristallographie et biotechnologies par la concentration de données fiables et précises en une seule entité.

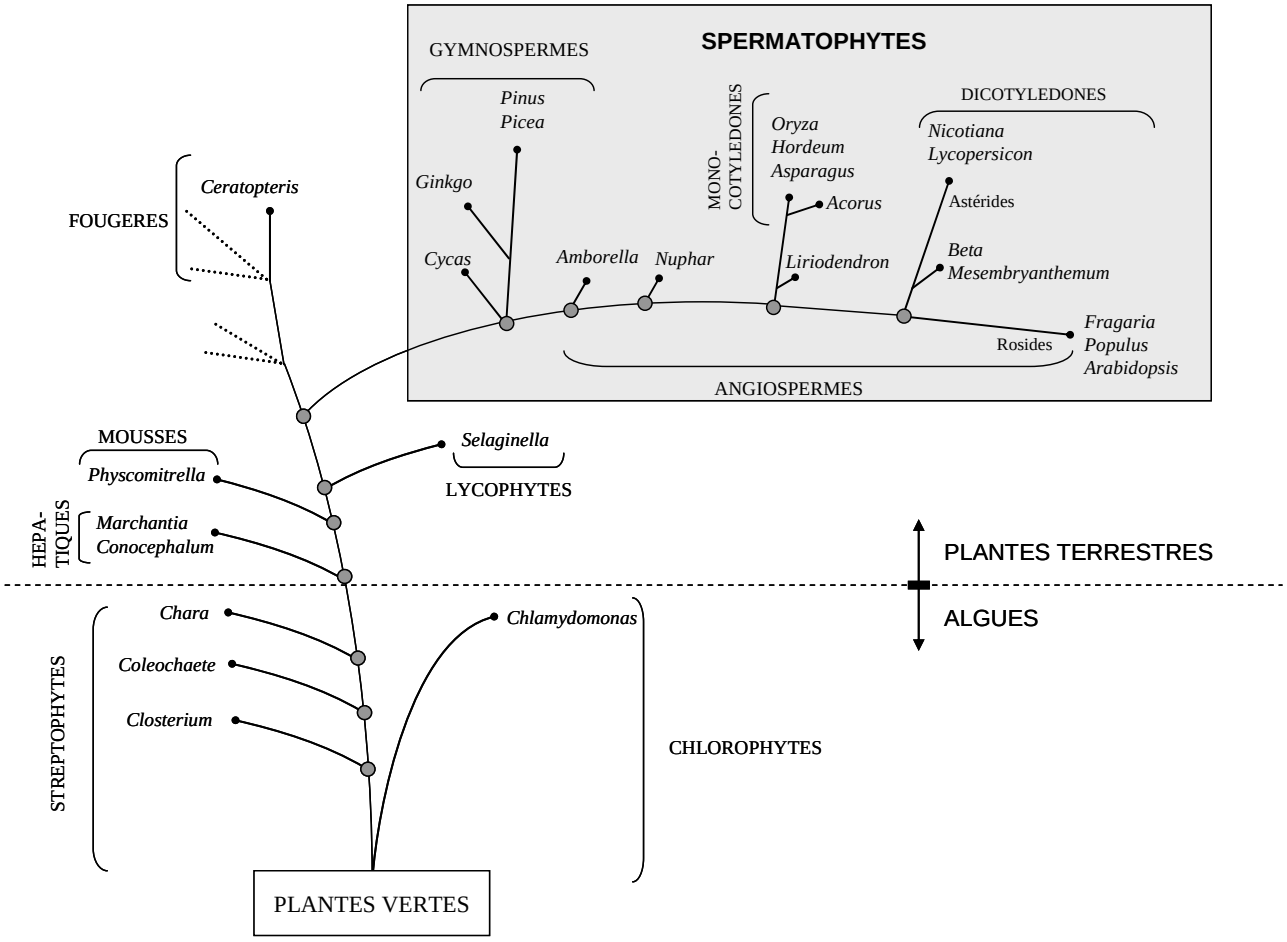


Figure 2

Taxonomie simplifiée des plantes vertes (Viridiplantae).

Quelques genres d'espèces décrites dans ce résumé et dans la thèse sont représentés sur l'arbre.

4. Phylogénie et étude du promoteur de la peroxydase Atprx42

Les séquences peptidiques des peroxydases de classe III sont en général bien conservées, en particulier dans les régions autour du site actif, mais aussi dans d'autres parties dont la fonction est moins bien définie. Ce degré de conservation élevé pose un problème majeur quant à l'identification de fonctions précises pour chaque peroxydase, comme énoncé précédemment. Il est en fait vraisemblable que plusieurs peroxydases remplissent des fonctions identiques. Alors pourquoi y en a-t-il autant dans chaque plante ? Un élément de réponse semble venir de leurs séquences promotrices : aucune homologie n'a été observée entre les promoteurs de peroxydases, y compris ceux appartenant à des isoformes extrêmement proches, tels qu'AtPrx33 et AtPrx34. Un profil d'expression différent, et ainsi une localisation spatio-temporelle différente, expliquerait et justifierait l'apparente redondance fonctionnelle des peroxydases. Dans notre laboratoire, Luisa Valério a d'ailleurs confirmé, en utilisant la technique du « macroarray », un profil d'expression très varié pour les 73 peroxydases d'*Arabidopsis* (Valério et al., 2004). Une peroxydase semble toutefois avoir une expression forte et stable, quels que soient l'organe considéré et l'âge de la plante : *AtPrx42*. Cette peroxydase est d'ailleurs utilisée dans notre laboratoire comme contrôle lors de réactions de RT-PCR. Luisa Valério l'a remarqué dans son étude, ainsi que d'autres expérimentateurs indépendants (Kjaersgard et al., 1997; Welinder et al., 2002).

Suite au nombre important d'entrées dans la PeroxiBase et au fait que les groupes principaux du règne végétal y étaient représentés, j'ai entrepris la recherche de protéines paralogues à la peroxidase AtPrx42 chez tous les organismes disponibles. Les résultats sont surprenants et démontrent que cette peroxydase est non seulement présente chez tous les spermatophytes (plantes à graine), à l'exception des céréales, mais que son niveau d'expression est toujours aussi fort : il s'agit, dans chaque organisme, de la peroxydase la plus exprimée. La disparition de cette peroxydase chez les céréales reste un mystère. Il semblerait que cette absence ne soit pas due à l'influence de l'homme, puisque des céréales sauvages (*Aegilops speltoides*, *Triticum monococcum*), ainsi que des plantes de groupes proches des céréales, ont aussi perdu la protéine paralogue de la peroxydase AtPrx42. L'absence de paralogues dans les plantes cryptogames (hépatiques, mousses, fougères) indique que cette peroxydase serait liée à l'apparition de la graine. D'ailleurs, dans notre laboratoire, de nombreux collaborateurs ont essayé d'isoler des mutants ayant perdu l'expression d'AtPrx42, mais sans succès. Si le rôle de cette peroxydase est, entre autres, de permettre la formation de la graine ou sa germination, alors il est probable que son élimination soit létale.

Suite à l'étude phylogénétique d'AtPrx42, j'ai essayé de comprendre comment cette peroxydase pouvait avoir une expression aussi forte et constitutive. L'analyse des promoteurs des paralogues chez d'autres plantes a permis de trouver, pour la première fois chez des séquences régulatrices de peroxydases, un motif commun à tous : TTATTTAAGxAG. En dehors de ce motif, aucune autre homologie n'a été trouvée dans les promoteurs étudiés. La séquence TTATTTAAGxAG n'a jamais été décrite jusqu'à présent. Elle s'apparente à une « TATA box » dans sa partie 5', mais la conservation de la partie 3' est remarquable. Afin de mieux comprendre la fonction de ce motif, j'ai construit des mutants de délétion, ainsi qu'un mutant comportant plusieurs mutations ponctuelles, du promoteur d'*Atprx42* fusionné au gène rapporteur *GUS*. Les plantes transgéniques sont sur le point d'être obtenues, et les premiers résultats devraient être disponibles prochainement. Les expériences futures permettront de détailler le profil d'expression de cette peroxydase, en particulier dans la graine. De plus, si la séquence TTATTTAAGxAG est véritablement responsable d'une expression constitutive, il s'agira de la première isolation chez une plante d'un tel motif, ce qui pourrait conduire à de nombreuses applications passionnantes.

5. Au-delà des classes : la superfamille

L'introduction de séquences de peroxydases appartenant à la superfamille dite des « plantes, champignons et bactéries » (Welinder, 1992) m'a conduit, avec Christophe Dunand et d'autres collaborateurs externes, à effectuer des travaux de recherche phylogénétiques étendus à d'autres peroxydases que celles de classe III. Mon premier travail a consisté à étudier la distribution des catalases-peroxydases chez tous les procaryotes dont le génome était séquencé (environ 500 organismes disponibles sur NCBI). Les résultats ont été récemment soumis à un journal scientifique. En particulier, ils démontrent que les catalases-peroxydases sont réparties de manière assez chaotique et imprévisible chez les bactéries, et que les bactéries s'échangeraient fréquemment ces peroxydases entre elles par transfert horizontal, même si elles appartiennent à des groupes taxonomiques très différents. Il en résulte une distribution phylogénétique qui suit rarement la distribution taxonomique des organismes concernés, au contraire des peroxydases de classe III.

Une deuxième étude a été effectuée sur la superfamille entière des peroxydases, avec pour but de mieux comprendre l'apparition et l'évolution de la superfamille chez tous les organismes vivants. Les résultats de ce travail ont été récemment envoyés à un journal

scientifique. Les conclusions principales indiquent que les premières peroxydases seraient apparues dans les ancêtres des mitochondries sous forme de cytochrome c peroxydase. La CCP originelle aurait ensuite évolué en APX chez les ancêtres des chloroplastes. Quant aux catalases-peroxidases, elles seraient apparues très tôt également chez un organisme procaryote primitif sous une forme dupliquée, ce qui explique l'absence d'une forme non-dupliquée dans les êtres vivants actuels. Cette dernière conclusion s'oppose à de précédents travaux (Klotz et Loewen, 2003) qui préconisaient que les CP (sous forme dupliquée) étaient à l'origine des APX et des CCP. Dans cette même étude, les classes II et III sont clairement des « produits dérivés » de peroxydases de classe I, respectivement une CCP et une APX. Enfin, la découverte de nombreuses peroxydases « inclassables » chez des protistes suggère aussi qu'il sera bientôt nécessaire de décrire une quatrième classe de peroxydases, ce qui constituerait une « révolution » au sein de la superfamille.

6. Conclusions

Les études effectuées à l'aide de plusieurs plantes transgéniques ont permis de déterminer que les deux proches orthologues AtPrx33 et AtPrx34 ont un rôle *in vivo* dans l'élongation cellulaire de la racine principale. AtPrx33 participe aussi à la différenciation des parties aériennes d'*Arabidopsis*, alors que pour AtPrx34, les conclusions indiquent qu'elle ne joue aucune fonction particulière. Cette dernière différence est probablement due à deux séquences promotrices très différentes, qui contrastent avec le très fort degré d'identité (92%) des séquences peptidiques. L'étude des promoteurs de peroxydases est probablement une priorité à l'avenir, qui nous permettra de mieux comprendre pourquoi ces enzymes forment des familles multigéniques pouvant atteindre plus de 100 isoformes (chez le riz, par exemple), alors qu'elles ont toutes *in vitro* des activités très similaires. Mon travail de thèse a montré que la fonction de chaque peroxydase *in vivo* peut être identifiée par mutagénèse et transgénèse. En revanche, le mécanisme d'action est bien plus difficile à caractériser, comme j'ai pu le constater avec les expériences sur les plantes GH3 ::*GUS*. Ce travail indique aussi qu'il est nécessaire, à l'avenir, de développer des techniques plus pointues que celles disponibles actuellement, telles que le dosage *in situ* de molécules, pour pouvoir enfin décrire le mécanisme d'action précis d'une peroxydase.

La création de la PeroxiBase a contribué à une meilleure compréhension de l'évolution des peroxydases et a mis en place un outil de recherche qui, nous l'espérons, facilitera

l’orientation de futurs projets destinés à trouver la fonction spécifique de chaque peroxydase. Les recherches phylogénétiques ont ainsi permis d’établir de nouvelles stratégies en laboratoire, comme l’a montré le travail sur la peroxydase AtPrx42. Dans le futur, grâce à l’apport constant de nouvelles données sur de nombreux organismes, des études phylogénétiques sur des familles multigéniques entières permettront probablement d’identifier des groupes de peroxydases spécifiques à des organismes et, idéalement, à certaines structures clé telles que les graines, les racines ou les tissus vasculaires. Cette approche pourrait remplacer des stratégies de biologie moléculaire en particulier dans des organismes où aucun protocole de transformation n’a encore été mis en place.

Enfin, certains organismes ne possèdent aucune peroxydase, comme il a été montré chez de nombreuses bactéries. Un projet envisageable serait d’introduire des peroxydases dans de telles espèces, afin de « mimer » des étapes qui auraient pu se produire au cours de l’évolution. Par exemple, les bactéries du genre *Clostridium* sont incapables de proliférer dans des conditions aérobies. En supposant que les catalases-peroxidases sont apparues pour éliminer les excès d’espèces actives de l’oxygène, on pourrait envisager de transformer *Clostridium* avec une CP. De la même manière, il serait intéressant de transformer des Chlorophytes, telles que *Chlamydomonas*, avec une peroxydase de classe III (de plante terrestre primitive ou d’algue streptophyte) et de voir quels seraient les changements induits par son expression. L’absence de peroxydases dans certains organismes est ainsi une chance d’approcher la problématique de la fonction des peroxydases au cours de l’évolution, une problématique qui, malgré les progrès obtenus jusqu’à présent, demeure une énigme encore non résolue.

7. Bibliographie

Auderset Joye D, Castella E, Lachavanne JB (2002) Occurrence of Characeae in Switzerland over the last two centuries (1800-2000). Aquatic Botany **72:** 369-385

Berglund GI, Carlsson GH, Smith AT, Szoke H, Henriksen A, et al. (2002) The catalytic pathway of horseradish peroxidase at high resolution. Nature **417:** 463-468

Cary A, Uttamchandani SJ, Smets R, Van Onckelen HA, Howell SH (2001) *Arabidopsis* mutants with increased organ regeneration in tissue culture are more competent to respond to hormonal signals. Planta **213:** 700-707

Dunand C, Tognolli M, Overney S, von Tobel L, De Meyer M, et al. (2002) Identification and characterisation of Ca^{2+}-pectate binding peroxidases in *Arabidopsis thaliana*. J Plant Physiol **159:** 1165-1171

Duroux L, Welinder KG (2003) The peroxidase gene family in plants: a phylogenetic overview. J Mol Evol **57:** 397-407

Edwards SL, Raag R, Wariishi H, Gold MH, Poulos TL (1993) Crystal structure of lignin peroxidase. Proc Natl Acad Sci U S A **90:** 750-754

Erman JE, Vitello LB (2002) Yeast cytochrome c peroxidase: mechanistic studies via protein engineering. Biochim Biophys Acta **1597:** 193-220

Gold MH, Alic M (1993) Molecular biology of the lignin-degrading basidiomycete Phanerochaete chrysosporium. Microbiol Rev **57:** 605-622

Kjaersgard IV, Jespersen HM, Rasmussen SK, Welinder KG (1997) Sequence and RT-PCR expression analysis of two peroxidases from *Arabidopsis thaliana* belonging to a novel evolutionary branch of plant peroxidases. Plant Mol Biol **33:** 699-708

Klotz MG, Loewen PC (2003) The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota. Mol Biol Evol **20:** 1098-1112

Liszkay A, Kenk B, Schopfer P (2003) Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth. Planta **217:** 658-667

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends Ecol Evol **19:** 661-666

Obinger C, Regelsberger G, Furtmuller PG, Jakopitsch C, Ruker F, et al. (1999) Catalase-peroxidases in cyanobacteria--similarities and differences to ascorbate peroxidases. Free Radic Res **31 Suppl:** S243-249

Ricard J, Nari J (1966) Contribution to the study of mechanisms of indolylacetic acid degradation by horseradish peroxidase. Biochim Biophys Acta **113:** 57-70

Ruiz-Dueñas FJ, Camarero S, Perez-Boada M, Martinez MJ, Martinez AT (2001) A new versatile peroxidase from *Pleurotus*. Biochem Soc Trans **29:** 116-122

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, et al. (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot **53:** 1305-1319

Skoog F, Miller CO (1957) Chemical regulation of growth and organ formation in plant tissues cultured *in vitro*. Symp Soc Exp Biol **54:** 118-130

Skulachev VP (1998) Cytochrome c in the apoptotic and antioxidant cascades. FEBS Lett **423:** 275-280

Smulevich G, Jakopitsch C, Droghetti E, Obinger C (2006) Probing the structure and bifunctionality of catalase-peroxidase (KatG). J Inorg Biochem **100:** 568-585

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene **288:** 129-138

Valério L, De Meyer M, Penel C, Dunand C (2004) Expression analysis of the *Arabidopsis* peroxidase multigenic family. Phytochemistry **65:** 1331-1342

Welinder KG (1992) Superfamily of plant, fungal and bacterial peroxidases. Curr Opin Struct Biol **2:** 388-393

Welinder KG, Justesen AF, Kjaersgard IV, Jensen RB, Rasmussen SK, et al. (2002) Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. Eur J Biochem **269:** 6063-6081

Yamazaki H, Yamazaki I (1973) The reaction between indole 3-acetic acid and horseradish peroxidase. Arch Biochem Biophys **154:** 147-159

INDEX

1. General introduction

2. Functional study of AtPrx33 and AtPrx34

3. Class III peroxidases in other plants

4. AtPrx42 phylogeny and promoter studies

5. Beyond classes: the peroxidase superfamily

6. Conclusions

7. Literature

Appendix

1

General introduction

A. The peroxidase superfamily

The iron form Fe(III) is probably the first electron acceptor, and hence the first source of energy storage and redistribution, incorporated and utilised by the most ancient living organisms (Vargas et al., 1998). Fe(III) is closely associated with porphyrin, which is even thought to have appeared before any life form on Earth (Aylward and Bofinger, 2005, 2006). The combination of the metal and the organic ring, together with evolution of proteins, gave rise to a wide variety of hemoproteins, such as cytochromes, globins, catalases and peroxidases. This latter category comprises different forms of hydrogen peroxide reducing proteins, with very different donor substrates (EC classification 1.11.1.x), and was clearly separated into three classes by Karen Welinder, in 1992, which formed the so-called "superfamily of plant, fungal and bacterial peroxidases" (Welinder, 1992). Peroxidases all share a heme formed by protoporphyrin IX and Fe(III); their three-dimensional structure is very similar (Welinder, 1992; Edwards et al., 1993; Smulevich et al., 2006), despite low identity in the primary amino acid sequence. Classes II and III both have four disulphide bridges, but not at the same sites (Welinder et al., 2002), and two calcium ions; they are glycosylated and bear a signal peptide for extracellular (through endoplasmic reticulum) secretion. Whereas class III is quite homogeneous and regroups plant peroxidases, class II is divided into manganese peroxidases, lignin peroxidases and versatile peroxidases (Ruiz-Dueñas et al., 2001); class II peroxidases are exclusively encoded by fungi. Class I peroxidases lack calcium ions and disulphide bridges; they are not glycosylated and do not have a signal peptide. The class I is split into three distinct groups: cytochrome c peroxidases (CCP), ascorbate peroxidases (APX) and catalase-peroxidases (CP). The latter group probably evolved from gene duplication, with an active moiety (N-terminal) and a second moiety without any defined function (Yamada et al., 2002). Class I peroxidases can be found in most of living organisms (except animals). Their widespread distribution, and particularly their presence in prokaryotes, indicate that class I peroxidases are at the origin of the two other classes.

B. Three classes, but multiple functions

Whereas they share the same general structure, the three classes of peroxidases have distinct functions and reaction mechanisms. Class I peroxidases have quite precise roles: APX have a very high affinity for ascorbate; they are generally found in photosynthetic organisms. In higher plants, they are subdivided according to their subcellular localisation: chloroplastic (stroma and thylakoid-bound), peroxisomal and cytoplasmic (Teixeira et al., 2004). Their main function in the cell is detoxification of excess H_2O_2 (Shigeoka et al., 2002). Regarding CCP, they are located in the mitochondrial intermembrane space: like APX, they scavenge excess H_2O_2, and (indirectly) superoxide by oxidising cytochrome c (Skulachev, 1998; Erman and Vitello, 2002). Curiously, removal of CCP from *Saccharomyces cerevisiae* did not result in a significant degradation of cell viability or respiration rate, although cells were more sensitive to H_2O_2 (Kwon et al., 2003). Catalase-peroxidases, similarly to APX and CCP, also scavenge H_2O_2. They can oxidise another molecule of H_2O_2 (catalase activity) or many other substrates (Obinger et al., 1999).

Class II peroxidases have a major role in degradation of soil debris, particularly lignin (Piontek et al., 2001; Martinez et al., 2005). No other haem peroxidase is able to degrade lignin. Due to their unique capacity to oxidise high redox potential molecules, class II peroxidases are also used for bioremediation of various pollutants, such as industrial dyes, xenoestrogens or antifouling compounds, as well as for chlorine-free bleaching of pulp (Katagiri et al., 1995; Tsutsumi et al., 2001; Wesenberg et al., 2003; Ogawa et al., 2004).

In striking contrast with the two first peroxidase classes, assigning a function to a class III peroxidase is a rather complex task. Class III peroxidases form multigenic families in plants, and within one organ or tissue, several isoforms are expressed at the same time. *In vitro*, any isoform is able to oxidise an innumerable amount of different phenolic substrates, and cannot hence give precise information on *in vivo* substrate specificity. Nevertheless, particular functions can be assigned to groups of peroxidases, although also in that case, they are very diverse and probably not all known. Regarding single peroxidases, their expression profile has been extensively characterised in specific organs and in response to various internal and external factors. Following accumulation of very diverse data on the many facets of class III peroxidases, several reviews were published. Particularly, one review focused mainly on application studies of plants transformed with engineered peroxidase genes (Yoshida et al., 2003); another recent work consisted of an historical overview of horseradish

peroxidase isoenzymes discovery, characterisation and properties (Veitch, 2004). Two other reviews discussed the multiple roles of peroxidases (Hiraga et al., 2001) and their relationship with reactive oxygen species (ROS) (Kawano, 2003).

The characterisation of a second peroxidase cycle involving formation of compound III (hydroxylic cycle) (Chen and Schopfer, 1999; Berglund et al., 2002; Liszkay et al., 2003) led us (Christophe, Claude and myself) to write the review "**Performing the paradoxical: how plant peroxidases modify the cell wall**". This is the first review about the interrelationships between ROS and class III peroxidases in light of the recent works on the hydroxylic cycle. A previous study (Kawano, 2003) also discussed roles of ROS production by peroxidases, but focused on signalling pathways. We did not discuss this last aspect but rather concentrated on influence of ROS generation and consumption on cell wall remodelling during a plant's life.

Concerning the role of class III peroxidases, the fast amount of studies recently published prompted us (Claudia, Christophe, Claude and myself) to perform an update on the functional importance of peroxidases. In particular, we wanted to highlight the implication of peroxidases during the whole lifespan of a plant, from the first hours of germination until senescence. We hence undertook the writing of a second review entitled "**Peroxidases have more functions than a Swiss army knife**".

The two articles mentioned here above are cited as follows:

Passardi F, Cosio C, Penel C, Dunand C.
Peroxidases have more functions than a Swiss army knife.
Plant Cell Rep. 2005 Jul; 24(5):255-65.
DOI: http://dx.doi.org/10.1007/s00299-005-0972-6

Passardi F, Penel C, Dunand C.
Performing the paradoxical: how plant peroxidases modify the cell wall.
Trends Plant Sci. 2004 Nov; 9(11):534-40.
DOI: http://dx.doi.org/10.1016/j.tplants.2004.09.002

Plant Cell Rep (2005) 24: 255–265
DOI 10.1007/s00299-005-0972-6

REVIEW

F. Passardi · C. Cosio · C. Penel · C. Dunand

Peroxidases have more functions than a Swiss army knife

Received: 15 December 2004 / Revised: 21 March 2005 / Accepted: 22 March 2005 / Published online: 22 April 2005

Abstract Plant peroxidases (class III peroxidases) are present in all land plants. They are members of a large multigenic family. Probably due to this high number of isoforms, and to a very heterogeneous regulation of their expression, plant peroxidases are involved in a broad range of physiological processes all along the plant life cycle. Due to two possible catalytic cycles, peroxidative and hydroxylic, peroxidases can generate reactive oxygen species (ROS) ($^{\bullet}OH$, $HOO^{\bullet}$), polymerise cell wall compounds, and regulate H_2O_2 levels. By modulating their activity and expression following internal and external stimuli, peroxidases are prevalent at every stage of plant growth, including the demands that the plant meets in stressful conditions. These multifunctional enzymes can build a rigid wall or produce ROS to make it more flexible; they can prevent biological and chemical attacks by raising physical barriers or by counterattacking with a large production of ROS; they can be involved in a more peaceful symbiosis. They are finally present from the first hours of a plant's life until its last moments. Although some functions look paradoxical, the whole process is probably regulated by a fine-tuning that has yet to be elucidated. This review will discuss the factors that can influence this delicate balance.

Keywords Evolution · ROS · (abiotic and biotic) stress · Cell wall loosening and cross-linking · Senescence · Fruit ripening · Symbiosis

Multigenic family, evolution and homology

Heme peroxidases specific to plants belong to a superfamily that contains three different classes of peroxidases (Welinder 1992b): the intracellular class I (EC 1.11.1.5/.6/.11), the class II released by fungi (EC 1.11.1.13/.14), and the secreted class III plant peroxidases (EC 1.11.1.7).

In their regular catalytic cycle, class III plant peroxidases (Welinder 1992a) catalyse the reduction of H_2O_2 by taking electrons to various donor molecules such as phenolic compounds, lignin precursors, auxin, or secondary metabolites (Hiraga et al. 2001). Their encoding genes form large multigenic families (Tognolli et al. 2002; Duroux and Welinder 2003; Passardi et al. 2004a) that are present in all land plants, but were not detected in the unicellular green algae (Passardi et al. 2004a). From the appearance of the first class III peroxidase, probably around the emergence of the terrestrial plants 450 MY ago, to the most evolved plants, the number of gene copies has largely increased. This increase seems to be correlated with the evolution of the plant architecture and complexity, as well as with the diversity of the biotopes and the pathogens.

The class III peroxidases gene structure, as well as key amino acid residues and protein size, are highly conserved between orthologs and paralogs. In spite of this conservation, their isoelectric points largely differ (anionic and cationic forms). Until now, no correlation has been established between their pI and their putative enzymatic function (Welinder 1992a).

The elevated number of paralogs found for example in *Arabidopsis thaliana* (Tognolli et al. 2002) and in *Oryza sativa* (Passardi et al. 2004a) can be related to the high duplication and conservation rate of the peroxidase genes. The conservation of duplicated genes can be explained by the acquisition of either a new expression profile (subfunctionalisation) or a novel function (neofunctionalisation). These two acquisitions can explain the rationale for the existence of multigenic families. Additionally, the diversity of the reactions catalysed by plant peroxidases accounts for the implication of these proteins in a broad range of physiological processes (Penel et al. 1992; Hiraga et al. 2001). Furthermore, the recent description of a separate hydroxylic cycle, which leads to the formation of various radical species, opens a new range of implications for

Communicated by P. Kumar

F. Passardi · C. Cosio · C. Penel · C. Dunand (✉)
Laboratory of Plant Physiology, University of Geneva,
Quai Ernest-Ansermet 30,
CH-1211 Geneva 4, Switzerland
e-mail: christophe.dunand@bota.unige.ch
Tel.: +41-22-379-3012
Fax: +41-22-379-3017

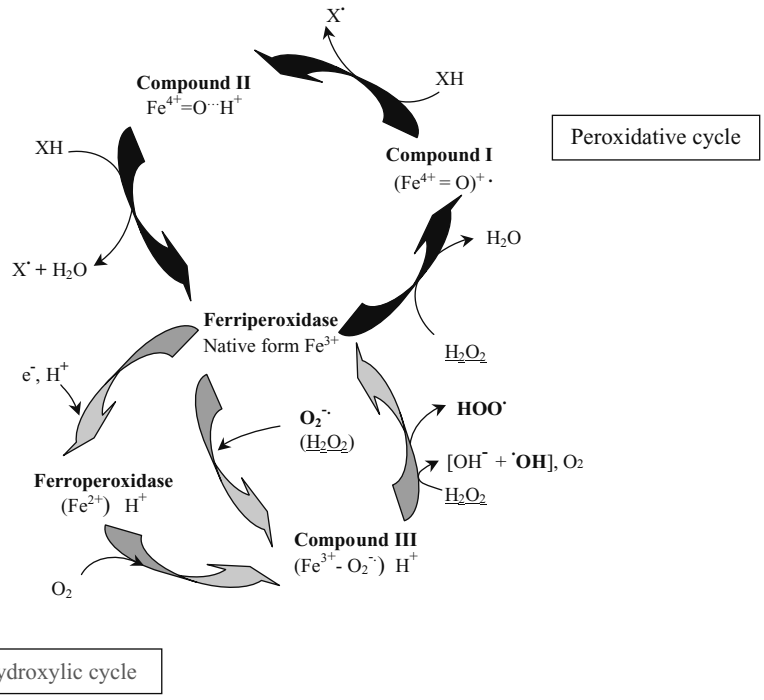


Fig. 1 Class III plant peroxidase cycles. The hydroxylic cycle (with grey arrows) can regulate the H_2O_2 level and release ROS (•OH, HOO•). The peroxidative cycle (with black arrows) can oxidise various substrates (XH) and release their oxidised form (X)

the peroxidases (Fig. 1) (Liszkay et al. 2003; Passardi et al. 2004b). Indeed, the reactive oxygen species (ROS) are considered as very reactive compounds, which can possess an intrinsic activity, or can act as part of signal transduction pathways.

Due to their implication in a broad range of physiological processes such as auxin metabolism, lignin and suberin formation, cross-linking of cell wall components, defence against pathogens or cell elongation, peroxidase activity can be easily detected in the whole lifespan of various plants: from the early stage of germination to the final step of senescence, through the control of cell elongation, defence mechanisms, and several other roles (Fig. 2).

Germination

Release of peroxidases and ROS during germination in the medium surrounding the seed has been reported in radish (*Raphanus sativus*) (Scialabba et al. 2002). To our knowledge, no in vivo direct proof has demonstrated the role of peroxidases in the generation of H_2O_2 during germination. Peroxidases have the capacity to generate H_2O_2 and subsequently •OH radicals via the hydroxylic cycle. Reactive oxygen species released during this cycle could play a defence role for the seed against pathogens, although they are also secreted in absence of pathogenic organisms (Scialabba et al. 2002). Class III peroxidases would then play a critical role in the very first days of life of the

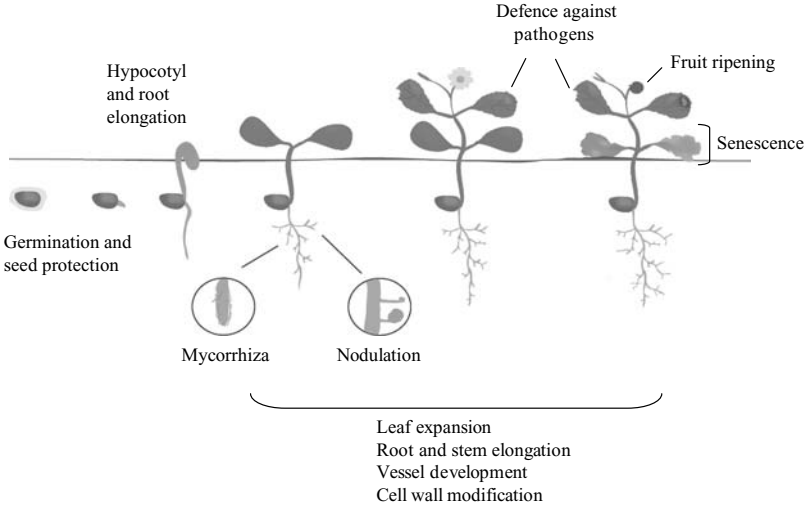


Fig. 2 From germination to senescence: implication of the Class III peroxidases through the whole plant lifespan

germinating seed by defending against pathogenic attacks and by cleaving the cell wall compounds around the radicle protrusion area. In agreement with this hypothesis, a study on tomato (*Lycopersicon esculentum*) seeds demonstrated that peroxidase genes start to be expressed as soon as the germination begins (Morohashi 2002). This event breaks down the endosperm, thus creating a wounded region vulnerable to pathogens: a burst of $^{\bullet}OH$ radicals at this precise time would certainly confer a great protective advantage and help the protrusion.

Cellular growth and cell wall loosening

Growth and elongation are related to cell extensibility, which occurs with a loosening of the cross-linking of cell wall compounds. Several enzymatic mechanisms are involved in cell wall loosening: the cleavage and the reassembly of xyloglucan polymers by xyloglucan endotransglycosylase or the suppression of hydrogen bonds between cellulose and xyloglucan by expansins (Cosgrove 2001). In addition, peroxidases, through their catalytic and hydroxylic cycles, regulate directly or indirectly the cell wall architecture.

The endogenous hydrogen peroxide (H_2O_2) level can be related to the elongation process. For example, in the Dicotyledons, such as soybean, the apoplastic H_2O_2 level is lower in the hypocotyl elongation zone (Schopfer 1994). In *A. thaliana*, the elongation during root curvature is also regulated by a variation in H_2O_2 concentration (Joo et al. 2001). On the contrary, in Monocotyledons such as onion, the location of H_2O_2 in the cell wall has been correlated with lignification and elongation during growth (Cordoba-Pedregosa et al. 2003). Peroxidases could participate in the control of the H_2O_2 level and therefore be directly related to the control of the elongation process. For example, transcripts of APRX, an anionic peroxidase from zucchini, accumulated strongly in the elongation zone of the hypocotyl and their accumulations were inversely correlated with lignin level (Dunand et al. 2003). Their specific localisations could be related to the high elongation rate observed in this region.

Besides H_2O_2, ascorbate has been shown to control elongation and expansion processes via the inhibition of enzymes involved in the cell wall stiffening. For example, the activity of apoplastic and cell wall isolated peroxidase involved in root elongation control was inhibited in presence of ascorbate (Cordoba-Pedregosa et al. 1996). Considering the peroxidase-driven cross-linking reaction, ascorbate treatment stimulates root elongation through the peroxidase inhibition. The same effect was observed during the mechanism of gravitropism in maize. Ascorbate can reduce the root curvature, which is directly related to the elongation process (differential elongation) (Schopfer 1994).

Other compounds such as hydroxyl radical ($^{\bullet}OH$) could be involved in the processes of elongation and expansion. This highly reactive radical produced by the Fenton reaction is capable of cleaving cell wall polysaccharides such as pectin and xyloglucan (Fry 1998). The production of $^{\bullet}OH$ near the cell wall could be related to non-enzymatic wall loosening mechanisms (Chen and Schopfer 1999). Hydroxyl radicals can be produced at the cell wall level from superoxide radical ($O_2^{\cdot -}$) and hydrogen peroxide (H_2O_2) by peroxidase (Schweikert et al. 2000; Schopfer et al. 2002; Liszkay et al. 2003). Auxin is also involved in this process, since it promotes the release of $O_2^{\cdot -}$ and production of $^{\bullet}OH$ when inducing elongation growth (Schopfer et al. 2002).

The analysis of these results suggests that peroxidases can regulate the growth elongation in different manners due to their two cycles and the various numbers of isoforms. They can be controlled by ascorbate and they can also induce the elongation process by generating $^{\bullet}OH$ (hydroxylic cycle) or by regulating the local concentration of H_2O_2 (owing to both cycles). However, peroxidases are not confined to this single elongation role: they can also physically inhibit the elongation process by creating cell wall cross-linking (peroxidative cycle).

Cell wall cross-linking

Peroxidases can create a physical barrier by catalysing cross-linking of cell wall compounds in response to different stimuli such as wounding, pathogen interactions, and climatic aggressions. This process also occurs simply as a normal cell wall evolution during the growth and senescence. The implication of the peroxidases in cross-linking is related to the oxidative capacity of the peroxidative cycle (Passardi et al. 2004b).

Lignification and suberisation

Cross-linking of the phenolic monomers in the formation of suberin and the oxidative coupling of lignin subunits has been associated with reduction of cell extensibility and growth. Peroxidases (Lewis and Yamamoto 1990) and laccases (EC 1.10.3.2) (Kiefer-Meyer et al. 1996) are both candidates for lignin units assembly by oxidative polymerisation. Using H_2O_2 as oxidant, peroxidases can generate monolignol phenoxy radicals that spontaneously form lignin polymers (Lewis and Yamamoto 1990). Down-regulation of the tobacco peroxidase TP60 and of the aspen anionic peroxidase prxA3a expressions lead to the reduction of lignin content (Blee et al. 2003; Li et al. 2003). Other peroxidases such as HRP A2 from horseradish, AtPrx21 (ATP2) from *A. thaliana*, various anionic peroxidases from poplar, and cationic peroxidases from tomato (TPX1) and *Zinnia* have also been shown to be involved in the lignification process (Ostergaard et al. 2000; Quiroga et al. 2000; Christensen et al. 2001; Nielsen et al. 2001; Lopez-Serrano et al. 2004).

The in vitro oxidation of lignin monomers by particular peroxidases and their specific localisation in lignifying plant tissues both suggest their role in the lignin biosynthetic pathways. Peroxidases are also involved in suberin formation through a similar process (Bernards et al. 1999; Keren-Keiserman et al. 2004).

Diferulic bonds and extensin network

The lignification step is not limited to the assembly of phenolic units. The cell wall network is further stiffened by association of monolignols to polysaccharides and by internal cross-linking of polysaccharides (mainly cellulose and pectin) through formation of diferulate bonds (Iiyama et al. 1994). The relationship between formation of diferulates and peroxidases is well-known in vitro. Different *Pinus elliotti* peroxidase fractions (soluble, ionically- and covalently-bound) were isolated from the cytoplasm and the wall of callus or cambial cells: only the cell wall fraction of cambium exhibited a ferulic acid oxidation activity (Whitmore 1976). Further studies confirmed this fact (Sanchez et al. 1996; MacAdam and Grabber 2002) and additionally showed that the proportion of different diferulic isomers varied according to the plant species (Ralph et al. 1994). Growth and cell wall extensibility are therefore inversely correlated with the increase in the content of diferulic acids in the cell wall (Sanchez et al. 1996; Kawamura et al. 2000; MacAdam and Grabber 2002).

Peroxidases are also major players in the building up of a dense extensin network destined to rigidify the cell wall. Peroxidases devoted to extensin cross-linking have been isolated from a large variety of plant species (Schnabelrauch et al. 1996; Magliano and Casal 1998; Jackson et al. 2001); they are known to act on the phenolic moiety of tyrosine, and maybe also on lysines, hence creating Tyr–Tyr or Tyr–Lys bonds (Schnabelrauch et al. 1996). Tyrosines and lysines are evenly spaced on extensins within conserved motifs, hence contributing to the formation of a very uniform mesh within the cell wall structure (Kieliszewski and Lamport 1994). Pectin is supposed to be involved in bringing together peroxidases and extensins, due to the affinity of some peroxidases to pectate–calcium complexes, and to a possible covalent bond between pectins and extensins (Qi et al. 1995; Dunand et al. 2002).

Abiotic and biotic stresses

Plants exposed to acute stress are known to up-regulate their overall peroxidase activity. This reaction happens equally with various abiotic and biotic stresses such as chemical (heavy metal, industrial, or agriculture pollution), biological (pathogens), or physical (wounding) assaults. Peroxidase expression results in plant defence either passively (building up of stronger walls) or actively (production of ROS against attacking organisms). When the stress factor manages to overcome the plant barriers and penetrates inside the plant, peroxidases may play a major role by isolating or eliminating the foreign body.

Chemical stresses

Toxicity of heavy metals in plants is probably due to their ability to promote damaging oxidative reactions (Schützendübel and Polle 2002). Cu and Fe ions can directly generate ROS from oxygenated molecules through the Haber–Weiss and Fenton reactions. However, the mechanisms of ROS production by metals such as Pb, Cd, and Zn ions is less clear, and may be mediated, for instance, by activation of lipoxygenase or binding to membrane proteins, thus triggering electron leakages responsible for formation of ROS (Seregin and Ivanov 2001). A possible function of peroxidases in treatment of heavy metals is their contribution in accumulating plants. For example, the waterlily *Nymphaea* probably uses peroxidases to produce phenolic polymers that trap Cd and isolate it in the form of Ca–Cd crystals in specific glands situated on the aquatic side of its leaves (Lavid et al. 2001b). When another aquatic plant, *Nymphoides peltata*, which contains a lower basal level of phenols, was subjected to the same growth condition and Cd exposure, it showed severe damages, thus supporting the implication of phenolic polymers and peroxidase in phytoaccumulation of metals (Lavid et al. 2001a).

Besides their participation to the entrapment of heavy metals, peroxidases can also degrade toxic molecules. Hairy root cultures of turnip (*Brassica napus*), are very efficient in degrading the toxic pesticide 2,4-dichlorophenol (2,4-DCP) (Agostini et al. 2003). The resulting products have not yet been precisely identified both in terms of chemical nature and toxicity, but they are probably a mixture of polymers created by the action of peroxidases. These enzymes can indeed attack phenolic moieties and form radicals that will then non-enzymatically polymerise. A further proof of the importance of peroxidases in the process was the appearance of an isozyme on an IEF gel upon repeated exposure to 2,4-DCP.

Although peroxidases are clearly induced by heavy metals and other toxic chemicals, it is still difficult to assign them a precise role in phytoremediation. Despite the poor knowledge about their function, the peroxidase up-regulation as a response of the plant to pollutants can already be used for the phytomonitoring of industrial or densely urbanised areas. Peroxidases have been shown to be quite sensitive to atmospheric pollution, with a response that can be stronger than the one of other classical biomarkers, such as superoxide dismutase, glutathione reductase levels (Wu and von Tiedemann 2002) or ascorbate concentration (Moraes et al. 2002). Similar observations have been made with heavy metals present in soil (Cho and Park 2000; Geebelen et al. 2002), although the response may be lower than that observed with atmospheric stress (Klumpp et al. 2000). The major drawback of this method is that it does not allow to distinguish between pollutants. However, it still remains one of the most sensitive methods to biologically evaluate the impact of pollution.

Biological stresses

Peroxidases are known to be activated in response to pathogen attacks and several roles have been attributed to plant peroxidases in host–pathogen interaction. They can

have a cell wall cross-linking activity (formation of lignin, extensin cross-links, dityrosine bonds) and create a highly toxic environment by massively producing ROS (oxidative burst), which results in adverse growth conditions for microorganisms.

In the defence of cotton against bacterial blight (*Xanthomonas campestris* pv. *malvacearum*), the resistance phenotype is characterised by a rapid localised tissue collapse resulting in necrotisation and immobilisation of the intruding pathogen at the sites of attack. Production and accumulation of superoxide anions and hydrogen peroxide in cell wall has been related to the presence of pathogens (Martinez et al. 1998). Total peroxidase activity is highly increased in the infected region 12 h after treatment and mainly localised in the apoplast and close to the bacterial infection site (Martinez et al. 1998; Delannoy et al. 2003). Similarly, rice leaves infected by rice blight (*Xanthomonas oryzae* pv. *oryzae*) strongly upregulated a single peroxidase isoform in xylem parenchyma. This peroxidase was then secreted to the xylem vessels, resulting in secondary wall thickening and reducing access of the pathogen to the pit membrane (the pathogen's contact point with living cells) (Hilaire et al. 2001). Lignin is not the only phenolic subunit that is used by peroxidase in defensive polymerisation reactions. Several other phenolics are involved and result in different morphological "barriers" against pathogens. In onion, for instance, granular deposits formed upon *Botrytis allii* infection contained tyramine derivatives, thought to strengthen the cell wall against fungal intrusion. These phenolics were probably polymerised by peroxidases (McLusky et al. 1999).

During germination, the aleurone layer of radish seeds functions as a secretory tissue. The seeds release ROS and peroxidases in the apoplastic space, respectively, between 6 to 12 h and between 24 and to 36 h after sowing to prevent pathogenic attack (Schopfer 2001). Reactive oxygen species and peroxidase act probably as a constitutive defence reaction against possible infections. A strong increase in H_2O_2 has also been reported following bacterial inoculation of lettuce (*Lactuca sativa* L.), together with a rise in apoplast peroxidasic activity. Localisation of the protective response was confined to the site of pathogenic intrusion (Bestwick et al. 1998).

Besides their functions during symbiosis, peroxidases from legumes have been associated with the plant's defence mechanisms. For example, FBP1, a French bean peroxidase, is responsible for the apoplastic oxidative burst and can generate hydrogen peroxide using cysteine (Blee et al. 2001). LEP1 from lupine is highly efficient in the cross-linking of extensin and could therefore form a physical barrier against invading organisms (Price et al. 2003). Other examples are reported, although no specific role has been assigned to the peroxidase: GMIPER1, from soybean, is induced in response to infection with *Phytophthora sojae* and other external stresses (Yi and Hwang 1998); Msprx1B, and the similar 1A and 1C peroxidases from *Medicago sativa*, were identified in alfalfa leaves after infection with *Pseudomonas syringae* (el-Turk et al. 1996).

Plant parasitism, the special case of Orobanche

Orobanche sp. (broomrape) are obligate root holoparasitic plants (Musselman 1980). They are one of the most important agricultural pests in several major cropping systems worldwide (Goldwasser and Yoder 2001). Several control strategies have been proposed and employed, but none have provided complete protection (Rubiales et al. 2003).

Several reports reveal that typical plant defence responses against pathogenic microorganisms are also induced upon parasitic plant infection. These include an increase of peroxidase activity, lignification, and cell-wall phenolic deposition (Goldwasser et al. 1999; Vieira Dos Santos et al. 2003a). For example, among the genes induced after infection of *A. thaliana* by *Orobanche ramose*, a gene coding for a class III peroxidase was upregulated (Vieira Dos Santos et al. 2003a). The expression of this gene was transient, early induced, and the transcript accumulated during the first 24 h after infection. A second induction was further detected at 7 days when the first parasite attachment was observed (Vieira Dos Santos et al. 2003b). A time-dependent increase of peroxidase activity was also reported in *O. aegyptiaca*-infected vetch (*Vicia sativa* cv. Yovel) plants (Goldwasser et al. 1999). However, *O. aegyptiaca* inoculation resulted in a relatively small increase in the enzyme activity in roots of susceptible vetch plants, whereas in roots of resistant plants (*Vicia atropurpurea* Desf. cv. Popany) the activity was four-fold increased compared to the non-inoculated treatments (Goldwasser et al. 1999). In addition, a sharp increase in free and bound phenolics was reported as well as an increase in lignin for resistant vetch, whereas only low increase in lignin and phenolics in response to infection were found for susceptible vetch (Goldwasser et al. 1999). Interestingly, another study revealed that most of the *Orobanche*-pea (*Pisum* spp.) resistant genotypes showed a higher constitutive peroxidase activity than susceptible ones (Castillejo et al. 2004). It can, therefore, be suggested that peroxidase proteins play a crucial role in plants' resistance against *Orobanche* by increasing the cell wall cross-linking or the generation of ROS.

Physical stresses

During a pathogenic attack, plant tissues can be damaged. Alternatively, wounding can occur through meteorological adversities or larger animals. Peroxidase expression during wounding is probably triggered in order to repair the damaged tissue, but also as a preventive defence mechanism against foreign attacks.

A study in sweet potato showed that the acidic peroxidase gene *swpa4* is not expressed in any tissue of the plant during normal growth, but that it is strongly up-regulated upon wounding of leaves, as well as incubation with high concentrations of H_2O_2 or NaCl (Park et al. 2003). This would mean either that the stress factors wounding and some chemicals follow the same pathway leading to peroxidase activation or that *swpa4* promoter is involved in

several response pathways. The latter hypothesis would not be surprising considering the high number of different regulatory sequences generally present in promoter sequences of peroxidase genes. Similarly, *ZmPox2* gene from *Zea mays* was both induced by wounding and ethylene, but not by methyl-jasmonate, another well-known wounding marker (Reymond et al. 2000; de Obeso et al. 2003). In striking contrast, two isozymes from northern red oak tree (*Quercus rubra* L.) were specifically induced by wounding and jasmonate, whereas another isozyme was induced by wounding, but not by jasmonate (Allison and Schultz 2004). This discrepancy may reflect the existence of different pathways in wounding signalling, that could also be due to the nature of the wounding. Indeed, a more detailed analysis performed on rice showed that the nature of the wound induces different peroxidase isoforms: rubbing or cutting leaves or tips did not induce the same peroxidases (Hiraga et al. 2000).

Symbiosis

During pathogenic interaction (fungi or bacteria), plants have the capability to respond with a set of defence reactions that include several peroxidases (as described above). In the case of symbiosis, the question arises whether symbionts influence peroxidase activities and isozyme patterns in a similar or different manner than pathogen attacks. Indeed, in symbiotic interactions, defence mechanisms have to be controlled and limited to allow the association of two different organisms.

Nodulation

After inoculation with compatible rhizobia, certain leguminous plants form nodules on their roots, providing conditions necessary for bacterial conversion of atmospheric dinitrogen to ammonia. Rhizobia soil bacteria colonise host cells and tissues through infection threads, which are tubular in-growths of the plant cell wall (Rathbun et al. 2002). The initiation of infection threads is apparently preceded by localised cell wall degradation (van Spronsen et al. 1994) and may be accompanied by modifications in cell wall peroxidase activities (Salzwedel and Dazzo 1993; Cook et al. 1995). Prior to infection, flavonoids or other plant metabolites contained in root exudates induce the expression of a set of bacterial genes (*nod* genes) (Cook et al. 1995; Mathesius 2001). The *nod* genes are required and are essential for the synthesis of rhizobial signal molecules called "Nod factors" and then for nodulation.

A *Rhizobium*-induced peroxidase, *rip1*, is rapidly and transiently expressed in the very early interactions of *Medicago truncatula*, with compatible *R. meliloti* (Cook et al. 1995). The *R. meliloti* Nod factor is both necessary and sufficient for the *rip1* induction (Peng et al. 1996) and for ROS production (Ramu et al. 2002). The *rip1* transcript and the ROS accumulations are both localised in the root epidermal cells and the level of *rip1* is maximal by 3 h post-inoculation and decreases by 48 h (Peng et al. 1996; Ramu et al. 2002). In contrast, maximal induction of other early genes by Nod factors is associated with rhizobial infection and early nodule morphogenesis, and these genes continue to be expressed in mature nodules (Cook et al. 1995). The *rip1* gene is also activated by exogenous H_2O_2 treatment in absence of Nod factors (Ramu et al. 2002), with a dose response similar to that of other characterised ROS-responsive genes (Chen et al. 1996). The RIP1 peroxidase could, therefore, mediate cell wall alteration in the early nodule development and at sites of infection thread formation (Salzwedel and Dazzo 1993). The increasing level of RIP1 and ROS could contribute to cell wall loosening facilitating early infection events. In addition, other peroxidases could be involved in cell wall cross-linking rapidly after rhizobia infection.

Several investigators highlighted the involvement of auxin in nodulation by demonstrating that the external application of auxin transport inhibitors lead to the formation of nodule-like structures (Allen et al. 1953; Hirsch et al. 1989). Moreover, auxin accumulated at the site of nodule initiation and subsequently diminished during nodule primordium formation (Mathesius et al. 1998; Penmetsa et al. 2003). The increase of peroxidases during nodulation could be related to the control of auxin level in roots. Indeed, some class III peroxidases are known to oxidise auxin and could then indirectly control nodule formation through the auxin catabolism (Savitsky et al. 1999; Mathesius 2001).

Mycorrhization

Ectomycorrhizal (ECM) and arbuscular mycorrhizal (AM) fungi are ubiquitous symbiotic microorganisms associated with plants of most families of Angiosperms and Gymnosperms (Koide and Schreiner 1992). They play an essential role in plant nutrition and stress tolerance and may be related to expression of various oxidative enzymes. Mycorrhiza are known to modify root enzymes (activity, isoforms, etc.), including peroxidase activity. As during nodulation, the mycorrhizal symbiotic interaction is also probably associated with a modification of the cell wall structure of the infected cells.

Arbuscular mycorrhizal fungal inoculation of leeks (*Allium porrum*) resulted in a transiently higher level of peroxidase activity compared to uninoculated controls (Spanu and Bonfante-Fasolo 1988). Ultrastructural localisation of peroxidase activity further revealed an accumulation in the middle lamella of cortical cells penetrated by intercellular hyphae, whereas no activity was found in the walls of cortical cells already containing arbuscular intracellular structures (Spanu and Bonfante-Fasolo 1988). Similarly, peroxidase activity in the fine roots of *Pinus sylvestris* seedlings planted in forest humus was highest at the beginning of the experiment and decreased during growing season (Tarvainen et al. 2004). This decrease of peroxidase can be associated with the increase of the number of ECM morphotypes as well as root biomass. These and several other results (Albrecht et al. 1994) suggest that the mycorrhizal

fungus generally induces peroxidase expression during the initial stage of symbiotic interaction (Günther et al. 1998; Salzer et al. 1999) and that this response is later either controlled by the plant or circumvented in mature mycorrhiza (Münzenberger et al. 1997). Nevertheless, some contradictory results have also been reported. In roots of *Medicago sativa* L. colonised by a mature AM fungi (*Glomus mossae*), increased peroxidase levels were observed (Criquet et al. 2000). A similar amount of peroxidases was found in roots of Scots pine (*Pinus sylvestris*) colonised and not colonised by different ECM fungi (Timonen and Sen 1998). However, the differences reported between the different mycorrhizal fungi are likely to be due to potential species-specific enzyme activities. Eucalyptus roots inoculated with different ECM strain of *Pisolithus* spp. showed a transient increase of peroxidase activity during the first days of inoculation in relation to the colonisation efficiency of the strain. Peroxidase activity of "poor" root colonisers remained relatively low, whereas activity in roots inoculated with "good" root colonisers increased sharply (Albrecht et al. 1994). Besides acting on the total peroxidase activity, mycorrhyza interaction can also regulate specific isoforms. Changes in root peroxidase profile following mycorrhiza colonisation were observed indicating a specific differential isozyme expression in roots (Münzenberger et al. 1997; Timonen and Sen 1998).

It has been suggested that auxins produced by the ECM fungi reduce peroxidase-catalysed cross-linking of cell-wall constituents (Salzer and Hager 1993; Charvet-Candela et al. 2002). This mechanism may make the cell walls in roots less rigid and allow colonisation of the intercellular space by symbiotic fungi. However, other investigators have shown that IAA levels were inversely related to peroxidase activity (Mitchell et al. 1986), suggesting that in mycorrhiza like in nodules, peroxidases could be involved in auxin catabolism. The role of auxin and peroxidases in mycorrhization is still difficult to define, because of the lack of information about the relative importance of these two mechanisms.

From these results it appears that symbionts, mycorrhizal fungi, as well as bacteria, may not trigger the full host defence response particularly where peroxidase activity is concerned. In successful infections, these defence responses are either controlled by the plant or overcome by the successful symbiont. The contradictory reports concerning the interactions of peroxidase with auxins afford only a limited first insight into mutual interactions of the symbiotic partners and need further exploration to be clarified. Expression of peroxidases in host roots might nevertheless limit and regulate the growth of the symbiont. During the early infection, peroxidase production can generate ROS implicated in the cell wall loosening to facilitate the symbiont entrance. During the time of colonisation and development of symbiosis, it is also likely that defence-related peroxidases are expressed to limit the intracellular spreading of the fungus but also to protect against soil microbial pathogens, through the establishment of a highly toxic environment by massively producing ROS (oxidative burst).

Senescence

It is generally accepted that cell death, senescence, ripening, necrosis, and lesions are correlated and depended on overlapping mechanisms. The first three events can be considered as normal evolution for a plant, the last ones being related to foreign actions. One difference could be the notion of reversibility: until a specific threshold, cell death, senescence, and ripening can be considered as reversible processes.

Plant senescence

Leaf senescence is the last stage before complete cell death. A number of cellular changes have been associated with this mechanism. For example, increase of ethylene synthesis and peroxidase activity have been reported (Jimenez et al. 1998). A change in gene expression has also been observed during senescence: a decrease for the photosynthesis-associated genes and an increase for the senescence-associated genes (Abarca et al. 2001). Reactive oxygen species such as $O_2^{\cdot -}$ have been shown to be involved in the induction and development of the senescence stage. The superoxide radical $O_2^{\cdot -}$ could be generated by extracellular peroxidase activity following salicylic acid treatment (Kawano et al. 1998) and could act on the senescence induction pathway. On the other hand, total peroxidase activity increased during *A. thaliana* leaf expansion as well as in the senescent tissues, with a stronger increase at the end of the leaf expansion (Abarca et al. 2001). It seems difficult to dissociate the leaf senescence from the cell wall stiffening and cross-linking. In *A. thaliana* leaf, the peroxidase activity is mainly associated with cell extension between the 1st and the 4th week of growth, the 4th to 5th weeks of growth correspond to a transition phase, and after the 5th week of growth peroxidase activity is associated in majority with the cross-linking of cell-wall compounds (Abarca et al. 2001). In agreement with this observation, the increase of syringaldazine-peroxidase activity in plants exposed to ozone may be related to the induction of early senescence through the activation or acceleration of lignification process (Ranieri et al. 2000).

Fruit growth and ripening

Fruit growth occurs by cell expansion rather than division. During tomato fruit development as in other cell elongation processes described previously, cell expansion is dependent on cell wall loosening (Andrews et al. 2000). Hence the implication of peroxidases in such a process is related to the fine balance between cell wall loosening and stiffening.

However, fruit ripening syndrome, which involves certain hormonal and enzyme classes, is commonly promoted by ethylene in climacteric plant (Alexander and Grierson 2002). The final fruit evolution is correlated with an increased peroxidase activity acting on cell-wall

modification. The cessation of tomato fruit growth follows a decrease of wall extensibility at the fruit skin level (Andrews et al. 2000). Interestingly, peroxidases associated to the cell wall of the epidermis are induced during fruit maturity. These peroxidases are probably not directly related to the fruit ripening but could control the growth arrest by changing the mechanical properties of cell wall and consequently produce a protective barrier in the epidermis.

Conclusions

In plant cells, class III peroxidases are present in soluble, ionically- and covalently-cell wall bound forms. They are involved in many functions as a consequence of their two cycles (Fig. 1): ROS generation and regulation, H_2O_2 level regulation, and oxidation of various substrates (which leads to the catabolism or the polymerisation of substrates) (Fig. 3). Using these three mechanisms, plants have at their disposal a "Swiss army knife-like" multifunctional tool that they can use in every tissue at any time throughout their life cycle. The high number of isoforms further allows a fine balance between antagonistic peroxidase functions such as cell wall cross-linking and loosening. The variation of the number of isoforms between species could explain the numerous discrepancies reported on the regulation of the different peroxidases related to physiological processes.

The comparison of the different situations described in this review shows that peroxidases up-regulation is generally transient. Peroxidases are often strongly induced at the beginning of an event, and then slowly decrease with time. This type of response has indeed been shown in chemical stress, where peroxidase expression was only significantly increased in acute stress (Klumpp et al. 2000). Plants respond to pathogens and symbionts similarly, with a rapid oxidative burst for the former (Blee et al. 2001), and the increase in peroxidase activity being mainly expressed in the first moments of colonisation in the latter (Cook et al. 1995; Tarvainen et al. 2004). This transient induction is, however, not common to all peroxidases. There is always a basal level of peroxidase activity in plants, probably to perform "housekeeping" functions, such as growth by elongation and lignification. It is possible that the appearance of the first peroxidase during evolution has allowed plants to build up a cell-wall structure that is able to stand out of the water and hence adapted to land colonisation. Later on, with the advent of insects, climatic changes, human impact, and other stresses, a quickly inducible peroxidase was advantageous to respond to these novel factors, and, therefore, was naturally selected and further duplicated. However, these assumptions remain purely speculative as long as the in vivo functions of single peroxidases have not been determined. To our knowledge, no studies have reported such information on peroxidases yet. When identified, this crucial point will certainly help us to better understand the evolution, the roles, and the regulations of this multifunctional enzyme.

Acknowledgements We thank William Deakin for his critical reading, his constructive suggestions, and for his original title. We thank the University of Geneva and the Swiss National Science Foundation (grant 31-068003.02) for research support to C.P. and C.D. The authors F. Passardi and C. Cosio have both contributed equally to this work

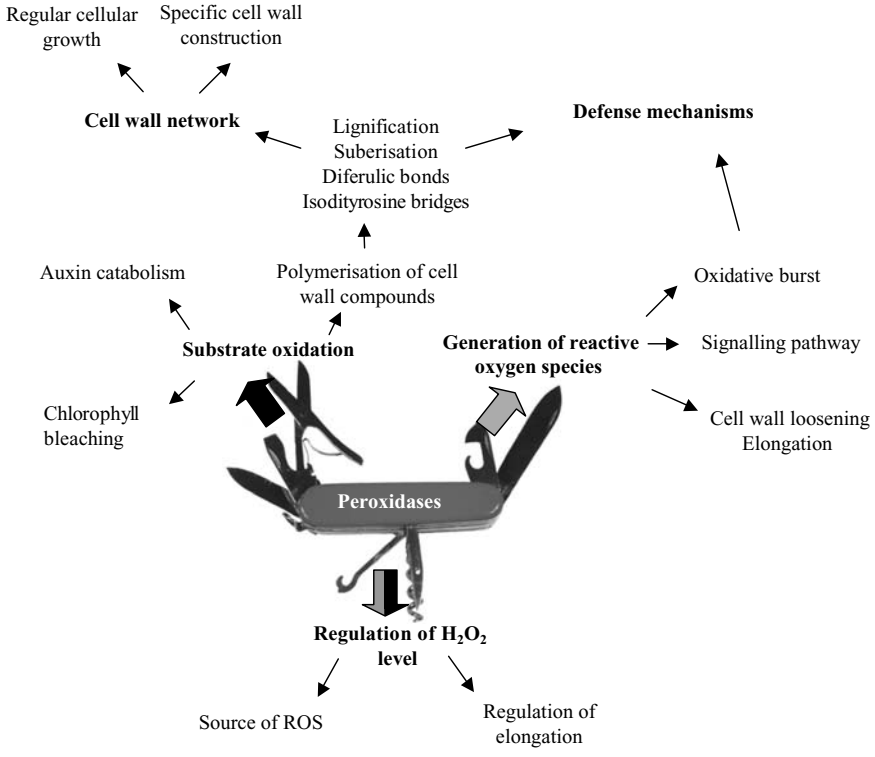


Fig. 3 Overview of the putative multiple functions of the class III peroxidases through both of their cycles. Peroxidases oxidise various subtrates through the peroxidative cycle (black arrow). Reactive oxygen species are generated through the hydroxylic cycle (grey arrow). H_2O_2 level can be regulated through both catalytic cycles (grey and black arrows)

References

Abarca D, Martin M, Sabater B (2001) Differential leaf stress responses in young and senescent plants. Physiol Plant 113:409–415

Agostini E, Coniglio MS, Milrad SR, Tigier HA, Giulietti AM (2003) Phytoremediation of 2,4-dichlorophenol by *Brassica napus* hairy root cultures. Biotechnol Appl Biochem 37:139–144

Albrecht C, Burgess T, Dell B, Lapeyrie F (1994) Chitinase and peroxidase activities are induced in eucalyptus roots according to aggressiveness of Australian ectomycorrhizal strains of *Pisolithus* sp. New Phytol 127:217–222

Alexander L, Grierson D (2002) Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. J Exp Bot 53:2039–2055

Allen EK, Allen ON, Newman AS (1953) Pseudonodulation of leguminous plants induced by 2-bromo-3,5-dichlorobenzoic acid. Am J Bot 40:429–435

Allison SD, Schultz JC (2004) Differential activity of peroxidase isozymes in response to wounding, gypsy moth, and plant hormones in northern red oak (*Quercus rubra* L.). J Chem Ecol 30:1363–1379

Andrews J, Malone M, Thompson DS, Ho LC, Burton KS (2000) Peroxidase isozyme patterns in the skin of maturing tomato fruit. Plant Cell Environ 23:415–422

Bernards MA, Fleming WD, Llewellyn DB, Priefer R, Yang X, Sabatino A, Plourde GL (1999) Biochemical characterization of the suberization-associated anionic peroxidase of potato. Plant Physiol 121:135–146

Bestwick CS, Brown IR, Mansfield JW (1998) Localized changes in peroxidase activity accompany hydrogen peroxide generation during the development of a nonhost hypersensitive reaction in lettuce. Plant Physiol 118:1067–1078

Blee KA, Jupe SC, Richard G, Zimmerlin A, Davies DR, Bolwell GP (2001) Molecular identification and expression of the peroxidase responsible for the oxidative burst in French bean (*Phaseolus vulgaris* L.) and related members of the gene family. Plant Mol Biol 47:607–620

Blee KA, Choi JW, O'Connell AP, Schuch W, Lewis NG, Bolwell GP (2003) A lignin-specific peroxidase in tobacco whose antisense suppression leads to vascular tissue modification. Phytochemistry 64:163–176

Castillejo MA, Amiour N, Dumas-Gaudot E, Rubiales D, Jorrin JV (2004) A proteomic approach to studying plant response to crenate broomrape (*Orobanche crenate*) in pea (*Pisum sativum*). Phytochemistry 65:1817–1828

Charvet-Candela V, Hitchin S, Reddy MS, Cournoyer B, Marmeisse R, Gay G (2002) Characterization of a *Pinus pinaster* cDNA encoding an auxin up-regulated putative peroxidase in roots. Tree physiol 22:231–238

Chen SX, Schopfer P (1999) Hydroxyl-radical production in physiological reactions. A novel function of peroxidase. Eur J Biochem 260:726–735

Chen W, Chao G, Singh KB (1996) The promoter of a H_2O_2-inducible, *Arabidopsis* glutathione S-transferase gene contains closely linked OBF- and OBP1-binding sites. Plant J 10:955–966

Cho U, Park J (2000) Mercury-induced oxidative stress in tomato seedlings. Plant Sci 156:1–9

Christensen JH, Overney S, Rohde A, Diaz WA, Bauw G, Simon P, Van Montagu M, Boerjan W (2001) The syringaldazine-oxidizing peroxidase PXP 3-4 from poplar xylem: cDNA isolation, characterization and expression. Plant Mol Biol 47:581–593

Cook D, Dreyer D, Bonnet D, Howell M, Nony E, Van den Bosch K (1995) Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. Plant Cell 7:43–55

Cordoba-Pedregosa M, Gonzalez-Reyes JA, Canadillas M, Navas P, Cordoba F (1996) Role of apoplastic and cell-wall peroxidases on the stimulation of root elongation by ascorbate. Plant Physiol 112:1119–1125

Cordoba-Pedregosa M, Cordoba F, Villalba JM, Gonzalez-Reyes JA (2003) Zonal changes in ascorbate and hydrogen peroxide contents, peroxidase, and ascorbate-related enzyme activities in onion roots. Plant Physiol 131:697–706

Cosgrove DJ (2001) Wall structure and wall loosening. A look backwards and forwards. Plant Physiol 125:131–134

Criquet S, Joner E, Leglize P, Leyval C (2000) Anthracene and mycorrhiza affect the activity of oxidoreductase in the roots and rhizosphere of lucerne (*Medicago sativa* L.). Biotechnol Lett 22:1733–1737

Delannoy E, Jalloul A, Assigbetse K, Marmey P, Geiger JP, Lherminier J, Daniel JF, Martinez C, Nicole M (2003) Activity of class III peroxidases in the defense of cotton to bacterial blight. Mol Plant Microbe Interact 16:1030–1038

de Obeso M, Caparros-Ruiz D, Vignols F, Puigdomenech P, Rigau J (2003) Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. Gene 309:23–33

Dunand C, Tognolli M, Overney S, von Tobel L, De Meyer M, Simon P, Penel C (2002) Identification and characterisation of Ca2+-pectate binding peroxidases in *Arabidopsis thaliana*. J Plant Physiol 159:1165–1171

Dunand C, de Meyer M, Crévecoeur M, Penel C (2003) Expression of a peroxidase gene in zucchini in relation with hypocotyl growth. Plant Physiol Biochem 41:805–811

Duroux L, Welinder KG (2003) The peroxidase gene family in plants: a phylogenetic overview. J Mol Evol 57:397–407

el-Turk J, Asemota O, Leymarie J, Sallaud C, Mesnage S, Breda C, Buffard D, Kondorosi A, Esnault R (1996) Nucleotide sequences of four pathogen-induced alfalfa peroxidase-encoding cDNAs. Gene 170:213–216

Fry SC (1998) Oxidative scission of plant cell wall polysaccharides by ascorbate-induced hydroxyl radicals. Biochem J 332:507–515

Geebelen W, Vangronsveld J, Adriano DC, Van Poucke LC, Clijsters H (2002) Effects of Pb-EDTA and EDTA on oxidative stress reactions and mineral uptake in *Phaseolus vulgaris*. Physiol Plant 115:377–384

Goldwasser Y, Yoder JI (2001) Differential induction of *Orobanche* seed germination by *Arabidopsis thaliana*. Plant Sci 160:951–959

Goldwasser Y, Hershenhorn J, Plakhine D, Kleifeld Y, Rubin B (1999) Biochemical factors involved in vetch resistance to *Orobanche aegyptiaca*. Physiol Mol Plant Pathol 54: 87–96

Günther T, Perner B, Gramss G (1998) Activities of phenol oxidizing enzymes of ectomycorrhizal fungi in axenic culture and in symbiosis with Scots pine (*Pinus sylvestris* L.). J Basic Microbiol 38:197–206

Hilaire E, Young SA, Willard LH, McGee JD, Sweat T, Chittoor JM, Guikema JA, Leach JE (2001) Vascular defense responses in rice: peroxidase accumulation in xylem parenchyma cells and xylem wall thickening. Mol Plant Microbe Interact 14:1411–1419

Hiraga S, Yamamoto K, Ito H, Sasaki K, Matsui H, Honma M, Nagamura Y, Sasaki T, Ohashi Y (2000) Diverse expression profiles of 21 rice peroxidase genes. FEBS Lett 471: 245–250

Hiraga S, Sasaki K, Ito H, Ohashi Y, Matsui H (2001) A large family of class III plant peroxidases. Plant Cell Physiol 42: 462–468

Hirsch AM, Bhuvaneswari TV, Torrey JG, Bisseling T (1989) Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitor. Proc Natl Acad Sci USA 86:1244–1248

Iiyama K, Lam T, Stone BA (1994) Covalent cross-links in the cell wall. Plant Physiol 104:315–320

Jackson PA, Galinha CI, Pereira CS, Fortunato A, Soares NC, Amancio SB, Pinto Ricardo CP (2001) Rapid deposition of extensin during the elicitation of grapevine callus cultures is specifically catalyzed by a 40-kilodalton peroxidase. Plant Physiol 127:1065–1076

Jimenez A, Hernandez JA, Pastori G, del Rio LA, Sevilla F (1998) Role of the ascorbate-glutathione cycle of mitochondria and peroxisomes in the senescence of pea leaves. Plant Physiol 118:1327–1335

Joo JH, Bae YS, Lee JS (2001) Role of auxin-induced reactive oxygen species in root gravitropism. Plant Physiol 126:1055–1060

Kawamura Y, Wakabayashi K, Hoson T, Yamamoto R, Kamisaka S (2000) Stress-relaxation analysis of submerged and air-grown rice coleoptiles: correlation with cell wall biosynthesis and growth. J Plant Physiol 156:689–694

Kawano T, Sahashi N, Takahashi K, Uozumi N, Muto S (1998) Salicylic acid induces extracellular superoxide generation followed by an increased in cytosolic calcium ion in tobacco suspension culture: the earliest events in salicylic acid signal transduction. Plant Cell Physiol 39:721–730

Keren-Keiserman A, Tanami Z, Shoseyov O, Ginzberg I (2004) Peroxidase activity associated with suberization processes of the muskmelon (*Cucumis melo*) rind. Physiol Plant 121:141–148

Kiefer-Meyer MC, Gomord V, O'Connell A, Halpin C, Faye L (1996) Cloning and sequence analysis of laccase-encoding cDNA clones from tobacco. Gene 178:205–207

Kieliszewski MJ, Lamport DT (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. Plant J 5:157–172

Klumpp G, Furlan CM, Domingos M, Klumpp A (2000) Response of stress indicators and growth parameters of *Tibouchina pulchra* Cogn. exposed to air and soil pollution near the industrial complex of Cubatao, Brazil. Sci Total Environ 246:79–91

Koide RT, Schreiner RP (1992) Regulation of the vesicular-arbuscuar mycorrhizal symbiosis. Annu Rev Plant Physiol Plant Mol Biol 43:557–581

Lavid N, Schwartz A, Lewinsohn E, Tel-Or E (2001a) Phenols and phenol oxidases are involved in cadmium accumulation in the water plants Nymphoides peltata (*Menyanthaceae*) and Nymphaeae (*Nymphaeaceae*). Planta 214:189–195

Lavid N, Schwartz A, Yarden O, Tel-Or E (2001b) The involvement of polyphenols and peroxidase activities in heavy-metal accumulation by epidermal glands of the waterlily (*Nymphaeaceae*). Planta 212:323–331

Lewis NG, Yamamoto E (1990) Lignin: occurrence, biogenesis and biodegradation. Annu Rev Plant Physiol Plant Mol Biol 41:455–496

Li Y, Kajita S, Kawai S, Katayama Y, Morohoshi N (2003) Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. J Plant Res 116:175–182

Liszkay A, Kenk B, Schopfer P (2003) Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth. Planta 217:658–667

Lopez-Serrano M, Fernandez MD, Pomar F, Pedreno MA, Ros Barcelo A (2004) *Zinnia elegans* uses the same peroxidase isoenzyme complement for cell wall lignification in both single-cell tracheary elements and xylem vessels. J Exp Bot 55:423–431

MacAdam JW, Grabber JH (2002) Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. Planta 215:785–793

Magliano TMA, Casal JJ (1998) In vitro cross-linking of extensin precursors by mustard extracellular isoforms of peroxidase that respond either to phytochrome or to wounding. J Exp Bot 49:1491–1499

Martinez C, Montillet J, Bresson E, Agnel J, Dai G, Daniel J, Geiger J, Nicole M (1998) Apoplastic peroxidase generates superoxide anions in cells of cotton cotyledons undergoing the hypersensitive reaction to *Xanthomonas campestris* pv. *malvacearum* race 18. Mol Plant Microbe Interact 11:1038–1047

Mathesius U (2001) Flavonoids induced in cells undergoing nodule organogenesis in white clover are regulators of auxin breakdown by peroxidase. J Exp Bot 52:419–426

Mathesius U, Schlaman HRM, Spaink HP, Sautter C, Rolfe BG, Djorhjevic MA (1998) Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. Plant J 14:23–34

McLusky SR, Bennett MH, Beale MH, Lewis MJ, Gaskin P, Mansfield JW (1999) Cell wall alterations and localized accumulation of feruloyl-3′-methoxytyramine in onion epidermis at sites of attempted penetration by *Botrytis allii* are associated with actin polarisation, peroxidase activity and suppression of flavonoid biosynthesis. Plant J 17:523–534

Mitchell RJ, Garrett HE, Cox GS, Atalay A (1986) Boron and ectomycorrhizal influences on indole-3-acetic acid levels and indole-3-acetic acid oxidase and peroxidase activities of *Pinus echinata* Mill. roots. Tree Physiol 1:1–8

Moraes RM, Klumpp A, Furlan CM, Klumpp G, Domingos M, Rinaldi MC, Modesto IF (2002) Tropical fruit trees as bioindicators of industrial air pollution in southeast Brazil. Environ Int 28:367–374

Morohashi Y (2002) Peroxidase activity develops in the micropylar endosperm of tomato seeds prior to radicle protrusion. J Exp Bot 53:1643–1650

Münzenberger B, Otter T, Wüstrich D, Polle A (1997) Peroxidase and laccase activities in mycorrhizal and non-mycorrhizal fine roots of Norway spruce (*Picea abies*) and larch (*Larix decidua*). Can J Bot 75:932–938

Musselman LJ (1980) The biology of *Striga*, *Orobanche*, and other root-parasitic weeds. Ann Rev Plant Physiol 18:463–489

Nielsen KL, Indiani C, Henriksen A, Feis A, Becucci M, Gajhede M, Smulevich G, Welinder KG (2001) Differential activity and structure of highly similar peroxidases. Spectroscopic, crystallographic, and enzymatic analyses of lignifying *Arabidopsis thaliana* peroxidase A2 and horseradish peroxidase A2. Biochemistry 40:11013–11021

Ostergaard L, Teilum K, Mirza O, Mattsson O, Petersen M, Welinder KG, Mundy J, Gajhede M, Henriksen A (2000) *Arabidopsis* ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. Plant Mol Biol 44:231–243

Park SY, Ryu SH, Kwon SY, Lee HS, Kim JG, Kwak SS (2003) Differential expression of six novel peroxidase cDNAs from cell cultures of sweet potato in response to stress. Mol Genet Genomics 269:542–552

Passardi F, Longet D, Penel C, Dunand C (2004a) The plant peroxidase multigenic family in rice and its evolution in green plants. Phytochemistry 65:1879–1893

Passardi F, Penel C, Dunand C (2004b) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci 9:534–540

Penel C, Gaspar T, Greppin H (1992) Plant peroxidases. University of Geneva, Geneva, Switzerland

Peng HM, Dreyer DA, VandenBosch KA, Cook D (1996) Gene structure and differential regulation of the Rhizobium-induced peroxidase gene rip1. Plant Physiol 112:1437–1446

Penmetsa RV, Frugoli JA, Smith LS, Long SR, Cook D (2003) Dual genetic pathways controlling nodule number in *Medicago truncatula*. Plant Physiol 131:998–1008

Price NJ, Pinheiro C, Soares CM, Ashford DA, Ricardo CP, Jackson PA (2003) A biochemical and molecular characterization of LEP1, an extensin peroxidase from Lupin. J Biol Chem 278:41389–41399

Qi X, Behrens BX, West PR, Mort AJ (1995) Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin. Plant Physiol 108: 1691–1701

Quiroga M, Guerrero C, Botella MA, Barcelo A, Amaya I, Medina MI, Alonso FJ, de Forchetti SM, Tigier H, Valpuesta V (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. Plant Physiol 122:1119–1127

Ralph J, Quideau S, Grabber JH, Hatfield RD (1994) Identification and synthesis of new ferulic dehydrodimers present in grass cell walls. J Chem Soc Perkin Trans 1:3485–3498

Ramu SK, Peng HM, Cook DR (2002) Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene rip1 in *Medicago truncatula*. Mol Plant Microbe Interact 15:522–528

Ranieri A, Petacco F, Castagna A, Soldatini GF (2000) Redox state and peroxidase system in sunflower plants exposed to ozone. Plant Sci 159:159–167

Rathbun EA, Naldrett MJ, Brewin NJ (2002) Identification of a family of extensin-like gltcoproteins in the lumen of *Rhizobium*-induced infection threads in pea root nodules. Mol Plant Microbe Interact 15:350–359

Reymond P, Weber H, Damond M, Farmer EE (2000) Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. Plant Cell 12:707–720

Rubiales D, Pérez-de-Luque A, Joel DM, Alcantara C (2003) Characterization of resistance in chickpea to crenate broomrape (*Orobanche crenata*). Weed Sci 51:702–707

Salzer P, Hager A (1993) Effects of auxins and ectomycorrhizal elicitors on wall-bond proteins and enzymes of spruce (*Picea abies* (L.) Karst.) cells. Trees 8:49–55

Salzer P, Corbière H, Boller T (1999) Hydrogen peroxide accumulation in *Medicago truncatula* roots colonized by the arbuscular mycorrhiza-forming fungus *Glomus intaradices*. Planta 208:319–325

Salzwedel JL, Dazzo FB (1993) pSym nod gene influence on elicitation of peroxidase activity from white clover and pea roots by rhizobia and their cell-free supernatants. Mol Plant Microbe Interact 6:127–134

Sanchez M, Pena MJ, Revilla G, Zarra I (1996) Changes in dehydrodiferulic acids and peroxidase activity against ferulic acid associated with cell walls during growth of *Pinus pinaster* hypocotyl. Plant Physiol 111:941–946

Savitsky PA, Gazaryan IG, Tishkov VI, Lagrimini LM, Ruzgas T, Gorton L (1999) Oxidation of indole-3-acetic acid by dioxygen catalysed by plant peroxidases: specificity for the enzyme structure. Biochem J 340:579–583

Schnabelrauch LS, Kieliszewski M, Upham BL, Alizedeh H, Lamport DT (1996) Isolation of pl 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link site. Plant J 9:477–489

Schopfer P (1994) Histochemical demonstration and localization of H_2O_2 in organs of higher plants by tissue printing on nitrocellulose paper. Plant Physiol 104:1269–1275

Schopfer P (2001) Hydroxyl radical-induced cell-wall loosening in vitro and in vivo: implications for the control of elongation growth. Plant J 28:679–688

Schopfer P, Liszkay A, Bechtold M, Frahry G, Wagner A (2002) Evidence that hydroxyl radicals mediate auxin-induced extension growth. Planta 214:821–828

Schützendübel A, Polle A (2002) Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by micorrhization. J Exp Bot 53:1351–1365

Schweikert C, Liszkay A, Schopfer P (2000) Scission of polysaccharides by peroxidase-generated hydroxyl radicals. Phytochemistry 53:565–570

Scialabba A, Bellani LM, Dell'Aquila A (2002) Effects of ageing on peroxidase activity and localization in radish (*Raphanus sativus* L.) seeds. Eur J Histochem 46:351–358

Seregin IV, Ivanov VB (2001) Physiological aspects of cadmium and lead toxic effects on higher plants. Russ J Plant Physiol 48:523–544

Spanu P, Bonfante-Fasolo P (1988) Cell-wall-bound peroxidase activity in roots of mycorrhizal *Allium porrum*. New Phytol 109:119–124

Tarvainen O, Markkola AM, Ahonen-Jonnarth U, Jumpponen A, Strömmer R (2004) Changes in ectomycorrhizal colonization and root peroxidase activity in *Pinus sylvestris* nursery seedlings planted in forest humus. Scan J For Res 19:400–408

Timonen S, Sen R (1998) Heterogeneity of fungal and plant enzyme expression in intact Scots pine- *Suillus bovinus* and *-Paxillus involutus* mycorrhizospheres developed in natural forest humus. New Phytol 138:355–366

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288:129–138

van Spronsen PC, Bakhuizen R, van Brussel AAN, Kijne JW (1994) Cell wall degradation during infection thread formation by the root nodule bacterium *Rhizobium leguminosarum* is a two-step process. Eur J Cell Biol 64:88–94

Vieira Dos Santos C, Delavault P, Letousey P, Thalouarn P (2003a) Identification by suppression subtractive hybridization and expression analysis of *Arabidopsis thaliana* putative defence genes during *Orobanche ramosa* infection. Physiol Mol Plant Pathol 62:297–303

Vieira Dos Santos C, Letousey P, Delavault P, Thalouarn P (2003b) Defense gene expression analysis of *Arabidopsis thaliana* parasited by *Orobanche ramosa*. Phytopathology 93:415–457

Welinder KG (1992a) Plant peroxidases: structure-function relationships. In: Penel C, Gaspar T, Greppin H (eds) Plant peroxidases. University of Geneva, Geneva, Switzerland, pp 1–24

Welinder KG (1992b) Superfamily of plant, fungal and bacterial peroxidases. Curr Opin Struct Biol 2:388–393

Whitmore FW (1976) Binding of ferulic acid to cell-walls by peroxidases of *Pinus-elliottii*. Phytochemistry 15:375–378

Wu YX, von Tiedemann A (2002) Impact of fungicides on active oxygen species and antioxidant enzymes in spring barley (*Hordeum vulgare* L.) exposed to ozone. Environ Poll 116:37–47

Yi SY, Hwang BK (1998) Molecular cloning and characterization of a new basic peroxidase cDNA from soybean hypocotyls infected with *Phytophthora sojae* f.sp. glycines. Mol Cell 8:556–564

Review TRENDS in Plant Science Vol.9 No.11 November 2004

Performing the paradoxical: how plant peroxidases modify the cell wall

Filippo Passardi, Claude Penel and Christophe Dunand

Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland

Since their appearance in the first land plants, genes encoding class III peroxidases have been duplicated many times during evolution and now compose a large multigene family. The reason for these many genes is elusive, and we are still searching for the specific function of every member of the family. Nevertheless, our current understanding implicates peroxidases as key players during the whole life cycle of a plant, and particularly in cell wall modifications, in roles that can be antagonistic depending on the developmental stage. This diversity of functions derives in part from two possible catalytic cycles of peroxidases involving the consumption or release of H_2O_2 and reactive oxygen species (e.g. $O_2\cdot^-$, H_2O_2, $\cdot OH$).

Plants contain two classes of HEME PEROXIDASES (see Glossary) – the class I and class III peroxidases (EC 1.11.1.7), according to the classification proposed by Karen Welinder [1]. Class I peroxidases are intracellular, whereas class III peroxidases are secreted into the cell wall or the surrounding medium. Both classes have a common prokaryotic origin: class I has preserved the bacterial sequence, whereas class III has evolved differently and is related only structurally to this ancestor [2]. All plant peroxidase enzymes share the same general structure, consisting of ferriprotoporphyrin IX as a prosthetic group and ten α-helices. The class III peroxidases also contain three extra α-helices, a few highly conserved amino acids and four disulfide bridges [1].

Genes encoding class III peroxidases form a large multigene family [3–5] that is present in all land plants but absent from the unicellular green algae [5] (Figure 1). From the appearance of the first class III peroxidase, just before the emergence of terrestrial plants, to Angiosperms, the number of gene copies has increased greatly [4,5]. The advent of these peroxidases can be associated with the adaptation of plants to terrestrial life in the presence of elevated oxygen concentrations. The evolution of a multigene family seems to be correlated with the increasing complexity of plant architecture and to the diversification of their biotopes and pathogens. In some plants, a high duplication rate has led to large multigene families, as suggested by the presence of 138 peroxidase-encoding genes in *Oryza sativa* [5], whereas *Arabidopsis* encodes 73 peroxidases [3]. SUBFUNCTIONALIZATION or NEOFUNCTIONALIZATION could explain the conservation of duplicated genes and the presence of such large multigene families in which each PARALOG has become specialized for a determinate task [6].

The metal center iron (III) protoporphyrin IX and two amino acid residues ('distal' arginine and histidine) are essential for the catalytic activity of class III peroxidases involving exchanges of electrons and protons. In the standard peroxidative cycle (Figure 2), the enzymes catalyze the reduction of H_2O_2 by taking electrons from various donor molecules such as phenolic compounds, lignin precursors, auxin or secondary metabolites [7]. The diversity of the substrates oxidized by the regular peroxidative cycle explains why these proteins are implicated in a broad range of physiological processes, such as auxin catabolism, formation of lignin and SUBERIN, cross-linking of cell wall components, defense against pathogens, and cell elongation [7,8]. However, little is known about their specific functions *in planta*.

Recently, the precise description of a separate hydroxylic cycle *in vivo* (Figure 2) has opened up the possibility of a new range of functions for peroxidases [9,10]. In this alternative cycle, native peroxidases can be converted to oxyferroperoxidases (also known as compound III) in two different ways, which can then lead to the formation of various reactive oxygen species (ROS) [11]; ROS are compounds that can possess an intrinsic activity or can act as part of signal transduction pathways.

Peroxidases can therefore be considered as bifunctional enzymes that can oxidize various substrates in the presence of H_2O_2 but also produce ROS. Their involvement in many physiological and developmental processes can be detected in plants from germination to senescence. One of the noteworthy features of peroxidases is that they are associated with cell elongation processes but also with reactions that restrict growth. This apparent paradox has been extensively studied and is reviewed here in the

Glossary

Heme peroxidases: oxidoreductases containing usually ferriprotoporphyrin IX as prosthetic group. Able to oxidize various substrates via the reduction of hydrogen peroxide.
Neofunctionalization: acquisition of a novel function in a protein owing to mutations or duplications at the DNA level during evolution.
Paralog: homologous genes produced by duplication within a genome.
Suberin: wax-like lipid polymer resulting from the polymerization of fatty acids and phenolics. It is formed in the outer cell wall of some tissues and after injury.
Subfunctionalization: acquisition of a new expression profile (e.g. distinct developmental stage or specific tissue) for a duplicated gene.

Corresponding author: Christophe Dunand (christophe.dunand@bota.unige.ch).
Available online 7 October 2004

 1360-1385/$ - see front matter doi:10.1016/j.tplants.2004.09.002

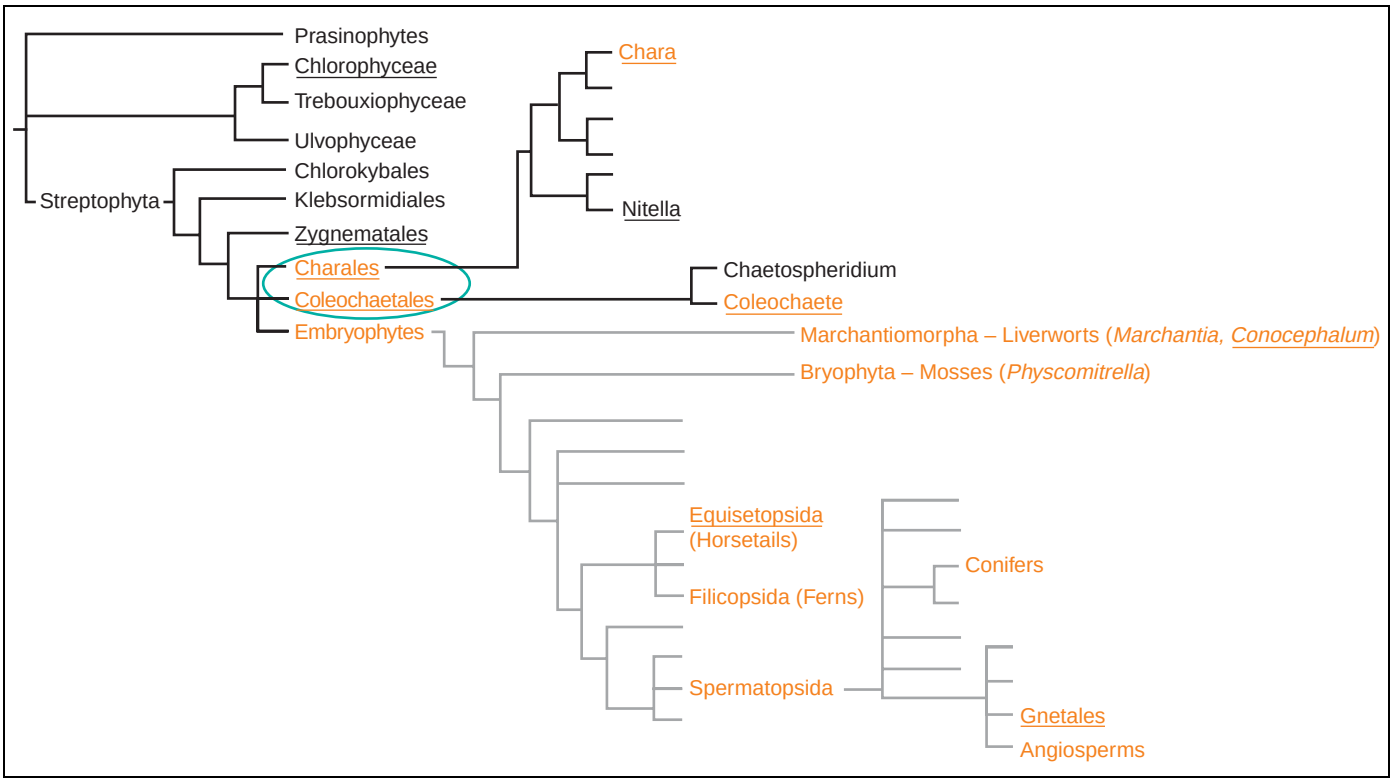


Figure 1. Plant taxonomic tree; the branches are not drawn to scale. The orange families include species that contain at least one sequence coding for a class III peroxidase or show a guaiacol peroxidase specific activity. Underlined families indicate that the presence or absence of peroxidase activity has been verified with the guaiacol peroxidase assay [5,75]. The unlabelled branches correspond to families in which no peroxidase encoding sequences have been reported to date. The circled area includes the organisms probably containing the first class III peroxidases. Modified from the tree of life (http://tolweb.org/).

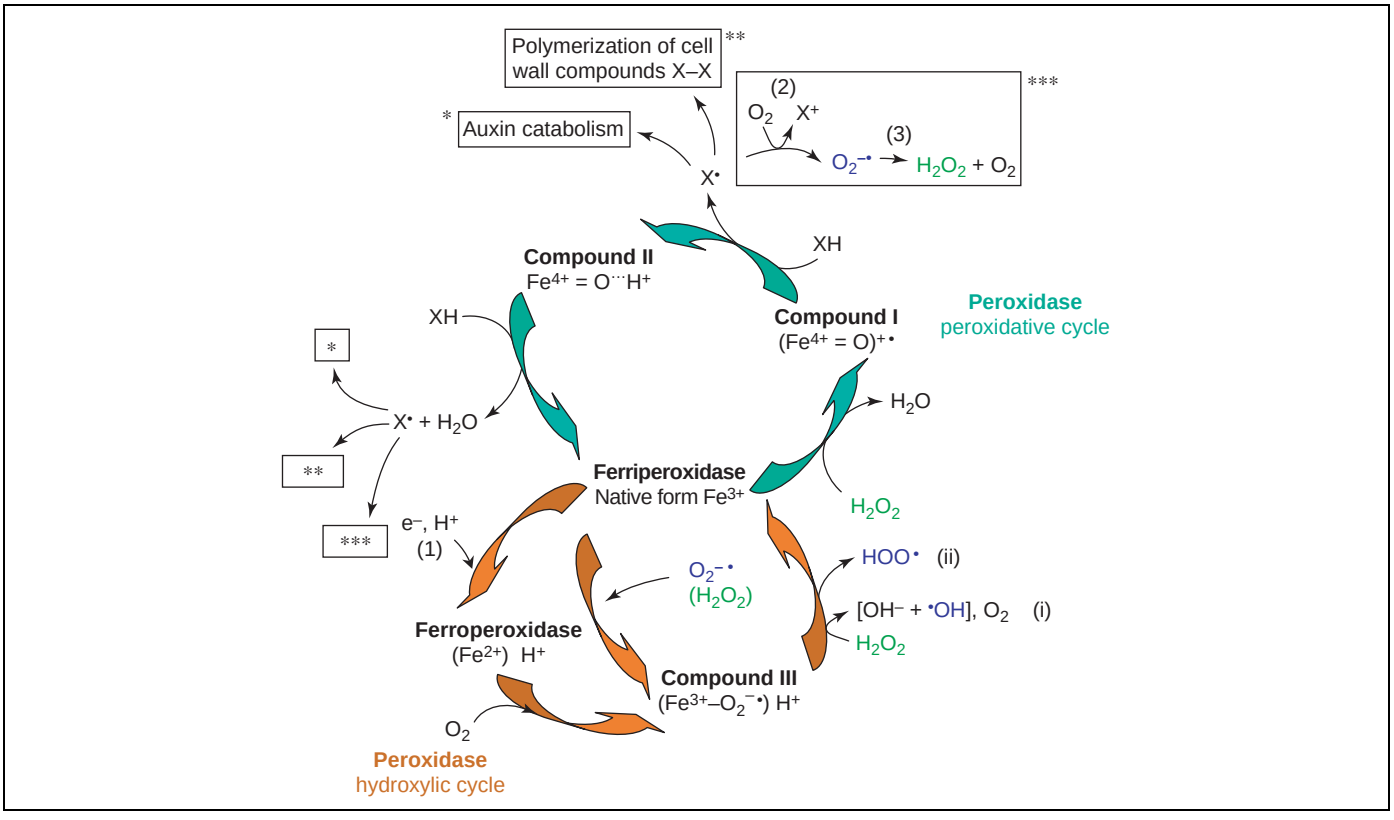


Figure 2. Class III peroxidase reaction cycles. The hydroxylic cycle (represented with orange arrows) can generate reactive oxygen species (ROS) such as ·OH and HOO· by two different routes (i,ii). The source of e^- and H^+ (1) can be auxin and other reducing molecules. The peroxidative cycle (represented in green) can oxidize various substrates, represented by XH and X· for the reduced and oxidized forms, respectively. X· has three major fates: auxin catabolism (*), cell wall component polymerization (**) and NAD(P)H oxidation (***) via a non-catalytic reaction (2). NAD(P)H oxidation produces superoxide, which is immediately converted either spontaneously or by superoxide dismutase to H_2O_2 and O_2 (3). Hydroxylic and peroxidative cycles can both regulate the H_2O_2 level. Modified from Refs [9,76].

context of growth and development through the lifespan of various plants.

Germination

Peroxidase activity can be detected in seeds and their mucilage before germination. By generating hydroxyl radicals (·OH), peroxidases could play a crucial role in seed protection as well as in the first days of germination by reducing pathogenic attack. Indeed, ROS are believed to play a role in defending seeds against pathogens. In this respect, two studies of tomato (*Lycopersicon esculentum*) and radish (*Raphanus sativus*) seeds confirm that peroxidase production begins as soon as the radicle protrudes, breaks down the endosperm and creates a 'wounded' region vulnerable to pathogens. A burst of ·OH radicals at this precise time would confer a protective advantage. The defensive action of peroxidases is probably accompanied by the activation of other defense enzymes, such as β-1,3-glucanase and chitinase [12,13].

Seed peroxidase activity is not caused by a single enzyme but rather by several with overlapping activities [14,15]. The expression of various isozymes in the seed during imbibition and in the first days of seedling growth suggests that peroxidases play roles other than defense during germination. For example, a recent study showed that a peroxidase was strongly induced in artificially aged seeds of radish, thereby protecting seeds against ageing-related lipid peroxidation [16]. Peroxidases might also participate in the lignification of new xylem elements in the embryo, hypocotyl and radicle. The production of ·OH radicals by peroxidases could further contribute to the breakage of the seed coat and the subsequent cell elongation.

Peroxidases can even participate in processes that happen before germination: a cationic peroxidase was able to restore normal embryogenesis to tunicamycin-treated carrot embryogenic cultures [17]. Tunicamycin, an inhibitor of glycosylation, caused an abnormal expansion of protoderm cells, which stopped embryo development at an early stage, indicating that this peroxidase plays a crucial role in maintaining the correct size and shape of protoderm cells.

Cellular growth and cell wall loosening

Growth by cell elongation (as opposed to growth by cell division) leads to an irreversible increase in cell volume that occurs together with a relaxation of the cell wall. The biochemical processes involved in cell wall loosening during extension growth are only partially known. Several enzymatic mechanisms have been proposed, such as the cleavage and reassembly of xyloglucan polymers by xyloglucan endotransglycosylase or the suppression of hydrogen bonds between cellulose and xyloglucan by expansins [18]. Changes in cell wall internal structure can also be achieved by peroxidases through their peroxidative or hydroxylic cycles.

In higher plants, both cycles contribute to the regulation of the endogenous hydrogen peroxide (H_2O_2) level (Figure 2). The amount of apoplastic H_2O_2 has been shown to be related to cell elongation. For example, in soybean, the H_2O_2 level increases strongly from the tip of the hypocotyl (elongation zone) to the highly lignified base [19]. In *Arabidopsis*, root curvature and the cellular growth occurring during root gravitropism are also downregulated by extracellular H_2O_2 [20]. Another example of this putative relationship between peroxidase and H_2O_2 level is illustrated by the mRNA encoding APRX, an anionic peroxidase from zucchini (*Cucurbita pepo*), which accumulated strongly in the elongation zone of the hypocotyl where the H_2O_2 level needs to be low. APRX accumulation was inversely correlated with lignin level [21]. Surprisingly, another study (of onion root fragments) showed that the H_2O_2 level was high in cells undergoing elongation and lignification, together with a lower level of total peroxidase activity in the elongation zone [22].

In addition to H_2O_2, ascorbate is also implicated in cellular elongation mechanisms. This function is not due to degradation of ascorbate by cytoplasmic class I peroxidases (ascorbate peroxidases) but to the reducing activity of this molecule onto the radicals produced by class III peroxidases. In onion root fragments, ascorbate and its free radical stimulates root elongation and expansion [23]. Ascorbate also has an effect on the mechanism of gravitropism in maize. It can reduce root curvature, which is directly related to the differential elongation process [20]. Elongation results from the balance between cell wall loosening (ROS activity) and stiffening (cross-linking of polymers). This balance, which probably differs between species, might explain the variation observed during ascorbate treatment.

Cell wall polysaccharides such as pectin and xyloglucan can be cleaved *in vitro* by the extremely reactive hydroxyl radicals [24]. *In vitro* cleavage of various polysaccharides by peroxidase-generated hydroxyl radicals has been demonstrated [25]. The production of ·OH in cell walls and in apoplastic spaces could cause non-enzymatic wall loosening [26]. High levels of ·OH, $O_2\cdot^-$ and H_2O_2 are produced in the leaf expansion zone of maize [27]. The $O_2\cdot^-$ released during the oxidative cycle by NADPH oxidase can convert peroxidase into compound III, which catalyzes the generation of ·OH from H_2O_2 in the cell wall (Figure 2) [9,25,28]. Addition of H_2O_2 can reverse the elongation inhibition caused by a ROS production inhibitor. Moreover, auxin promotes the release of $O_2\cdot^-$ and the subsequent production of ·OH through the hydroxylic cycle when inducing elongation growth. Auxin-mediated $O_2\cdot^-$ production by peroxidases uses O_2 instead of H_2O_2. Alternatively, ·OH can be generated by a mixture of H_2O_2 and ascorbate [29]. The short-term effect of such a mixture on living coleoptiles or upon isolated coleoptile cell walls is irreversible wall extension [30].

Taken together, these results show that peroxidases can regulate growth by elongation in different ways because of the two distinct cycles (Figure 3). They can favor elongation by generating oxygen radicals (hydroxylic cycle) or by regulating the local concentration of H_2O_2 (peroxidative and hydroxylic cycles). Alternatively, peroxidase activity can be controlled by radical-generating compounds such as ascorbate and H_2O_2, confirming the complex relationship between ascorbate, H_2O_2 and peroxidases.

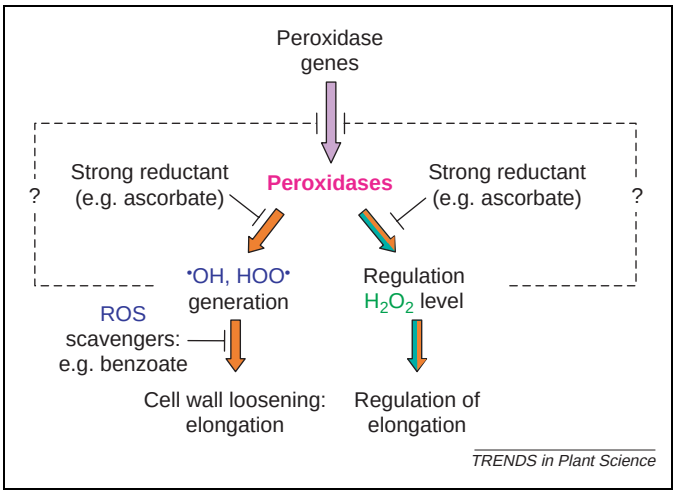


Figure 3. Role of peroxidases in cell elongation processes. Peroxidases are implicated in cell elongation via the generation of reactive oxygen species (ROS) through the hydroxylic cycle (orange arrows). Peroxidases can also regulate the H_2O_2 level through both catalytic cycles (orange and green arrows). The peroxidase activity and function are down regulated by various inhibitors and ROS scavengers and can be also controlled at the transcriptional level. Question marks indicate putative inhibition at the transcriptional level.

Cell wall cross-linking

Peroxidases can control the availability of H_2O_2 in the cell wall (Figure 2), which is a prerequisite for the cross-linking of phenolic groups, to inhibit cell elongation (Figure 4). Peroxidases catalyze this process in response to various external factors such as wounding, pathogen interactions and environmental constraints or just as a part of normal cell wall development during growth, differentiation and senescence. Indeed, dehydration and pathogen invasion can be limited by the formation of a physical barrier of lignin or suberin. Cell wall rigidification is the result of the peroxidase-mediated cross-linking of several compounds. Among them, polysaccharide-linked ferulates (bound to lignin or not), extensins (not bound to lignin) and, ultimately, lignin monomers form a complex network that solidifies the plant wall. The biochemical implication of peroxidases in such a mechanical process is related to the peroxidative cycle (Figure 4).

Diferulic bonds

The formation of diferulic linkages occurs at various stages during cell wall formation. Elongation and cell wall extensibility are inversely correlated to the increase in the content of ferulic and diferulic acid in the primary cell wall [31–33]. Indeed, during lignification, cell walls are further stiffened by the formation of diferulic linkages between polysaccharide-bound lignins or polysaccharides. Ferulate can also join a lignin moiety to a polysaccharide without forming a diferulic bond [34–36]. Peroxidases are the major players in the formation of diferulic bonds [37,38]. There are many possible isomers of diferulate units, the proportion of each isomer varying between different plant species. Diferulic bonds are also formed photochemically; the resulting cyclodimer, although present at a significant proportion, is not likely to be the dominant diferulic species [37]. Diferulic bridges between sugars are formed at two stages: in the protoplasm (probably in the Golgi body) and in the cell wall [39]. Polysaccharide-linked diferulates are exported from the Golgi apparatus to the cell wall, where they contribute to its stiffening. Peroxidases are only acting at the cell wall level. Therefore, rigidification of the cell wall will slow down plant growth and also participate in other functions, such as defense against pathogen intrusion.

Extensins

Extensins are hydroxyproline-rich proteins (HPRP) and are essential for primary cell wall structure and development in plants [40–42]. There is strong evidence that peroxidases work closely with extensins during the

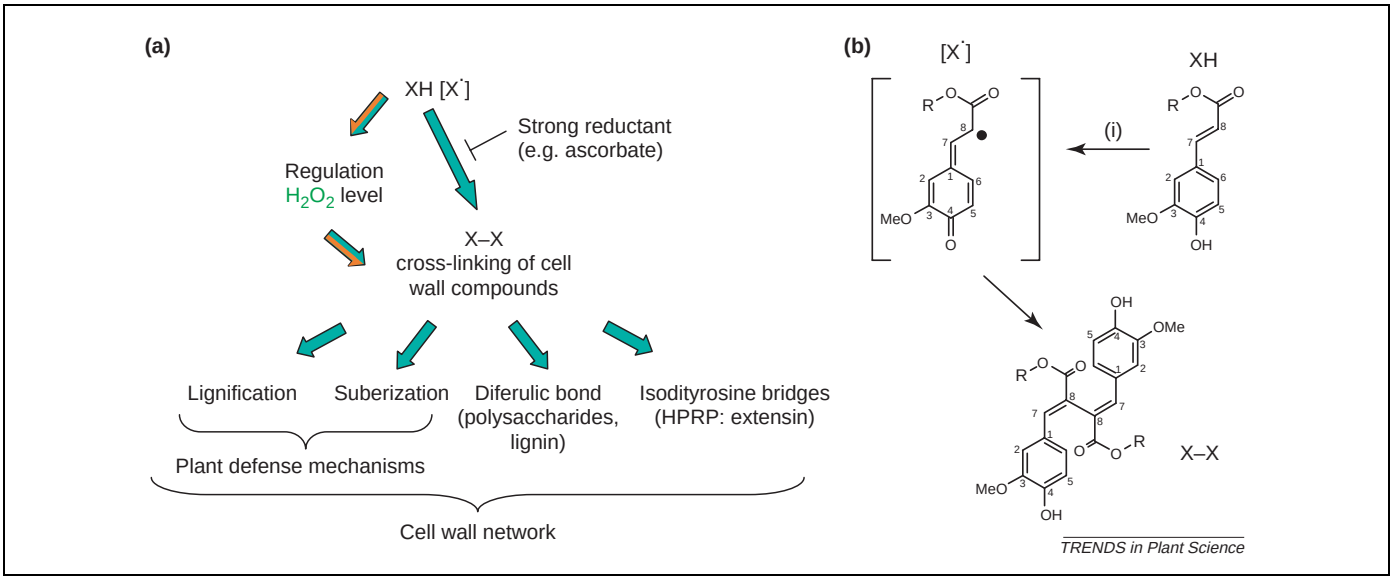


Figure 4. Peroxidases in the cell wall cross-linking mechanism. **(a)** Through the peroxidative cycle (green arrows), the peroxidases oxidize various substrates (XH) such as tyrosine residues, monolignols, suberin units and ferulic acids (some of these molecules being linked to various polymers). The radicals produced by the peroxidative cycle can form cross-links (X–X) between cell wall polymers and proteins implicated in plant defense reactions, and during elaboration of the cell wall network. Peroxidases can also regulate the H_2O_2 level through both catalytic cycles (orange and green arrows). **(b)** Example of a peroxidase-mediated cross-linking. Peroxidase is indicated by (i). Abbreviations: R, bound polysaccharide; XH, ferulate ester; X•, ferulate ester radical intermediate (only one isomer is shown); X–X, 8–8-coupled dehydroferulate ester. Coupling can occur at other positions, giving 5–5-, 8–5- or 8–*O*–4- coupled dehydroferulate esters.

process of cell wall stiffening. Peroxidases can create rigid extensin cross-links that block any further wall loosening and hence cell expansion [43]. The precise nature of these cross-links is still being debated: isodityrosine bonds or Tyr–Lys bonds have been implicated in this process [44–46]. So far, extensin cross-linking peroxidases have been isolated independently from tomato cell cultures [43,44], from mustard (*Sinapis alba*) [47], from grapevine (*Vitis vinifera*) [48] and from lupin (*Lupinus albus*) [49]. Tomato peroxidase exhibited a strong activity not only with tomato extensins but also with other plant extensins, such as those from *Ginkgo*, sugarbeet or carrot [44]. It is interesting that all these heterologous extensins had a Val–Tyr–Lys sequence; other extensins lacking this motif were not cross-linked by the isolated tomato peroxidase. Peroxidase activity might be modulated *in vivo*, as shown by a study that used a natural inhibitor of extensin cross-linking activity isolated from a cell wall compartment containing extensins and peroxidases [50,51]. In the presence of this inhibitor, tomato plants experienced significantly enhanced growth. The common relationship between extensins and peroxidases could be a pectin polymer. Indeed, some peroxidases can bind to calcium–pectate complexes [52,53] and there is growing evidence that covalent bonds exist between pectins and extensins [54]. Pectins would act as an anchor for peroxidases, which would then cross-link extensins to create a dense, solid network in the plant wall. The interaction of peroxidases with extensins has two main functions: structure and defense (by making the cell wall harder to penetrate). Defense is induced by elicitor molecules coming from pathogens in the initial steps of the attack [48]. One study showed that an extensin cross-linking peroxidase from mustard was under the control of a phytochrome, indicating a possible influence of light on peroxidase-mediated plant growth [47,55]. Other functions have been suggested for extensins, such as definition of cell morphology, embryogenesis or adaptation of cell wall to mechanical stress [49]. All these functions probably need the participation of peroxidases, underlining their role in virtually all stages of plant development.

Lignification and suberization

Cross-linking of phenolic monomers in the formation of suberin and the oxidative coupling of lignin subunits as part of lignin biosynthesis are related to secondary cell wall formation [56] and have been associated with reduction of extensibility and growth. Peroxidases [57] and laccases [58] are both candidates for monolignol unit oxidation, which leads to the final step of lignin assembly. Using H_2O_2 as an oxidant, peroxidases can generate monolignol phenoxy radicals that couple spontaneously to form lignin polymers. For example, the expression of the tobacco peroxidase TP60 is coordinated with lignin pathway-specific enzymes such as ferulate-5-hydroxylase, cinnamoyl CoA reductase and cinnamyl alcohol dehydrogenase [56]. Downregulation of TP60 expression using an antisense strategy leads to the reduction of lignin content, associated with a decrease of both guaiacyl and syringyl monolignols [56]. The promoter of the gene encoding prxA3a, an anionic peroxidase from an hybrid aspen (*Populus sieboldii* × *Populus grandidentata*), is active mainly in the lignifying cells of stem tissues, and the antisense construct induced a similar regulation to that observed with TP60 [59]. Additionally, HRP A2 from horseradish, AtPrx53 (ATPA2) from *Arabidopsis*, various anionic peroxidases from poplar and cationic peroxidases from tomato (TPX1) and *Zinnia* have also been shown to be involved in the lignification process [60–64]. The *in vitro* oxidation of syringaldazine by certain peroxidases and their specific expression in lignifying plant tissues both argue for their role in the lignin biosynthetic pathways.

It is interesting that the overexpression of a putative tobacco lignoperoxidase in tobacco resulted in plants that were smaller than the controls [65]. The same peroxidase overexpressed in tomato, sweetgum (*Liquidambar styraciflua*) and tobacco also showed a protective effect against insect attacks [66]. The most likely explanation for this observation is that the transgenic peroxidase was able to stiffen the tissues of the plants by accumulating lignins, rendering the plants tougher to eat.

Similar conclusions can be drawn about the relationship between peroxidases and suberization [67,68]: peroxidases participate in suberization through the same mechanism by the spontaneous coupling of peroxidase-generated phenoxy radicals.

Senescence

Senescence is the final stage before complete tissue death. It is dependent on various plant hormones, such as cytokinins, which prevent the onset of senescence, or ethylene and salicylic acid, which by contrast induce the senescent phenotype *in vivo* [69]. Several morphological and cellular changes are associated with this mechanism. For example, in some plants, leaves fall off the stem (abscission) before complete death. Moreover, changes in gene expression have been reported during senescence: a decrease of photosynthesis-associated genes and an increase in senescence-associated genes [70].

ROS such as $O_2\cdot^-$ have been shown to be involved in the induction and development of the senescence stage. $O_2\cdot^-$ could be generated by extracellular peroxidase activity after salicylic acid treatment of tobacco suspension cell cultures and thus could act in the senescence induction pathway [71]. In addition, total peroxidase activity rises during *Arabidopsis* leaf expansion and in the senescent tissues, with a strong increase at the end of leaf expansion [70]. It is difficult to dissociate leaf senescence from cell wall stiffening and cross-linking. Indeed, in the first 4 weeks of the life of *Arabidopsis* plants, peroxidase activity in leaves is associated with cell extension and, to a minor extent, the lignification of the vascular system. Weeks 4 and 5 correspond to the transition phase and, after 5 weeks, peroxidases are mainly associated with the cross-linking of the cell wall components [70]. The higher peroxidase activity could be associated with the reduction of H_2O_2 and the increase in $O_2\cdot^-$. In agreement with this hypothesis, the chemical induction of a delay in senescence in pepper plants (*Capsicum annuum*) provokes a decrease in total peroxidase activity [72]. In striking contrast to this, another

study showed that KCl-induced senescence of sunflower (*Helianthus annuus*) plants and calli is also accompanied by a decrease in total peroxidase activity [73]. This discrepancy is probably related to the non-physiological levels of KCl used in this study. Indeed, high KCl concentrations could have various effects independent of senescence, such as a strong stress, thus modifying the interpretation of the results. However, in our view, peroxidases are involved in senescence via their implication in cell wall control, first by generating ROS as possible signaling molecules and, second, by cross-linking the components in the secondary cell wall.

Conclusions

To date, studies of plant peroxidases have provided evidence that class III peroxidases are omnipresent enzymes, expressed throughout the plant life cycle from germination to senescence. They can induce cell wall loosening and growth by elongation as well as cross-linking of cell wall components. Indeed, the balance between cleavage and cross-linking is associated with ROS action and with H_2O_2 and ascorbate concentrations. These compounds can regulate enzyme activity as well as peroxidase gene expression. The specific function of each isoform remains unclear and can even vary following the modification of external conditions such as temperature, pH or the presence of various substrates. A peroxidase might shift *in vitro* from an extensin cross-linking activity to another role just by changing the H_2O_2 concentration or pH in the reaction solution [49]. The many facets of peroxidases are still one of the main obstacles to finding specific functions of this family of proteins *in vivo*.

Untangling the apparent complexity of peroxidase functions, when considered as a whole, is a challenge. Integrating the study of T-DNA mutant collections, the exhaustive and specific expression profiles using cDNA arrays [74], the building of reliable evolutionary relationships (ancestral peroxidase sequences) [5] and synthetic dedicated databases (http://www.unige.ch/LABPV/perox/ or http://peroxidase.isb-sib.ch/) appear to be a necessary strategy to one day deciphering the function of each unique peroxidase.

Acknowledgements

We thank William Deakin and Xavier Perret for their critical reading and constructive suggestions, and Olivier Lustenberger, Laurent Falquet and Vassilios Ioannidis for the realization of the peroxidase website. We thank the University of Geneva and the Swiss National Science Foundation (grant 31-068003.02) for research support to C.P. and C.D.

References

1 Welinder, K.G. (1992) Plant peroxidases: structure–function relationships. In *Plant Peroxidases* (Penel, C. *et al*., eds), pp. 1–24, University of Geneva, Switzerland
2 Schuller, D.J. *et al*. (1996) The crystal structure of peanut peroxidase. *Structure* 4, 311–321
3 Tognolli, M. *et al*. (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. *Gene* 288, 129–138
4 Duroux, L. and Welinder, K.G. (2003) The peroxidase gene family in plants: a phylogenetic overview. *J. Mol. Evol.* 57, 397–407
5 Passardi, F. *et al*. (2004) The plant peroxidase multigenic family in rice and its evolution in green plants. *Phytochemistry* 65, 1879–1893
6 Zhang, J. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298
7 Hiraga, S. *et al*. (2001) A large family of class III plant peroxidases. *Plant Cell Physiol.* 42, 462–468
8 Penel, C. *et al*. (1992) *Plant Peroxidases*, University of Geneva, Switzerland
9 Liszkay, A. *et al*. (2003) Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth. *Planta* 217, 658–667
10 Berglund, G.I. *et al*. (2002) The catalytic pathway of horseradish peroxidase at high resolution. *Nature* 417, 463–468
11 Halliwell, B. (1977) Generation of hydrogen peroxide, superoxide and hydroxyl radicals during the oxidation of dihydroxyfumaric acid by peroxidase. *Biochem. J.* 163, 441–448
12 Morohashi, Y. and Matsushima, H. (2000) Development of beta-1,3-glucanase activity in germinated tomato seeds. *J. Exp. Bot.* 51, 1381–1387
13 Wu, C.T. *et al*. (2001) Class I beta-1,3-glucanase and chitinase are expressed in the micropylar endosperm of tomato seeds prior to radicle emergence. *Plant Physiol.* 126, 1299–1313
14 Bellani, L.M. *et al*. (2002) Differences in the activity and distribution of peroxidases from three different portions of germinating *Brassica oleracea* seeds. *Physiol. Plant.* 114, 102–108
15 Gijzen, M. (1997) Seed peroxidases. *Plant Perox. Newslett.* 10, 30–33
16 Scialabba, A. *et al*. (2002) Effects of ageing on peroxidase activity and localization in radish (*Raphanus sativus* L.) seeds. *Eur. J. Histochem.* 46, 351–358
17 Cordewener, J. *et al*. (1991) Tunicamycin-inhibited carrot somatic embryogenesis can be restored by secreted cationic peroxidase isoenzymes. *Planta* 184, 478–486
18 Cosgrove, D.J. (2001) Wall structure and wall loosening. A look backwards and forwards. *Plant Physiol.* 125, 131–134
19 Schopfer, P. (1994) Histochemical demonstration and localization of H_2O_2 in organs of higher plants by tissue printing on nitrocellulose paper. *Plant Physiol.* 104, 1269–1275
20 Joo, J.H. *et al*. (2001) Role of auxin-induced reactive oxygen species in root gravitropism. *Plant Physiol.* 126, 1055–1060
21 Dunand, C. *et al*. (2003) Expression of a peroxidase gene in zucchini in relation with hypocotyl growth. *Plant Physiol. Biochem.* 41, 805–811
22 Cordoba-Pedregosa, M. *et al*. (2003) Zonal changes in ascorbate and hydrogen peroxide contents, peroxidase, and ascorbate-related enzyme activities in onion roots. *Plant Physiol.* 131, 697–706
23 Cordoba-Pedregosa, M. *et al*. (1996) Role of apoplastic and cell-wall peroxidases on the stimulation of root elongation by ascorbate. *Plant Physiol.* 112, 1119–1125
24 Fry, S.C. (1998) Oxidative scission of plant cell wall polysaccharides by ascorbate-induced hydroxyl radicals. *Biochem. J.* 332, 507–515
25 Schweikert, C. *et al*. (2000) Scission of polysaccharides by peroxidase-generated hydroxyl radicals. *Phytochemistry* 53, 565–570
26 Chen, S.X. and Schopfer, P. (1999) Hydroxyl-radical production in physiological reactions. A novel function of peroxidase. *Eur. J. Biochem.* 260, 726–735
27 Rodriguez, A.A. *et al*. (2002) Reactive oxygen species in the elongation zone of maize leaves are necessary for leaf extension. *Plant Physiol.* 129, 1627–1632
28 Schopfer, P. *et al*. (2002) Evidence that hydroxyl radicals mediate auxin-induced extension growth. *Planta* 214, 821–828
29 Miller, J.G. and Fry, S.C. (2001) Characteristics of xyloglucan after attack by hydroxyl radicals. *Carbohydr. Res.* 332, 389–403
30 Schopfer, P. (2001) Hydroxyl radical-induced cell-wall loosening *in vitro* and *in vivo*: implications for the control of elongation growth. *Plant J.* 28, 679–688
31 Kawamura, Y. *et al*. (2000) Stress-relaxation analysis of submerged and air-grown rice coleoptiles: correlation with cell wall biosynthesis and growth. *J. Plant Physiol.* 156, 689–694
32 MacAdam, J.W. and Grabber, J.H. (2002) Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. *Planta* 215, 785–793
33 Sanchez, M. *et al*. (1996) Changes in dehydrodiferulic acids and peroxidase activity against ferulic acid associated with cell walls during growth of *Pinus pinaster* hypocotyl. *Plant Physiol.* 111, 941–946

34 Grabber, J.H. *et al*. (2000) Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *J. Agric. Food Chem.* 48, 6106–6113
35 Grabber, J.H. *et al*. (2002) Model studies of ferulate–coniferyl alcohol cross-product formation in primary maize walls: implications for lignification in grasses. *J. Agric. Food Chem.* 50, 6008–6016
36 Iiyama, K. *et al*. (1994) Covalent cross-links in the cell wall. *Plant Physiol.* 104, 315–320
37 Ralph, J. *et al*. (1994) Identification and synthesis of new ferulic dehydrodimers present in grass cell walls. *J. Chem. Soc., Perkin Trans.* 1, 3485–3498
38 Veitch, N.C. (2004) Horseradish peroxidase: a modern view of a classic enzyme. *Phytochemistry* 65, 249–259
39 Fry, S.C. *et al*. (2000) Intraprotoplasmic and wall-localised formation of arabinoxylan-bound diferulates and larger ferulate coupling-products in maize cell-suspension cultures. *Planta* 211, 679–692
40 Showalter, A.M. (1993) Structure and function of plant cell wall proteins. *Plant Cell* 5, 9–23
41 De Tullio, M.C. *et al*. (1999) Changes in onion root development induced by the inhibition of peptidyl-prolyl hydroxylase and influence of the ascorbate system on cell division and elongation. *Planta* 209, 424–434
42 Cooper, J.B. *et al*. (1994) 3,4-Dehydroproline inhibits cell wall assembly and cell division in tobacco protoplasts. *Plant Physiol.* 104, 747–752
43 Brownleader, M.D. *et al*. (1995) Purification and partial characterization of tomato extensin peroxidase. *Plant Physiol.* 109, 1115–1123
44 Schnabelrauch, L.S. *et al*. (1996) Isolation of pl 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val–Tyr–Lys as putative intermolecular cross-link site. *Plant J.* 9, 477–489
45 van Huystee, R.B. and Zheng, X. (1995) Plant peroxidase, its location and extensin, coniferyl oxidation. *Plant Physiol. Biochem.* 33, 55–60
46 Kieliszewski, M.J. and Lamport, D.T. (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J.* 5, 157–172
47 Magliano, T.M.A. and Casal, J.J. (1998) *In vitro* cross-linking of extensin precursors by mustard extracellular isoforms of peroxidase that respond either to phytochrome or to wounding. *J. Exp. Bot.* 49, 1491–1499
48 Jackson, P.A. *et al*. (2001) Rapid deposition of extensin during the elicitation of grapevine callus cultures is specifically catalyzed by a 40-kiloDalton peroxidase. *Plant Physiol.* 127, 1065–1076
49 Price, N.J. *et al*. (2003) A biochemical and molecular characterization of LEP1, an extensin peroxidase from lupin. *J. Biol. Chem.* 278, 41389–41399
50 Brownleader, M.D. and Dey, P.M. (1993) Purification of extensin from cell walls of tomato (hybrid of *Lycopersicon esculentum* and *L. peruvianum*) cells in suspension culture. *Planta* 191, 457–469
51 Brownleader, M.D. *et al*. (2000) Role of extensin peroxidase in tomato (*Lycopersicon esculentum* Mill.) seedling growth. *Planta* 210, 668–676
52 Dunand, C. *et al*. (2002) Identification and characterisation of Ca^{2+}–pectate binding peroxidases in *Arabidopsis thaliana*. *J. Plant Physiol.* 159, 1165–1171
53 Shah, K. *et al*. (2004) Purification and identification of a Ca^{2+}–pectate binding peroxidase from *Arabidopsis* leaves. *Phytochemistry* 65, 307–312
54 Qi, X. *et al*. (1995) Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin. *Plant Physiol.* 108, 1691–1701
55 Casal, J.J. *et al*. (1990) Phytochrome control of extracellular peroxidase activity in mustard internodes: correlation with growth, and comparison with the effect of wounding. *Photochem. Photobiol.* 52, 165–172
56 Blee, K.A. *et al*. (2003) A lignin-specific peroxidase in tobacco whose antisense suppression leads to vascular tissue modification. *Phytochemistry* 64, 163–176
57 Lewis, N.G. and Yamamoto, E. (1990) Lignin: occurrence, biogenesis and biodegradation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41, 455–496
58 Kiefer-Meyer, M.C. *et al*. (1996) Cloning and sequence analysis of laccase-encoding cDNA clones from tobacco. *Gene* 178, 205–207
59 Li, Y. *et al*. (2003) Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. *J. Plant Res.* 116, 175–182
60 Nielsen, K.L. *et al*. (2001) Differential activity and structure of highly similar peroxidases. Spectroscopic, crystallographic, and enzymatic analyses of lignifying *Arabidopsis thaliana* peroxidase A2 and horseradish peroxidase A2. *Biochemistry* 40, 11013–11021
61 Ostergaard, L. *et al*. (2000) *Arabidopsis* ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. *Plant Mol. Biol.* 44, 231–243
62 Christensen, J.H. *et al*. (2001) The syringaldazine-oxidizing peroxidase PXP 3-4 from poplar xylem: cDNA isolation, characterization and expression. *Plant Mol. Biol.* 47, 581–593
63 Quiroga, M. *et al*. (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. *Plant Physiol.* 122, 1119–1127
64 Lopez-Serrano, M. *et al*. (2004) *Zinnia elegans* uses the same peroxidase isoenzyme complement for cell wall lignification in both single-cell tracheary elements and xylem vessels. *J. Exp. Bot.* 55, 423–431
65 Lagrimini, L.M. *et al*. (1997) Characterization of antisense transformed plants deficient in the tobacco anionic peroxidase. *Plant Physiol.* 114, 1187–1196
66 Dowd, P.F. *et al*. (2000) Mechanisms of insect resistance in transgenic plants (over) expressing a tobacco anionic peroxidase. *Plant Perox. Newslett.* 14, 93–101
67 Bernards, M.A. *et al*. (1999) Biochemical characterization of the suberization-associated anionic peroxidase of potato. *Plant Physiol.* 121, 135–146
68 Keren-Keiserman, A. *et al*. (2004) Peroxidase activity associated with suberization processes of the muskmelon (*Cucumis melo*) rind. *Physiol. Plant.* 121, 141–148
69 Morris, K. *et al*. (2000) Salicylic acid has a role in regulating gene expression during leaf senescence. *Plant J.* 23, 677–685
70 Abarca, D. *et al*. (2001) Differential leaf stress responses in young and senescent plants. *Physiol. Plant.* 113, 409–415
71 Kawano, T. *et al*. (1998) Salicylic acid induces extracellular superoxide generation followed by an increase in cytosolic calcium ion in tobacco suspension culture: the earliest events in salicylic acid signal transduction. *Plant Cell Physiol.* 39, 721–730
72 Flors, V. *et al*. (2001) Effect of a novel chemical mixture on senescence processes and plant–fungus interaction in Solanaceae plants. *J. Agric. Food Chem.* 49, 2569–2575
73 Santos, C.L. *et al*. (2001) *In situ* and *in vitro* senescence induced by KCl stress: nutritional imbalance, lipid peroxidation and antioxidant metabolism. *J. Exp. Bot.* 52, 351–360
74 Valério, L. *et al*. (2004) Expression analysis of the *Arabidopsis* peroxidase multigenic family. *Phytochemistry* 65, 1331–1342
75 Penel, C. and Greppin, H. (1996) Pectin binding proteins: characterization of the binding and comparison with heparin. *Plant Physiol. Biochem.* 34, 479–488
76 Kawano, T. (2003) Roles of the reactive oxygen species-generating peroxidase reactions in plant defense and growth induction. *Plant Cell Rep.* 21, 829–837

2

Functional study of AtPrx33 and AtPrx34

A. Introduction

AtPrx33 peroxidase was identified for the first time in *Arabidopsis thaliana* from a genomic DNA library, and was named prxCa (Intapruk et al., 1991). In a later study performed in our laboratory, several clones encoding peroxidase sequences were identified from an *Arabidopsis* cDNA library. Eleven peroxidase cDNA sequences were further transfected into insect cells through baculovirus infection. Among these peroxidases, some were found to strongly bind to a Ca^{2+}-pectate gel (Dunand et al., 2002). In a previous study, a pectate-binding domain had been identified in the zucchini peroxidase *APRX*: $R[X]_5R[X]_2R$ (Carpin et al., 2001). All 73 *Arabidopsis* sequences were hence screened for presence of this motif, and the sequence $K[X]_5K[X]_2K$ was identified in peroxidases AtPrx32, 33, 34, 37 and 38. This motif was strongly suspected to play the same role as the APRX motif. Confirming this hypothesis, one of the peroxidases used during pectate-binding assays was *AtPrx32*; furthermore, AtPrx34 was purified from a Ca^{2+}-pectate column in our laboratory (Shah et al., 2004). Other peroxidases that had a strong pectate-binding affinity did not possess the $K[X]_5K[X]_2K$ motif, thus showing that other amino acids can come into play. Besides presence of the $K[X]_5K[X]_2K$ motif, AtPrx32, 33, 34, 37 and 38 showed a high degree of identity between each other, forming a well-supported cluster in an AtPrx phylogenetic tree (Tognolli et al., 2002). Finally, these peroxidases also bore a C-terminal extension (rarely found in the other *AtPrx* genes), like APRX, but without any relevant sequence identity to *APRX*. C-terminal extensions are suspected to direct proteins to the vacuole (Neuhaus, 1996; Carter et al., 2004), however there is still a debate on whether once in the vacuole, the peroxidase can be secreted or remains intracellular (Matsui et al., 2003; Kis et al., 2004; Matsui et al., 2006).

Following these observations, a part of my thesis work consisted in identifying the cellular localisation of AtPrx33 in *Arabidopsis thaliana*, in order to confirm the *in silico* predicted pectate-binding activity. Ca^{2+}-pectate is indeed a key constituent of primary cell walls, in particular the middle lamella (Minor, 1991; Carpita and Gibeaut, 1993; Wi et al., 2005). At the same time, cellular localisation of AtPrx34 was also assayed by Christophe. When this aim would have been achieved, a second project was to monitor expression of the *AtPrx33* promoter *in planta*. Finally, I would have to express and purify AtPrx33 in *Escherichia coli* or in the baculovirus-insect cell system (Carpin et al., 2001), in order to perform pectate-binding assays.

In parallel to pectate-binding assays, the genomes of T-DNA insertion libraries in *Arabidopsis thaliana* were screened for peroxidase mutants by a former PhD student (Michael Tognolli) and, among three others, the knock-out mutant *atprx33* was isolated. *atprx33* was analysed for particular phenotypic traits: Michael observed that a callus spontaneously formed from leaves of a 6 week-old *atprx33* plant grown on MS medium. This result suggested that the hormonal balance of *atprx33* was deregulated, in particular concerning auxin, a known substrate of peroxidases. Formation of calli is indeed dependent on the balance between auxins and cytokinins (Skoog and Miller, 1957). The *atprx33* callus was further cultured on a medium containing the synthetic cytokinin 6-benzylaminopurine (BAP) and the synthetic auxin dichlorophenoxyacetic acid (2,4D) in order to sustain callus growth. When the callus was transferred to hormone-free medium, it developed roots: this differentiation was considered as due to a non-natural increase in endogenous auxin concentration and was attributed to the loss of *AtPrx33* expression.

Consequently to the observation of the properties of *atprx33* callus formation and differentiation, one of the goals of my thesis was to better understand differentiation mechanisms of *atprx33* and wild type calli, in order to verify and refine Michael's observation. Moreover, following the hypothesis that auxin regulation could be perturbed in the *atprx33* mutant, I transformed wild type and *atprx33 Arabidopsis* plants with the glucuronidase reporter gene fused to the auxin-responsive GH3 promoter.

B. AtPrx33 and AtPrx34 expression, cellular localisation, and function in root elongation

B.1. Isolation and preliminary studies on various transgenic lines related to AtPrx33 *and* AtPrx34

During his thesis, Michael isolated, from a screen of two T-DNA insertion mutant libraries, four distinct peroxidase knock-out lines: *atprx10* and *atprx33* (Wassilewskija ecotype, Ws); *atprx3* and *atprx70* (Columbia ecotype, Col). He then meticulously observed the two first mutants for phenotypic differences compared to a wild type control. He focused his attention on roots, as he had previously shown, in agreement with the literature (Hiraga et al., 2000), that peroxidases were mostly expressed in roots. He saw no difference in root architecture during vertical growth of seedlings on MS agar. When incubating transversal cuts of 2 week-old roots with the lignin-specific dye toluidine blue, he noted that the *atprx33* xylem cells of the stele (stained by toluidine blue) were fewer and bigger than in wild type controls, suggesting a role for AtPrx33 in root cell growth. The most striking phenotype was spontaneous callus formation, but he could only see it happen once (see previous section).

Concerning the *atprx34* mutant, it was isolated by Christophe in the Columbia ecotype of the SALK T-DNA insertion lines library. Christophe also made the double-mutant *atprx33/34* by transforming *atprx33* with the RNAi construct *43xrPtA-linker-AtPrx34*. As soon as he obtained heterozygous lines of the double-mutant, he observed a dramatic drop in root length of various *atprx33/34* lines. One of my tasks was hence to study if there were truly differences in root length between wild type controls and *atprx33*, *atprx34*, and *atprx33/34*. After having set up proper conditions (no antibiotics, use of control with corresponding ecotype, monitoring of light conditions), I noted that *atprx33* roots were shorter than a Ws wild type control in 7 day-old seedlings, and that indeed, *atprx33/34* roots were even shorter.

I further made two new transgenic lines in order to verify the hypothesis for a role of AtPrx33 in root elongation. One line was 35S::*AtPrx33*::*GFP*, and the other *atprx33* complemented with AtPrx33::*AtPrx33*::*GFP* (promoter-gene-GFP). After introduction of the transgenes by the spraying technique into a Ws wild type and the *atprx33* mutant, respectively, I obtained homozygous lines through antibiotic selection. Overexpression was demonstrated through GFP fluorescence: the 35S::*AtPrx33*::*GFP* line was used, in parallel

with the 35S::*AtPrx34*::*GFP* line obtained by Christophe, to locate AtPrx33 in the cell wall. Complementation was assayed by RT-PCR reactions (Fig.2-1).

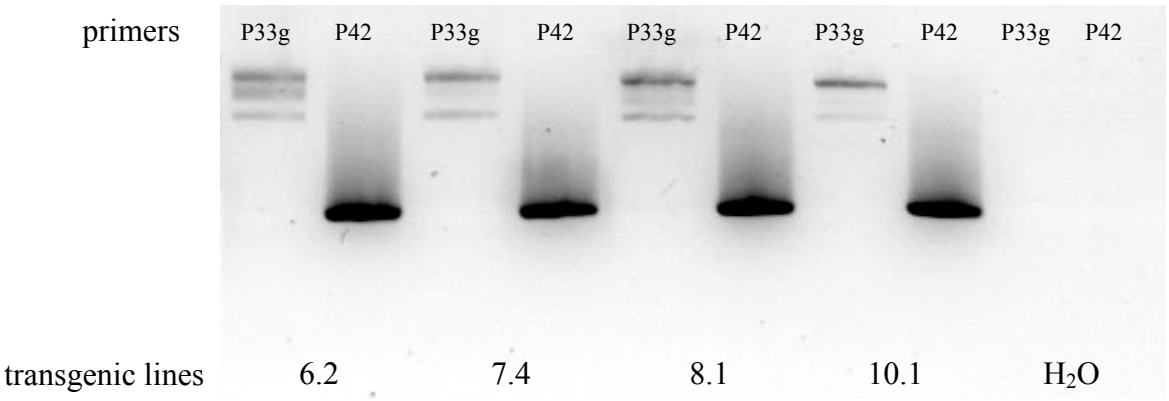


Figure 2-1

RT-PCR reactions with the T3 generation of *atprx33* complemented with AtPrx33::*AtPrx33*::*GFP*.

Four different lines (6.1, 7.4, 8.1 and 10.1) were assayed with primers specific for *AtPrx33*::*GFP* (P33g) or the constitutive peroxidase control *AtPrx42* (P42).

The detail of the results obtained is described in the article "**Two cell wall associated peroxidases from *Arabidopsis* influence root elongation**".

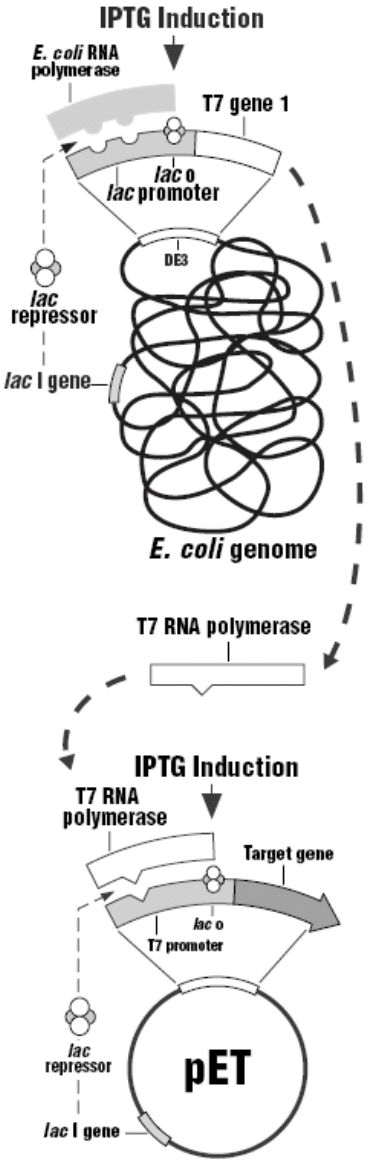
B.2. Expression of AtPrx33 in Escherichia coli *and baculovirus-infected insect cells*

In an earlier study from our laboratory (Dunand et al., 2002), it had been shown that peroxidase AtPrx32 had, among other peroxidases, a notable capacity to bind Ca^{2+}-pectate complexes, even in strong saline (800mM NaCl) conditions. As this peroxidase is a very close ortholog of AtPrx33, I tried to express AtPrx33 in order to compare the properties of both peroxidases in various *in vitro* experiments. The two proteins have a high level of identity (87%), hence this comparison would have brought valuable information on the role of specific amino acids in biochemical properties of peroxidases. This study would have been followed by expression of two other close paralogs: AtPrx34 and AtPrx37.

B.2.1 *Escherichia coli*

Two strategies were considered: expression through *Escherichia coli* or through the baculovirus-insect cell system (used for AtPrx32). We chose to start with prokaryotic production: this method generates non-glycosylated proteins, thus allowing to eventually making a crystal of AtPrx33. In terms of yield, it is generally equivalent to eukaryotic cell systems (Veitch, 2004). More crude protein is produced in bacteria: however, two more steps decrease the final yield. First, proteins must be extracted from bacterial inclusion bodies by solubilisation in high concentration of urea. Secondly, proteins are produced in their native form, and must be refolded by diluting urea and adding hemin (precursor of heme), oxidised glutathione (for disulphide bridges) and calcium (necessary for peroxidase structure).

In order to have a high rate of protein expression, I used the pET29a vector expression system in two BL21(DE3) *E.coli* lines. BL21(DE3) lines have been engineered in order to encode a T7 RNA polymerase under the control of the *lac* promoter. Moreover, they possess a *lac* I gene encoding a repressor of the *lac* promoter. This repressor binds to the *lac* operator and inhibits transcription of the T7 gene. The pET29a vector contains a T7 promoter under the control of a *lac* operator, an additional copy of *lac* I and a multiple cloning site (MCS) for insertion of a target gene (*AtPrx33*). When isopropyl-D-thiogalactopyranoside (IPTG) is added, *lac* repressor binds to IPTG, which relieves *lac* promoter inhibition, and the subsequently produced T7 RNA polymerase can induce transcription of the target gene (Fig.2-2). A problem often encountered when expressing eukaryotic proteins in bacteria is differential codon usage: bacteria rarely employ tRNAs for AGA and AGG (Arg) codons, as well as for AUA (Ile), CUA (Leu) and CCC (Pro) codons.



Except for codon AUA, all other codons are found between twice (CCC) and nine times (AGA) in *AtPrx33* cDNA sequence. Two different BL21(DE3) lines were hence transformed with pET29a::*AtPrx33*: "RIL" (encodes tRNAs for Arg, Ile, Leu) and "RP" (Arg, Pro) strains. Both supernatant and pellet of IPTG-induced bacterial cultures were collected. Despite several parameters were varied (IPTG concentration, growth temperature, induction duration), I could never see any induction of *AtPrx33* expression on SDS-PAGE gels. The decision was hence taken to produce AtPrx33 through the baculovirus-insect cell system.

Figure 2-2 (from pET System Manual, 11th ed., Novagen)

pET29a expression system in BL21(DE3) *E.coli*.

The *lac* repressor encoded by *lac* I targets the *lac* operator (*lac* o). A copy of *lac* I is also present on the plasmid pET29a. Once IPTG is added both *lac* and T7 promoters are de-repressed.

Failure to produce AtPrx33 in *E.coli* was quite puzzling: the main parameters known to have an influence on protein expression and recovery were all varied. Plasmids used to transform BL21 lines were analysed by PCR and restriction enzyme analysis, in order to verify that I had used proper pET29a plasmids containing the *AtPrx33* cDNA construct: PCR reactions were clearly positive, and restriction enzyme digestions (PvuI: one site in the vector, one site in AtPrx33 cDNA) gave unequivocal positive results. Sequencing was carefully checked again: the full *AtPrx33* cDNA was inserted correctly, without any frameshift or STOP codons. Among other non-tested possibilities, one could try and decrease temperature to 20 or 25°C. At low temperature, more soluble protein (i.e. not trapped in inclusion bodies) is produced, favouring its isolation and detection on SDS-PAGE gels. If *AtPrx33* was still

bound within inclusion bodies after boiling, it may have been retained in the loading wells or the stacking gel, and not migrated to its predicted size (36kDa, from Shah et al., 2004) on SDS-PAGE. Besides boiling, I could also have sonicated cell pellets: this would have increased the chances of freeing the peroxidase.

B.2.2 The baculovirus-insect cell system

Production of peroxidases by eukaryotic cells has the advantage over bacteria that proteins are correctly refolded around a heme residue, and are hence immediately functional. Purification is much easier, as no inclusion body or any other similar structure are formed. However, this technique is rather time-consuming and it does not allow having non-glycosylated proteins for crystallographic studies. In our laboratory, the baculovirus-insect cell system has already successfully been used to produce several peroxidases (Carpin et al., 2001; Dunand et al., 2002). The principle of this method (Fig. 2-3) is that insect cells are co-transfected with an insect-specific viral DNA (*Autographa californica* nuclear polyhedrosis virus, AcNPV) and a plasmid bearing the peroxidase sequence. The plasmid (pVL1392) contains a fragment of the viral polyhedrin locus, which is essential for virus viability (this fragment has been deleted from the wild type viral DNA). The peroxidase gene is inserted within the polyhedrin locus in pVL1392, just downstream from the strong polyhedrin promoter. When co-transfection occurs, a homologous recombination takes place, resulting in the insertion of the peroxidase sequence and the polyhedrin locus into the viral genome and restoring the capacity to form a fully functional virus. AcNPV will then proliferate, eventually lysing its host cell and infecting neighbouring cells. Once the cell is infected, expression of peroxidase can (theoretically) reach 50% of the total insect cell protein production. The AtPrx33 produced is easily collected by pelletting dead cells and recovering the culture medium, and pectate-binding activity can be readily assayed by addition of calcium and polygalacturonic acid (the major component of pectin polymers).

My work consisted in the cloning of *AtPrx33* cDNA into pVL1392. Mireille de Meyer, from our laboratory, infected insect cells and studied pectin binding activity of AtPrx33 and other related peroxidases. The results are detailed in the article "**Two cell wall associated peroxidases from *Arabidopsis* influence root elongation**".

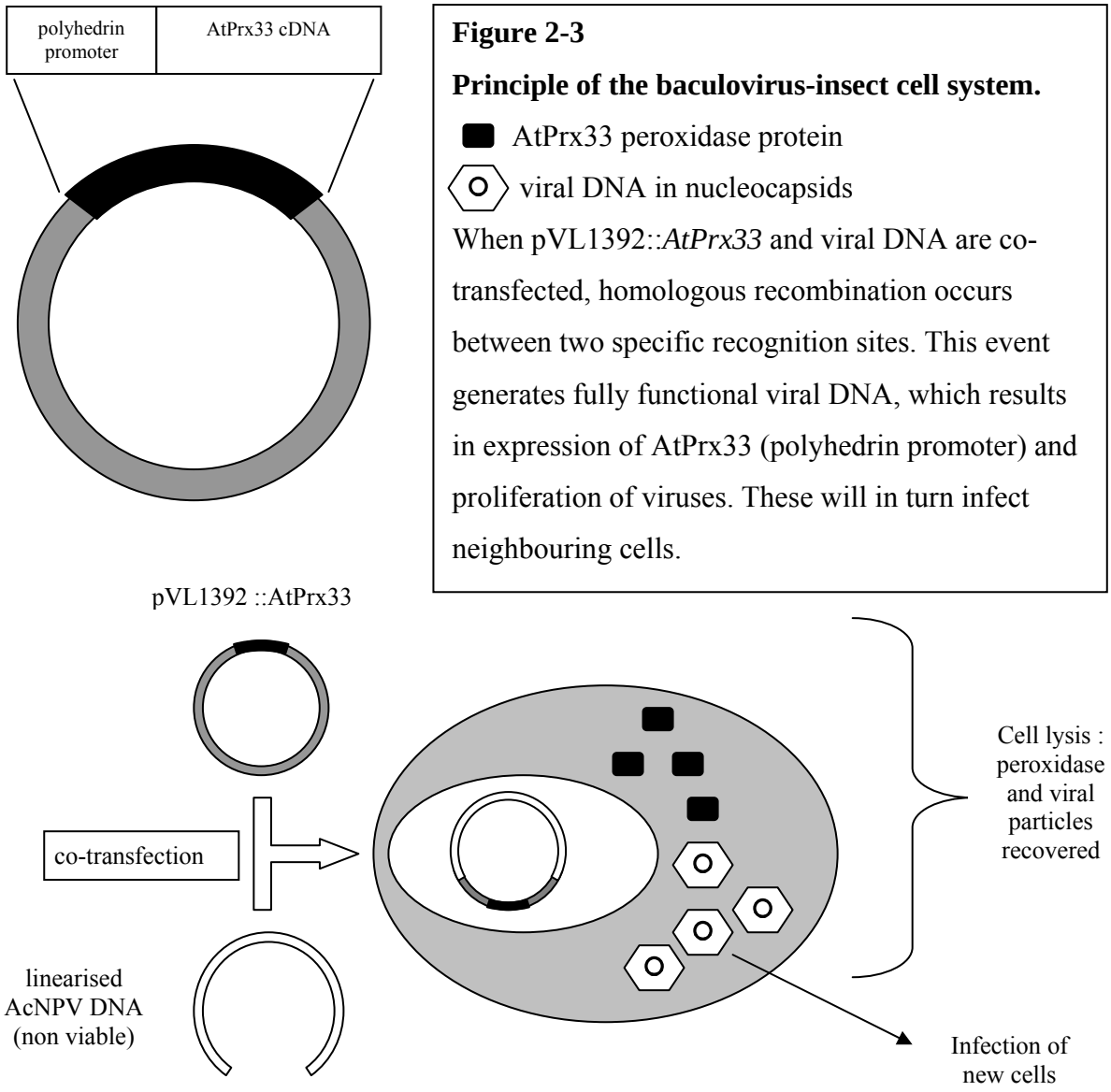


Figure 2-3

Principle of the baculovirus-insect cell system.

AtPrx33 peroxidase protein

viral DNA in nucleocapsids

When pVL1392::*AtPrx33* and viral DNA are co-transfected, homologous recombination occurs between two specific recognition sites. This event generates fully functional viral DNA, which results in expression of AtPrx33 (polyhedrin promoter) and proliferation of viruses. These will in turn infect neighbouring cells.

B.2.3 Materials and methods

B.2.3.a Cloning and expression of AtPrx33 with Escherichia coli

AtPrx33 cDNA was amplified by PCR with primers (engineered restriction sites are underlined) 5'CA KpnI (5'-GGGGTACCATGCAATTCTCTTCATCTTC-3') and 3'CA XhoI (5'-CCGTTCTCGAGACATAGAACTTACAAAGTC-3'), and then directly cloned into pGEM-T vector (Promega). The plasmid was then cloned, purified and *AtPrx33* cDNA was further inserted (KpnI/XhoI) into the expression vector pET29a (Novagen). Once the construct had been verified by sequencing, it was used to transform *E.Coli* strains BL21-CodonPlus®(DE3)-RIL and –RP competent cells (Stratagene).

Transgenic BL21(DE3) cultures were then grown overnight in 5ml LB medium supplemented with 50µg/ml kanamycin (pET29a) and 34µg/ml chloramphenicol (BL21 strains). An aliquot (1ml) of the culture was diluted the next day into 50ml fresh LB, and bacteria were grown until an absorbance (600nm) of 0.600 (exponential growth phase). At this point, the inducer isopropyl-D-thiogalactopyranoside (IPTG) was added at various concentrations (0.1 to 0.8mM), and bacterial growth was resumed during 1, 4, 5, 6 or 20 hours at 30°C or 37°C. After centrifugation, pellet and supernatant were separately collected, and proteins were denatured in 1.6M urea and 1x protein sample buffer (50mM TrisCl pH6.8, 1% sodium dodecyl sulphate [SDS], 5% β-mercaptoethanol [βme]) by boiling during 5 minutes. A loading solution was added to each sample (final concentration: 50mM TrisCl pH6.8, 0.1% bromophenol blue and 5% glycerol) and the mixture was run on a SDS-polyacrylamide (12.5%) gel. After migration at 120V, the gel was stained (0.3% Coomassie brilliant blue R250, 50% MeOH, 10% acetic acid) during 30 minutes and non-specific staining was washed out during several hours either with water or with a destaining solution (5% acetic acid, 5% methanol).

B.2.3.b Cloning and expression of AtPrx33 with the baculovirus-insect cell system

Two new primers were designed for amplifying *AtPrx33* cDNA, with two new restriction sites: 5'CA PstI (5'-AACTGCAGATGCAATTCTCTTCATCTTC-3') and 3'CA BamHI (5'-CGCGGATCCACATAGAACTTACAAAGTC-3'). *AtPrx33* cDNA was then

cloned (PstI/BamHI) into the expression vector pVL1392 (Pharmingen). Baculovirus-mediated transfer into insect cells (*Spodoptera frugiperda*) was then performed by Mireille de Meyer, in our laboratory. Briefly, cells from the *Sf*9 insect line were grown at 27°C in an appropriate medium (TC-100 1x insect medium, foetal bovine serum 1/10 v/v, antibiotic/antimycotic solution 1x [Invitrogen]) and $2x10^6$ cells were co-transfected with pVL1392::*AtPrx33* and BaculoGold® DNA (Pharmingen). Infected cells were then put into culture during one week, in parallel with a culture of non-infected cells. The virus and AtPrx33 peroxidase produced were collected by centrifuging dead cells every week (virus and peroxidase are in the supernatant), and virus was propagated by re-infecting new cells. After 6 weeks of virus propagation, pooled supernatants of infected cell cultures were assayed for peroxidase activity (guaiacol oxidation assay at 470nm). Pectin-binding activity was tested by adding 10µg of polygalacturonic acid and 2mM $CaCl_2$ in a total volume of 100µl, centrifuging and comparing peroxidase activities in the supernatant and the pellet fractions.

B.3 Article "Two cell wall associated peroxidases from Arabidopsis *influence root elongation"*

The article "**Two cell wall associated peroxidases from *Arabidopsis* influence root elongation**" describes a series of experiments showing that both AtPrx33 and AtPrx34 peroxidases:

- are mainly expressed in roots
- are localised in the cell wall and bind pectin
- promote elongation of root cells

The conclusions of this work indicate for the first time a precise function for two *Arabidopsis* peroxidases. Following the results obtained, several other experiments were performed in order to understand more precisely the mechanism of action of these peroxidases. These studies are detailed in the next chapters of section B.

This article is cited as follows:

Passardi F, Tognolli M, De Meyer M, Penel C, Dunand C.
Two cell wall associated peroxidases from *Arabidopsis* influence root elongation.
Planta. 2006 Apr; 223(5):965-74.
DOI: http://dx.doi.org/10.1007/s00425-005-0153-4

Planta (2006) 223: 965–974
DOI 10.1007/s00425-005-0153-4

ORIGINAL ARTICLE

Filippo Passardi · Michael Tognolli · Mireille De Meyer
Claude Penel · Christophe Dunand

Two cell wall associated peroxidases from *Arabidopsis* influence root elongation

Received: 9 June 2005 / Accepted: 19 September 2005 / Published online: 12 November 2005

Abstract Two class III peroxidases from *Arabidopsis*, AtPrx33 and Atprx34, have been studied in this paper. Their encoding genes are mainly expressed in roots; *AtPrx33* transcripts were also found in leaves and stems. Light activates the expression of both genes in seedlings. Transformed seedlings producing AtPrx33-GFP or AtPrx34-GFP fusion proteins under the control of the CaMV 35S promoter exhibit fluorescence in the cell walls of roots, showing that the two peroxidases are localized in the apoplast, which is in line with their affinity for the Ca^{2+}-pectate structure. The role they can play in cell wall was investigated using (1) insertion mutants that have suppressed or reduced expression of *AtPrx33* or *AtPrx34* genes, respectively, (2) a double mutant with no *AtPrx33* and a reduced level of *Atprx34* transcripts, (3) a mutant overexpressing *AtPrx34* under the control of the CaMV 35S promoter. The major phenotypic consequences of these genetic manipulations were observed on the variation of the length of seedling roots. Seedlings lacking *AtPrx33* transcripts have shorter roots than the wild-type controls and roots are still shorter in the double mutant. Seedlings overexpressing *AtPrx34* exhibit significantly longer roots. These modifications of root length are accompanied by corresponding changes of cell length. The results suggest that AtPrx33 and Atprx34, two highly homologous *Arabidopsis* peroxidases, are involved in the reactions that promote cell elongation and that this occurs most likely within cell walls.

Keywords Cell wall · Green fluorescent protein · Pectin · Peroxidase · Root elongation

F. Passardi · M. De Meyer · C. Penel · C. Dunand (✉)
Laboratory of Plant Physiology, University of Geneva,
Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland
E-mail: christophe.dunand@bota.unige.ch

M. Tognolli
Swiss Institute of Bioinformatics Swiss-Prot Group,
rue Michel Servet 1, 1211 Geneva 4, Switzerland

Abbreviations CaMV 35S: Cauliflower mosaic virus 35S promoter · CTPP: C-terminal propeptide · DAPI: 4',6-diamidino-2-phenylindole · GFP: green fluorescent protein · ROS: reactive oxygen species · RT-PCR: reverse-transcriptase polymerase chain reaction · WAK: wall associated kinase

Introduction

The plant specific heme peroxidases belong to a superfamily that contains three different classes of peroxidases (Welinder 1992): the intracellular class I (EC 1.11.1.5/.6/.11), the class II released by fungi during plant–fungi interaction (EC 1.11.1.13/.14), and the secreted class III plant peroxidases (EC 1.11.1.7).

Class III peroxidases are supposed to be involved in a broad range of processes in plants, due to their catalytical versatility and the great number of their isoforms (Passardi et al. 2005). In *Arabidopsis*, they constitute a multigenic family encoding 73 isoenzymes (Tognolli et al. 2002) implicated in diverse activities that are still poorly understood for each single peroxidase isoform. The great diversity of the promoter and intronic sequences partially explains that all kinds of internal or external stimuli regulate the gene expression. On the other hand, the large variability of the putative substrate access channel between the F and G α-helices justifies the diversity of activities catalyzed (Gajhede et al. 1997). The flexibility of the regulation and the substrate specificity within this multigenic family could explain the omnipresence of the peroxidases in the plant life cycle.

AtPrx33 and AtPrx34 are two *Arabidopsis* peroxidases belonging to a cluster of five enzymes containing a putative Ca^{2+}-pectate binding domain (Dunand et al. 2002) similar to the binding domain of APRX, an anionic peroxidase from zucchini (Carpin et al. 1999, 2001), which binds specifically to the Ca^{2+}-pectate complex within cell walls (Carpin et al. 1999). *APRX* transcripts are mainly localized in the elongation zone of

root and hypocotyl (Dunand et al. 2002). The Ca^{2+}-pectate binding property could be important in regulating the enzyme activity and localization within the cell wall.

AtPrx33 and AtPrx34 show nearly 95% homology at the protein level, but their promoter and intronic sequences are strongly divergent (Valério et al. 2004). They can be expected to have similar activity and cellular localization, although the control of their expression and tissular localization could be different. If their cell wall localization is confirmed *in planta*, what could be their function therein and does the pectin binding play a role in this function?

The plant cell wall is a very dynamic structure, which controls both cell shape and cell elongation. Various enzymatic processes cleave and reassemble the cell wall constituents during cell extension. Changes in the cell wall architecture can be achieved by class III peroxidases through their two catalytic cycles: peroxidative and hydroxylic (Passardi et al. 2004). They can stop elongation by forming bonds within the cell wall or favor it by regulating the local concentration of H_2O_2 or by generating reactive oxygen species (ROS), which breaks cell wall bonds (Passardi et al. 2004). Indirectly, peroxidases can also control the cell elongation through their auxin oxidase activity. IAA can be oxidized by following two separate mechanisms: a conventional hydrogen-peroxide-dependent pathway and a second one, which is hydrogen-peroxide-independent and requires oxygen (Savitsky et al. 1999). By this way, peroxidases might regulate locally auxin concentration. In parallel, peroxidase expression levels are dependent on the endogenous auxin concentrations (Gaspar et al. 1996).

In this work, we have combined two different approaches: the RNA silencing technique and the use of T-DNA insertion mutants to get an insight into the function of AtPrx33 and Atprx34. The growth of seedlings lacking *AtPrx33* or *Atprx34* transcripts was studied. In addition, the localization of the enzymes fused to the green fluorescent protein was analyzed.

Materials and methods

Plant material and growth conditions

The Columbia (Col) and Wassilewskija (Ws) ecotypes of *Arabidopsis thaliana* were used as wild-type controls. Plants were grown in soil or on ½ MS medium (Murashige and Skoog 1962) at 24°C under 16 h light/8 h dark and 60% humidity. For the in vitro culture, the light intensity was strictly controlled and fixed at 120 µmol photons m^{-2} s^{-1}. The T-DNA lines CS10885 mutated for *AtPrx33* (in Ws background) and SALK_051769 mutated for *AtPrx34* (in Col background) were isolated from the Feldmann collection (Feldmann et al. 1991) and from the SALK T-DNA insertion lines (Alonso et al. 2003), respectively. All seed lines are available from the stock center of TAIR (http://www.arabidopsis.org).

Mutant screening

The collection of T-DNA insertion mutants from K. Feldmann (Feldmann et al. 1991) was screened by PCR using the *AtPrx33* sense primer 5′-ATGC AATTCTCTTCATCTTC-3′, in combination with the T-DNA-specific primer CD5LB 5′-ATGCAATCGA TAT CAGCAGCCAATTTTA- 3′.

A search in the TAIR database (http://www.arabidopsis.org) revealed a T-DNA insertion within the promoter region of *AtPrx34* (SALK_051769). The precise location of the insert was determined by DNA sequence analysis of a PCR amplification product that was obtained with the *AtPrx34* reverse primer 5′-TCCAAGTGGACGATGTTGAAG-3′ and the T-DNA right border primer CD5RB 5′-GCTCAGGATCC-GATTGTCGTTTCCCGCCTT-3′. PCR amplification was carried out with 0.2 mM dNTPs, 0.4 µM of each primer, 10× *Taq* buffer, and 1 unit of *Taq* polymerase (Sigma) in a volume of 20 µl. The amplification program consisted of an initial denaturation at 95°C for 1 min followed by 35–40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, polymerization at 72°C for 90 s, and a final extension at 72°C for 10 min.

Homozygous insertion mutants in *AtPrx33* and *AtPrx34* were identified by PCR amplification with gene-specific primers (sense primer for *AtPrx33* and reverse primer for *AtPrx34* as described above and the reverse primer 5′-TGTTTGTGTGCCATC-3′ and sense 5′-TTGAAAGGTCGTCGAGGAGT-3′ for *AtPrx33* and *AtPrx34*, respectively).

Imaging of GFP expression

The following constructs (Fig. 1c, d, e) have been designed with the purpose of studying the localization *in planta* of the AtPrx33 and AtPrx34 proteins and the role of the promoter sequence in this localization. *AtPrx33* and *AtPrx34* coding sequences lacking the stop codon were cloned in frame in pAVA 393 (von Arnim et al. 1998), respectively, with *Kpn*I/*Xho*I and *Nco*I/*Xho*I. The *AtPrx33-GFP* and *AtPrx34-GFP* constructs were subcloned in pCHF3 binary vector with *Kpn*I and *Bgl*II. 3.3 kb PCR fragment containing 1.2 kb of *AtPrx33* promoter sequence and the genomic sequence of *AtPrx33* was cloned with *Kpn*I and *Xho*I in frame in pAVA 393. The AtPrx33::*AtPrxP33-GFP-Term* construct obtained was subcloned in pZP222 (Hajdukiewicz et al. 1994) with *Kpn*I and *Bgl*II. 3.3 kb PCR fragment containing 1 kb of *AtPrx34* promoter sequence and the genomic sequence of *AtPrx34* has been cloned with *Hind*III and *Nco*I in pAVA 393. The AtPrx34::*AtPrx34-GFP-Term* construct has been subcloned in pZP211 binary vector with *Kpn*I and *Pst*I.

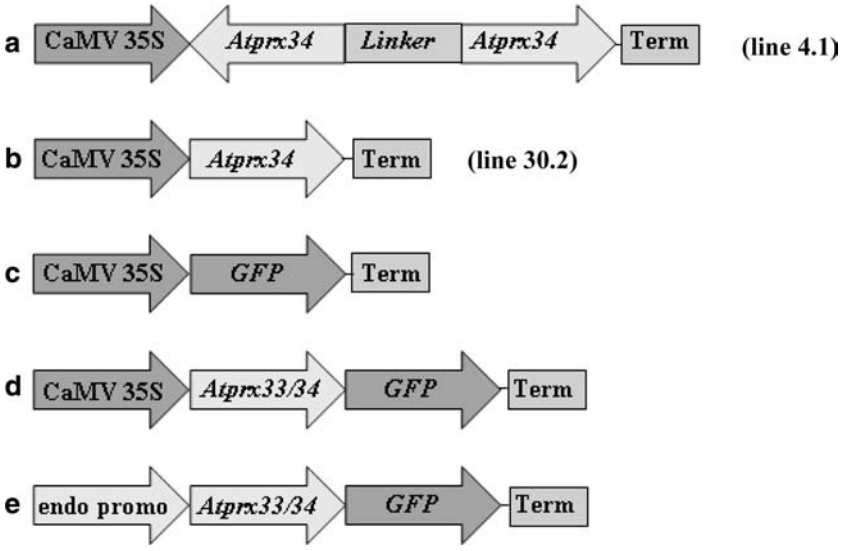


Fig. 1 Schematic structures of the gene cassettes. **a** CaMV 35S:: *43xrPtA-linker-AtPrx34* RNAi construct used to obtain the 4.1 line. **b** CaMV 35S::*AtPrx34* overexpression construct used to obtain the 30.2 line. **c** CaMV 35S::*GFP* contruct for the localization of GFP alone. **d** CaMV 35S::*AtPrx33/34-GFP* contructs for cellular localization of the peroxidases. (e) AtPrx33/34 endogenous promoter::*AtPrx33/34-GFP* contructs for subcellular localization of the peroxidases

A third construct used as control has been obtained by cloning 35S::*GFP-Term* in pCGN1547 (McBride and Summerfelt 1990).

The sequences of the individual primers are as follows (engineered restriction sites are underlined): *AtPrx33*, 5′-GGGGTACCGCTTGGTTTGGTTTCCATTG-3′, 5′-GGGGTACC**ATG**CAATTCTCTTCATCTTC-3′ and 5′- CCGTTCTCGAGACATAGAACTTACAAAGTC -3′ and *AtPrx34*, 5′-CCCAAGCTTTGGATTTCTTC-3′, 5′-CCGCTCGAG**ATG**CATTTCTCTTCGTCTT-3′ and 5′-CATGCCATGGGCATAGAGCTAACAAAG-TC-3′. For the GFP analysis, seedlings were counter-stained with 10 μg/ml propidium iodide and placed on slides in a drop of water. The GFP fluorescence was imaged with an Axioplan2 Zeiss microscope with narrow band excitation (470 ± 20 nm) and emission (510 ± 20 nm) filters (Chroma Technology Corp, VT, USA) and using the Metamorph software (Molecular Devices Corp, CA, USA).

Overexpression of Atprx34 and RNAi constructs

For the overexpression experiment, *AtPrx34* cDNA sequence under the control of CaMV 35S (Fig. 1b) was cloned in pCHF3 vector using *Kpn*I and *Bam*HI. From the transgenic plants that have been obtained, only the line "30.2" has been used for an extensive phenotype study.

The RNAi construct with the whole *AtPrx34* cDNA (Fig. 1a) has been obtained in three successive steps in pBluescript KS+. The cDNA has been cloned in the antisense orientation with *Eco*RI and *Bam*HI. A linker of 700 bp has been cloned into the previous plasmid with *Hind*III and *Xho*I. The final *43xrPtA-linker-AtPrx34* construct has been obtained by a three-way ligation: the *43xrPtA -linker* opened with *Xba*I and *Xho*I and the *AtPrx34* opened with *Xho*I and *Pst*I was subcloned into pCHF5 with *Xba*I and *Pst*I. The capacity of the construct to form a hairpin loop has been controlled before plant transformation using an exonuclease after denaturation and quick renaturation. The construct was used to transform the *atprx33* insertion knock-out mutant, thus resulting in a double-transgenic line called "4.1".

Complementation assays have been performed by transforming *atprx33* T-DNA mutant with the AtPrx33::*AtPrxP33-GFP-Term* construct. Root and root cell lengths were determined in homozygous complemented lines. RT PCR using specific primers spanning the junction site Atprx33-GFP verified the level of AtPrx33 transcript.

Plant transformation

All plasmids were introduced into *Arabidopsis* Col ecotype (except 4.1 double mutant in Ws) by *Agrobacterium-tumefaciens*-mediated transformation using the spraying technique (Bent 2000). *Agrobacterium tumefaciens* strain ASE was used in all cases. The transformants were selected on ½ MS medium containing 50 μg/ml kanamycin or 40 μg/ml Basta, depending on the binary vector used. The presence of the transgenes in all transgenic plants was confirmed by genomic PCR using primers specific in each case.

Root and cell length measurements

The length of the roots and cells was measured on one-week old seedlings grown in square boxes placed in vertical position. Root length was measured on 40–50 seedlings from several independent batches. Cell length was determined by measuring 200 cells in the root hair zone from six independent seedlings. For this purpose, roots were stained with 10^{-4} M DAPI and observed at a 10× magnification by fluorescence with an Axioplan2

Zeiss microscope (excitation: 359 ± 50 nm, emission: 460 ± 50 nm), using the Metamorph software for measurement. Differences between wild-type and transgenic seedlings were evaluated for every measurement assay using a 1-way between subjects ANOVA test (Analyse-it Software, v.1.71).

Reverse transcriptase PCR

Reverse transcriptase-PCR was used as a semiquantitiative method to assess the expression of peroxidase genes. Leaves, stems, flowers, and roots from 5-week old plants were harvested and frozen immediately in liquid nitrogen. Approximately, 100 mg of tissue sample was ground in liquid nitrogen, and total RNA was extracted with the Tri-reagent solution (Sigma) according to the instructions of the manufacturer. After quantification of the concentration by spectrophotometry and confirmation by electrophoresis, 1 µg of the crude RNA preparations was treated with one unit of RNase-free DNase I (Promega). The DNA-free RNA was then used as a template during reverse transcription according to the ImPromII RT protocol from Promega. PCR amplification was conducted for up to 40 cycles using the following thermal profile: denaturation at 95°C for 1 min, annealing at 55°C for 1 min, and polymerization at 72°C for 30 s, with a 10 min terminal extension step at 72°C. To determine whether comparable amounts of RNA from the different tissues had been used for RT-PCR, the level of *AtPrx42* transcript was used as a constitutive control. Indeed, *AtPrx42* is strongly expressed in all *Arabidopsis* organs (Tognolli et al. 2002; Valério et al. 2004), it is the most abundant peroxidase in *Arabidopsis* (505 ESTs were found from a total of 353 000 ESTs indexed in TIGR) and its expression was found strong and stable independently of the age of the plants, various treatments and growth conditions in reverse transcription assays (data not shown). Reactions without RT were used to rule out contamination by genomic DNA. Primers used to determine the homozygosity have also been used for RT-PCR. The following primers were used for *AtPrx42* cDNA: 5′-GGTCCATCGTTTGTACCCT-3′ and 5′-CCCCTGTCTTTCTCACTTTT-3′.

Quantification of the RT-PCR bands on agarose was performed with the software QuantityOne (BioRad). The results shown in Fig. 4 represent the sum of four bands resulting from four independent experiments for each condition (light/dark).

Expression in baculovirus-insect cell system and binding activity

The cDNA sequences encoding AtPrx33, AtPrx34 and AtPrx37 without the signal peptide have been cloned, respectively, with *Bam*HI/*Pst*I, *Bam*HI/*Eco*RI, and *Bam*HI/*Xba*I into pVL1392 vector (Pharmingen, San Diego, CA, USA) and expressed in baculovirus-insect cells (Carpin et al. 2001). After purification, the resulting recombinant proteins were tested for their binding capacities. The binding of peroxidases to the Ca^{2+}-pectate gel was assessed by centrifugation as already described (Penel and Greppin 1996). APRX, an anionic peroxidase from zucchini (Carpin et al. 2001) and AtPrx32 from *Arabidopsis* (Dunand et al. 2002) have been used as positive controls. A horseradish peroxidase mixture containing mainly the HRPC isoform (Fluka, Buchs, Switzerland) was used as a negative control for the binding activity. The assays were performed in a total volume of 100 µl containing the same level of guaicol peroxidase activity (0.025 $0D_{470nm}$/min).

Sequence analysis

The 1,000-bp region upstream each peroxidase gene was analyzed using PLACE and PlantCARE softwares (http://www.dna.affrc.go.jp/htdocs/PLACE and oberon.rug.ac.be:8080/PlantCARE).

Results

Identification of the atprx33 and atprx34 T-DNA tagged mutants

To obtain plant lines with a disruption in the *AtPrx33* and *AtPrx34* genes, we screened the collection of T-DNA insertion mutants of Feldmann (Feldmann et al. 1991) and the TAIR knockout facility (http://www.arabidopsis.org). We used a PCR strategy with primers specific for T-DNA borders and gene-specific sequence in order to identify T-DNA mutants for *AtPrx33* or *AtPrx34* genes. A homozygous line with insertions in intron 1 was obtained in the case of *atprx33*, leading to gene disruption 660 bp downstream of the ATG initiation codons (Fig. 2). A homozygous line with insertion 200 bp upstream of the ATG initiation codons was obtained for *atprx34* (Fig. 2). The position of the insert was verified using T-DNA specific primers, the wild-types (Col and Ws) being used as negative control. The homozygosity was confirmed by PCR with gene-specific primers flanking the T-DNA insert. No amplification was observed for the T-DNA mutants (Fig. 3a). The presence of a single insert and the homozygosity were confirmed with the ratio of resistant plants for both lines (data not shown).

Reverse transcription-PCR experiments on RNA from *atprx33* seedlings indicated the absence of transcripts from this gene in the mutant background, whereas amplification products were readily detectable in the wild-type seedlings (Fig. 3b). *Atprx34* transcripts could be slightly detected in the *atprx34* mutant, but their level was decreased when compared to the wild-type level in 2-week-old seedlings (Fig. 3c). This means that the transcription of *AtPrx34* was not completely

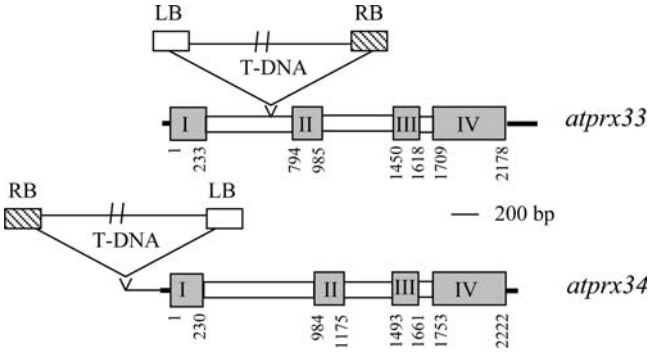


Fig. 2 *AtPrx33* (M58380) and *AtPrx34* (X71794) gene structures and insertion sites of the T-DNA. *Grey boxes* represent exons and *empty boxes* introns (numbered from I to IV); *bold lines* are the predicted and observed 5′and 3′UTR. *RB* right border, *LB* left border

abolished by the presence of the insert in the promoter region.

Expression profile of *AtPrx33* and *AtPrx34* peroxidase genes

High level of *AtPrx33* transcripts was detectable in the roots of 5-week-old plants and at a lower extent in leaves and stems. Some transcripts were also found in flowers. In contrast, low levels of *AtPrx34* transcripts were detected only in roots (Fig. 4a). *AtPrx33* transcripts increased with the age of seedlings, both in light and in the dark (Fig. 4b). *AtPrx34* transcripts, present at a much lower level seemed to increase as well. The expression of *AtPrx33* was largely lower in the dark than under light (Fig. 4b). In all cases, independently of the organs, the light quantity and the age of the plants, *AtPrx33* transcripts were more abundant than *AtPrx34* ones. The equivalent efficiency of *AtPrx33* and *AtPrx34* primer annealing was verified with genomic templates (data not shown).

Subcellular localization of the AtPrx33 and AtPrx34 in planta

Plants were transformed with *Atprx33-GFP* and *Atprx34-GFP* constructs under the control of the CaMV 35S promoter. Endogenous promoters were also used, but in that case the expression was too low and the fluorescence could not be observed. Under the control of CaMV 35S, AtPrx33-GFP, and AtPrx34-GFP fusion proteins accumulated in cell walls (Fig. 5c-g). GFP protein alone had a nuclear and cytoplasmic localization

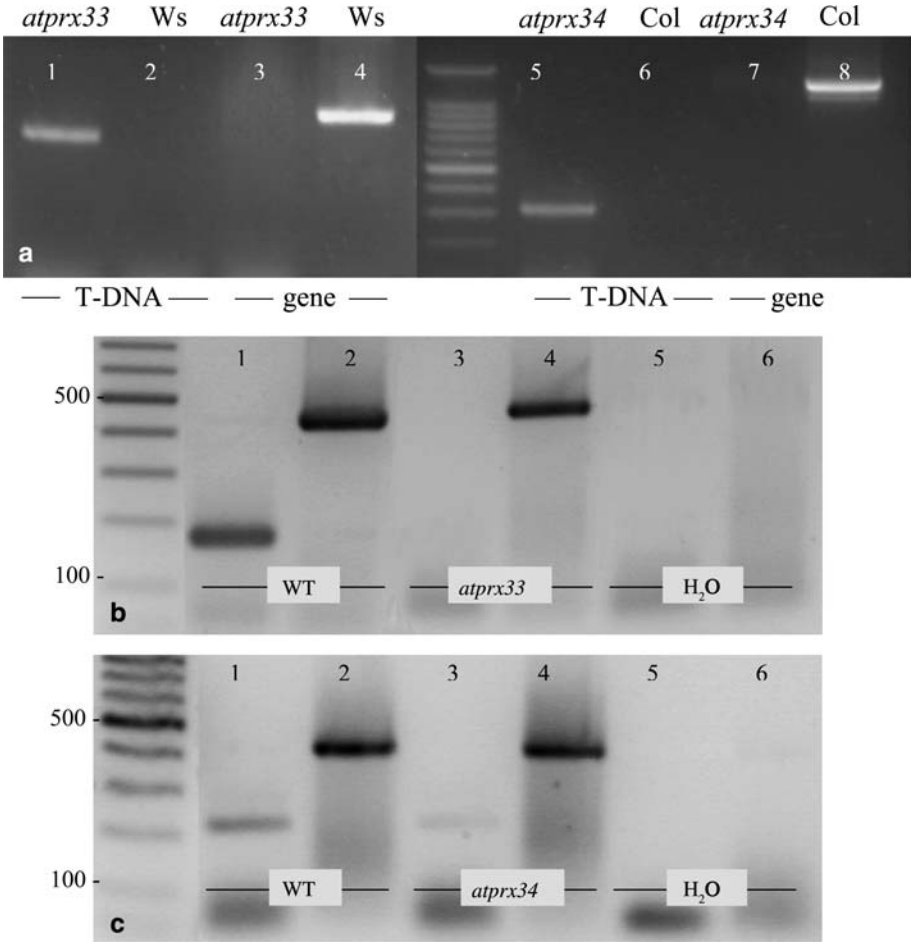


Fig. 3 Molecular characterization of the *atprx33* and *atprx34* T-DNA tagged mutants. **a** Control of the position and homozygosity of the insert. Genomic DNA from *atprx33* and *atprx34* mutants and from the corresponding wild-types (Ws and Col, respectively) was used for PCR amplification with primers specific for the T-DNA insert or for the peroxidase genes. **b** *AtPrx33* transcript levels in 2-week-old *atprx33* and wild-type (WT) seedlings. RT-PCR with *AtPrx33* gene-specific primers (*lanes 1, 3* and *5*); with *AtPrx42* primers (*lanes 2, 4* and *6*). **c** *AtPrx34* transcript levels in 2-week old *atprx34* and wild-type seedlings. RT-PCR with *AtPrx34* gene-specific primers (*lanes 1, 3* and *5*); with *AtPrx42* primers (*lanes 2, 4*, and *6*). RT-PCR products shown for *AtPrx33* and *AtPrx34* were amplified for 40 cycles from 1 µg of total RNA. In b and c, the *AtPrx42* transcript level was used as a constitutive control

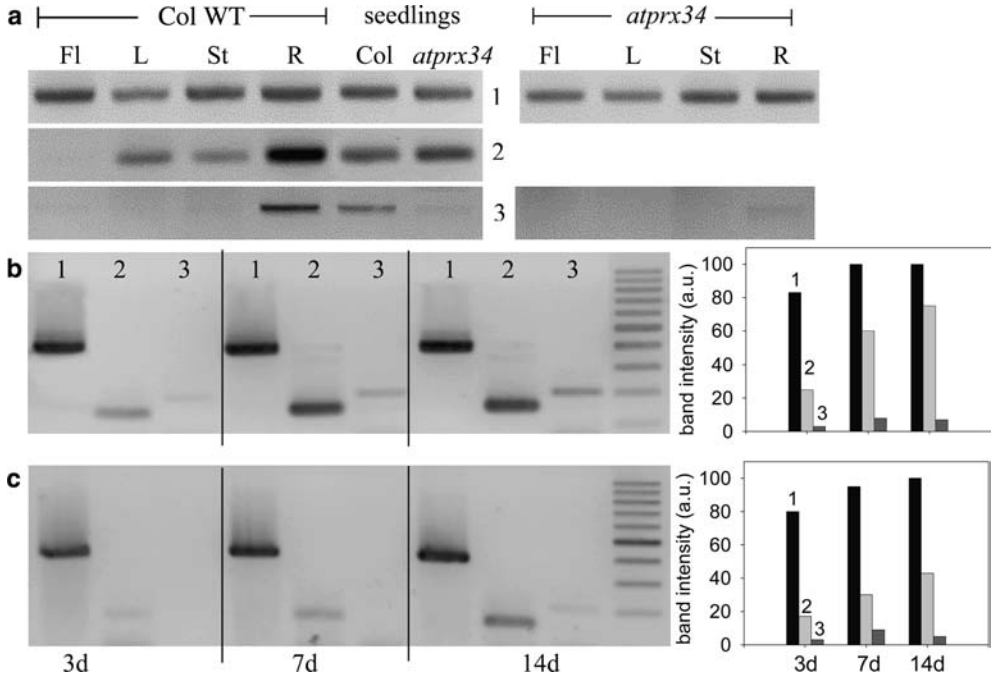


Fig. 4 Expression level of *AtPrx42* (1), *AtPrx33* (2), and *AtPrx34* (3) genes assessed by RT-PCR. **a** Flowers (Fl), leaves (L), stems (St), and roots (R) from 5-week-old plants; 7- day-old whole seedlings from Columbia wild-type (Col), and *atprx34* mutant. **b** Columbia seedlings grown for 3, 7, and 14 days under a 16 h photoperiod. **c** Columbia seedlings grown for 3, 7, and 14 days in darkness. In b and c, a representative gel and a graph corresponding to the mean of four independent experiments are shown. Total RNA (1 μg) was used for RT-PCR with primers specific for *AtPrx42* (1), *AtPrx33* (2), and *AtPrx34* (3) genes. The RT-PCR products shown were amplified for 40 cycles. The *AtPrx42* transcript level was used as a constitutive control

(von Arnim et al. 1998) (Fig. 5a, b). In vitro, recombinant AtPrx33, AtPrx34, and AtPrx37, produced by the baculovirus insect cell system, showed some capacity to bind to the Ca^{2+}-pectate complex, which is mainly found in the middle lamella. Their binding activities were much lower than the binding capacity of AtPrx32 and zucchini APRX, but significant when compared to the lack of binding of horseradish peroxidase (Fig. 6). AtPrx32 and HRPC both show high homology at the protein level with AtPrx33/34 (over 87%), but only the former bears the putative pectate binding domain.

Root length and peroxidase gene expression

The length of roots was measured on various transgenic seedlings, including *atprx33, atprx34*, the *AtPrx34* overexpressor 30.2, and the 4.1 RNAi double mutants. Since the mutants were obtained in two different backgrounds, Columbia and Wassilewskija *Arabidopsis* were also analyzed. The 4.1 double mutants do not contain *AtPrx33* transcripts and exhibit a reduction of *AtPrx34* mRNA (Fig. 7a). Seedlings containing a T-DNA inserted in the *AtPrx34* promoter sequence (*atprx34*) showed a reduced expression of *AtPrx34* and the seedlings containing the 35S::*AtPrx34* construct (30.2 line) an accumulation of *AtPrx34*. *atprx34* seedlings had no significant root length variation when compared to the wild-type plants, whereas the 30.2 mutant had clearly longer roots (Fig. 7b). In knocked-out *atprx33* plants and in the double mutant (4.1 line), a significant reduction of the root length could be observed (Fig. 7b). The root length reduction is additive when both genes are affected. The fully elongated cells present in the root hair zone were also measured. It appeared that the length of these cells was also significantly reduced in *atprx33* and 4.1 mutants and increased in 30.2 mutant.

In order to determine the heritability of the mutation in mutant lines, homozygous transgenic plants were crossed with corresponding wild-type ecotypes. The F1 (for overexpressor lines) and the F2 progenies (for the *atprx33* mutant) showed the appropriate root phenotype (data not shown). The co-segregation of the resistance and of the phenotype confirmed that the mutations related to *AtPrx33* and *AtPrx34* are directly responsible for the modification of the root length.

Finally, complementation assays were performed by introducing the AtPrx33::*AtPrxP33-GFP-Term* construct in knocked-out *atprx33* plants. A significant increase of root and root cell lengths was observed and compared to *atprx33* plants in a complemented line. There was no significant difference between this complemented line and the Ws wild-type control (Fig. 8).

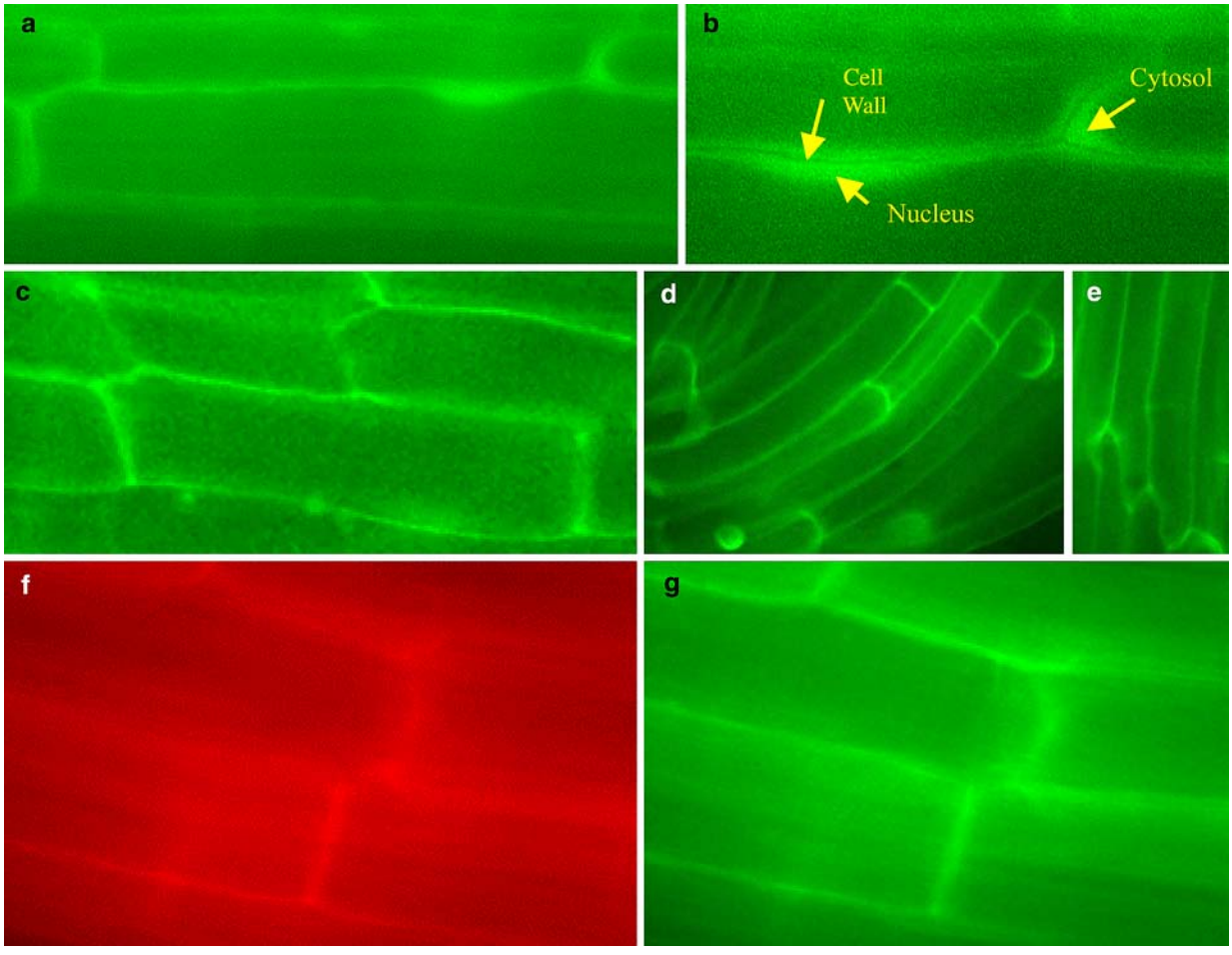


Fig. 5 Detection of GFP fluorescence in roots of seedlings transformed with a CaMV 35S::*GFP* construct (**a**, **b**), with a CaMV 35S::*AtPrx33-GFP* construct (**c**–**e**) and with a CaMV 35S::*AtPrx34-GFP* construct (**f**, **g**). Staining with 10 μg/ml propidium iodide (**f**). Cell wall (unstained zone), nucleus, and cytosol (stained zones) are indicated by *arrows* in B

Discussion

Analysis of the nucleotide sequences

The structure of *AtPrx33* and *AtPrx34* corresponds to a classical peroxidase gene with three introns and four exons (Tognolli et al. 2002). The first 100 bp upstream the ATG codon of the two genes are highly conserved. This short sequence contains the TATA box and is sufficient for a basal expression, but probably does not allow a fine and specific regulation of expression as shown by T-DNA insertion in the *AtPrx34* promoter. Indeed, the insertion 200 bp upstream of the ATG should remove light response and root localization. Known cis-elements found in the promoter of *AtPrx33*

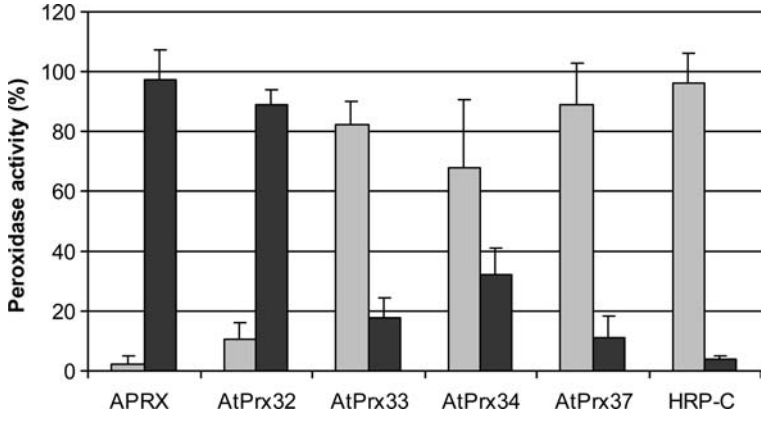


Fig. 6 Binding of recombinant peroxidases to Ca^{2+}-pectate. The assays were performed in the presence of 10 μg of polygalacturonic acid and 2 mM $CaCl_2$ in a total volume of 100 μl with the same level of peroxidase activity (Penel and Greppin 1996). After centrifugation, the peroxidase activity was measured in the Ca^{2+}-pectate pellets (() and in the corresponding supernatant (). The data are the means of three independent experiments ± SD

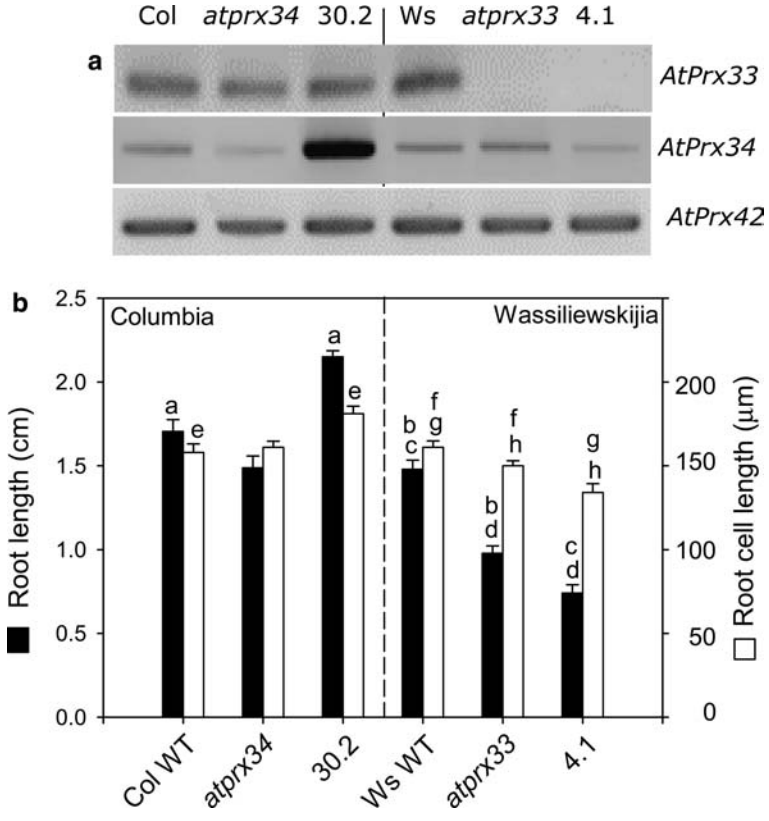


Fig. 7 Expression of *AtPrx33*, *AtPrx34* and *AtPrx42* genes and the measurement of root length and cell root length in various transgenic *Arabidopsis* seedlings. **a** Expression of the three peroxidase genes in *aprx34* and overexpressor 30.2 mutants in a Columbia (Col) background and in *atprx33* and RNAi 4.1 mutants in a Wassilewskija (Ws) background, assessed by RT-PCR. The *AtPrx42* transcript level was used as a constitutive control in the various seedlings. **b** Root lengths and cell root lengths of 7-day-old seedlings from the plant lines described in a. The differences between wild-type and corresponding mutant seedlings were analyzed by 1-way ANOVA test. Same letters indicate significant differences with *P* values: $P < 0.0001$ (**a,b,c,g**); $P = 0.0046$ (**d**); $P = 0.001$ (**e**); $P = 0.0297$ (**f**); $P = 0.0121$ (h)

and *Atprx34* genes have been listed in Table. 1. There is a similar distribution of cis-elements related to the control by light and the expression in root. This *in silico* analysis is in agreement with the actual expression pattern, which showed a preferential transcription in the roots and an up-regulation by light (Fig. 4). In contrast, large differences exist for the abscissic acid response elements (seven elements in *AtPrx33* and none in *AtPrx34*) and for the salicylic acid response elements (six elements in *AtPrx34* and none in *AtPrx33*), which is in line with the idea of a similar activity but a different regulation for the two peroxidases.

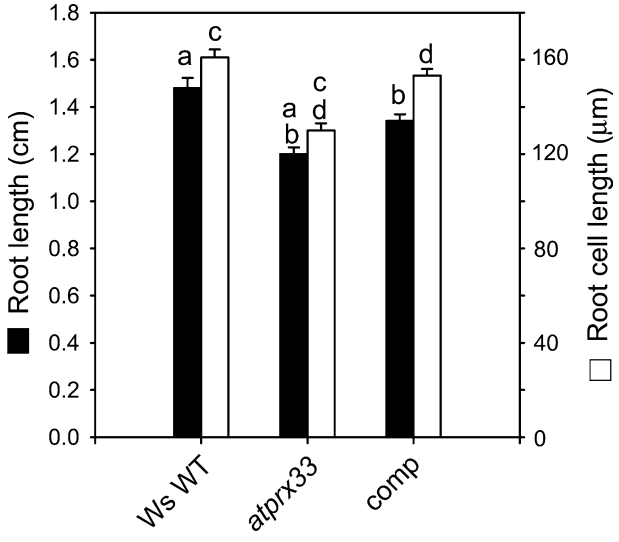


Fig. 8 Complementation assays. Root and cell root lengths were determined for 7-day-old seedlings from Wassilewskija (Ws), *atprx33* and *atprx33* complemented line (comp). The differences between the seedlings were analyzed by a 1-way ANOVA test. Same letters indicate significant differences with *P* values: $P < 0.0001$ (**a, c, d**); $P = 0.0007$ (**b**)

AtPrx33 and AtPrx34 are cell wall associated and involved in cell elongation

AtPrx33 and AtPrx34 are members of a same gene cluster and show a high level of homology with 90% of identity. Both peroxidases are localized in cell walls (Fig. 5). They have an amino acid motif including three lysine residues, which is similar to the well-characterized Ca^{2+}-pectate binding site of the zucchini peroxidase APRX (Carpin et al. 1999, 2001; Dunand et al. 2002). Purified AtPrx34 has also been shown to bind to Ca^{2+}-pectate (Shah et al. 2004) and recombinant AtPrx33 and AtPrx34 produced by insect cells exhibited some affinity for Ca^{2+}-pectate (Fig. 6). This means that, their localization within the cell walls could be modulated by Ca^{2+}, as already demonstrated for APRX (Carpin, 2001).

AtPrx33 and AtPrx34 sequences possess a signal peptide and a C-terminal extension. The signal peptide targets the protein to endoplasmic reticulum and the secretory pathway, whereas the C-term extension (CTPP) may address the protein to the vacuole

Table 1 Main *cis*-elements present in the 1,000 bp upstream the *AtPrx33* and *AtPrx34* genes

Motif	Sequence	Function	Position for *AtPrx33*	Position for *AtPrx34*
ABRE	YACGTGGC	Abscissic acid responsiveness	46(-), 47, 144(-), 222(-), 539(-), 540, 898(-)	
ARF/AuxREs	TGTCTC	Auxin response element	618(-)	247(-)
AuxRR	GGTCCAT	Auxin response	938(-)	
ERE	ATTTCaaa	Ethylene-responsive element	647	5, 295, 337
G-box	CACGTG	Element involved in light responsiveness	45, 538	
GA-motif	ATAGATAA	Part of a light responsive element	273(-), 477(-)	725
GATA-box	GATAGa	Part of a light responsive element	65(-), 652	244 (-), 636, 790
LAMP-element	CTTTATCA	Part of a light responsive element		303
LTR	CCGAaa	Low-temperature responsiveness	616, 76(-)	585
MBS	CAACTG	MYB binding site involved in drought-inducibility	247(-)	
P-box	CCTTTTG	Gibberellin responsive element		833(-)
Root motif	ATATT	Root localization	289(-), 407(-), 410, 645, 791(-), 792, 852	6(-), 406(-),407, 518(-), 522, 561, 616, 644(-)
TGA-element	AACGAC	Auxin response element	808	
W-box	TTGACC	Salicylic acid responsiveness	903	378(-), 579, 672(-), 813, 903, 979

The data have been obtained with PlantCare and PLACE. Indicated positions are relative to the ATG codon. Underlined positions correspond to the underlined part of the cis-element

(Neuhaus 1996). However, a recent study cast some doubt on the validity of this hypothesis in the case of a horseradish peroxidase (Kis et al. 2004), and so far there has been no evidence suggesting that the CTPP is cleaved from HRPC1a for the cell wall targeting (Matsui et al. 2003). Similarly to APRX, which also has a C-term extension but was found in apoplast (Carpin et al. 1999), AtPrx33 and AtPrx34 are localized in the cell wall (Fig. 5). Even if both proteins transit via the vacuole (Carter et al. 2004), they finally end in the apoplast. The same transitory targeting has been observed for a cell-wall-associated kinase 1 (WAK1) found in the vacuole (Carter et al. 2004) but known to bind the Ca^{2+}-pectate complex (Decreux and Messiaen 2005). All WAK family members (WAK1–5) exhibit a similar protein length without apparent C-term extension. Thus, vacuole targeting could be transient before a final localization in the cell wall regulated by the presence of the CTPP but not related to the cleavage of this extension. Regarding AtPrx33/34 GFP fusion protein, substantial evidence shows that the fusion proteins are correctly folded and active: GFP fluorescence was detected in the plants containing AtPrx33/34-GFP, and the AtPrx33-GFP construct can complement the *atprx33* mutant phenotype.

The peroxidase cellular localization could be associated with a particular activity in relation to cell wall elongation. AtPrx33 and AtPrx34 are highly homologous and most likely have similar enzymatic properties in the cell wall. On the other hand, we have observed important difference concerning the transcription profile. These variations might explain the presence of two very close isoforms, performing similar tasks, but differentially transcribed in response to various stimuli.

Knocking out *AtPrx33* gene expression resulted in the reduction of root elongation. However, this did not correspond with a significant modification of the total peroxidase activity (data not shown). The high number of peroxidase isoforms present in Arabidopsis can easily explain this. Even if the overall peroxidase activity was not affected in *atprx33* or 4.1 mutants, the AtPrx33 protein was most likely absent since there were no detectable transcripts. This absence resulted in a reduction of growth, thus suggesting that this peroxidase is involved in growth promoting reactions. This was confirmed by the additional growth reduction observed following *AtPrx34* downregulation in 4.1 double mutant, the restoration of normal root growth in *atprx33* mutant complemented with Atprx33*::Atprx33-GFP* construct and the stimulation of root growth brought about in the 30.2 mutant overexpressing *AtPrx34*. Another peroxidase, horseradish prxC1a peroxidase, was already reported to stimulate growth when expressed in hybrid aspen (Kawaoka et al. 2003). In the present work, there was a correlation between the root length and the length of root cells (Fig. 7), suggesting that the observed modifications of root growth are due to the changes in cell elongation. This means that AtPrx33 and AtPrx34 would be involved in reactions promoting cell elongation, for example free radical production that may loosen cell wall, as proposed by Schopfer (2001). The level of *AtPrx33* and *AtPrx34* transcripts was light- and age-dependent (Fig. 4), which can be related to the stimulating effect of light and of the stage of development on root growth. The light regulation of both genes is also in line with the presence of numerous light related elements in their promoter sequences (Table 1), and with observations reporting a phytochrome-dependent regulation of peroxidases in other plant species (Kim et al. 1989; Casal et al. 1994).

The assessment of the transcript levels by RT-PCR indicated that *AtPrx33* gene was expressed at a much higher rate than *AtPrx34*. This is not in agreement with a previous northern blot analysis (Shah et al. 2004). This

latter discrepancy could be explained by the specificity of the primers used for the PCR amplification versus the lower specificity of the cDNA probes used previously for the northern blot analysis. However, both techniques showed that *AtPrx33* and *Atprx34* mRNAs accumulated mainly in roots. Concerning the EST counts, it is probable that the developmental stage at which we studied our plants was not one in which *AtPrx34* expression was the highest. Further studies are needed to clarify this point.

The results obtained in this work as well as the presence of auxin related cis-elements in the promoter region of *AtPrx33* and *Atprx34* genes (Table 1) argue in favor of an active role of the two encoded peroxidases in root cell elongation.

Acknowledgements We thank Prof. Susan Gasser (Friedrich Miescher Institute, FMI, Basel, Switzerland) for the use of the Zeiss Axioplan II microscope, Dr. Thierry Laroche (FMI) for his scientific comments and technical support, and Prof. Christian Fankhauser (University of Lausanne, Switzerland) for the gift of the binary vectors pCHF3 and pCHF5. This work was supported by the Swiss National Science Foundation (grant 31-068003.02).

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653–657

Bent AF (2000) Arabidopsis in planta transformation. Uses, mechanisms, and prospects for transformation of other species. Plant Physiol 124:1540–1547

Carpin S, Crevecoeur M, de Meyer M, Simon P, Greppin H, Penel C (2001) Identification of a Ca^{2+}-pectate binding site on an apoplastic peroxidase. Plant Cell 13:511–520

Carpin S, Crevecoeur M, Greppin H, Penel C (1999) Molecular cloning and tissue-specific expression of an anionic peroxidase in zucchini. Plant Physiol 120:799–810

Carter C, Pan S, Zouhar J, Avila EL, Girke T, Raikhel NV (2004) The vegetative vacuole proteome of *Arabidopsis thaliana* reveals predicted and unexpected proteins. Plant Cell 16:3285–3303

Casal JJ, Mella RA, Ballare CL, Maldonado S (1994) Phytochrome-mediated effects on extracellular peroxidase activity, lignin content and bending resistance in etiolated *Vicia faba* epicotyls. Physiol Plant 92:555–562

Decreux A, Messiaen J (2005) Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. Plant Cell Physiol 46:268–278

Dunand C, Tognolli M, Overney S, von Tobel L, De Meyer M, Simon P, Penel C (2002) Identification and characterisation of Ca^{2+}-pectate binding peroxidases in *Arabidopsis thaliana*. J Plant Physiol 159:1165–1171

Feldmann KA, Wierzbicki AM, Reiter RS, Coomber SA (1991) T-DNA insertion in *Arabidopsis*. In: Hermann R, Larkins B (eds) Plant molecular biology. Plenum Press, pp 563–574

Gajhede M, Schuller DJ, Henriksen A, Smith AT, Poulos TL (1997) Crystal structure of horseradish peroxidase C at 2.15 A resolution. Nat Struct Biol 4:1032–1038

Gaspar T, Kevers C, Penel C, Greppin H, Reid D, Thorpe T (1996) Plant hormones and plant growth regulators in plant tissue culture. In Vitro Cell Dev-Plants 32:272–289

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol Biol 25:989–994

Kawaoka A, Matsunaga E, Endo S, Kondo S, Yoshida K, Shinmyo A, Ebinuma H (2003) Ectopic expression of a horseradish peroxidase enhances growth rate and increases oxidative stress resistance in hybrid aspen. Plant Physiol 132:1177–1185

Kim SH, Shinkle JR, Roux SJ (1989) Phytochrome induces changes in the immunodetectable level of a wall peroxidase that precede growth changes in maize seedlings. Proc Natl Acad Sci USA 86:9866–9870

Kis M, Burbridge E, Brock IW, Heggie L, Dix PJ, Kavanagh TA (2004) An N-terminal peptide extension results in efficient expression, but not secretion, of a synthetic horseradish peroxidase gene in transgenic tobacco. Ann Bot (Lond) 93:303–310

Matsui T, Nakayama H, Yoshida K, Shinmyo A (2003) Vesicular transport route of horseradish C1a peroxidase is regulated by N- and C-terminal propeptides in tobacco cells. Appl Microbiol Biotechnol 62:517–522

McBride KE, Summerfelt KR (1990) Improved binary vectors for Agrobacterium-mediated plant transformation. Plant Mol Biol 14:269–276

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Neuhaus JM (1996) Protein targeting to the plant vacuole. Plant Physiol Biochem 34: 217–221

Passardi F, Cosio C, Penel C, Dunand C (2005) Peroxidases have more functions than a Swiss army knife. Plant Cell Rep 24:255–265

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci 9:534–540

Penel C, Greppin H (1996) Pectin binding proteins: characterization of the binding and comparison with heparin. Plant Physiol Biochem 34:479–488

Savitsky PA, Gazaryan IG, Tishkov VI, Lagrimini LM, Ruzgas T, Gorton L (1999) Oxidation of indole-3-acetic acid by dioxygen catalysed by plant peroxidases: specificity for the enzyme structure. Biochem J 340:579–583

Schopfer P (2001) Hydroxyl radical-induced cell-wall loosening in vitro and in vivo: implications for the control of elongation growth. Plant J 28:679–688

Shah K, Penel C, Gagnon J, Dunand C (2004) Purification and identification of a Ca^{2+}-pectate binding peroxidase from *Arabidopsis* leaves. Phytochemistry 65:307–312

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288:129–138

Valério L, De Meyer M, Penel C, Dunand C (2004) Expression analysis of the *Arabidopsis* peroxidase multigenic family. Phytochemistry 65:1331–1342

von Arnim AG, Deng XW, Stacey MG (1998) Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene 221:35–43

Welinder KG (1992) Superfamily of plant, fungal and bacterial peroxidases. Curr Opin Struct Biol 2:388–393

B.4 Study of AtPrx33 expression: the AtPrx33::*AtPrx33*::*GFP transgene*

After having shown that *AtPrx33* is closely associated to the cell wall and is particularly found in roots, I started a series of experiments in order to gain information on the regulation of the *AtPrx33* gene expression. The construct AtPrx33::*AtPrx33*::*GFP* was already available, since it had been used to complement the knock-out mutant *atprx33*. I hence transformed wild type Columbia plants, and selected resistant seedlings on gentamicin. Three independent lines were obtained. RT-PCR reactions on 7 day-old seedlings confirmed expression of the transgene in homozygous lines (generations T2/T3). Some lines were silenced at generation T4. When observed under a fluorescence microscope, none of the plants showed any sign of GFP fluorescence. Growth conditions were modified (no antibiotics, darkness, horizontal/vertical growth), but results were still negative. In order to avoid a possible partial silencing of the transgene, I observed T1 and T2 generations of the Atprx33::*AtPrx33*::*GFP* transgenic plants, without any improvement. I then tried another method of fluorescence detection, in case the signal was too low to be detected by the microscope I was using. I ground roots of 7 day-old seedlings in HEPES buffer (20mM, pH 7), transferred the whole mixture to a glass cuvette, and measured fluorescence with a Kontron SFM 25 fluorometer (excitation: 470nm, emission: 510nm): no fluorescence was ever detected. Similar conclusions came out from studies performed with two independent lines obtained with the construct AtPrx34::*AtPrx34*::*GFP*.

The contrasting results obtained between RT-PCR and fluorescence show that although the transgene is transcribed, either the AtPrx33::GFP fusion protein cannot fold properly and is targeted for destruction within the cell, or fluorescence is prevented or at least decreased by quenching due to presence of the AtPrx33 protein very close to the GFP protein. Alternatively, I recently found out that in pAVA393, the plasmid used as a source of the GFP gene (see Appendix), the small sequence upstream from GFP is an enhancer, and not simply a linker as previously thought. The enhancer is effective when it is placed downstream from a promoter and upstream from a gene (Carrington and Freed, 1990; Nicolaisen et al., 1992), which was not the case of my AtPrx33::*AtPrx33*::*GFP* construct. The first hypothesis (protein is degraded) is less likely, as *atprx33* knock-out mutant lines complemented with AtPrx33::*AtPrx33*::*GFP* exhibited longer roots than *atprx33* (Passardi et al., 2006), showing that the peroxidase is still expressed and active. Assuming that the "quenching hypothesis" was the explanation, I designed a new transgene: AtPrx33::*GFP*. This strategy proved to be

the right one: the detailed experimental protocol and results are described in the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**".

*B.5 Searching for relationships between AtPrx33 and auxin: preliminary studies with the GH3::*GUS *transgene*

B.5.1 Transgenic plants: importance of the insertion site and generational differences

Besides studying localisation of AtPrx33 *in planta*, a parallel project was proposed to me since the beginning of my thesis: making transgenic plants that would express the *E.coli* marker gene *uidA* (glucuronidase-GUS) under the control of the auxin-responsive promoter GH3. This promoter was originally found in soybean, where *gh3* mRNA was a predominant sequence recovered during the screening of auxin-induced cDNA sequences (Hagen et al., 1984; Hagen et al., 1991). GH3 promoter has since been used by many different research groups, in various plants, such as *Trifolium repens* (clover), *Nicotiana tabacum* (tobacco) and *Physcomitrella patens* (a moss) (Mathesius et al., 1998; Mauro et al., 2002; Bierfreund et al., 2003). In *Arabidopsis thaliana*, the artificial promoter DR5 was used instead (Avsian-Kretchmer et al., 2002; Benkova et al., 2003).

As soon as the first transgenic GH3::*GUS* plants were obtained (from transformed wild type Columbia ecotype), I assayed rosette leaves of T0 plants for glucuronidase activity by fluorometry. In this technique, glucuronidase is extracted and assayed for hydrolysis of one of its substrates, 4-methylumbelliferyl β-D-glucuronide (MUG), into a fluorescent product. It clearly appeared that every line was giving quite a different signal (Fig.2-4). This observation has been already reported with both GH3 and the artificial DR5 promoters in *Arabidopsis* (Ulmasov et al., 1997). Curiously, in their study, Ulmasov and colleagues had claimed that their DR5 promoter was stronger than GH3, although they only showed a limited selection of transformants: with such a high variability of MUG hydrolysis activity, it is difficult to convincingly compare both promoters by only showing a few transgenic lines.

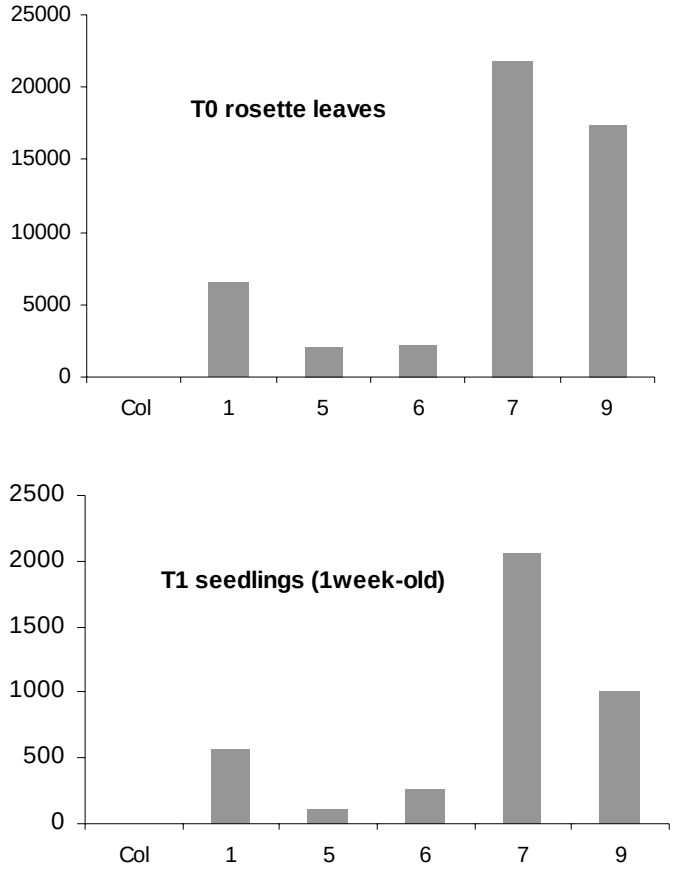


Figure 2-4
MUG hydrolysis activity (30min incubation) of Col GH3 ::*GUS* lines

Y axis: relative fluorescence (normalised to protein concentration)

X axis: non-transformed wild type (Col) and different lines (1 to 9) of GH3::*GUS* transformants

Seedlings of the T1 generation were also assayed, and gave very similar results to T0 leaves, indicating that variability of GUS expression was due to positional insertion of the transgene in the genome, and was hence maintained across generations. In parallel to Col GH3::*GUS* plants, I had also started transforming Ws wild type, *atprx33* (Ws) and *atprx34* (Col) with the GH3::*GUS* transgene. Given that variability was so strong, I then had to cross Col GH3::*GUS* (T1 lines) to *atprx34* and Ws GH3::*GUS* (T1 lines) to *atprx33* in order to be able to compare GUS activity between wild types and peroxidase mutants.

The first results obtained with Col GH3::*GUS* and *atprx34* x Col GH3::*GUS* showed a generational dependence of MUG hydrolysis activity: the later the generation, the lower the fluorescence (Fig. 2-5). The decrease observed with later generations was repeatedly reported in further experiments. The conclusions drawn from the comparison between wild types and peroxidase mutants could then be very different, depending on which generation was considered. As Col and Ws GH3::*GUS* T1 lines were crossed with the respective *atprx34* and *atprx33* mutants, the seeds obtained from the cross (T0) were equivalent to a T2 wild type generation. Hence, in figure 2-5, generation "Col T4" should be compared to generation "*atprx34* T2".

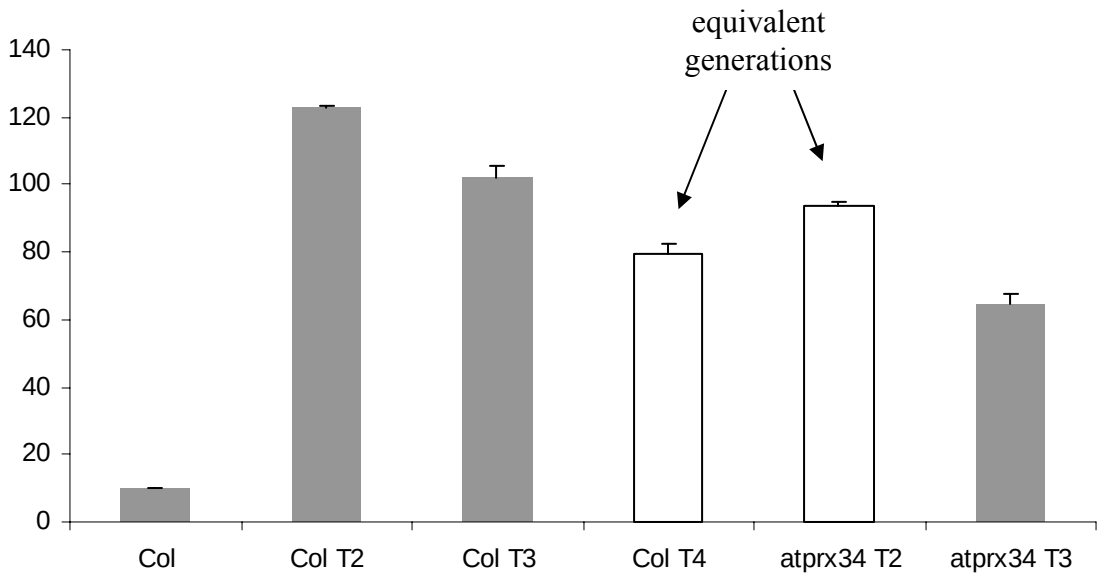


Figure 2-5

Comparison of MUG hydrolysis activity between Col GH3::*GUS* (Col) and Col GH3::*GUS* x *atprx34* (atprx34) lines of various generations (30' incubation).

Y axis: relative fluorescence (normalised to protein concentration)

Once it was clear that the generation used was a critical parameter to be considered, I made several repeats of experiments involving Col GH3::*GUS* and two independent lines of *atprx34* x GH3::*GUS*. After four different experiments, the results were rather confusing: *atprx34* x GH3::*GUS* lines had a higher glucuronidase activity than control in one occasion, a lower activity in two other experiments and the same activity in a last case. Concerning Ws, I repeatedly observed a higher fluorescence in *atprx33* x GH3::*GUS* compared to control.

Following results with the Columbia ecotype, I decided to make another type of control line: in order to be as close as possible to *atprx* x GH3::*GUS* lines, I crossed a wild type with its corresponding wild type GH3::*GUS* ecotype, giving Col x Col GH3::*GUS* and Ws x Ws GH3::*GUS* controls. Once obtained, the T2 lines of controls were compared to the T2 lines of mutants. Results were much more reproducible for Columbia than with the older controls, and hence much more trustable. Concerning Ws, I observed exactly the opposite as for the older controls: *atprx33* x GH3::*GUS* lines had a **lower** GUS activity than controls. Considering that these new controls have a closer resemblance to the peroxidase mutant GH3::*GUS* lines in terms of generation and historical background, and that higher reproducibility was obtained with the Columbia ecotype, I considered the last results obtained

as more trustable. The graphs and conclusions from these experiments are described in more detail in the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**".

A last important parameter to be considered was the tissue observed. The first MUG experiments were performed on 7 day-old seedlings. When separating roots and shoots during extraction of glucuronidase, and then assaying both fractions separately in MUG assays, I noticed that glucuronidase activity in shoots was around ten times higher that in roots (Fig.2-6a). Therefore, the measurements performed on whole seedlings are only representative of GUS activity in shoots, but they mask what is happening in roots. From then on, roots and shoots were always analysed separately. Importantly, roots were cut below the hypocotyl, as experiments with another GUS substrate, X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt), indicated a high GUS activity at the base of the hypocotyl region (Fig.2-6b).

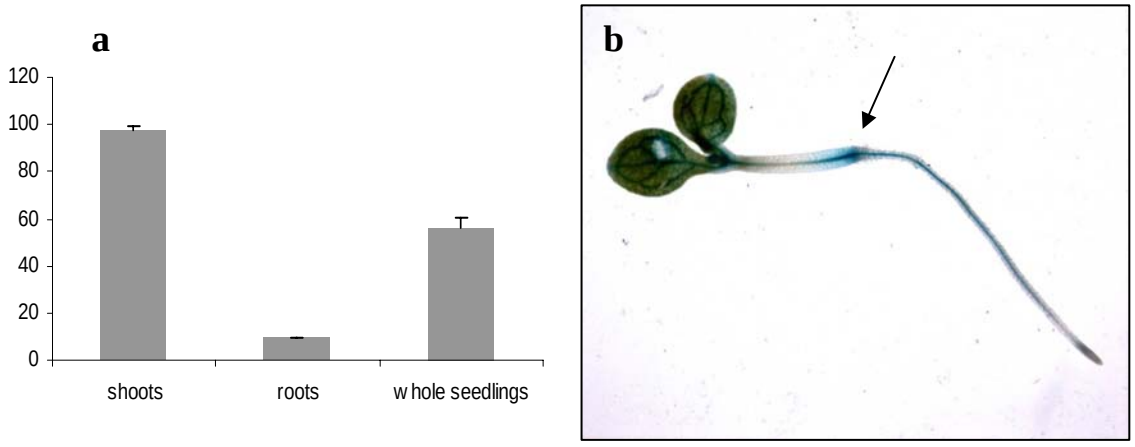


Figure 2-6

Variability of *GUS* expression in *Arabidopsis* seedlings.

(a) Y axis: absolute fluorescence (not normalised to protein concentration) after 30'.

(b) X-Gluc staining of a Ws GH3::*GUS* 7 day-old seedling. Arrow marks the region of the hypocotyl showing the highest X-Gluc staining.

B.5.2 Methodology of MUG fluorometry experiments

In order to quickly and precisely quantify glucuronidase activity in GH3::*GUS* seedlings, I first had to establish a protocol, as this technique was new in the laboratory. I found in the internet several variations of a protocol derived from previous literature (Gallagher, 1992). Glucuronidase was extracted from GH3::*GUS* lines by grinding plant material (10-50mg) in 400µl of extraction buffer (KPO_4 buffer 100mM pH 7.8) freshly supplemented with dithiotreitol (DTT) 1mM and leupeptin 20µg/ml. After centrifugation at 12'000 g, the supernatant was collected and split into two 150µl aliquots that were frozen in liquid nitrogen and stored at -70°C. For MUG hydrolysis activity, 150µl of extract were then mixed with 100µl of MeOH and 250µl of "MUG 2x" solution ($NaPO_4$ buffer 50mM pH 7.0, 4-methylumbelliferyl β-D-glucuronide [MUG, Sigma] 2mM, EDTA 10mM; Triton X-100 0.1%, sodium lauryl sarkosyl 0.1%, DTT 10mM). A SFM 25 fluorometer (Kontron Instruments AG, Switzerland) was used, with excitation at 365nm and detection at 445nm. All experiments were performed at room temperature, and started in the fluorometer room, which was warmer than the laboratory by sometimes more than 5°C: this detail can have a significant importance when measuring little activity differences.

After 30 minutes of incubation at room temperature, 50µl of solution were taken, and Na_2CO_3 0.3M was added to a final volume of 600µl. Sodium carbonate is used in order to raise the pH of the final solution to more than 10: at such an alkaline pH, the fluorescence signal is strongly increased. I also took another aliquot after 110': I noticed indeed that the ratio 110'/30' was always close to 3 if the reaction was taking place normally. It happened that this ratio was lower: the reasons were either a silencing occurred in a transgenic line (Col x Col GH3::*GUS* line "A"), or a too strong activity resulting in complete hydrolysis of the substrate.

As samples had quite different activities, I did not take at first the same amount of extract for analysis. However, I then noted that if I was taking 20µl and then 40µl of MUG incubation mixture from the same sample, the fluorescent signal was not doubled: the ratio was rather lower than 2, usually around 1.5. I hence decided to try several dilutions from three different Col GH3::*GUS* control lines (Fig.2-7). The results clearly show a non-directly proportional relationship between volume analysed and signal observed. The curves obtained are quite similar between lines; therefore, different lines (mutants and wild types) are still comparable between them. Following these conclusions, I decided to always take the same volume (50µl) of MUG incubation mixture for each sample in the next experiments.

Protein concentration was also another key parameter: depending on the extraction yield and on the tissue analysed, total protein amounts were quite variable. Fluorescence was hence always divided by protein concentration measured by a classical BioRad assay (595nm). Finally, some transgenic lines had such a strong GUS activity that taking 150µl of extract for incubation with MUG resulted in fluorescence values out of range for the fluorimeter and a total consumption of MUG before 30 minutes. Therefore, I had to dilute extract in a final volume of 150µl with fresh extraction buffer. In order to avoid variability due to pipetting errors, I made three dilutions of each sample, and three measurements of each dilution: with 9 measurements per sample, I can consider my experiments to me trustable. Although it might seem trivial, this precaution was important when analysing small activity differences.

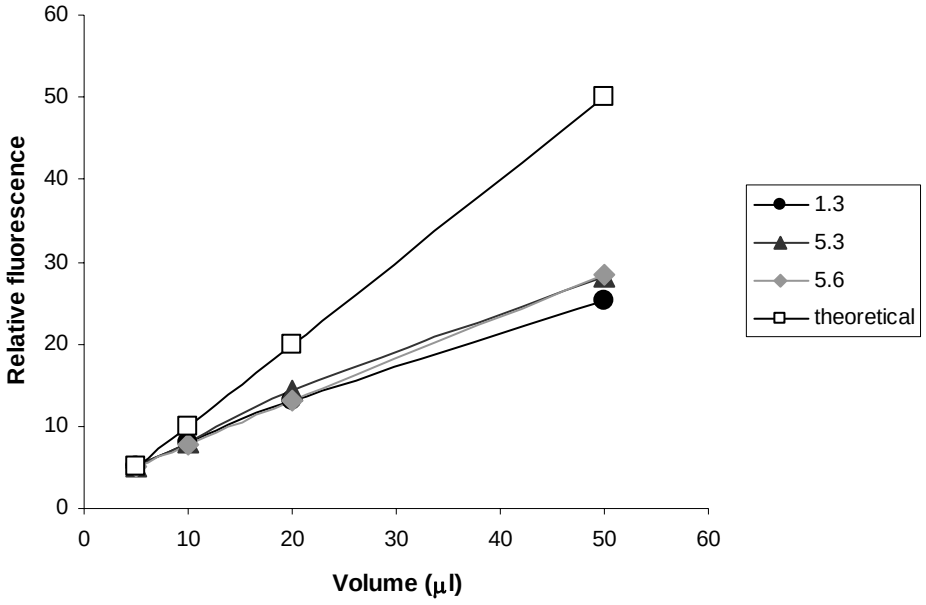


Figure 2-7

Study of fluorescence (MUG hydrolysis) for various sample volumes

Col GH3::*GUS* lines "1.3", "5.3" and "5.6" were studied. Activity was measured after 30 min. Fluorescence was standardised to the value "5" for volumes of 5 µl. The theoretical curve corresponds to a direct proportionality relationship.

B.6 Article project : "The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?"

In the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**", we have tried to better understand the interrelationship existing between auxin and the peroxidase AtPrx33. The main result at the origin of this project was the co-localisation observed between the peroxidase and the GH3::*GUS* transgene expression in the root elongation zone. Although there is no definitive evidence that both players interact *in vivo*, absence of co-localisation of AtPrx33, auxin and the superoxide anion in this zone correlates with a decreased root length. We hence propose the following model: the triad AtPrx33-auxin-O_2 may be responsible for production of superoxide, which in turn softens the cell wall and hence favours elongation.

The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?

Filippo Passardi[1], Claire Kevers[2], Claude Penel[1] and Christophe Dunand[1]*

[1]Laboratoire de Physiologie Végétale, Université de Genève, Quai Ernest-Ansermet 30, CH-1211 Genève 4, Switzerland. [2] Unité de Biologie Moléculaire et Biotechnologie Végétales, Université de Liège, Sart Tilman B22, 4000 Liège, Belgium.

*author for correspondence; Tel.: +41223793012; Fax +41223793017; e-mail: christophe.dunand@bota.unige.ch

Abstract

Auxin is an essential hormone for plant development: depending on its concentration and on the organ, it can promote or inhibit cell elongation. Auxin catabolism is strongly suspected to be regulated, at least in part, by class III peroxidases, as most of these enzymes rapidly oxidise auxin *in vitro*. Recently, we have shown that two closely related *Arabidopsis* peroxidases (AtPrx33 and AtPrx34) are involved in root cell elongation. In the following study, we performed a more detailed analysis of *AtPrx33* expression through fusion of *GFP* to the *AtPrx33* promoter. Parallel experiments with GH3::*GUS* lines suggested a co-localisation of auxin and AtPrx33 in the elongation zone. However, when perturbing elongation through different chemical treatments, auxin and AtPrx33 distributions did not vary always the same way. Auxin dosage was further conducted on single and double knock-out, as well as over-expressor lines of *AtPrx33* and *AtPrx34*: results indicate that both peroxidases have a similar influence on auxin content, but that their effect is different depending on their localisation in roots or shoots. In roots, AtPrx33 and AtPrx34 actively participate in cell elongation. In shoots, both peroxidases may be involved in auxin homeostasis. AtPrx33 (and probably AtPrx34) has many occasions to meet auxin within a seedling: do the peroxidase and the hormone only look at each other, or do they sometimes dance together? We favour the dance, but we also show that other players can break this close relationship and move each partner to separate places.

Key words : Auxin, elongation, GFP, GH3 promoter, glucuronidase, peroxidase, superoxide

Introduction

Since their probable appearance during land colonisation, the number of peroxidase paralogs per organism has constantly increased throughout evolution: in higher plants, they form large multigenic families (Duroux and Welinder, 2003; Passardi et al., 2004; Bakalovic et al., 2006). Up to 138 peroxidase-encoding genes were found in *Oryza sativa* (rice), and 73 in *Arabidopsis thaliana* (Tognolli et al., 2002). In addition, within a tissue, several peroxidases are usually expressed at the same time (Valério et al., 2004). The redundancy of expression (and probably of activity) has made characterisation of a single peroxidase function very complex. Rises in expression of specific peroxidases have been reported in various conditions, leading to the attribution of several general roles: protection of germinating seeds, growth by elongation, building of cell walls, defence against pathogens or senescence (Penel et al., 1992; Hiraga et al., 2001; Passardi et al., 2005). Recently, we have identified a function *in vivo* for two *Arabidopsis* peroxidases, *AtPrx33* and *AtPrx34*. Both peroxidases are close paralogs, and are mainly found in roots. Knocking out or over-expressing *AtPrx33* provoked, respectively, the shortening or lengthening of roots. The same results were obtained with *AtPrx34*. The double knock-out mutant *atprx33*/*atprx34* showed an additive effect (Passardi et al., 2006).

Root elongation is dependent on the concentration of auxin present in the elongation zone: root growth is promoted at 10^{-10} to 10^{-9} M, whereas at 10^{-6} M, the process is stopped (Taiz and Zeiger, 2002). Despite the whole mechanism is not yet fully understood, a widely accepted hypothesis for initiation of cell elongation by auxin is the acid growth theory (Rayle and Cleland, 1977): auxin activates H^+-ATPases, resulting in acidification of the cell wall. Changes in pH lead to a cascade of events promoting cell expansion, such as increase of turgor pressure and activation of expansins (Cosgrove, 2000; Becker and Hedrich, 2002). In parallel, many genes bearing auxin-responsive elements (AuxRE) are activated, including H^+-ATPases and members of the SAUR, Aux/IAA and GH3 families (Abel and Theologis, 1996; Hagen and Guilfoyle, 2002; Goda et al., 2004). Auxin does not directly bind to AuxRE: its targets are the proteins members of the Aux/IAA family, which repress the action of AUXIN RESPONSE FACTOR (ARF) proteins. ARFs are bound to AuxRE independently of the auxin status in the cell. However, their action is inhibited by the heterodimer formed with Aux/IAA proteins. When auxin levels increase, the hormone first binds to the TIR subunit of the SCF complex (a ubiquitin protein ligase). Auxin-bound TIR then interacts with Aux/IAAs, resulting in ubiquitination of Aux/IAAs and their subsequent degradation through the proteasome pathway (Worley et al., 2000; Gray et al., 2001; Dharmasiri et al., 2005).

DNA-bound ARFs can then exert their activator function, which is further stimulated by additional non-DNA bound ARFs (Ulmasov et al., 1999). Whereas the complex pathway leading to auxin activation of genes starts to be elucidated, the proteins issued from this up-regulation are poorly characterised. Microarray analyses performed on *Arabidopsis* seedlings treated with exogenous auxin showed that other genes than the "classical" H^+-ATPases, SAUR, Aux/IAA and GH3 families were induced (Sawa et al., 2002; Tian et al., 2002; Goda et al., 2004). Peroxidases were not detected in early time points (less than 2 hours), however one was induced after a 2-hour auxin incubation (Tian et al., 2002), and five in seedlings exposed to exogenous auxin during periods ranging from 3 to 24 hours (Goda et al., 2004). These results suggest that peroxidases are involved in later responses to auxin stimulation, such as cell and tissue remodelling (Passardi et al., 2005).

In the microarray analyses performed by Tian *et al.* (2002) and Goda *et al.* (2004), *AtPrx33* and *AtPrx34* were not particularly up-regulated. Following our findings that both peroxidases have an influence on root elongation, we transformed plants with the *AtPrx33* promoter fused to the green fluorescent protein gene (*GFP*) in order to determine the localisation of *AtPrx33* in the root, and hence to better explain its function. In addition, we obtained lines bearing the glucuronidase (*GUS*) reporter gene under the control of the auxin-responsive promoter GH3 (Liu et al., 1994). Our results indicate that *AtPrx33* expression as well as auxin localisation are the highest in the elongation zone of the root. Treatments affecting root length were also able to modify both *AtPrx33* and GH3::*GUS* repartition, although not always in a similar way. Finally, auxin dosage of knock-out and over-expressor lines of *AtPrx33* and *AtPrx34* points to an implication of both peroxidases in auxin homeostasis, particularly in shoots.

Materials and methods

Plant material

The Columbia (Col) and Wassilewskija (Ws) ecotypes of *Arabidopsis thaliana* were used as wild type controls. Stable homozygous *Arabidopsis thaliana* T-DNA lines *atprx33* (Wassilewskija ecotype) and *atprx34* (Columbia ecotype) were obtained and characterized as described previously, together with the over-expressor line (Col) CaMV 35S::*AtPrx34* (Passardi et al., 2006). The *atprx33* line was further transformed with a sense-antisense construct *43xrPtA-linker-AtPrx34* (RNAi strategy) for generation of the *atprx33*/*34* double knock-out line (Passardi et al., 2006).

Transgenic lines

In order to obtain the AtPrx33::*GFP-Term* construct, 1000 bp of the *AtPrx33* promoter were first amplified from genomic DNA with primers 5'pCA KpnI 5'-GGGGTACCGCTTGGTTTGGTTTCCATTG-3' and 3'pCA NcoI 5'-CATGCCATGGTTTTCACAAGGAC-3' (engineered restriction sites are underlined). The PCR product was directly ligated to pCR II plasmid (Invitrogen Ltd, United Kingdom), and then subcloned into pAVA 393 with KpnI and NcoI (von Arnim et al., 1998) upstream from a *GFP-Term* sequence. The AtPrx33::*GFP-Term* cassette was further subcloned (KpnI/BamHI) into the binary vector pZP222 (Hajdukiewicz et al., 1994). A second construct used as a control has been obtained by cloning CaMV 35S::*GFP-Term* in pCGN1547 (McBride and Summerfelt, 1990).

GH3::*GUS-Term* constructs used for colorimetry and fluorimetry experiments were derived from the plasmid pJJ430 GH3::*GUS-Term* (Larkin et al., 1996). The GH3::*GUS-Term* cassette was excised with EcoRI and cloned into EcoRI-digested binary vector pZP222.

atprx33 x Ws GH3::*GUS-Term* transgenic lines were obtained by manually fertilizing *atprx33* homozygous carpels with pollen issued from Ws GH3::*GUS-Term* flowers, and then selecting double transgenic plants with kanamycin 50µg/ml (*atprx33*) and gentamicin 75µg/ml (GH3::*GUS*). Ws wild type carpels were also fertilized with Ws GH3::*GUS-Term* pollen in order to obtain the control line Ws wild type x Ws GH3::*GUS-Term*. The same procedure was applied to obtain *atprx34* GH3::*GUS* lines. In figure 6, the control lines and the respective mutant contain hence the transgene in the same insertion site within the genome. Control lines issued from the cross Col/Ws wild type with Col/Ws GH3::*GUS* are noted Col x Col GH3::*GUS* and Ws x Ws GH3::*GUS*.

Plant transformation

pZP222 constructs were introduced into *Arabidopsis thaliana* by *Agrobacterium tumefaciens* strain ASE-mediated transformation using the spraying technique (Bent, 2000). Transformants were selected on gentamicin 75µg/ml. The presence of the transgene was confirmed by GFP fluorescence and X-Gluc staining. Homozygous lines were then obtained through antibiotic selection.

Fluorescence experiments

AtPrx33::*GFP* homozygous seedlings were observed for GFP fluorescence after a 1 week growth on ½ MS medium under a light intensity of 180µE/m^2/sec. GFP detection was performed with an Axioplan 2 Zeiss microscope, a monochromator-mediated narrow band excitation (475±2nm), a narrow band EGFP emission (510±10nm) filter (Filter set 41020, Chroma Technology Corp, USA), and using the Metamorph software (Molecular Devices Corp, USA).

Quantification of GFP fluorescence was performed with Adobe Photoshop CS2 software v.9.0 (Adobe Systems Inc., USA). Average luminosity was reported on three zones: root tip, cell division zone and elongation zone. Cell division zone was used as a reference in order to detect variations of fluorescence in the root tip and the elongation zone.

Colorimetry experiments

X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt) was purchased from X-GLUC DIRECT (United Kingdom). Staining was performed as described elsewhere (Blazquez et al., 1997; Lariguet et al., 2003). Briefly, whole seedlings were soaked in cold (4°C) acetone 90%, then rinsed and stained (5' at 4°C and 20' at 37°C) in a solution containing 2mM X-Gluc, 2mM ferrocyanide, 2mM ferricyanide and 10mM sodium phosphate buffer. Seedlings were then mounted on 50% glycerol and observed with a Leica MZ16 microscope (Leica Mycrosystems GmbH, Germany).

For NBT (Nitro blue tetrazolium, Sigma, Switzerland) staining, seedlings were incubated during 15 minutes in a 2mM solution (20mM pH 6.1 phosphate buffer), and then visualized under a Leica MZ16 microscope.

Fluorimetry

For fluorimetry experiments, glucuronidase was extracted from GH3::*GUS* lines by grinding seedlings in an appropriate extraction buffer (KPO_4 buffer 100mM pH 7.8, DTT 1mM, leupeptin 20µg/ml), and collecting the supernatant after centrifugation at 12'000 g. 150µl of extract were then mixed with 100µl of MeOH and 250µl of "MUG 2x" solution ($NaPO_4$ buffer 50mM pH 7.0, 4-methylumbelliferyl β-D-glucuronide [MUG, Sigma] 2mM, EDTA 10mM; Triton X-100 0.1%, sodium lauryl sarkosyl 0.1%, DTT 10mM). The mixture was incubated at 25°C and measurements were made after 30' (not shown) and 120' with a SFM 25 fluorimeter (Kontron Instruments AG, Switzerland), under an excitation wavelength of 365nm and a detection at 445nm. Values obtained were then normalised to protein concentration determined with Bio-Rad reagent (see section "peroxidase activity").

Reverse transcriptase PCR

Reverse transcriptase PCR was used as a semi-quantitative method to assess the expression of peroxidase genes. Whole seedlings were ground in liquid nitrogen at the following time points: 7 day-old (grown at 25°C); 2, 5 and 9 days after incubation at 4°C; 18, 24 and 48 hours back at 25°C after cold treatment. Total RNA was then extracted and purified with the Tri-reagent solution (Sigma). After quantification of the concentration by spectrophotometry and confirmation by electrophoresis, 1µg of the crude RNA preparations was treated with one unit of RNase-free DNase I (Promega). The DNA-free RNA was then used as a template during reverse transcription according to the ImPromII RT protocol from Promega. PCR amplification was conducted for up to 40 cycles using the following thermal profile: denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and polymerization at 72°C for 30 sec, with a 5 min terminal extension step at 72°C. The primers used were specific for each peroxidase, as described previously (Valério et al., 2004).

Quantification of the RT-PCR bands on agarose was performed with the software QuantityOne (BioRad Laboratories Inc., Switzerland) on at least three independent RT-PCR reactions for each time point. In order to compare the level of transcription of each peroxidase RNA, the same peroxidase-specific primers were used to amplify *Arabidopsis* genomic DNA. Intensity of the resulting bands was quantified, thus giving a measurement of the efficiency of each primer pair. RT-PCR intensities were then divided by the corresponding primer efficiency, normalised to *AtPrx42* transcription level, and were reported as "level of RNA transcripts". *AtPrx42* is indeed known to have a stable and strong expression at many stages

of *Arabidopsis* development (Kjaersgard et al., 1997; Welinder et al., 2002; Valério et al., 2004; Passardi et al., 2006).

Root length experiments

For the study of treatments influencing root length, *Arabidopsis* (Columbia) seeds were sown on square (12x12mm) Petri dishes either containing 40ml ½ MS medium, or on ½ MS supplied with at least 3 different concentrations of indole-3-acetic acid (Sigma), sirtinol (Sigma), brassinolide (Super-Grow, Canada), 2,3,5-triiodobenzoic acid (Serva, Germany), diphenyleneiodonium (Acros Organics, Belgium) or potassium iodide (Merck AG, Switzerland). After two days of vernalisation at 4°C, plates were grown vertically under a light intensity of 180µE/m^2/sec. Length of roots was measured on 7 day-old seedlings.

Peroxidase activity

For guaiacol peroxidase activity, the seedlings used for root length measurement were ground (three separate extractions per condition) in HEPES 20mM, EGTA 2mM (200µl per 100mg fresh weight), centrifuged, and the supernatant was assayed for guaiacol oxidase activity (1mM phosphate buffer pH 6.1, guaiacol 0.5%, 1mM H_2O_2) at 470nm (Ultrospec 2000 spectrophotometer, Pharmacia Biotech, Siwtzerland).

Protein concentration was then measured at 595nm with Bio-Rad reagent (ref. 500-0205, Bio-Rad Laboratories GmbH, Germany) and calculated by comparison with a bovine serum albumin standard titration. Nanokatal (nkat-mol of tetra-guaiacol formed per second) activity was then calculated using the Beer-Lambert law with a molar absorption coefficient of 26.6 [mol l^{-1} cm^{-1}] for tetraguaiacol formation (Chance and Maehly, 1955).

Auxin extraction and determination

The extraction and analytical methods have been performed as described (Nordström and Eliasson, 1991). Materials were homogenised in liquid nitrogen. The powder was extracted with 5mM phosphate buffer (pH 6.5) containing 3-[5(n)-^{3}H]indolylacetic acid as internal standard and butylated hydroxytoluene as antioxidant. After incubation in darkness for 1h, the extract was filtered through a glass-fibre filter under vacuum. The filtrates were purified through Bond-Elute C18 column conditioned at pH 6.5. The pH of the eluates was adjusted to 2.5 using 2.5 M phosphoric acid and then applied to C18 columns (chromabond) pre-conditioned at pH 2.5. The columns were washed with distilled water, followed by acidic ethanol (ethanol/glacial acetic acid/water, 20/2/78 v/v). A second purification of the last

eluates was performed on a second C18 column at pH 2.5. Auxins were eluted from the second C18 columns with 2 x 300µl aliquots of 80% methanol. 50µl of the methanolic extract were injected in an automated Merck-Hitachi HPLC system in the same conditions of elution pattern as those previously described (Heloir et al., 1996): Lichrospher 100-RP18 column, 12.5cm x 4mm internal diameter, 5µm particle size; column and solvent at 30°C; flux 1 cm^3 min^{-1}; mobile phase acetonitrile: glacial acetic acid: water (10:2:88, v/v); detection by fluorescence detector (absorbance 292nm, emission 360nm).

Sequence analysis

The 1,000-bp region upstream each peroxidase gene was analyzed using PLACE software (http://www.dna.affrc.go.jp/htdocs/PLACE).

Statistical analysis

1-way between subjects ANOVA tests were performed in order to evaluate statistical differences with the software Analyse-it (v.1.73, Analyse-it software Ltd., United Kingdom).

Results

Both AtPrx33 expression and auxin levels are particularly high in the elongation zone

Once obtained, two independent and homozygous AtPrx33::*GFP* transgenic lines were observed for GFP fluorescence after 7 days of growth on agar. GFP was present ubiquitously in the whole seedling (Fig. 1a, b) with a distribution comparable to a control line CaMV 35S::*GFP*. In the root end however, *AtPrx33* expression was strongly localised in the elongation zone (Fig. 1c), in contrast with the control 35S::*GFP* line, where fluorescence was more diffuse (Fig.1f). The level of expression of *AtPrx33* was much weaker than 35S::*GFP* (Fig. 1c, h).

In parallel with this study, *Arabidopsis* lines transformed with the glucuronidase gene under the control of an auxin-inducible promoter (GH3::*GUS*) were grown on the same conditions and the GUS substrate X-Gluc was added. The results show that staining is mainly present in the elongation zone, similarly to AtPrx33::*GFP* (Fig. 1e). The *atprx33* x Ws GH3::*GUS* line was further analysed with X-Gluc, but no significant difference was observed compared to the control Ws x Ws GH3::*GUS*, indicating that loss of *AtPrx33* does not alter auxin patterning (data not shown).

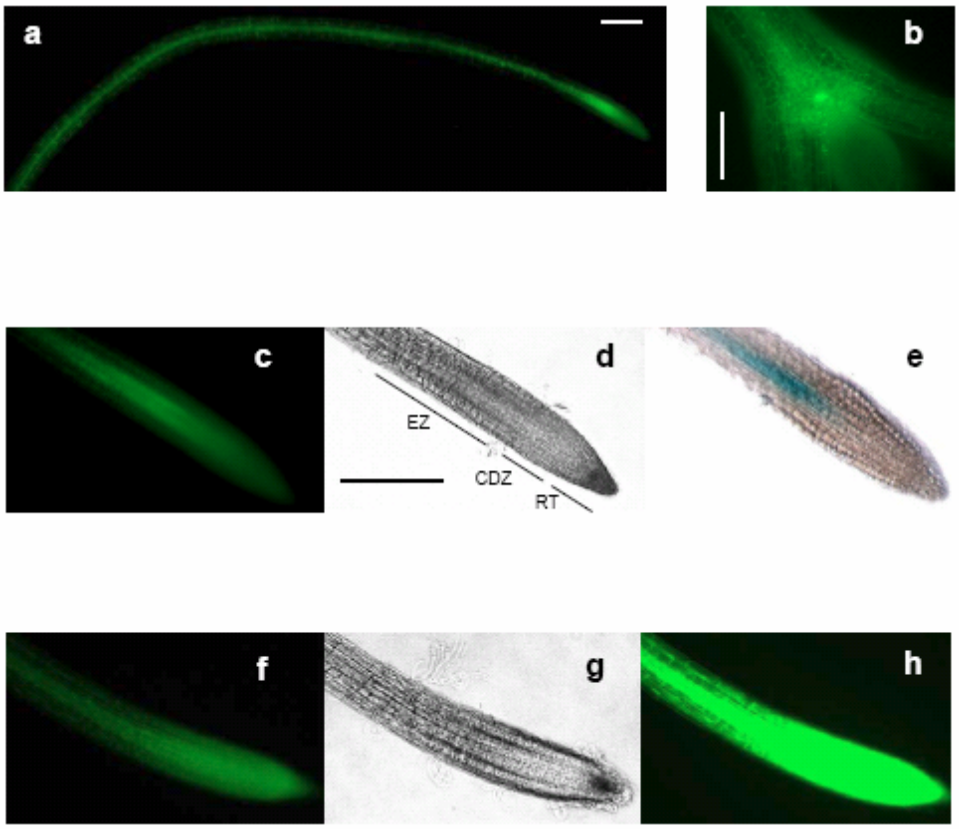


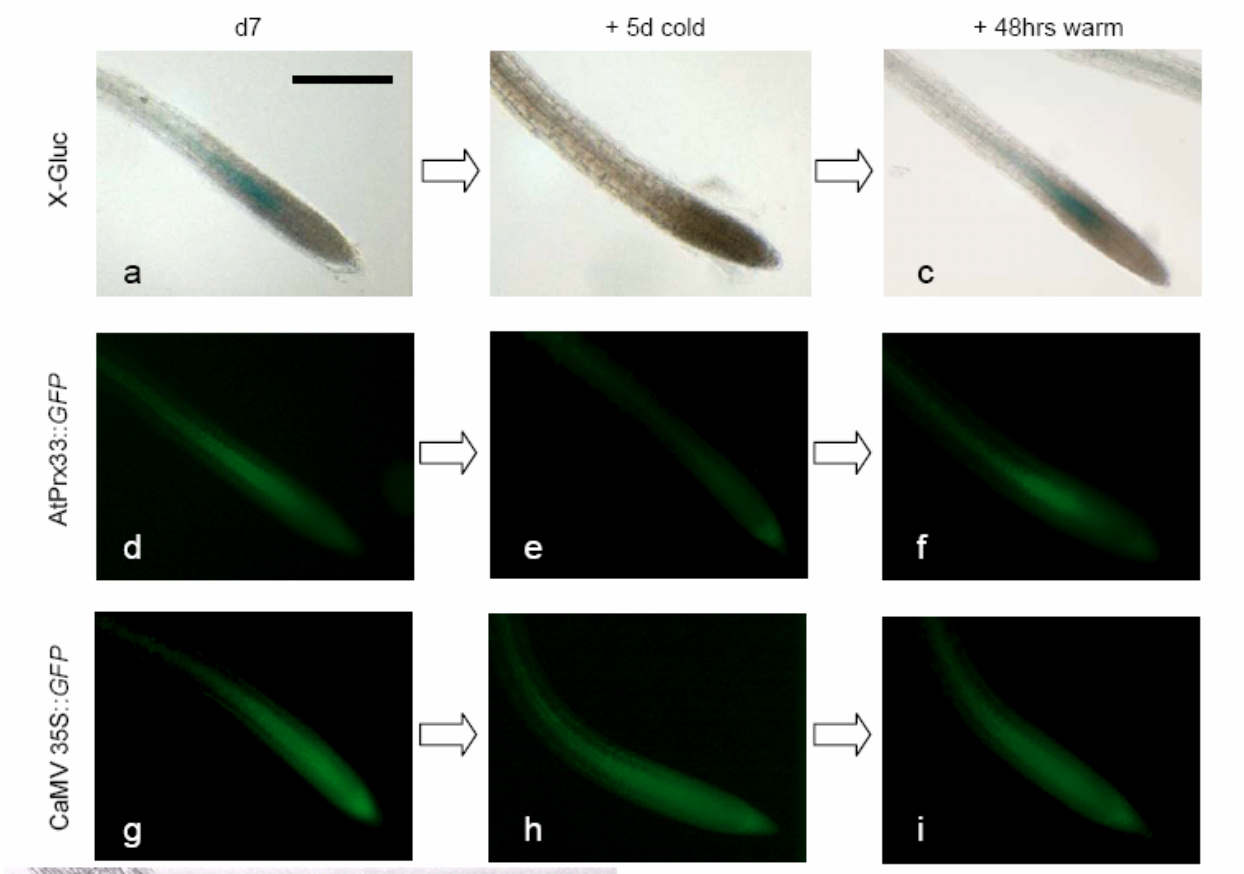
Figure 1 (see legend in next page)

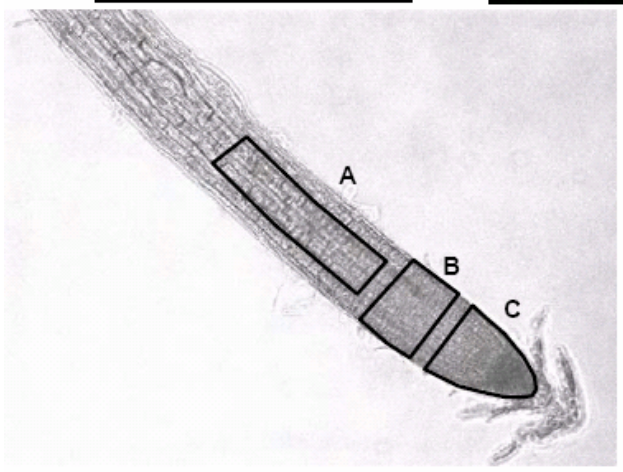
Figure 1 (previous page)
Detection of GFP fluorescence and GUS activity (X-Gluc staining) in 7 day-old seedlings transformed with AtPrx33::*GFP* (**a-c**), CaMV 35S::*GFP* (**f, h**) and GH3::*GUS* (**e**). **a** whole root. **b** apical meristem. **c-h** root end. **d** is the phase contrast picture of **c**, and **g** of **f** and **h**. All GFP pictures were taken with the same exposure time (1000ms), except for **f** (100ms). EZ: elongation zone; CDZ: cell division zone ; RT : root tip. Scale bars represent 250µm.

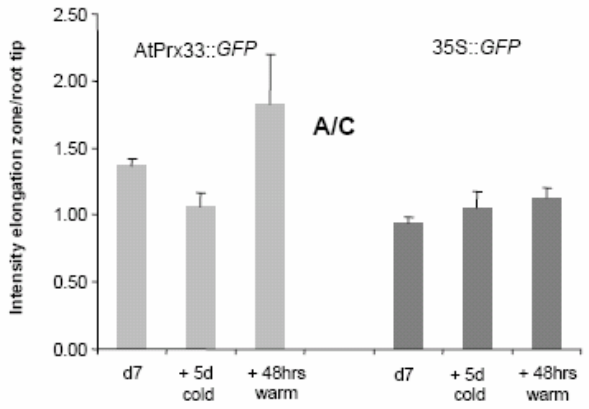
Cold differentially affects expression of peroxidases

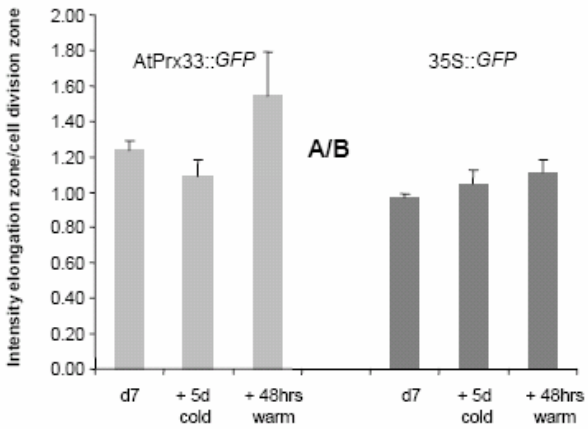
Further analysis for X-Gluc staining and AtPrx33::*GFP* expression was performed on seedlings subjected to cold treatment for up to 9 days, and then put back at 25°C for up to 48 hours. After 5 days at 4°C, the level of auxin decreased to undetectable levels in the GH3::*GUS* lines, according to X-Gluc experiments (Fig. 2a, b). Staining was still comparable to 25°C-grown seedlings in the hypocotyl and the aerial parts (data not shown). In parallel, AtPrx33::*GFP* expression also decreased in the elongation zone. In contrast with X-Gluc staining, we observed a significant increase of fluorescence in the root extreme tip (Fig. 2d, e, k, l, m). This increase was not detected in the control line 35S::*GFP*, where the fluorescence intensity was not significantly changing in any of the regions observed (Fig. 2g, h, k, l, m). When seedlings were taken out from 4°C and grew again at 25°C, AtPrx33::*GFP* expression rose in the elongation zone and decreased at the root tip within 48 hours, which was not the case for the control line 35S::*GFP* (Fig. 2f, i, k, l, m). The level of fluorescence in the elongation zone of AtPrx33::*GFP* roots was even stronger than at d7. In parallel with the increase of fluorescence, GH3::*GUS* expression recovered the same levels as before cold exposure in the elongation zone (Fig. 2c).

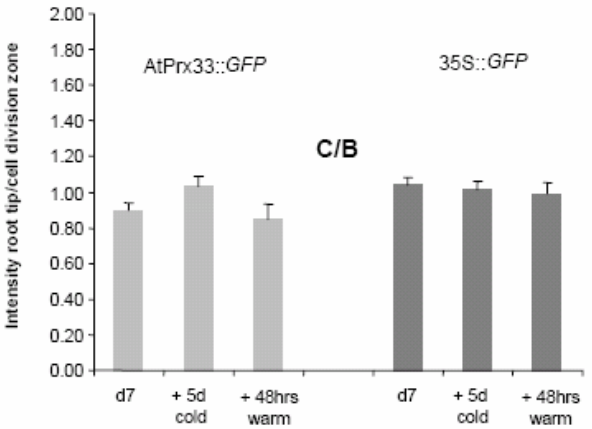
Figure 2 (next page)
Detection of GFP fluorescence and GUS activity (X-Gluc staining) after cold exposure in roots of seedlings transformed with GH3::*GUS* (**a-c**), AtPrx33::*GFP* (**d-f**) and CaMV 35S::*GFP* (**g-i**). 7 day-old seedlings were stored at 4°C during 5 days, and were then put back to 25°C during 48 hours. At each time point, seedlings were either incubated in X-Gluc (**a-c**) or directly observed for fluorescence under a microscope (**d-i**). **j** Quantification of fluorescence was performed on at least 10 roots for each transgenic construct at each time point on three limited areas: elongation zone (A), cell division zone (B) and root tip (C). Ratios A/C (**k**), A/B (**l**) and C/B (**m**) are reported for both AtPrx33::*GFP* and CaMV 35S::*GFP*. Scale bar represents 250µm.


d7
+ 5d cold
+ 48hrs warm
X-Gluc
AtPrx33::GFP
CaMV 35S::GFP
a
c
d
e
f
g
h
i


A
B
C

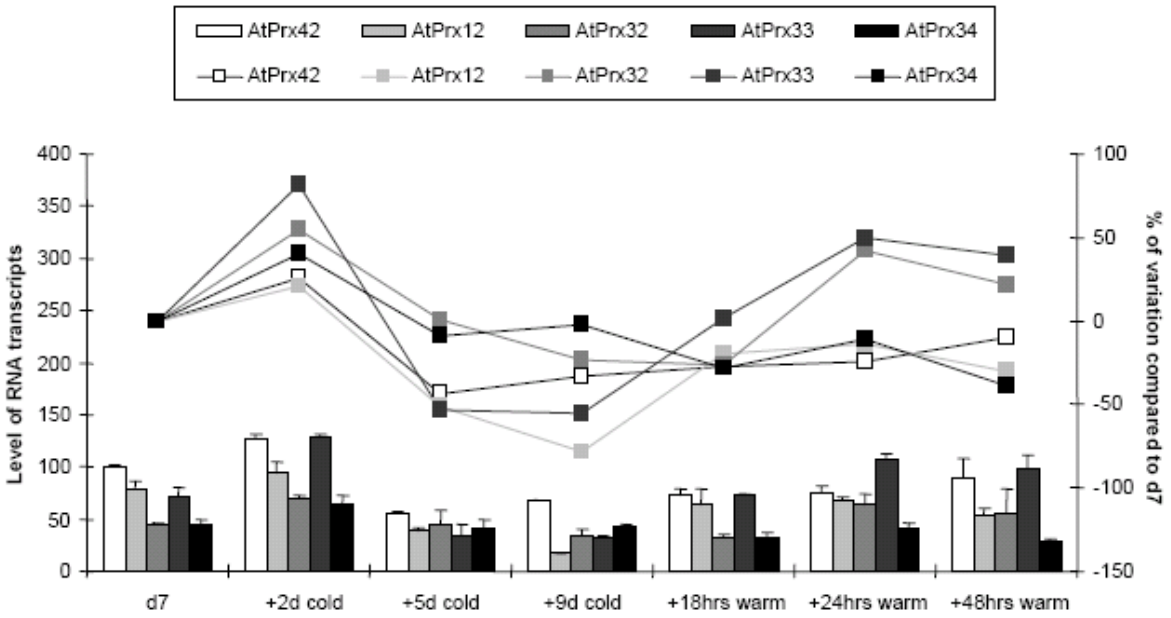

Intensity elongation zone/root tip
AtPrx33::GFP
35S::GFP
A/C
2.50
2.00
1.50
1.00
0.50
0.00
d7
+ 5d cold
+ 48hrs warm


Intensity elongation zone/cell division zone
AtPrx33::GFP
35S::GFP
A/B
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
d7
+ 5d cold
+ 48hrs warm


Intensity root tip/cell division zone
AtPrx33::GFP
35S::GFP
C/B
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
d7
+ 5d cold
+ 48hrs warm

RT-PCR reactions were then performed in order to check the level of *AtPrx33* RNA transcripts at various time points of cold exposure and warm recovery, together with its closest orthologs *AtPrx32* and *AtPrx34* (Tognolli et al., 2002). *AtPrx42*, a peroxidase known to have a stable and strong level of RNA transcription in all organs, as well as the highest number of EST counts of all *Arabidopsis* peroxidases (Kjaersgard et al., 1997; Welinder et al., 2002; Valério et al., 2004; Passardi et al., 2006), was also analysed and used as a reference. After 2 days of cold exposure, the level of transcription of all peroxidases increased, but at different amplitudes (Fig. 3). Whereas *AtPrx42* expression remained quite stable, as expected (+26% compared to d7), *AtPrx33* responded strongly (+81%) to the cold stress, reaching a similar level of expression as *AtPrx42*. 5 days and 9 days of cold exposure induced a general drop in transcription, the more resistant peroxidases being *AtPrx32* and *AtPrx34* (respectively +1% and -9% compared to d7), and the more sensitive being again *AtPrx33* (-53%). Replacing seedlings at 25°C induced a rapid rise in *AtPrx32* and *AtPrx33* expression, which was not the case for *AtPrx42* (almost no change) or *AtPrx34* (slight decrease). *AtPrx33* levels were even higher than *AtPrx42*.

We then screened the first 1000bp of each peroxidase promoter for presence of cold-responsive elements. Two motifs were found: LTRE-1 (Dunn et al., 1998) and AtMYC-consensus (Chinnusamy et al., 2003). The latter motif was found in all the peroxidase promoters analysed, from once (*AtPrx32*) to 10 times (*AtPrx33*). As *AtPrx33* showed the strongest induction at the time point "+2d cold", this motif could then be responsible for the variation observed after 2 days at 4°C. In order to verify this hypothesis, we searched various other peroxidase promoters for presence of the AtMYC-consensus motif, and found it four times in the first 1000bp of *AtPrx12*, a peroxidase unrelated to the cluster *AtPrx32*/*33*/*34* (Tognolli et al., 2002). Following RT-PCR reactions from the same cDNA samples used for the other peroxidases, *AtPrx12* showed to be not particularly induced after 2 days at 4°C (Fig. 3). Whereas we cannot exclude a role for AtMYC-consensus, it appears that other factors are playing a determinant role in cold induction. The RT-PCR results suggest nonetheless that *AtPrx33* is very sensitive to both cold temperature and recovery from cold conditions, particularly when compared to its close orthologs *AtPrx32* and *AtPrx34*.



		Position for				
Motif	Sequence	AtPrx42	AtPrx12	AtPrx32	AtPrx33	AtPrx34
LTRE-1	CCGAAA	654(+)	371(+)		924(-)	414(+)
AtMYC-consensus	CANNTG	546, 620	255, 622, 919, 988	683	122, 372, 388, 461, 617, 667, 753, 810, 864, 954	333, 737

Figure 3

Expression levels of *AtPrx12*, *32*, *33*, *34* and *42* mRNA assessed by RT-CPR on 7 day-old seedlings ("d7"), after 2, 5 or 9 days of exposure to 4°C ("+2d cold", "+5d cold", "+9d cold"), and 18, 24 or 48 hours after recovery from cold at 25°C ("+18hrs warm", "+24hrs warm", "+48hrs warm"). Histogram (left axis): level of RNA transcripts, standardised (set to 100) to the level of *AtPrx42* at d7; standard deviation is indicated by t-bars. Line plot (right axis): percentage of variation of the cDNA level compared to the initial level at d7. Each measurement is the result of two independent experiments and three independent RNA extractions for each experiment. Table: Position of cold-responsive motifs in the first 1000bp of the peroxidase promoters. Position indicated corresponds to distance from the START codon. (+): sense strand; (-) reverse strand.

Treatments modifying root length influence peroxidase activity

Following our finding that *AtPrx33* is mainly expressed in the elongation zone, *Arabidopsis* seeds were germinated and grown during 7 days on agar containing different substances in order to alter root growth. In a first series, indole-3-acetic acid (auxin-IAA) and sirtinol (SIR), a product known to induce a similar set of genes as auxin (Zhao et al., 2003), were used at concentrations ranging from 0.1 to 10µM. Root elongation was similarly reduced for both treatments, with a stronger effect of IAA at 1 and 2µM (Fig. 4a). In parallel, total peroxidase activities were analysed through guaiacol oxidation reactions: curiously, at 2µM, IAA induced a marked decrease, but SIR still had no significant effect on peroxidase activity; at 10µM, both treatments resulted in the same reduction (Fig. 4b).

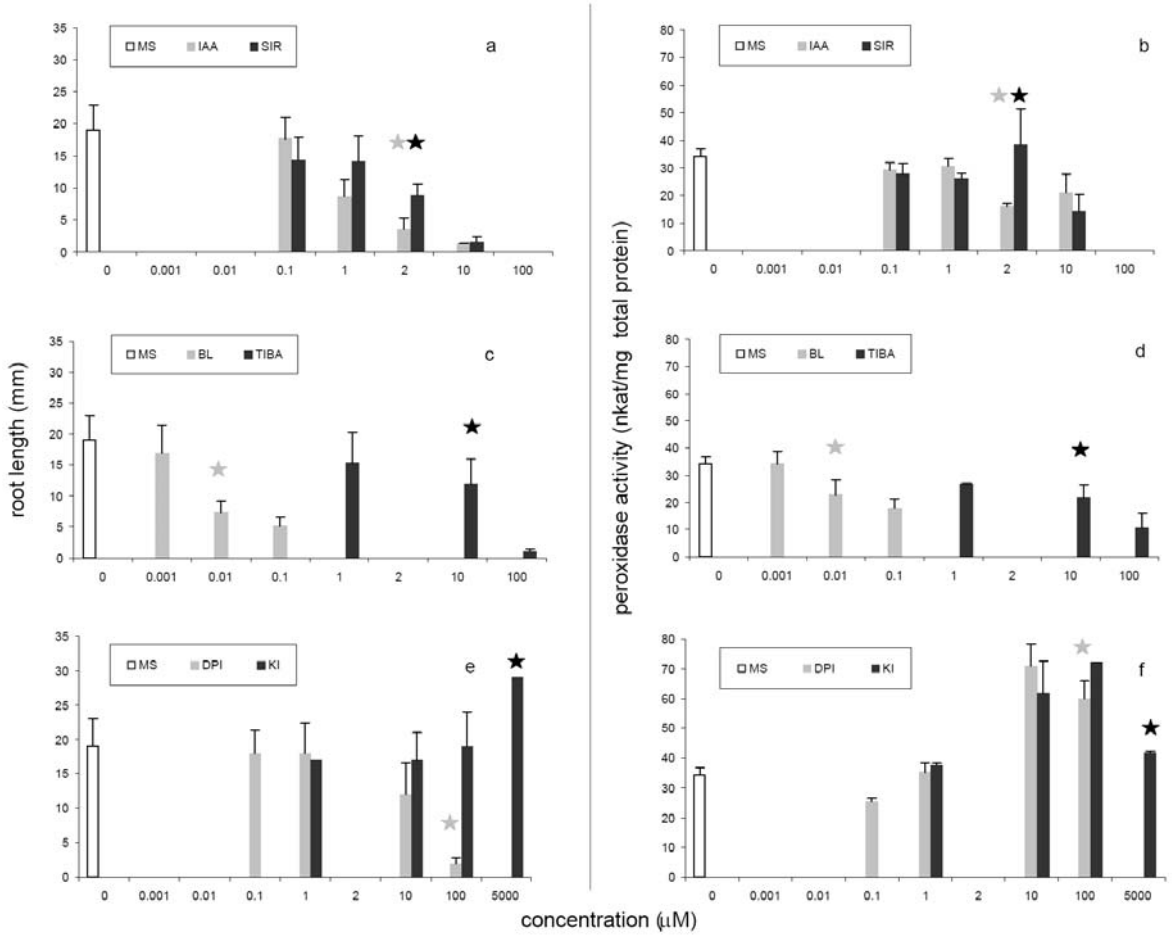


Figure 4

Measurements of root length (**a, c, e**) and total peroxidase activity (**b, d, f**) of 7 day-old seedlings grown on various media. Indole-3-acetic acid (IAA), sirtinol (SIR) (**a, b**); brassinolide (BL), 2,3,5-triiodobenzoic acid (TIBA) (**c, d**); diphenyleneiodonium (DPI), potassium iodide (KI) (**e, f**). ★★: concentrations chosen for further experiments (Fig. 5). Standard deviation is represented by t-bars.

Two products that modulate auxin transport were next assayed: brassinolide (BL), which has been recently proposed to provoke the accumulation of the PIN2 auxin carrier in the elongation zone (Li et al., 2005), and the known auxin transport inhibitor 2,3,5-triiodobenzoic acid (TIBA). In contrast with IAA and SIR treatments, both BL and TIBA induced a reduction in root elongation paralleled with a drop in peroxidase activity (Fig, 4c, d). The range of concentration inducing a similar effect was one order of magnitude lower for BL compared to TIBA. The last two substances studied were chosen for their action on radical oxygen species (ROS), since ROS are suspected to have a role on cell elongation (Schopfer et al., 2002; Liszkay et al., 2003). Diphenyleneiodonium (DPI), a NADPH oxidase inhibitor, can prevent formation of superoxide; potassium iodide (KI), which is readily oxidised by H_2O_2 to iodine, leads to a strong decrease in H_2O_2 levels *in planta*. Whereas increasing amounts of DPI reduced root elongation, high concentration (5mM) of KI resulted in longer roots (Fig. 4e). Guaiacol oxidase activities changes were opposite to the effect on root length: shorter roots induced higher peroxidase activity with DPI, and longer roots issued in weaker peroxidase activity with 5mM KI (Fig. 4f).

Modulation of auxin distribution and AtPrx33 patterning in elongation-altered roots

The results obtained in figure 4 demonstrate that perturbing root elongation generally affects global peroxidase activity. We could not however establish a clear relationship between the two parameters. Decrease of root length usually correlates with a decrease in peroxidase activity, but in some cases the opposite effect was observed (SIR 2µM, DPI, KI). We hence decided to study the localisation pattern of one peroxidase, *AtPrx33*, in a limited region of the seedling (root end), using the same treatments. For each substance, we chose a concentration where both root length and peroxidase activity were affected (Fig. 4a-f). For IAA/SIR, the concentration 10µM was not considered, as root morphology was too strongly affected. In parallel, auxin distribution was analysed with GH3::*GUS* transgenic lines and X-Gluc staining.

As a first experiment, we exposed 7 day-old 25°C-grown seedlings to 5 days of cold (4°C). The drop in AtPrx33::*GFP* expression previously observed (Fig. 2e) correlated with the disappearance of the X-Gluc signal (Fig. 5b, d). The shift of GFP fluorescence to the root extreme tip was not paralleled by any X-Gluc staining in this region. IAA treatment did not provoke any substantial change in GFP fluorescence (Fig. 5c, g) or X-Gluc staining (Fig. 5a, e). Sirtinol is usually expected to activate the same genes as auxin (Zhao et al., 2003): surprisingly though, a growth on equimolar concentrations of sirtinol gave a radically

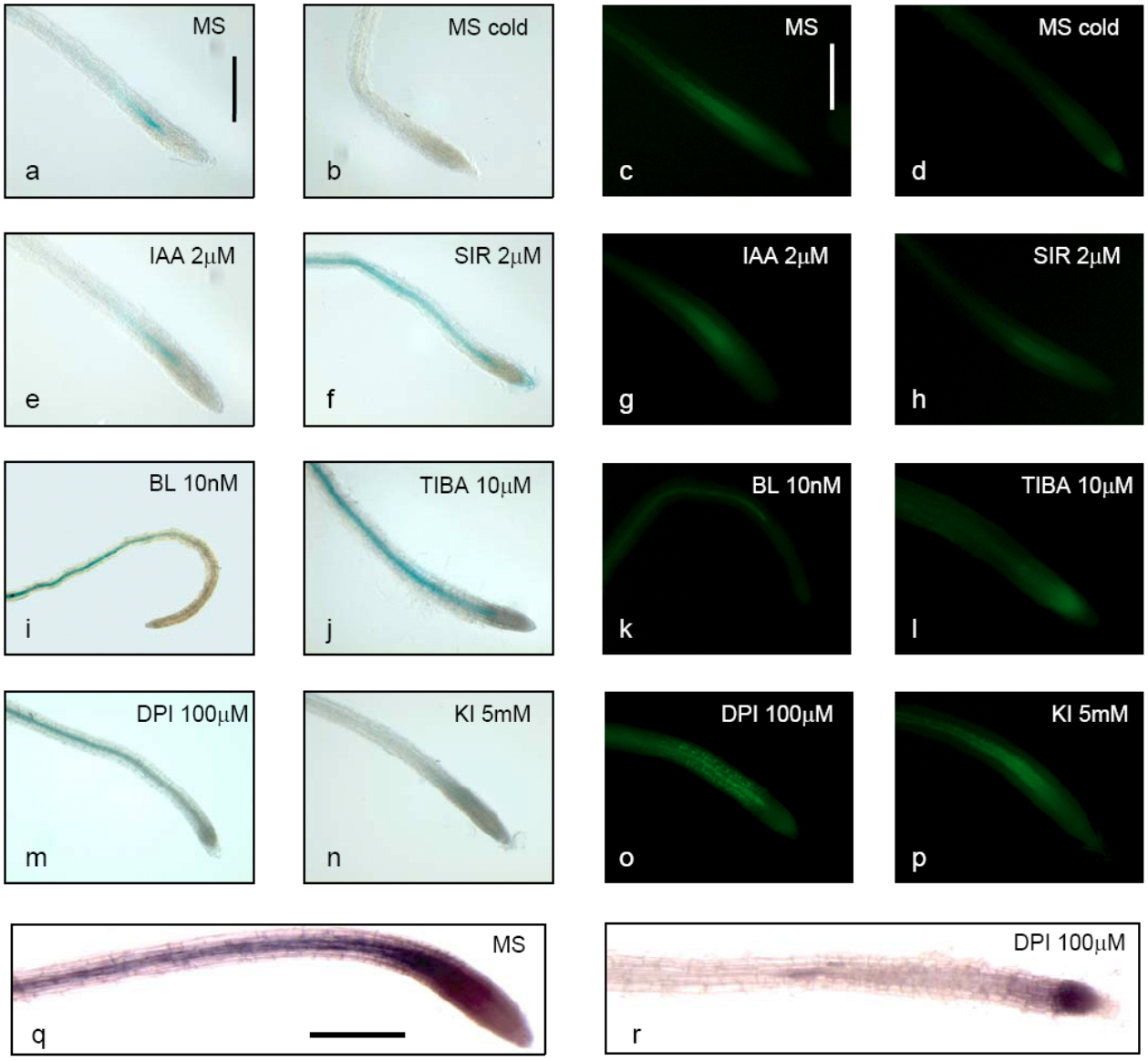


Figure 5

Detection of GFP fluorescence and GUS activity (X-Gluc staining) in roots of 7 day-old seedlings grown on the indicated concentrations of each treatment, or incubated during 5 days at 4°C (**b, d**). Columbia *Arabidopsis* lines transformed with the construct GH3::*GUS* were observed by incubating seedlings in X-Gluc (**a, b, e, f, i, j, m, n**). Col AtPrx33::*GFP* lines were directly observed for GFP fluorescence (**c, d, g, h, k, l, o, p**). NBT staining of seedlings grown on MS or DPI 100µM (**q, r**). Each picture is representative of at least 8 roots observed. Scale bars represent 250µm.

different GUS patterning than IAA. In sirtinol-grown roots, the elongation zone was not particularly marked, suggesting a perturbation of auxin transport. GFP expression was not modified: despite presence of an auxin-response element in the *AtPrx33* promoter (Passardi et al., 2006), it seems that, at least at a concentration of 2µM, this element is not targeted by auxin- and sirtinol-responsive transcription factors. RT-PCR experiments performed on the same auxin-treated seedlings confirmed that there was no change in expression of endogenous *AtPrx33* (data not shown).

The roots of brassinolide-grown seedlings exhibited a loss of *GUS* expression in a large part of the root end, in perfect correlation with the loss of *GFP* fluorescence (Fig. 5i, k). The auxin transport inhibitor TIBA induced a ubiquitous GUS staining in the whole stele, without any gradient in the elongation zone (Fig. 5j). GFP fluorescence decreased in the elongation zone and increased in the root tip (Fig. 5l), similarly to roots of cold-exposed seedlings (Fig. 5d).

Growth on the NADPH oxidase inhibitor DPI resulted in roots with a X-Gluc staining decreasing acropetally and almost no signal in the elongation zone (Fig. 5m). The pattern of *GFP* expression was rather different: staining in the elongation zone appeared to diffuse out of the central stele to the whole root section (Fig. 5o). When looking more carefully at the GFP-stained cells, we observed that they were all of the same length. The elongation zone was hence probably reduced to a small region very close to the root tip. Recent works (Dunand *et al.*, submitted) suggest that superoxide levels are particularly high in the elongation zone. We hence stained roots with the superoxide-specific dye nitro blue tetrazolium (NBT). NBT staining showed indeed a marked displacement towards the root tip in DPI-treated seedlings (Fig. 5q, r). The diffuse pattern observed for GFP after DPI treatment corresponds in fact to the staining usually observed in upper root parts (Fig. 1a), and the "reduced" elongation zone is not fluorescent. Similarly, X-Gluc patterning resembles the one observed in the basal part of the root, and elongation zone is not marked. In the last experiment, X-Gluc-stained roots of 5mM KI-grown seedlings showed a strong decrease in the elongation zone signal, but no change in *GFP* fluorescence (Fig. 5n, p).

Following application of various treatments affecting root elongation, we observed that changes in auxin localisation and *AtPrx33* expression were identical only in two cases (brassinolide and DPI) out of seven. Regarding DPI, the difference in patterning between GUS and GFP above elongation zone has been repeatedly seen in untreated roots (data not shown). IAA treatment can also be considered as a third case of GUS-GFP concordance, showing no difference compared to control in both X-Gluc and GFP observation.

Intermediate correlations were found for cold and TIBA treatments: auxin decreased in the elongation zone together with a decrease in *AtPrx33* expression. However, the root tip became fluorescent, whereas no GUS was detected in this region. For KI and SIR, changes in X-Gluc staining were not followed by any modification in GFP localisation or intensity.

Loss or over-expression of AtPrx33 and AtPrx34 modify auxin content

Quantitative analysis of auxin content was conducted following two different protocols. First, auxin was extracted from roots and shoots of 7 day-old seedlings. Two T-DNA insertion mutants for the peroxidase genes *AtPrx33* and *AtPrx34* were studied, together with their respective wild type ecotypes Wassilewskija (Ws) and Columbia (Col). Two other transgenic lines were also analysed: the double-mutant *atprx33/atprx34*, obtained by transforming the *atprx33* T-DNA mutant with a RNAi construct for *AtPrx34*, and the over-expressor line 35S::*AtPrx34* (Passardi et al., 2006). Two forms of auxin were assayed: free auxin (IAA) and aspartate-conjugated auxin (IAA-Asp). The latter species is usually found at high levels in parallel with IAA in expanding tissues (Kowalczyk and Sandberg, 2001); it is probably inactive and represents a major product of auxin degradation (Woodward and Bartel, 2005). In the samples tested, the level of IAA-Asp was 10 to 50 times higher than free IAA (data not shown). When compared to free auxin (Fig. 6a), the pattern of variation of IAA-Asp appeared quite similar for most of the cases (Fig. 6b). In roots for instance, no significant difference in auxin content was observed between Ws wild type and *atprx33*. In shoots, the level of auxin was largely decreased for all four peroxidase mutants. A marked discrepancy was observed with 35S::*AtPrx34* in roots, where free IAA was strongly increased, whereas IAA-Asp showed a slight decrease. Concerning *atprx34* roots, they contained more auxin (free and aspartate-bound) than the wild type. The higher root free auxin levels of the double-mutant *atprx33/atprx34* suggests a cooperation of AtPrx33 and AtPrx34 on free auxin content regulation, or the non-specific down-regulation of other peroxidases through RNAi. Peroxidase homology is quite high indeed, and the latter possibility can not be ruled out.

The second approach used was to study auxin levels *in planta* with GH3::*GUS* transgenic lines. In order to compare results with auxin dosage, the T-DNA mutants *atprx33* and *atprx34* (Columbia background) were crossed with GH3::*GUS* lines (Ws and Col ecotypes, respectively). The control lines were issued from the cross Col x Col GH3::*GUS* and Ws x Ws GH3::*GUS*, as described in the materials and methods section. For each GH3::*GUS* transgenic plant (Col wild type, *atprx34*, Ws wild type, *atprx33*), at least two independent homozygous lines were obtained and assayed for hydrolysis of the GUS substrate

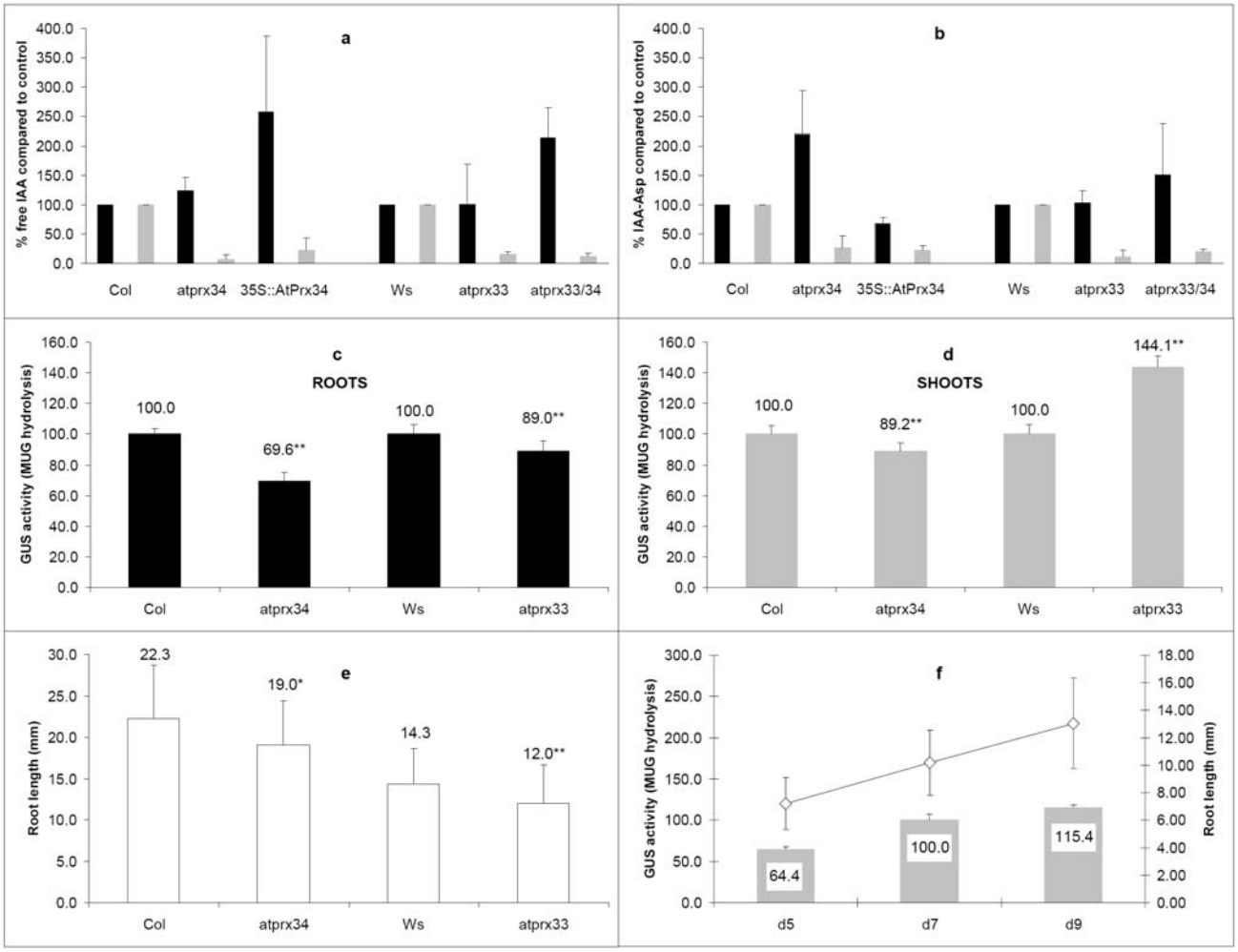


Figure 6

Quantification of auxin content by two different methods. **a, b** direct dosage of auxin in roots (black bars) and shoots (grey bars) of 7 day-old seedlings. IAA-Asp: aspartate-conjugated auxin. Graphs represent three independent experiments. **c, d** GUS activity measured by fluorescence on GH3::*GUS* lines incubated with the GUS substrate 4-methylumbelliferyl β-D-glucuronide (MUG) and then separated into roots and shoots. Control line Col is issued from the cross between Col and Col GH3::*GUS* (idem for Ws). Two control lines of each ecotype were tested in each experiment, and always gave similar results (data not shown). All seedlings are of the same generation. Relative values (wild types set to 100) are shown. **e** Root length of the seedlings of **c** and **d** before GUS extraction. **f** GUS activity (histogram) and root length (line plot) of GH3::*GUS* seedlings grown in normal conditions between 5 and 9 days (d5, d7, d9). Relative values (d7 was set to 100) of GUS activity are written in each bar. All p values (root length and MUG) between d5 and d7, as well as d7 and d9, are equal to or lower than 0.0005. Standard deviation is indicated by t-bars in all graphs. p values (1-way ANOVA test, mutant versus control): ** $p < 0.0001$; * $p = 0.0025$.

4-methylumbelliferyl β-D-glucuronide (MUG) by fluorimetry. Roots and shoots were analysed separately. The results for roots showed at first that GUS activity was lower in the peroxidase mutants than in controls (Fig. 6c). Concomitantly, *atprx34* and *atprx33* roots were shorter than their control counterparts (Fig. 6e). In order to check whether the decrease in MUG activity was due to loss of the peroxidase or decreased root length (or both), we extracted glucuronidase from roots of 5 day- to 9 day-old seedlings (Fig. 6f): fluorescence increased regularly in parallel with the age of the plants and the root length. Finally, we extracted glucuronidase from root ends, including elongation zone: the dependence on root length was lost, but there was still no difference between wild types and peroxidase mutants (data not shown). We hence conclude that there is essentially no change in GH3::*GUS* expression in roots between controls and the *atprx33* and *atprx34* knock-out mutants, or that the lower activity of the mutants is at least in part due to root length.

Concerning shoots, the results obtained with *atprx34* are in agreement with the decrease observed for the mutant in auxin dosage, though the drop in GUS activity is not as important (Fig. 6d). In contrast, *atprx33* showed a marked increase in fluorescence compared to control. No correlation could be found with root length.

The lines tested for MUG hydrolysis were then incubated with the GUS substrate X-Gluc and observed from roots to shoots, in order to detect qualitative changes in GH3::*GUS* expression. However, no significant difference was reported (data not shown). Our results suggest two distinct conclusions, depending on the method used to measure auxin content. Direct dosage of auxin indicates a clear drop in auxin content for peroxidase mutants shoots, a rise in auxin content for the *atprx33/34* double-mutant roots and probably also for the over-expressor 35S::*AtPrx34* roots. On the other hand, MUG assays favour the conclusion that no or little change is occurring in roots, and that loss of *AtPrx33* has an opposite effect than loss of *AtPrx34* in shoots.

Discussion

Perturbing elongation modifies expression of several peroxidases as well as global peroxidase activity

Exposure of seedlings to cold provoked a rise in peroxidase cDNA levels after two days, followed by a drop when cold incubation was extended (Fig. 3). When considered separately, each peroxidase studied behaved differently: despite the high degree of homology generally found among all peroxidases in their coding sequence, peroxidases have unique and very distinct promoter sequences. The presence of cold-induced motifs suggested that they may indeed play a role in the strong response of *AtPrx33* to a short cold stress. However, regulation appears to be more complex, as exemplified by the relative "insensitivity" of *AtPrx12* despite the presence of five cold-responsive elements. Differential regulation leads to variation in the levels of each peroxidase *in vivo* and hence to the proportion of each peroxidase. Within an organ, a tissue or even a cell type, several peroxidases are expressed at the same time (Valério et al., 2004; Jones et al., 2006). For instance, *AtPrx33* and *AtPrx34* are found together in root cell walls of 7 day-old seedlings (Shah et al., 2004; Passardi et al., 2006). Modifying the ratio of each peroxidase may then cause subtle changes in peroxidase-mediated mechanisms. Peroxidases can indeed have different substrate specificities or affinities. Therefore, if an internal (auxin homeostasis, ROS levels) or an external (cold, chemical treatment) parameter is altered, the plant will have at hand a set of peroxidases whose expression can be controlled: the peroxidases that are better adapted to the new conditions will be up-regulated.

Growth of seedlings on different chemicals also provoked a change in total peroxidase activity. In many cases, a direct correlation was observed between variation in root length and peroxidase activity (Fig. 4). For DPI however, the relationship was inversed. By inhibiting NADPH oxidase, DPI decreases superoxide levels. Superoxide, in turn, is a substrate for the hydroxylic cycle of peroxidases (Liszkay et al., 2003; Passardi et al., 2004). If $^{\bullet}O_2^-$ levels are decreased, the hydroxylic cycle may be slowed down, thus favouring the peroxidative cycle implicated in the guaiacol oxidation reaction that was used to measure total peroxidase activity. Alternatively, the modified activity of peroxidases may have provoked a lignification of the tissues instead of the promotion of elongation: peroxidases can indeed play opposite roles depending on the surrounding conditions (Passardi et al., 2004). Finally, presence of DPI may have triggered the induction of a set of peroxidases involved in lignification, and down-regulated the peroxidases usually present in the elongation zone. For KI 5mM, the drop

in peroxidase activity (compared to 10 and 100μM) may be due to the high salt concentration used more than the KI itself. We hence propose that unless the substance used directly interferes with peroxidase oxidation mechanisms, root elongation and peroxidase activity are closely associated.

Localisation of AtPrx33 expression in roots indicates a possible role in root elongation, but does not generally follow auxin distribution

In a previous study, we showed that the *atprx33* mutant had a short root phenotype, which was due to shortened cells (Passardi et al., 2006). The finding of a maximal AtPrx33::*GFP* expression in the elongation zone (Fig. 1) supports the hypothesis for a role of AtPrx33 in cell elongation. Furthermore, when elongation is slowed down by exposure to cold, GFP fluorescence decreases in the elongation zone. Interestingly, after 5 days of cold exposure, a significant level of GFP appeared at the root tip (Fig. 2). This localisation was also found when seedlings were grown on the auxin transport inhibitor TIBA, which non-specifically inhibits the activity of PIN proteins. It seems then that when auxin transport is globally hindered, *AtPrx33* expression is displaced to the root tip. On the other hand, when using brassinolide, which only perturbs localisation of PIN2 (Li et al., 2005), GFP fluorescence was not found in the root tip. Furthermore, displacement of *AtPrx33* expression was not observed in brefeldin-treated seedlings (data no shown). Brefeldin prevents the polar relocation of PIN proteins at the surface of cells, thus globally affecting auxin transport in root (Geldner et al., 2001; Baluska et al., 2002; Boutte et al., 2006): the mechanism triggering displacement of *AtPrx33* expression to the root tip, as well as the function of AtPrx33 at this site are still unclear. Monitoring of auxin repartition in root ends with the GH3::*GUS* reporter construct and comparison with AtPrx33::*GFP* fluorescence during various chemical treatments suggest that peroxidase expression does not usually follow regulation of the GH3 promoter (Fig. 5). Is AtPrx33 promoter responsive to auxin? A search in the first 1000bp of the promoter showed that one copy of the auxin-response (AuxRE) motif TGTCTC was present. This AuxRE is also present in two copies in the GH3 promoter, and plays a major role in auxin responsiveness of GH3 (Liu et al., 1994). Studies by Guilfoyle and colleagues demonstrated that TGTCTC alone is not able to respond to auxin, and must either be directly preceded by the sequence CCTCG or by a short-distant sequence CACGCAAT (Ulmasov et al., 1995). None of these sequences could be found close to the *AtPrx33* AuxRE, thus questioning the capacity of the peroxidase to be sensitive to auxin changes.

Auxin and AtPrx33 probably interact in vivo

Auxin is a known substrate of peroxidases (Ricard and Nari, 1966; Yamazaki and Yamazaki, 1973). Co-localisation of GUS and GFP (Fig. 1) in the elongation zone of untreated roots indicates that AtPrx33 can oxidise auxin *in vivo*. However, IAA concentration does not seem to be significantly modified by AtPrx33 auxin oxidase activity, as the roots of *atprx33* GH3::*GUS* lines showed no difference in MUG hydrolysis compared to control lines (Fig. 6c), and loss of *AtPrx33* had no consequences on root auxin levels (Fig. 6a, b). Auxin may then be oxidised by AtPrx33 in negligible amounts, or the level of auxin destroyed may be immediately restored through polar transport or local production (Ljung et al., 2005). The auxin radical intermediates formed during the reaction lead to generation of radical oxygen species or other radical species. Interestingly, we found that superoxide anions are present in large amounts in the elongation zone (Fig. 5q). It has been already proven *in vitro* that in absence of H_2O_2, the triad peroxidase-auxin-O_2 dissociates into peroxidase, auxin radical and superoxide (Gazaryan and Lagrimini, 1996; Gazaryan et al., 1996; Savitsky et al., 1999). Preliminary results (unpublished data) indicate that the major part of H_2O_2 in root is located in the cell differentiation region, and not in the cell elongation region. We hence favour the following model: the combination of AtPrx33, auxin and O_2 *in vivo* results in formation of superoxide anions; high $^{\bullet}O_2^-$ levels in the elongation zone cause a softening of the cell wall, thus promoting cell elongation. When cells reach the differentiation zone, auxin concentration, *AtPrx33* expression and $^{\bullet}O_2^-$ levels decrease, thus letting cell wall-rigidifying mechanisms take over.

AtPrx33 may also participate in auxin homeostasis in shoots, as knocking out the peroxidase had a dramatic effect on both IAA and IAA-Asp contents (Fig. 6a, b). Loss of *AtPrx34*, or both *AtPrx33* and *AtPrx34* and, more intriguingly, over-expression of *AtPrx34* all induced a drop in shoot auxin levels. Both peroxidases are the closest paralogs in *Arabidopsis* with 92% identity, hence the similar effect on auxin content is not surprising. The result with over-expressor line 35S::*AtPrx34* suggests that peroxidases are important players in auxin homeostasis: if peroxidase expression is strongly up- or down-regulated, the consequence is a sharp decrease in auxin concentration in the shoot. No phenotype could however be detected in shoots for any of the mutants. The rise in free IAA observed in roots of *atprx34*, 35S::*AtPrx34* and *atprx33/34* may be a reaction of the root to the shortage of apex-supplied auxin, since several sites in the root are able to produce IAA *de novo* (Ljung et al., 2005). The GH3::*GUS* reporter was not able to confirm these results: whereas it has proved useful for monitoring changes in auxin localisation, it appears not to be an ideal marker for

quantification of auxin variation. In this respect, a promoter such as DR5, constituted of seven AuxRE repeats and a 35S promoter (Ulmasov et al., 1997), proved in many occasions to have a direct correlation with auxin variation *in situ*, down to the cellular level (Avsian-Kretchmer et al., 2002; Benkova et al., 2003). DR5 could thus be used in future experiments in parallel with GH3, in order to better understand the *in vivo* activation of AuxRE-containing promoters, and to better define *in situ* the effect of loss or over-expression of peroxidases. Both promoters have been already compared in one study, and showed to give quite distinct expression patterns throughout plant development and also in response to exogenous auxin (Bierfreund et al., 2003). The most precise way of sensing auxin changes is probably direct auxin dosage through chromatography. Immunochemical staining may also prove useful, although it cannot detect free auxin from the various forms of conjugated auxin (Leverone et al., 1991; Veselov et al., 1992). The GH3 promoter has been shown to respond to exogenously applied auxin in many different plants (Liu et al., 1994; Ulmasov et al., 1997; Mathesius et al., 1998; Schwalm et al., 2003). Therefore, in our study, variation of GH3::*GUS* expression indicates that auxin localisation has been modified, but it does not always reveal to what extent, and in one case, MUG assays have led to opposite conclusions than auxin dosage: the promoter sequence used is 750 bp long, and many other regulator elements may come into play in order to modulate *GUS* expression.

Auxin and AtPrx33 dance together in *Arabidopsis*, but auxin also has many other partners. And the nature of the dance differs from one ballroom to another. In shoots, they cannot leave each other, and in root cells, they are bound for a long, long story.

Acknowledgements

The authors thank Dr. Ulrike Mathesius (Australian National University, Canberra, Australia) for the generous gift of plasmid pJJ402 GH3::*GUS-Term*. The financial support of the Swiss National Science Foundation (grant 31-068003.02) to C.P. and C.D. is gratefully acknowledged.

B.6.1 Supplementary data: brefeldin A treatment

When observing root ends of 7day-old AtPrx33::*GFP* seedlings treated with either cold or triiodobenzoic acid (TIBA), I noted a marked shift in fluorescence from elongation zone to root tip (article Fig.2, 5). The common link between these treatments was that auxin transport was globally down-regulated (cold) or inhibited (TIBA). No other treatment provoked this shift, and no other treatment was acting so extensively on auxin transport. In this respect, it has been shown, in this article project (article Fig.5q) and by Claude in other studies, that the predominant reactive oxygen species (ROS) in elongation zone is superoxide. Expression of *AtPrx33* is also the strongest in elongation zone. As previously mentioned in the article project, in absence of H_2O_2, the triad peroxidase-auxin-O_2 dissociates into peroxidase, auxin radical and superoxide (Gazaryan and Lagrimini, 1996; Gazaryan et al., 1996; Savitsky et al., 1999). It was hence very tempting to propose the following model: generation of superoxide by AtPrx33 could contribute to maintaining the elongating cell wall soft enough to continue the elongation process. When cells arrive in the upper part of the root, there is no more superoxide. Claude showed in recent studies that this upper region contains a high concentration of H_2O_2: instead of softening cell wall, H_2O_2 would be used by other peroxidases in the "classical" peroxidative cycle to create lignin radicals and hence rigidify the cell wall. When auxin transport was completely inhibited or disturbed, root elongation was slowed down because at least one peroxidase (AtPrx33 in our case) was removed from the elongation zone and its expression was displaced to the root tip. In root tip, there is no superoxide, hence AtPrx33 cannot contribute any more to elongation.

In order to verify this model, it was necessary to use another global inhibitor of auxin transport: brefeldin A (BFA). This substance inhibits the polar relocation of PIN proteins to the surface of cells. A concentration of 15µM was chosen, following previous experiments mentioned in the literature (Geldner et al., 2001). Two different protocols were tested:

- seeds plated on MS agar, then growth during 4 days, then transfer on MS agar + BFA 15µM and growth during 24, 48 and 72 hours.
- seeds directly plated on MS agar + BFA 15µM and grown during 7 days.

Both GH3::*GUS* (Col and Ws ecotypes) and AtPrx33::*GFP* plants were used (Figure 2-8a). The results obtained for AtPrx33::*GFP* roots show that expression of *AtPrx33* progressively disappears from the elongation zone with longer BFA exposure times. BFA induced a significant shortening of root length (data not shown), confirming a role of AtPrx33 in root elongation. Displacement of AtPrx33 to the root tip was not observed.

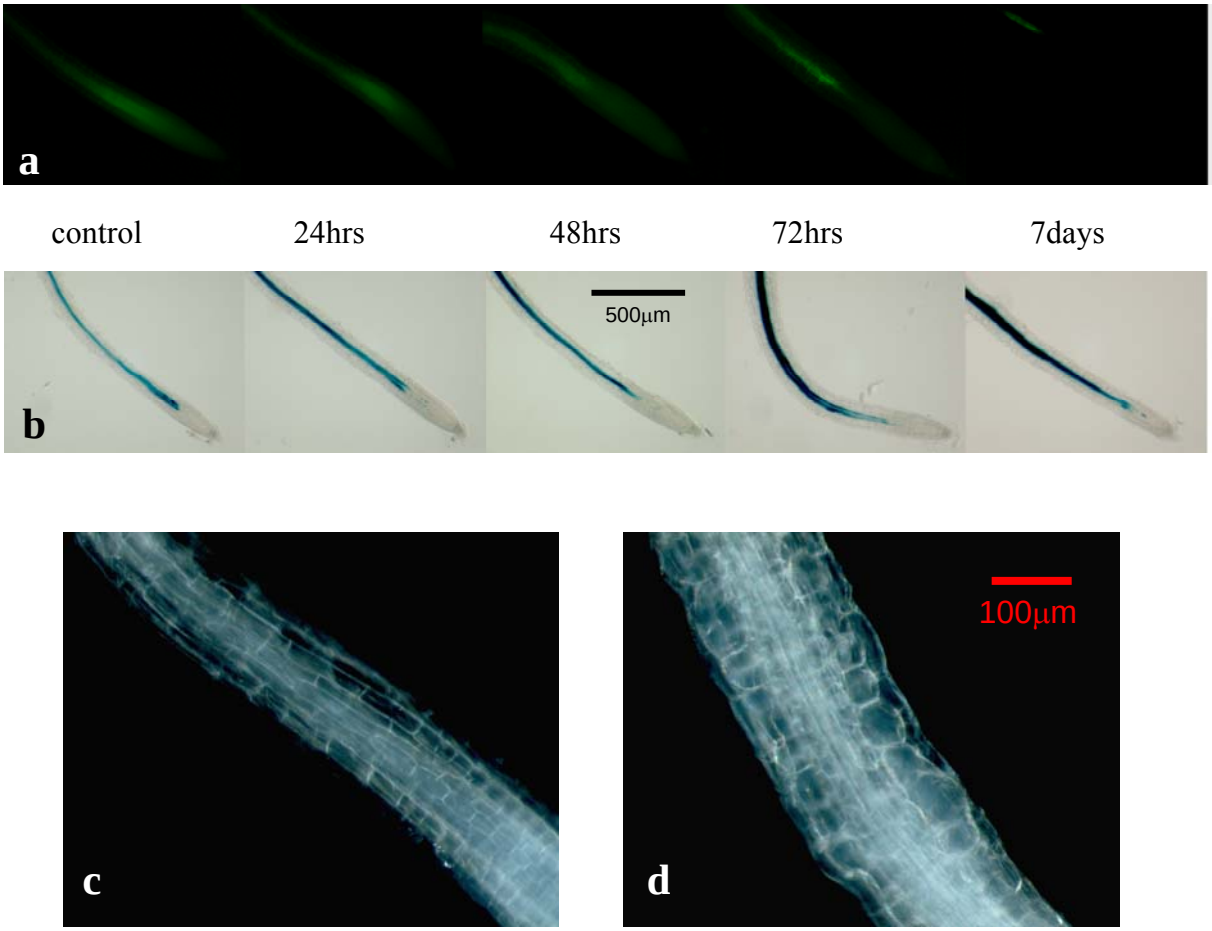


Figure 2-7
Root ends of brefeldin A-grown seedlings (15µM BFA)

(a) AtPrx33::*GFP* roots

(b) GH3::*GUS* (Columbia ecotype) roots

"control": root of a seedling grown on MS agar during 6 days

Protocol: - seedlings were grown on MS agar during 4 days, then on MS+BFA during 24, 48 and 72 hrs

or

- seedlings were directly grown on MS+BFA and observed after 7 days

(c-d) Root ends of seedlings 48 hours after transfer on MS (c) or MS+BFA 15µM (d)

The model proposed for a displacement of AtPrx33 expression to the root tip in case of a total loss of auxin directional transport is hence not valid. Regarding *GUS* expression, its intensity constantly increased along with BFA-exposure period, and the elongation zone was not particularly marked after BFA treatments (Fig. 2-8b). This implies a probable disorganisation of auxin transport, as it had been expected. The effect on *GUS* expression is quite similar to TIBA (article Fig.5j), supporting the use of the GH3 promoter as marker of auxin changes *in vivo*. When using higher magnification, I noted that the shape of cells in the elongation zone was severely affected, even in 24 hour-treated roots: cells were bigger and rounded, as if they had lost the capacity to orient themselves along the root axis (Fig. 2-8c, d). This observation is in agreement with loss of polar PIN distribution.

C. Involvement of AtPrx33 and AtPrx34 in organogenesis

As mentioned in the introduction, Michael observed spontaneous formation of a callus from adult leaves of an *atprx33* mutant, and subsequent generation of roots from the same callus on hormone free medium. He suggested in his thesis that hormonal balance might have been deregulated in *atprx33*, in particular for auxin metabolism. Most peroxidases are indeed able to oxidise auxin, at least *in vitro* (Ricard and Nari, 1966; Yamazaki and Yamazaki, 1973).

C.1 Preliminary studies: conditions for callus formation and organogenesis

When Michael obtained the *atprx33* callus, he maintained it in culture by using agar medium supplemented with 0.45µM BAP (cytokinin) and 0.45µM 2,4D (auxin). Generating a callus on simple MS medium is however a very rare event. I chose hence to provoke callus formation and differentiation by following (with slight modifications) a protocol (Fig.2-9) established by Stephen Howell's laboratory (Cary et al., 2001): seeds are sown on MS agar and grown during 7 days; roots are then excised just below the hypocotyl and transferred on CIM; 4 days later, roots are transferred to shoot inducing medium (SIM). As mentioned in figure 2-9, the passage on CIM is critical for acquisition of competency of root cells to "de-differentiate" (to form a callus).

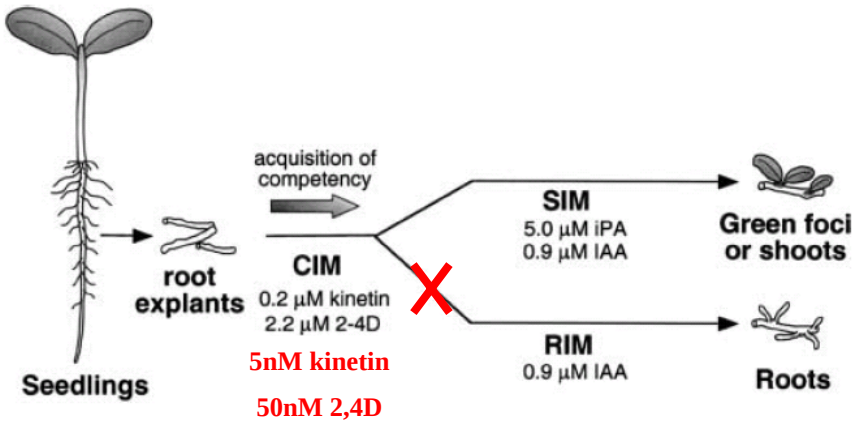


Figure 2-9 (from Cary et al., 2001)

Protocol used for callus formation and differentiation

Concentrations used during my experiments for CIM are shown in red. Studies with root inducing medium (RIM) were not performed.

Calli will appear later, on SIM, and will soon become green ("green foci") and eventually differentiate into aerial organs. After 30 days on SIM, whole plants are formed, with flowers and even siliques. The ratio of 2,4D to kinetin was however kept the same as in the original protocol. Despite the lower concentrations, I could soon obtain a high rate of callus formation (>90%). A major problem occurred when the laboratory in "Bastions" closed: I had to move in a short time to three different phytotron chambers. In the last one, I could not reproduce any longer callus differentiation. The major changes being mainly the use of a new type of neon tubes, I set light intensity to the same intensity as in a previous phytotron chamber, where differentiation was still taking place. However, the problem remained: light quality (different neon tubes) and probably other parameters (different batches of hormones, humidity) were having an influence on callus differentiation. By studying various series of calli on different concentrations of IAA (SIM medium) and light intensities, I finally succeeded in finding optimal conditions, as described in the article project "***Arabidopsis* peroxidase AtPrx33: a new actor on the scene of cell differentiation**": SIM 2x IAA and 30 $\mu E/m^2/sec$ for the Ws ecotype (SIM 10x IAA for Columbia). Moving to a last phytotron in September 2005 did not require additional changes, although the lamps were, again, different.

*C.2 Article project: "*Arabidopsis *peroxidase AtPrx33: a new actor on the scene of cell differentiation"*

In the article project "***Arabidopsis* peroxidase AtPrx33: a new actor on the scene of cell differentiation**", we have characterised a second possible function for the peroxidase AtPrx33. Wild type, *atprx33* (knock-out), *atprx33* x AtPrx33::*AtPrx33* (complemented) and 35S::*AtPrx33* (overexpressor) lines were compared for their aptitude to form shoots from callus structures. The study was extended to twelve T-DNA insertion lines in other *Arabidopsis* peroxidase genes, where we found that other peroxidases participate in the process of callus differentiation. Finally, observation of calli from AtPrx33::*AtPrx33* and GH3::*GUS* lines suggests that AtPrx33 is more specifically involved in the initiation of callus differentiation, but only in root elongation.

Arabidopsis peroxidase AtPrx33: a new actor on the scene of cell differentiation

Filippo Passardi, Claude Penel and Christophe Dunand

Laboratoire de Physiologie Végétale, Université de Genève, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland.

Abstract

Organogenesis from *in vitro* cultured plant calli is a process known for more than a century. The setting up of the conditions for inducing callus differentiation through a controlled regulation of auxin to cytokinin ratio opened the door to large-scale plant propagation, but also yielded valuable information about the processes underlying the so-called "de-differentiation" phenomenon (formation of calli) and the decision of cell fate for generation of organs. While regulation mechanisms start to be understood at the molecular level, it remains unclear which are the actors that participate in downstream processes. In this article, we have focused on the importance of plant peroxidases in shoot organogenesis from *Arabidopsis thaliana* calli. The role of the peroxidase AtPrx33 and 34 have been studied in more detail, through knock-out, complemented and overexpressor transgenic lines. We show here for the first time that this (these) peroxidase (peroxidases) participates (participate) in shoot differentiation. So far, the majority of proteins known to be involved in callus differentiation are transcription factors or hormone-sensing and signalling proteins: by characterizing an effector protein, this study opens new insights on downstream processes, and identifies novel targets for regulator proteins. AtPrx33::*GFP* and GH3::*GUS* lines, as well as H_2O_2 and superoxide-specific staining allowed better defining the localisation of AtPrx33, auxin and reactive oxygen species in calli, and giving indications on their relationships during callus formation and organogenesis. The role of peroxidases appears to be complex, with dependency on hormone levels, light quantity and quality. The latter point suggests a novel regulation of callus-expressed peroxidases by photoreceptor proteins. Antagonistic and overlapping peroxidase functions have also been observed with the double mutant *atprx33*/*34* and four other T-DNA insertion mutants that present a similar delay in shoot regeneration.

Introduction

One of the turning-points of evolution has been the capacity of living organisms to control the localisation of gene expression, which led to more and more complex patterns of embryogenesis and organogenesis. When the adult stage is reached, only plants keep the ability to change the fate of their cells. They can form again any, or most of, parts of the organism. This phenomenon can be quite easily studied by growing tissues on appropriate concentrations of hormones (cytokinins, auxin): tissues can be induced to "de-differentiate", hence forming calli. When choosing various ratios of cytokinins versus auxin, calli can regenerate either root or shoot structures (Skoog and Miller, 1957). The mechanism of differentiation is still incompletely understood, and with the advent of molecular biology, the problem is even more elaborate, as hundreds of genes seem to be involved (Che et al., 2006). Mutation studies have so far identified transcription regulators (Cary et al., 2001; Sieberer et al., 2003; Che et al., 2006), as well as genes involved in auxin and cytokinin metabolism (Boerjan et al., 1995; Frank et al., 2000). The genes that directly contribute to the formation of calli and to differentiation have been less studied. Among possible candidates, peroxidases have been shown to vary their expression in experiments involving calli (Kay and Basile, 1987; Tournaire et al., 1996; Kim et al., 2004), however their precise role in callus growth and differentiation has never been investigated.

Plant peroxidases are a class of proteins that can be found in any terrestrial green plant, and also in some algae (Passardi et al., 2004). During evolution, their number has considerably increased along with the appearance of new and more complex organisms. Whereas their number can reach 10 or 15 genes in early Embryophytes, they are often forming multigenic families of more than 100 paralogs in Angiosperms. The large number of peroxidase isoforms within higher plant genomes suggests subfunctionalization, and indeed, it is well-known that peroxidases can perform many diverse tasks, such as building cell walls, defending the plant against pathogenic attacks or participating to tissue senescence (Hiraga et al., 2001; Passardi et al., 2005). The presence of these large multigenic families leads to functional redundancies: it is still now very difficult to attribute one role to a single peroxidase *in vivo*. A few studies have been conducted already in order to answer this question, mainly by expressing a peroxidase from a plant into another plant (Dowd et al., 1998; Kawaoka et al., 2003). More recently, we have reported that twoa close paralogous peroxidases (92% identity), AtPrx33 and 34, participate in root elongation (Passardi et al.,

2006). In this study, we have suggested that a correlation may exist between peroxidase activity and control of the intracellular auxin level.

In order to determine whether these peroxidases might have a role in the control of auxin level and they could have more than one function *in vivo*, the dedifferentiation and redifferentiation capacity of the two T-DNA mutants were assayed. By inducing roots to form calli and then to differentiate into shoots, we show here that AtPrx33 and 34 can influence organ regeneration in an antagonistic manner. The colocalization of auxin and the AtPrx33, could point the putative implication of peroxidases in the auxin control. Moreover, participation to callus differentiation is probably not restricted to one or two peroxidases: three peroxidase knock-out lines and one overexpressor line also showed a significant delay in callus differentiation. Further analysis finally indicates that reactive oxygen species (and auxin), two peroxidase substrates, play a role in callus formation and differentiation.

Results

AtPrx33 absence or overexpression affects callus differentiation

Studies comparing differentiation of (Ws) wild type and *atprx33* calli (on 2x IAA) showed that *atprx33* had a delay in differentiation rate (Fig. 1A). The *atprx33* mutant line was then complemented with the *AtPrx33* gene under the control of the endogenous *AtPrx33* promoter. Two independent and already characterized (Passardi et al., 2006) homozygous lines called "comp 8.1" and "comp 10.1" were tested and gave an intermediate phenotype: they partially recovered the delay in differentiation of the *atprx33* line (Fig. 1A). The overexpression of AtPrx33 in a (Ws) wild type line did not lead to an increase in callus

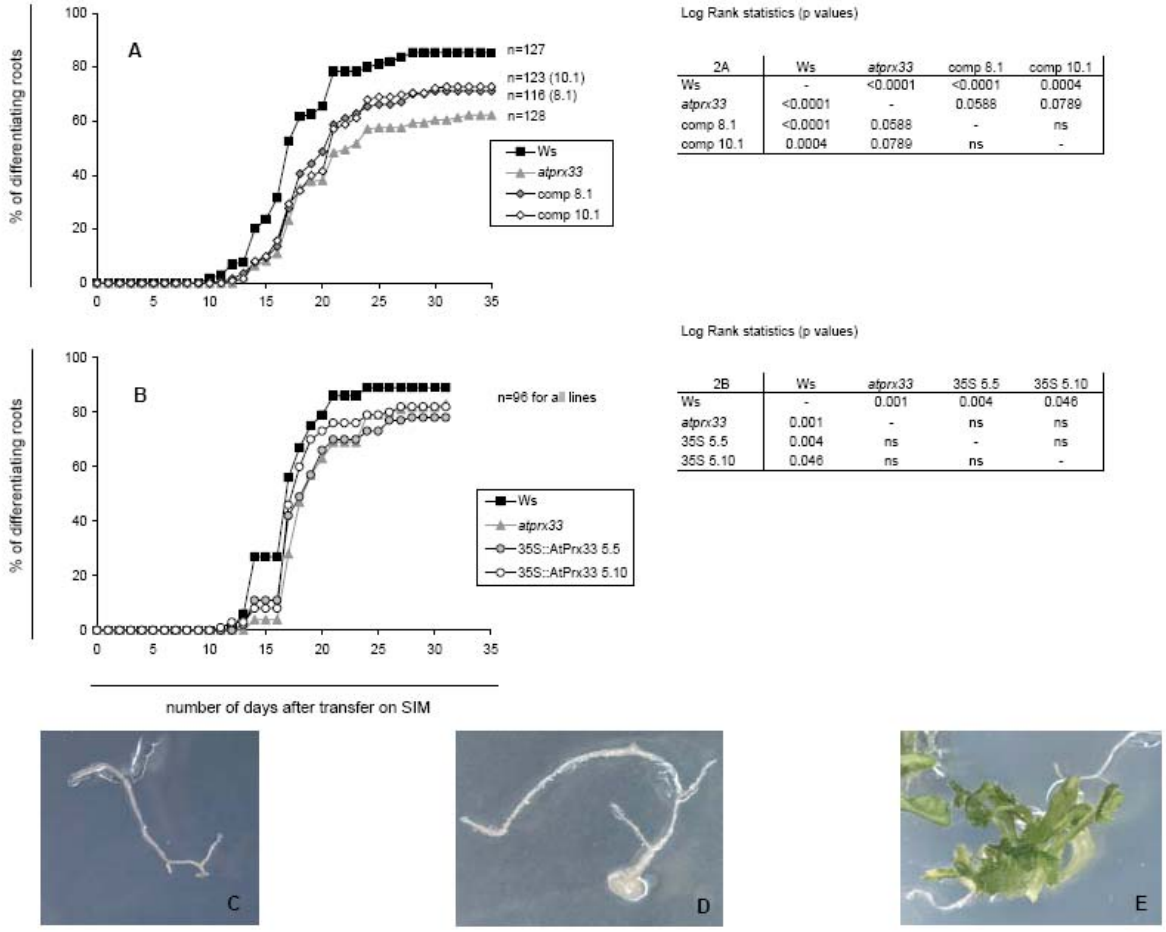


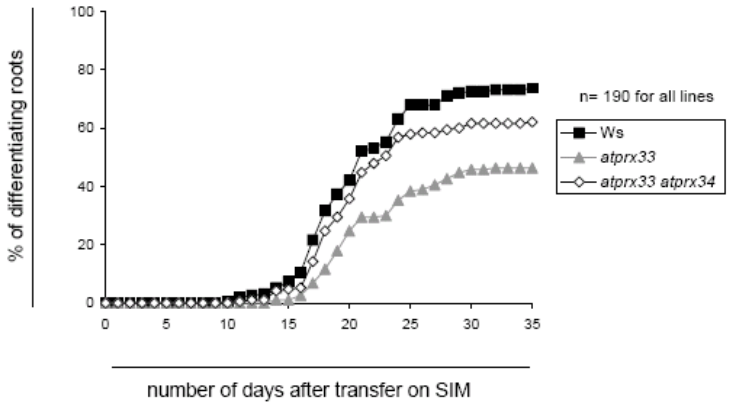
Figure 1

Callus differentiation in Ws wild type calli versus *atprx33* complemented and overexpressor lines, on SIM 2x IAA under white light (30µmol m^{-2} s^{-1}). (**A**) Comparison of callus differentiation between Ws wild type, Ws *atprx33* and two lines of Ws *atprx33* complemented with AtPrx33::*AtPrx33-GFP* (comp 8.1 and comp 10.1). (continued next page)

Figure 1 (continued)
The result is the sum of two independent experiments. (**B**) Comparison of callus differentiation between Ws wild type, Ws *atprx33* and two lines of Ws with the construct CaMV 35S::*AtPrx33-GFP*. The total number of roots observed (n) is indicated on each graph. Log-Rank statistics (Kaplan-Meier survival analysis) are shown in the table enclosed with this figure. (**C**) Undifferentiated root. (**D**) Callus-forming root after a 4-day incubation on CIM. (**E**) Fully differentiated callus after 25 days on SIM 2x IAA.

differentiation, but reduced it to a similar rate as the *atprx33* line, indicating that a higher level of AtPrx33 does not necessarily enhance this physiological response (Fig. 1B). In parallel, an *atprx34* line was also compared with (Col) a wild type for its capacity to form calli and redifferentiate. Columbia calli are not able to differentiate on SIM 1x IAA (Passardi et al., 2006). Various concentrations of IAA (from 0 to 10x IAA) were tested for callus formation and differentiation in order to obtain a satisfactory rate of differentiation. Differentiation was only observed on 5x or 10x IAA after 16 and 17 days, respectively (data not shown). In these conditions, no significant delay was observed for *atprx34* (data not shown). The double mutant, *atprx33* knock-out line transformed with a sense-antisense construct *43xrPtA-linker-AtPrx34* (RNAi strategy) was finally tested on 2x IAA regarding the callus differentiation. Down-regulation of *AtPrx34* expression surprisingly did not show an additive phenotype, but an intermediate differentiation rate between (Ws) wild type and *atprx33* (Fig. 2).

The results obtained here show that the delay in callus differentiation observed with line *atprx33* is indeed due to absence of the AtPrx33 peroxidase. Partial phenotype recovery of complemented lines is maybe due to the lack of regulatory elements upstream of the promoter fragment used (1000 bp). Alternatively, the insertion site of the transgene in the genome may be a less transcribed region than the endogenous region of *AtPrx33*. The phenotype observed in the overexpressor line suggests that the function of AtPrx33 in differentiation depends on a fine balance of its expression, which must not be too high or too low. The absence of differentiation delay in the *atprx34* is not surprising: in our previous study, the same mutant did not have significantly shorter roots. The most probable explanation is that *AtPrx33* expression is decreased by the T-DNA insert, but not totally knocked-out (Passardi, 2006).



Log Rank statistics (p values)

	Ws	*atprx33*	*atprx33-34*
Ws	-	<0.0001	0.0275
atprx33	<0.0001	-	0.0016
atprx33-34	0.0275	0.0016	-

Figure 2

Comparison of callus differentiation between Ws wild type, *atprx33* and *atprx33 atprx34* double mutant (T-DNA for *AtPrx33*, RNAi for *AtPrx34*), on SIM 2x IAA under white light (30µmol m^{-2} s^{-1}). The total number of calli observed is indicated as "n". Log-Rank statistics (Kaplan-Meier survival analysis) are shown in the table enclosed with this figure. The curves represent the sum of two independent experiments.

Other peroxidases participate in callus differentiation

Peroxidases are known to often have redundant functions in a same tissue (Passardi et al., 2005). In addition to AtPrx33 and 34, we hence searched for other peroxidase candidates that would be able to influence callus differentiation. We performed a search for more T-DNA insertion lines in other peroxidase genes. Homozygous lines were obtained by selection on appropriate antibiotics, and 7 day-old seedlings were checked for expression or absence of the peroxidase by RT-PCR reactions. Most of the Arabidopsis lines analysed have lost expression of the peroxidase, except two lines, which in contrast over-express the peroxidase gene (Fig. 3). The over-expression is probably due to the position of the T-DNA in the promoter region of AtPrx61. Concerning *AtPrx36*, it was not expressed in 7 day-old Col wild type seedlings, but it was clearly expressed in the *atprx36* insertion line. It seems that insertion of a T-DNA in the first exon caused a misregulation in the expression of the peroxidase, which is probably transcribed in a truncated form. All the lines described in Figure 3 were made in the Columbia ecotype of *Arabidopsis thaliana*, whereas the *AtPrx33*-related lines studied so far were made in a Wassilewskija background. For the reasons previously described, peroxidase mutant lines were screened on SIM 5xIAA.

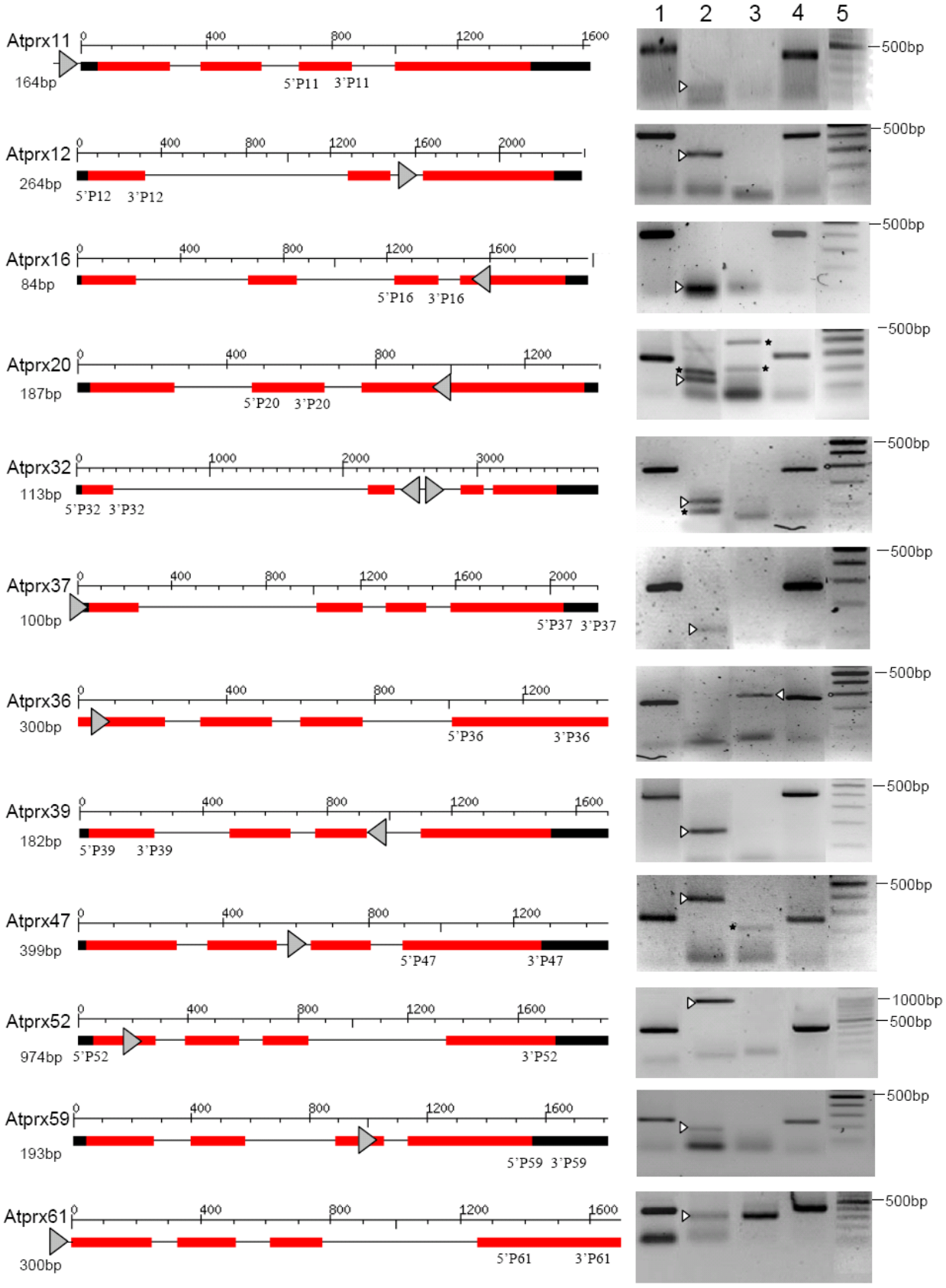
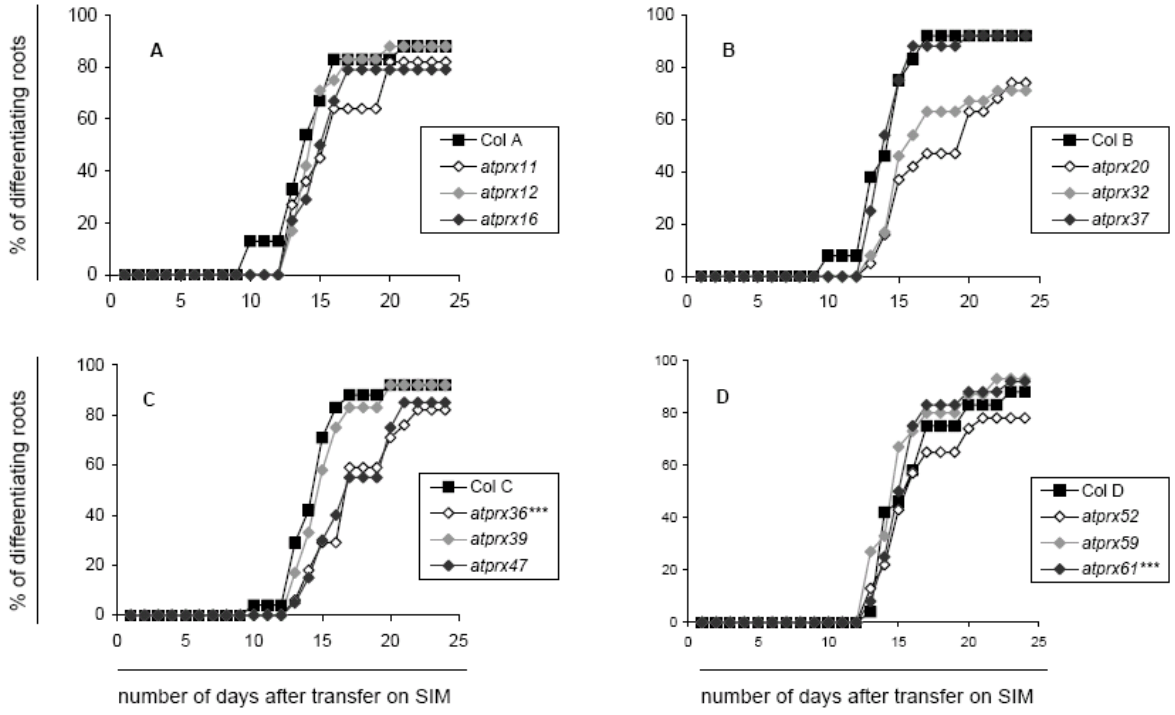

1 2 3 4 5
Atprx11
164bp
5'P11 3'P11
500bp
Atprx12
264bp
5'P12 3'P12
500bp
Atprx16
84bp
5'P16 3'P16
500bp
Atprx20
187bp
5'P20 3'P20
500bp
Atprx32
113bp
5'P32 3'P32
500bp
Atprx37
100bp
5'P37 3'P37
500bp
Atprx36
300bp
5'P36 3'P36
500bp
Atprx39
182bp
5'P39 3'P39
500bp
Atprx47
399bp
5'P47 3'P47
500bp
Atprx52
974bp
5'P52 3'P52
1000bp
500bp
Atprx59
193bp
5'P59 3'P59
500bp
Atprx61
300bp
5'P61 3'P61
500bp

Figure 3 (previous page)
Schematic description of the peroxidase T-DNA insertion lines screened (modified from http://www.tigr.org). The terminology used to name each peroxidase is related to the work of Tognolli et al. (2002), Valério et al. (2004) and to the peroxidase repository (http://peroxidase.isb-sib.ch/). Dark blocks correspond to untranslated regions, grey blocks to exons and dark lines to intronic regions of each peroxidase gene. Position of primers used is indicated on each gene, as well as position of the T-DNA (grey triangle), where the tip of the triangle stands for the T-DNA left border. The size of expected bands (RT-PCR) is indicated below the name of each peroxidase.
RT-PCR results (7 day-old seedlings) are shown on the right of each peroxidase gene, with column 1: P42 primers, control (wild type) cDNA; column 2: Peroxidase-specific primers (e.g. 5'P11-3'P11), control (wild type) cDNA; column 3: Peroxidase-specific primers, T-DNA insertion mutant cDNA; column 4: P42 primers, T-DNA insertion mutant cDNA; column 5: molecular weight (100bp-spaced bands). Two different P42 primers were used depending on the RT-PCR, which give either a 400bp or a 270bp band. ▷: band corresponding to the expected size; *: non-specific bands.

Among 12 T-DNA mutant lines screened, three knock-out mutants had a significant delay in differentiation rate: *atprx20*, *atprx32* and *atprx47* (Fig. 4). Interestingly, AtPrx32 is very closely related to both AtPrx33 and AtPrx34: in a phylogenetic tree, the three peroxidases form a well-defined cluster (Tognolli et al., 2002). Concerning AtPrx20 and AtPrx47 however, they are not particularly close paralogs to one another or to the AtPrx32-33-34 cluster. An overexpressor line (AtPrx36) also had a delay in differentiation, similarly to the CaMV 35S::*AtPrx33* transgenic Ws line. In none of the mutant screened we observed an increase in differentiation rate.

Expression of AtPrx33 in calli and distribution of auxin, superoxide and H_2O_2

In order to support a possible function for AtPrx33 in callus differentiation, we monitored its expression in calli at different developmental stages from 0 to 20 days on SIM (SIMd0 to 20). To this purpose, we transformed Arabidopsis plants with a construct bearing the *AtPrx33* promoter fused to a GFP gene (AtPrx33::*GFP*). *GFP* is not expressed at SIMd0, but five days later, a rise in expression is observed; with a strong staining on the whole callus structure (Fig. 5A, B). At SIMd5, GFP fluorescence is mainly detected on the site of root excision, which is also the region where most of calli initiate. *GFP* is then constantly



number of roots (n)		number of roots (n)		Log Rank statistics (p values)			
Col A	24	Col B	24	Col A vs.		Col B vs.	
atprx11	11	*atprx20*	19	*atprx11*	ns	*atprx20*	0.006
atprx12	24	*atprx32*	24	*atprx12*	ns	*atprx32*	0.009
atprx16	24	*atprx37*	24	*atprx16*	ns	*atprx37*	ns
Col C	24	Col D	24	Col C vs.		Col D vs.	
*atprx36****	17	*atprx52*	23	*atprx36****	0.016	*atprx52*	ns
atprx39	12	*atprx59*	15	*atprx39*	ns	*atprx59*	ns
atprx47	20	*atprx61****	24	*atprx47*	0.017	*atprx61****	ns

Figure 4

Comparison of callus differentiation between Columbia wild type and various peroxidase mutants (SIM 5x IAA). Each figure represents one independent experiment (with independent Col wild type lines). The total number of roots (n) observed is indicated at the bottom of the figures. p values of Log-Rank tests are indicated only when significant. ***: over-expressor mutant lines (in 7 day-old seedlings); ns: not significant.

expressed in the whole callus (Fig. 5C-E). The intensity of the fluorescence is not homogeneous in the callus, which is explained by differential tissue density (data not shown). Differentiating organs, which appeared distinctively at SIMd20, show no expression of GFP, in marked contrast with the callus part (Fig. 5E).

In parallel with the observation of *AtPrx33* expression profile, calli were also stained for two reactive oxygen species: O_2^- and H_2O_2, which are both involved in the two catalytic cycles of peroxidases (Kawano, 2003; Liszkay, 2003; Passardi, 2004). Two molecules were

used for specific staining: nitroblue tetrazolium (NBT, for superoxide) and hydroxyphenyl fluorescein (HPF, for H_2O_2). Superoxide was found in the same regions as GFP (Fig. 5F-J). Concerning H_2O_2 however, staining was more limited to the borders of the callus, the root hairs, the trichomes, as well as to the central stem of the root at the base of the callus (Fig. 5K-N). For both O_2^- and H_2O_2, staining was present since SIMd0 in the cutting site before callus formation, in two close yet distinct regions: H_2O_2, as in the callus, is limited to the border, whereas O_2^- is located more in the interior of the root. In contrast with GFP observations, both NBT and HPF staining were quite strong in differentiating tissues (Fig. 5J,O).

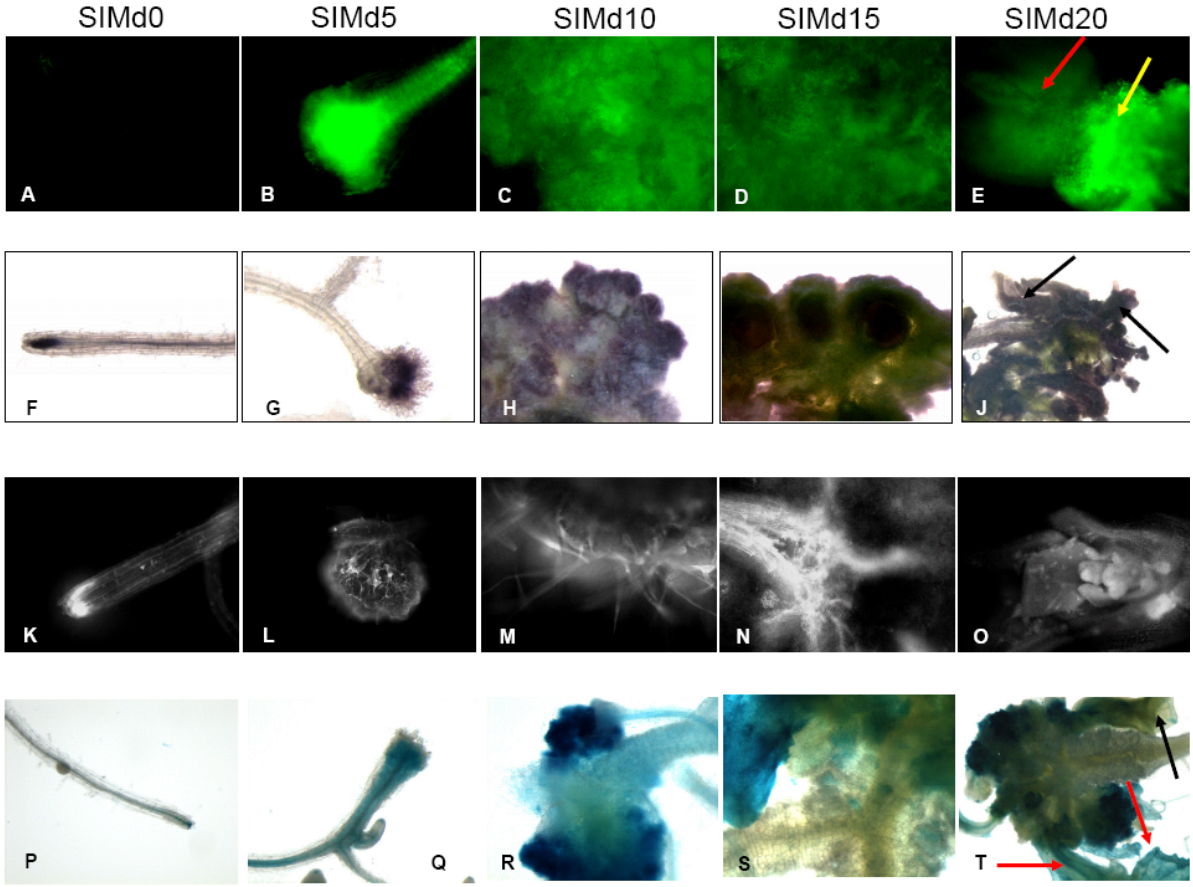


Figure 5

Pictures representing roots and callus structures from SIMd0 (no transfer on SIM) to SIMd20 (20 days after transfer on SIM medium). Pictures were all taken with a 100x magnification, except for J (16x) and T (32x). At SIMd15 and SIMd20, calli were cut into 300um-thick slices.

(**A–E**) GFP fluorescence of a Col AtPrx33::*GFP* line. (**G-J**) NBT (superoxide-specific) staining of a Col wild type line. (continued next page)

Figure 5 (continued)
(**K-O**) HPF (H_2O_2-specific) staining of a Col wild type line. (**P-T**) X-Gluc staining (auxin-specific) of a Col GH3::*GUS* line.
SIMd0: region where the root was cut from the seedlings before transfer on CIM; SIMd5 to SIMd15: calli formed from the excised roots. SIMd20, (**E**) differentiating tissue (red arrow) from a callus (yellow arrow); (**J**) Callus with differentiating regions (black arrows); (**O**) Differentiating tissues (issued from a callus); (**T**)Callus with differentiating regions stained (red arrow) or unstained (black arrow) by X-Gluc.

In a separate experiment, localisation of auxin was monitored with Arabidopsis lines expressing a glucuronidase reporter gene under the control of the auxin-responsive promoter GH3 (GH3::*GUS* lines). Staining was absent when compared to Col wild type at SIMd0, however it markedly increased at SIMd5 until the end of the observation period. Localisation was usually stronger on the borders of the callus, but staining was much deeper into the callus than for H_2O_2. The root stele was stained at SIMd5, but not in later stages: this suggests a possible role of auxin in the building of the first callus structures. Regarding differentiating regions, no general rule was observed, with strong staining to total absence of staining. Both auxin and AtPrx33::*GFP* co-localize quite well, which can argue in favour of the possible implication of peroxidase in auxin level regulation.

Callus differentiation depends on light intensity and wavelength

In order to determine the effect of light quality and quantity on callus differentiation, wavelength and intensity of light were modified. Wavelengths were filtered through blue- and red-coloured Plexiglas. Light intensity was varied from total darkness to 60 µmol m^{-2} s^{-1}. Concerning darkness conditions, seedlings and roots were first kept under an intensity of 12 µmol m^{-2} s^{-1} before transfer on SIM. For the experiments in blue and red light, the initial light intensity of 12 µmol m^{-2} s^{-1} was reduced to 5 µmol m^{-2} s^{-1} and 1.2 µmol m^{-2} s^{-1}, respectively, after crossing the blue and red filters. The results show that a low light intensity combined or not with a blue or red filter abolish the difference observed in white light between Ws wild type and *atprx33* (Fig. 6B, C, E). In striking contrast, for calli grown in total darkness (Fig. 6D), the delay in differentiation rate between Ws wild type and *atprx33* was still present and

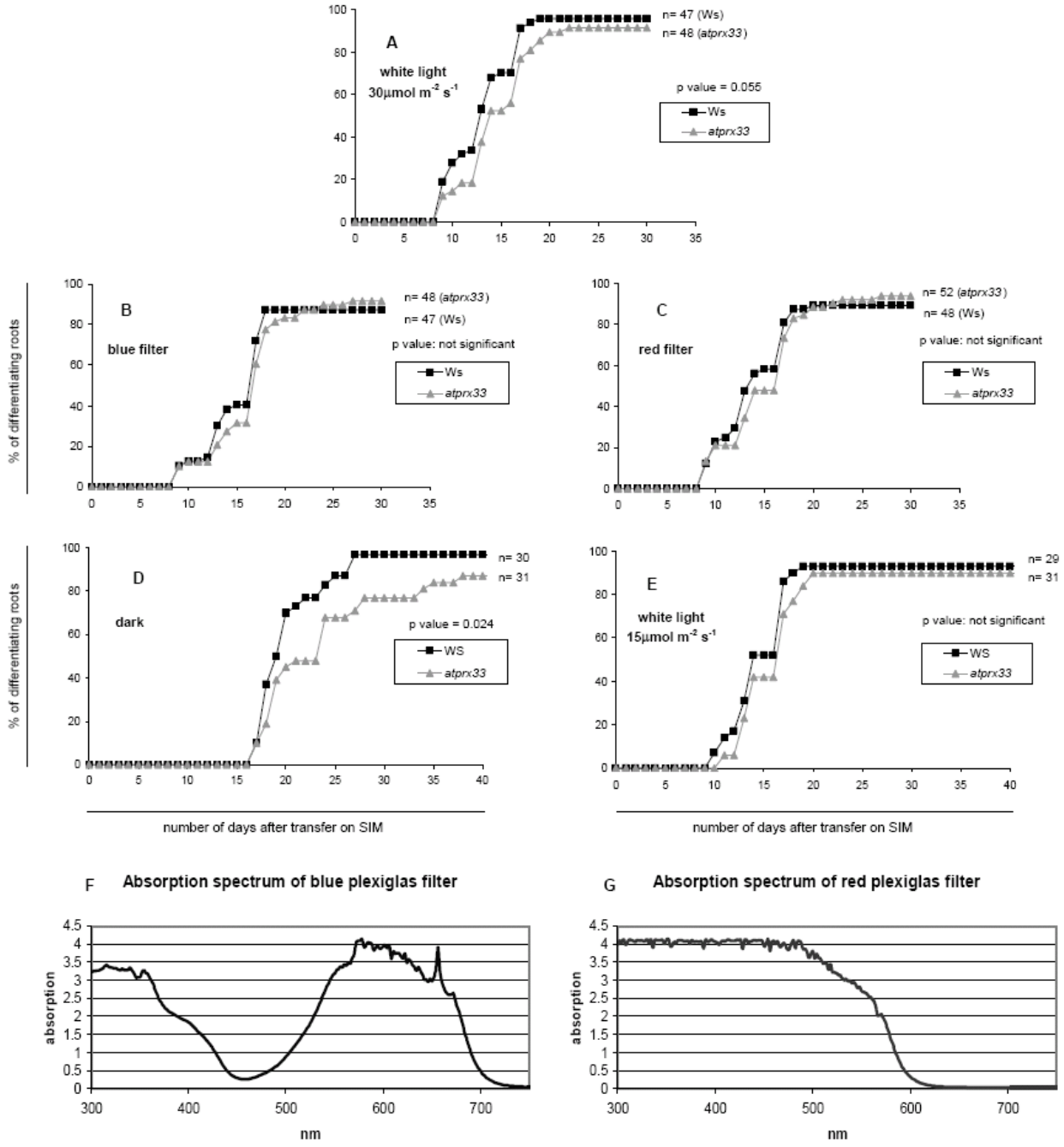


Figure 6

Ws wild type and *atprx33* callus differentiation on SIM 2x IAA, with various light conditions: (**A**) white light, 30 µmol m^{-2} s^{-1}, (**B**) blue filter, (**C**) red filter, (**D**) darkness and (**E**) white light, 15 µmol m^{-2} s^{-1}. (**F-G**) Absorption spectra of (**F**) blue and (**G**) red Plexiglas filters. Each differentiation curve represents a 32 to 40 day-observation. The number of roots (n) observed is indicated on each graph. p values are the result of Log Rank statistical analysis from a Kaplan-Meier survival plot.

similar to the one reported with white light. At 60 μmol m^{-2} s^{-1}, calli formed well from roots, but they never differentiated over the 40-day observation period (data not shown).

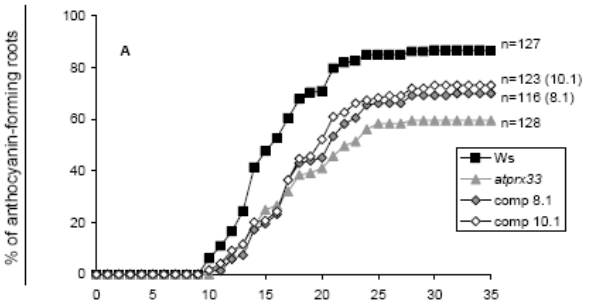
AtPrx33 expression influences two additional parameters on callus-forming roots

While observing callus differentiation, two other features were analyzed. A first one was formation of anthocyanins. A day before appearance of differentiating parts, coloured spots of anthocyanins appear on most greening calli. The percentage of anthocyanin-forming roots was hence also counted daily. Whenever there was a distinct differentiation rate between a wild-type control and a transgenic line (Fig. 1A, B), this difference was also observed for anthocyanin formation (Fig. 7A, B).

Secondly, during experiments performed with different qualities of light (Fig. 6A-D), we noticed that clumps of undifferentiated cells ("puffs") were forming quite frequently in various parts of the root, starting from SIMd10. The number of "puff-forming" roots has been quantified for each light condition (Fig. 7C). In Ws wild type roots, the amount of "puffs" observed increased in the following order: white, red, blue and dark. The most interesting result came from the comparison between Ws wild type and *atprx33* mutant. Except for white light, "puffs" were forming more frequently in the *atprx33* mutant, suggesting an inclination of this mutant towards formation of undifferentiated tissues, in straight correlation with the delay in callus differentiation rate.

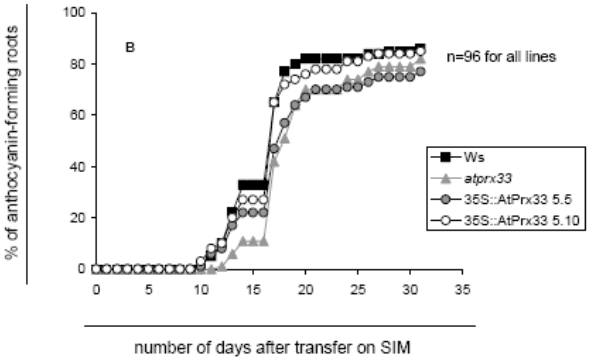
Figure 7 (next page)

Anthocyanin accumulation in Ws wild type calli versus *atprx33* calli measured in parallel (same experiments) with differentiation events. (**A**) Comparison of anthocyanin accumulation between Ws wild type, Ws *atprx33* and two lines of Ws *atprx33* complemented with AtPrx33::*AtPrx33-GFP*. The result is the sum of two independent experiments. (**B**) Comparison of anthocyanin accumulation between Ws wild type, Ws *atprx33* and two lines of Ws with the construct CaMV 35S ::*AtPrx33-GFP*. (**C**) Percentage of "puff-forming" roots in Ws and *atprx33*, under various conditions of light, at SIMd25. (**D**) Detail of a differentiating region in a callus: anthocyanins appear dark at the base of the differentiating tissue. (**E**) "puff" structure (undifferentiated cells). The total number of roots observed (n) is indicated on each graph. Log-Rank statistics (Kaplan-Meier survival analysis) are shown in the table enclosed with this figure.



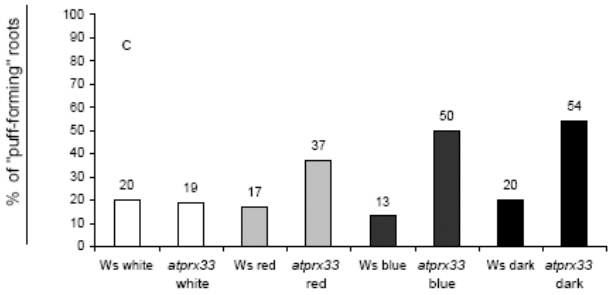
Log Rank statistics (p values)

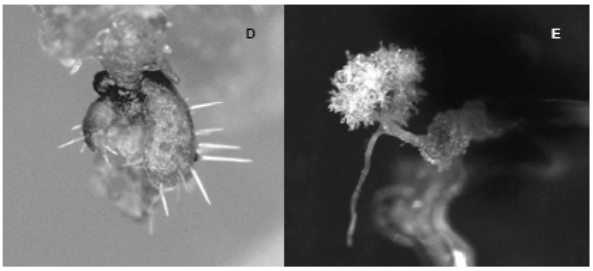
3A	Ws	*atprx33*	comp 8.1	comp 10.1
Ws	-	<0.0001	<0.0001	<0.0001
atprx33	<0.0001	-	ns	0.043
comp 8.1	<0.0001	ns	-	ns
comp 10.1	<0.0001	0.043	ns	-



Log Rank statistics (p values)

3B	Ws	*atprx33*	35S 5.5	35S 5.10
Ws	-	0.005	0.017	ns
atprx33	0.005	-	ns	0.034
35S 5.5	0.017	ns	-	ns
35S 5.10	ns	0.034	ns	-





Discussion

A significant delay in callus differentiation was repeatedly observed between Ws wild type and *atprx33* lines. This observation, together with complementation studies (Fig. 1A), gives a new role to AtPrx33 and 34 in the process of organogenesis. The intermediate phenotype observed for the double mutant suggests that AtPrx33 and 34 might act antagonistically (Fig. 2). Overexpression of *AtPrx33* gave rise to a delay of differentiation (Fig. 1B), indicating that a too high level of expression of *AtPrx33* is as detrimental for shoot formation as absence of the peroxidase. Although it might seem surprising, overexpressing and knocking-out a same gene do not always lead to opposite effects. For instance, phytochrome kinase substrate (PKS) 1 and 2 overexpressor and mutant (*pks 1* or *pks 2*) lines

had an identical phenotype: they both showed a marked inhibition of hypocotyl growth. (Lariguet, 2003). The explanation proposed by Lariguet and colleagues was that PKS1 and PKS2 were mutually inhibiting each other *in vivo*. Peroxidases are not known to interact *in vivo*. We hence propose that either the overexpressor or the *atprx33* lines perturb an optimal equilibrium of substrates related to differentiation, the most relevant being auxin. In one or the other way, destroying this balance results in reduced differentiation rate. One could oppose that any peroxidase could have the same effect, as most peroxidases share the same substrates (Penel, 1992; Hiraga, 2001). However, the screening of 12 different T-DNA insertion lines demonstrated that in eight cases, differentiation rate was not affected. Although it was already known that peroxidase expression was varying during root, callus, and differentiation growth (Kay and Basile, 1987; Tournaire et al., 1996; Kim et al., 2004), this is, to our knowledge, the first time a direct relationship is demonstrated for an intervention of a peroxidase in callus differentiation. Our results open a new field of research on effector proteins involved in organogenesis. Peroxidases represent a novel target for the many transcription regulators and hormone-sensing proteins that are induced during callus differentiation.

The general rule of plant peroxidases is that within a same tissue, several isoenzymes are expressed (Valério et al., 2004; McInnis et al., 2005; Veljovic-Jovanovic et al., 2006). A similar situation was observed in calli, where it has been shown that 3 peroxidases are among the 20 most up-regulated genes in root-forming calli (Che et al., 2006). In the study of Che and co-workers however, no peroxidase was found among the top 20 up-regulated genes in shoot-forming calli. We hence chose to screen several peroxidase T-DNA insertion lines, and the results show that other peroxidases than AtPrx33 are also affecting shoot organogenesis rate (Fig. 4). Similarly to the results obtained with *atprx33*, differentiation was delayed in the three knock-out mutant lines *atprx20*, *atprx32*, *atprx47* and also in the over-expressor line *atprx36*. In these four mutant lines, the delay was always of a similar order of magnitude. Regarding *atprx36*, the phenotype can be explained similarly to 35S::*AtPrx33*. These results suggest a redundant and equivalently important function of each peroxidase in callus differentiation.

Peroxidases catabolise a wide variety of substrates through reduction of H_2O_2. Among those substrates, auxin decrease often correlates with an increase in global peroxidase activity (Mitchell et al., 1986; Mathesius, 2001), and many peroxidases are able to oxidise auxin *in vitro* (Ricard and Nari, 1966; Yamazaki and Yamazaki, 1973). IAA catabolism can generate superoxide anions, which will in turn be converted to hydroxylic ions and H_2O_2 through the

peroxidase hydroxylic cycle (Liszkay et al., 2003; Passardi et al., 2004). We observed distribution of peroxidase, H_2O_2, superoxide and auxin in calli in order to check for possible co-localisations (Fig. 5). The results tend to demonstrate that peroxidase and ROS co-localise in definite regions, but not in the whole callus. Whereas AtPrx33 peroxidase expression is not associated with a particular zone of the callus, H_2O_2 distribution is limited to the periphery, where differentiated tissues form (trichomes, absorbing roots, shoot structures); in contrast, superoxide repartition is mainly limited to the inner part of the callus. Both H_2O_2 and superoxide may hence interact with AtPrx33, leading to two possible distinct consequences: differentiation of tissues for H_2O_2 and callus growth for superoxide. The underlying mechanism has still to be investigated, but we favour the model that presence of reactive oxygen species can trigger softening of cell walls and elongation or remodelling of cells and tissues. H_2O_2 and superoxide might be only intermediate substrates, which would subsequently generate hydroxyl radicals through the peroxidase hydroxylic cycle (Passardi et al., 2004). Hydroxyl radicals are indeed suspected to play a major role in cell wall cleavage, extension and remodelling (Schweikert et al., 2000; Miller and Fry, 2001; Liszkay et al., 2003). Auxin distribution is quite strong at SIMd5 from the root stele to the inner part of the forming callus, suggesting that auxin is transported to the callus in order to promote morphogenesis. Once the callus is formed, auxin disappears from the root central cylinder, and is mainly located to the periphery of the callus, where differentiation happens. It is then also probable that auxin stimulates differentiation in the callus, confirming its already described function in organogenesis (Friml, 2003). Distribution of auxin shows, like for ROS, that it may be a substrate in reactions involving AtPrx33. Finally, the patterning of GFP expression at SIMd20 (Fig. 5E) confirms that AtPrx33, besides participating to callus growth, is likely to trigger differentiation. However, its role stops once differentiation has started, and expression is lost in the nascent shoot structures.

Photoreceptors are a known family of proteins upstream in the cascade of events leading to organogenesis (Casal et al., 2003). According to our studies, it seems that the role *AtPrx33* in callus differentiation can be influenced by light-sensing proteins (Fig. 6). In red light, the major photoreceptors present in *Arabidopsis thaliana* are the phytochromes PhyA and PhyB, whereas in blue light, Cry1 and PhyA are predominant (Casal, 2000; Gyula et al., 2003; Sullivan and Deng, 2003). It seems then that PhyB and Cry1 alone (possibly in combination with PhyA) prevent the appearance of the delayed *atprx33* phenotype. When PhyB and Cry1 are both activated (white light), they may reciprocally inhibit each other, which also explains that the differentiation phenotype is still present in dark conditions (both

PhyB and Cry1 are inactive). Photoreceptor-mediated regulation of peroxidase has been already investigated in two experiments on mustard stems and cucumber hypocotyls (Casal et al., 1990; Shinkle et al., 1992), where it has been shown that specific peroxidase isoforms can change their expression and that total peroxidase activity can be modified according to light conditions. Peroxidases are known to be involved in a wide variety of functions, particularly in the building of cell wall and elongation (Hiraga et al., 2001; Passardi et al., 2005; Passardi et al., 2006). The relationship between photoreceptors and peroxidases may therefore be extended to many processes in plant growth and differentiation, opening an exciting field of research on a novel and probably essential regulation pathway in plants. Other factors were found to have an impact on callus differentiation rate and *atprx33* delay, such as light intensity (Fig. 6A, E and 7C) and hormonal balance (data not shown). This series of observations suggests a complex regulation of *AtPrx33* expression.

Whereas there were already some hints that peroxidases could be involved in callus formation and differentiation (Cordewener et al., 1991; Takeda et al., 2003), we have shown in this study that they play a significant role in organogenesis. They represent a novel actor on the scene of cell differentiation, at the end of transcription factor- and hormone-regulated processes. Moreover, following the results of a previous study on the role of AtPrx33 and 34 in root elongation (Passardi et al., 2006), we have now demonstrated that *in vivo*, peroxidases can have two very distinct functions. The mechanism by which AtPrx33 are 34 are able to contribute to organogenesis is still unclear, but there are chances that some interplay exists between peroxidases, auxin, H_2O_2 and superoxide. This topic has still to be thoroughly investigated, and will bring fascinating insights into downstream processes of morphogenesis, as well as to the development and improvement of plant propagation techniques.

Materials and methods

Plant material

The Columbia (Col) and Wassilewskija (Ws) ecotypes of *Arabidopsis thaliana* were used as wild type controls. Stable homozygous *Arabidopsis thaliana* T-DNA lines *atprx33* (Wassilewskija ecotype-Ws) and *atprx34* (Columbia ecotype) lines, as well as 35S::*AtPrx33* lines, were obtained and characterized as described previously (Passardi et al., 2006). The *atprx33* line was further transformed with either AtPrx33::*AtPrx33* for complementation or a sense-antisense construct *43xrPtA-linker-AtPrx34* (RNAi strategy) (Passardi et al., 2006).

Other peroxidase T-DNA insertion lines were obtained from the SALK T-DNA insertion collection (Alonso et al., 2003) and from Syngenta Inc. (San Diego, CA, USA). After selection on antibiotics, homozygous lines for T-DNA insertion were verified by PCR on genomic DNA with a primer specific for the peroxidase sequence ("amplicons primers", Valério et al., 2004) and another one for the left border primer 3'LB3 5'-TAGCATCTGAATTTCATAACCAATCTCGATACAC-3' (Syngenta lines) or 3'LBb1 5'-GCGTGGACCGCTTGCTGCAACT-3' (SALK lines). For *atprx12*, the specific primer 13.3R (5'-CAAGAGACGACTTGACC-3') was used instead of an amplicon due to T-DNA orientation. Lack of expression (or overexpression, in some cases) was further assessed by RT-PCR on total seedling RNA, with primers specific for each peroxidase.

Obtention of transgenic lines

In order to obtain the AtPrx33::*GFP-Term* construct, 1000 bp of the *AtPrx33* promoter were first amplified from genomic DNA with primers 5'pCA KpnI 5'-GGGGTACCGCTTGGTTTGGTTTCCATTG-3' and 3'pCA NcoI 5'-CATGCCATGGTTTTCACAAGGAC-3' (engineered restriction sites are underlined). The PCR product was directly ligated to pCR II plasmid (Invitrogen Ltd, Paisley, United Kingdom), and then subcloned (KpnI/NcoI) into pAVA 393 (von Arnim et al., 1998) in frame with a *GFP-Term* sequence. The AtPrx33::*GFP-Term* cassette was further subcloned (KpnI/BamHI) into the binary vector pZP222 (Hajdukiewicz et al., 1994).

Ws GH3::*GUS-Term* transgenic lines were obtained starting from the construct pJJ430 GH3::*GUS-Term* (Larkin et al., 1996). The GH3::*GUS-Term* cassette was excised with EcoRI and cloned into EcoRI-digested binary vector pZP222.

atprx33 x Ws GH3::*GUS-Term* transgenic lines were obtained by manually fertilizing *atprx33* homozygous carpels with pollen issued from Ws GH3::*GUS-Term* flowers, and then selecting double transgenic plants with kanamycin 50µg/ml (*atprx33*) and gentamicin 75µg/ml (GH3::*GUS*). Ws wild type carpels were also fertilized with Ws GH3::*GUS-Term* pollen in order to obtain the control line Ws wild type x Ws GH3::*GUS-Term*.

Plant transformation

pZP222 constructs were introduced into *Arabidopsis thaliana* Col ecotype (AtPrx33::*GFP-Term*) or Ws ecotype (GH3::*GUS*) by *Agrobacterium tumefaciens* strain ASE-mediated transformation using the spraying technique (Bent, 2000). Transformants were selected on gentamicin 75µg/ml. The presence of the transgene was confirmed by GFP fluorescence and X-Gluc staining. Homozygous lines were then obtained through antibiotic selection.

Callus generation : media and growth conditions

Generation of calli from Arabidopsis roots was made according to previous literature (Cary et al., 2001), but with several modifications, described hereafter. Arabidopsis seeds were sterilized (Bleach 2.4%, Triton 0.02%), sown in circular 9x9 mm Petri dishes on ½ MS medium and put at 4°C during 2 days for vernalisation. If not stated otherwise in the text, Petri dishes were then placed in a Percival I-30 BL4C8 culture chamber (Percival Scientific, Inc.; Perry, IA, U.S.A) under 16h light /8 hour dark photoperiod, at 22°C under four mercury lamps delivering a total of 30 µmol m^{-2} s^{-1} (measured with a radiometer/photometer International light IL1400A and a filter White # 9540). After one week, roots were cut and transferred onto callus inducing medium (CIM: 5 nM kinetin; 50 nM 2.4D, 2,4-dichlorophenoxyacetic acid). 4 days later, roots were transferred onto shoot inducing medium ("SIM2x": 1.8 µM IAA, indole-3-acetic acid; 5 µM iPA, isopentenyladenine). Callus differentiation was then checked daily for each callus, during at least 25 days (Fig. 1A-C).

In order to determine the influence of light quality (Fig. 1F,G), 3mm-thick Plexiglas filters (Röhm & Haas GmbH, Darmstadt, Germany) were used to cover the callus cultures. The absorption spectrum of each Plexiglas was verified by cutting a piece of the filter and scanning it in a diode array spectrophotometer (Hewlett Packard 8452A) between 300 and 750 nm.

Statistical analysis

Differences in the rate of differentiation between control and transgenic lines were assessed by a Kaplan-Meier survival analysis with SPSS v.13 software (LEAD Technologies, Inc.). At the end of the observation, undifferentiated calli were considered as "censored" for statistical analysis. The p values shown (Fig. 1-4) correspond to the result of the Log Rank statistics method.

Fluorescence experiments

AtPrx33::*GFP* homozygous calli were observed for GFP fluorescence after a 1-week growth on ½ MS medium under a light intensity of 80µmol m^{-2} s^{-1}. GFP detection was performed with an Axioplan 2 Zeiss microscope with a monochromator-mediated narrow band excitation (475±2nm), a narrow band EGFP emission (510±10nm) filter (Filter set 41020, Chroma Technology Corp, VT, USA), and using the Metamorph software (Molecular Devices Corp, CA, USA).

For HPF (Hydroxyphenyl Fluorescein-Alexis Biochemicals, Switzerland) staining (Setsukinai et al., 2003), calli were incubated during 10 minutes in an aqueous 10µM solution, and fluorescence was monitored with monochromator-mediated narrow band excitation (485±2nm) and a FITC-suited emission (>515nm) filter (Filter set 09, Carl Zeiss, Inc., USA).

For NBT (Nitrotetrazolium blue chloride, Sigma) staining, calli were incubated during 15 minutes in a 2mM solution (20mM pH 6.1 PO_4 buffer), and then visualized under a Leica MZ16 microscope (Leica Mycrosystems GmbH, Germany).

Colorimetry experiments

X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt) was purchased from X-GLUC DIRECT (United Kingdom). Staining was performed as described elsewhere (Blazquez et al., 1997; Lariguet et al., 2003). Briefly, calli were soaked in cold (4°C) acetone 90%, then rinsed and stained (5' at 4°C and 20' at 37°C) in a solution containing 2mM X-Gluc, 2mM ferrocyanide, 2mM ferricyanide and 10mM sodium phosphate buffer. After fixation, calli were observed directly under a Leica MZ16 microscope (Leica Mycrosystems GmbH, Germany) or cut into 300µm-thick slices as described hereafter.

At SIMd15 and SIMd20, some calli were stained for X-Gluc or NBT, embedded into 5% agarose (Sigma, A5093) and finally cut into 300µm thick slices with a Leica VT1000S

microtome (Leica Mycrosystems GmbH, Germany). For HPF, calli were first embedded and cut, and then immersed in HPF as described above.

Acknowledgements

The authors thank Dr. Ulrike Mathesius (Australian National University, Canberra, Australia) for the generous gift of plasmid pJJ402 GH3::*GUS*-*Term*.

C.2.1 Supplementary data: *atprx34* T-DNA mutant

During this study, the T-DNA mutant *atprx34* was compared to a Columbia wild type ecotype, in three separate experiments totalising more than 330 calli for each line. No difference in callus differentiation was noted. This data was not shown nor mentioned in the article project. In the *atprx34* line, it has been demonstrated that the level of *AtPrx34* mRNA was decreased, but was not completely abolished, probably because the T-DNA was inserted in the promoter, and not in the gene itself (Passardi et al., 2006). Despite being lower, the amount of AtPrx34 is apparently sufficient to prevent appearance of a phenotype. Alternatively, AtPrx34 may not be involved in differentiation.

C.2.2 Supplementary data: reverse transcription PCR of calli

Figure 5 of the article project shows that *AtPrx33* is strongly expressed since early formation of callus (at SIMd5, calli are still white, i.e. they do not form green foci). Previously to the fluorescence experiment, I had performed RT-PCR reactions at SIMd10, 15, 20, 25 and 30 in order to monitor endogenous *AtPrx33* expression in wild type (Ws ecotype) calli. During all these time points, I only collected roots with green foci (SIMd10) or with green foci, anthocyanins and differentiating calli (from SIMd15). This study was achieved in two independent experiments (separate Petri dishes, separate RNA extractions, series “A” and “B”), and gave the same result: endogenous *AtPrx33* expression can only be detected at SIMd10 (Fig.2-10). GFP fluorescence experiments showed, concordantly, that AtPrx33 is expressed at SIMd10. However, in contrast to what RT-PCR results suggest, AtPrx33 is still expressed at SIMd15 and SIMd20. This discrepancy could be due to stability of the GFP protein in callus cells. Another possibility is that GFP signal is indeed decreasing, but the region observed contains such a thickness of cells that accumulation of cell layers increases the fluorescence observed. Although not detailed in the article, I compared several pictures of calli taken at SIMd15 and SIMd20 on 300µm-thick slices (and not on whole calli). The general impression is a decrease in GFP signal from SIMd15 to SIMd20, which would then be in agreement with RT-PCR results (Fig. 2-11).

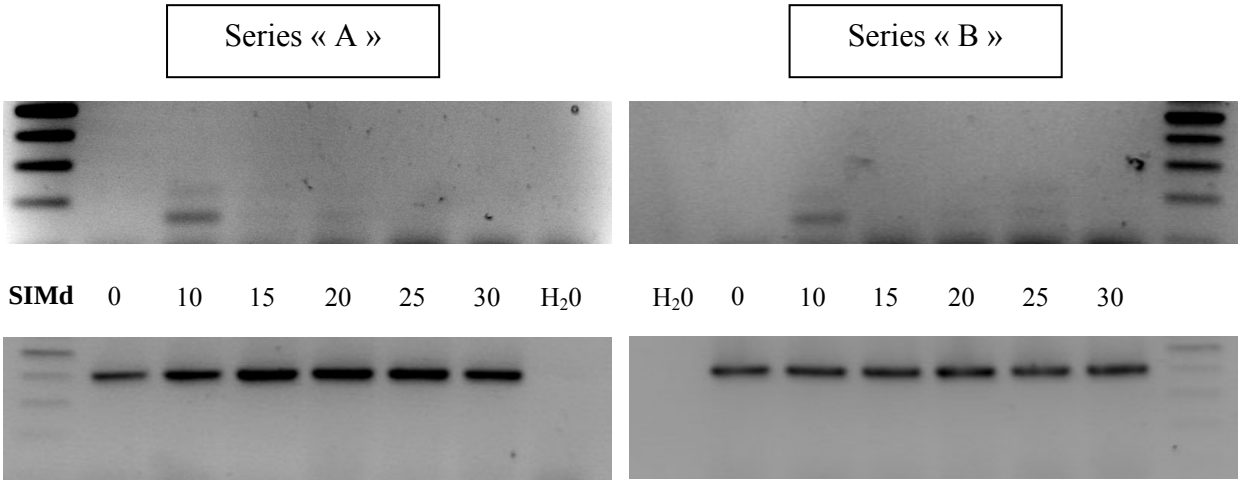


Figure 2-10

RT-PCR reactions on calli at different time points (SIMd = days after transfer on SIM).

Upper row: primers against *AtPrx33*

Lower row: primers against *AtPrx42* (control)

Extreme right and left lanes: molecular weight (from top to bottom: 500, 400, 300, 200 bp)

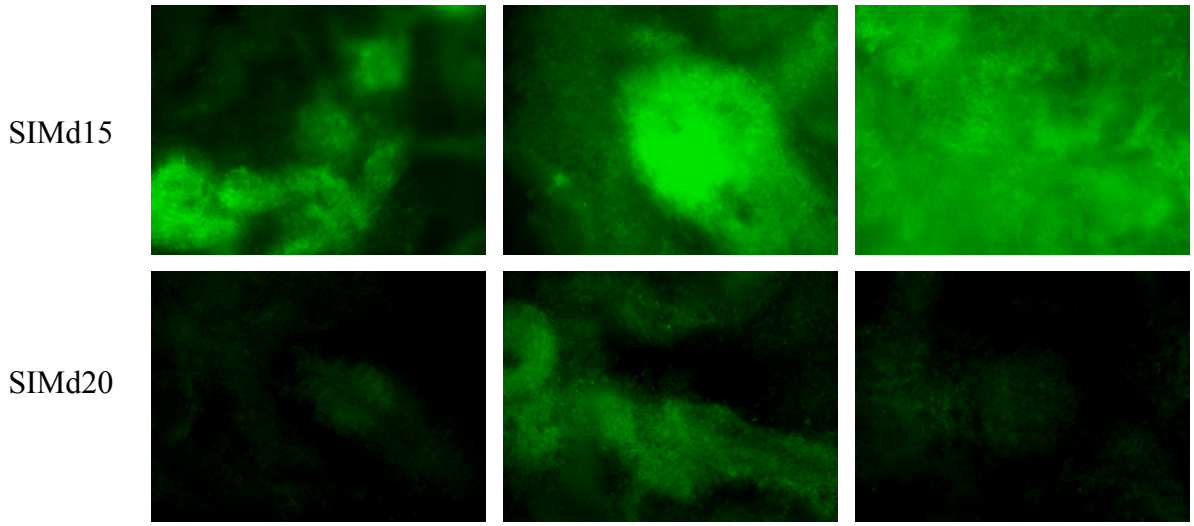


Figure 2-11

Fluorescence of AtPrx33 ::*GFP* calli at SIMd15 and SIMd20.

All pictures were taken with identical settings on 300µm-thick slices of calli. The region observed in each case is the middle of a callus.

C.2.3 Supplementary data: xanthine oxidase experiments

Following callus staining with NBT and HPF, it appeared that superoxide was located quite ubiquitously within the callus, whereas H_2O_2 was more limited to the borders. AtPrx33::*GFP* expression seemed hence to co-localise with superoxide rather than H_2O_2. This result confirms the idea proposed earlier (section B.5: supplementary data) for a role of AtPrx33 in generation of superoxide. As for elongation zone, production of superoxide would maintain the cell wall soft enough to allow elongation, or expansion (in calli) of the cell. The next question was: is there any regulation of AtPrx33 expression by superoxide? The callus model is rather time-consuming and not as easy to standardise (variability between calli) as the root end model. I hence chose to study root ends of seedlings treated with a mixture of xanthine (X) and xanthine oxidase (XO). The substrate and its enzyme are known to produce, when mixed, a burst of superoxide anions. Following the literature (Overmyer et al., 2000; Mazel and Levine, 2001), I chose to use 1mM of xanthine and 0.075 U/ml of enzyme. As soon as xanthine and xanthine oxidase were mixed, optical density (560nm) increased by a rate of 0.320/min. When the X-XO solution was let at room temperature during 45 minutes, and then assayed for NBT reduction, no more superoxide was detected. The half life of superoxide is a few milliseconds (Saran and Bors, 1994): therefore, after 45 minutes, either xanthine oxidase has lost its activity, or xanthine has been entirely consumed.

Seedlings were incubated during 30 minutes in a X-XO solution just after addition of xanthine oxidase, and then immersed in a 2mM NBT solution (20mM pH 6.1 PO_4 buffer) during 15 minutes: the elongation zone was strongly stained, as already observed by Claude, however there was no difference between treated and non-treated roots. This observation could be explained by the short duration of superoxide production with the X-XO reaction system, as mentioned in the previous paragraph. I hence changed the protocol and infiltrated seedlings by exposing them to 5 minutes of vacuum in a X-XO solution, and then directly observed them. This method still did not show any difference in NBT staining between treated and non-treated roots. If superoxide is produced by X-XO, why is it not visible, at least at the root periphery? The explanation came from staining with guaiacol for peroxidase activity: when incubating seedlings in a guaiacol solution without addition of H_2O_2, guaiacol was oxidised only after X-XO treatment (Fig.2-12). Superoxide is hence probably converted to H_2O_2 by superoxide dismutases, and does not enter the root. The system X-XO cannot therefore be used for monitoring effects of superoxide on *AtPrx33* expression in roots. A possible solution would be the use of extremely fine micropipettes for direct injection of X-

XO into the root. Alternatively, the herbicide Paraquat (1,1'-Dimethyl-4,4'-bipyridinium dichloride) could be used. This molecule can indeed penetrate within plants through membrane protein transporters (DiTomaso et al., 1993). Once arrived in thylakoids, the Paraquat cation (PQ^{2+}) is reduced to PQ^{+} by photosystem I electrons: PQ^{+}, in presence of molecular oxygen, reacts very quickly to form superoxide (Hassan and Fridovich, 1979). By this means, we could generate superoxide intracellularly, and analyse possible changes in expression of *AtPrx33*. The only drawback is that such studies could only be performed on aerial organs, as presence of photosystem I is required. In agreement with the failure of X-XO incubation to induce a burst of superoxide in roots, AtPrx33::*GFP* expression in root ends of X-XO-treated seedlings was monitored 6, 24 and 48 hours after incubation in X-XO, but did not vary.

Parallel experiments with DPI, an inhibitor of superoxide production, where conducted: 7 day-old seedlings (grown on normal MS) were transferred on MS agar containing 10 or 100µM DPI. Superoxide localisation did not change within up to 48 hours. DPI 100µM has however an effect on Atprx33::*GFP* expression and superoxide localisation, as demonstrated in figures 5 (o, q, r) of the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**". In these figures, expression of AtPrx33::*GFP* is down-regulated in the "reduced" elongation zone, whereas superoxide is

Figure 2-12
Guaiacol staining (without addition of H_2O_2) of *Arabidopsis* roots.

upper root: control
lower root: pre-treated with a solution of X-XO during 30 min.

still strongly present. Co-localisation of superoxide and AtPrx33 in calli and the root elongation zone is hence not due to induction of AtPrx33 by superoxide. Moreover (article project Fig. 5m), generation of superoxide does not necessarily depend on the presence of auxin and AtPrx33.

A last assay was performed with differentiating calli: staining with ortho-dianisidine and H_2O_2. Calli were analysed at SIMd20. The results show a predominant peroxidase activity in differentiating shoots and in roots, but not in the callus structures (Fig. 2-13), in striking contrast with AtPrx33::*GFP* expression (callus article project Fig. 5E). AtPrx33 seems hence to have a major role in initiation of callus differentiation, which is not shared by most of the other peroxidases (callus article project Fig. 4).

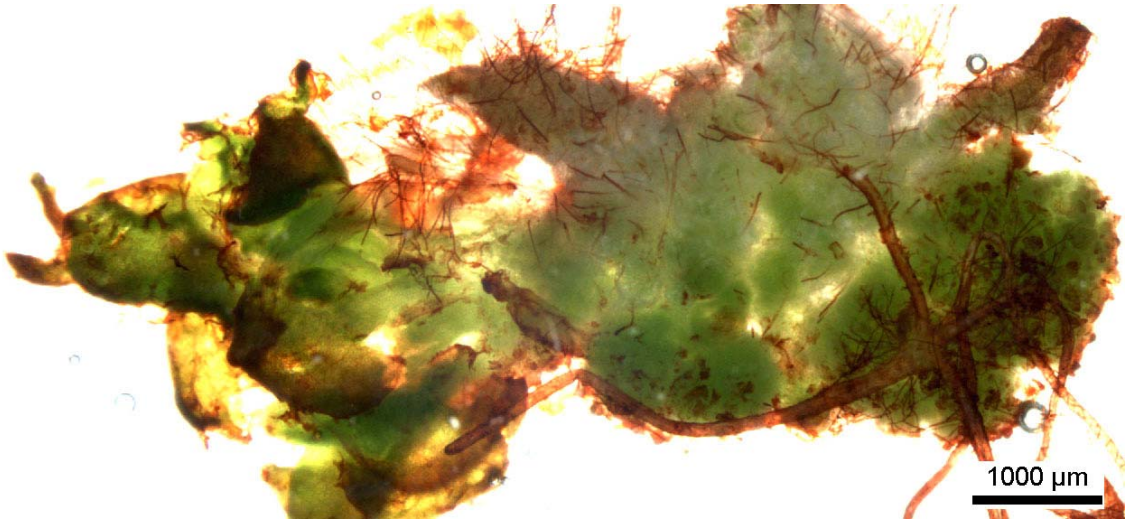


Figure 2-13

Ortho-dianisidine staining of a callus at SIMd20.

C.2.4 Supplementary data : phenotype of adult plants over-expressing *AtPrx33*

In the article project "***Arabidopsis* peroxidase AtPrx33: a new actor on the scene of cell differentiation**", I showed that peroxidases may well be involved in tissue differentiation from calli. During other experiments, I also noticed a possible role of AtPrx33 in tissue differentiation of adult aerial structures, as described hereafter.

When selecting for homozygous T2 Ws 35S::*AtPrx33* lines issued from the T0 transformant line "5", antibiotic-resistant seedlings (7 day-old) were transferred on soil. 25 days after transfer, a dramatic phenotype appeared (Fig. 2-14): many 35S::*AtPrx33* plants were much smaller than controls, and a few plants showed stem fasciation. Both phenotypes

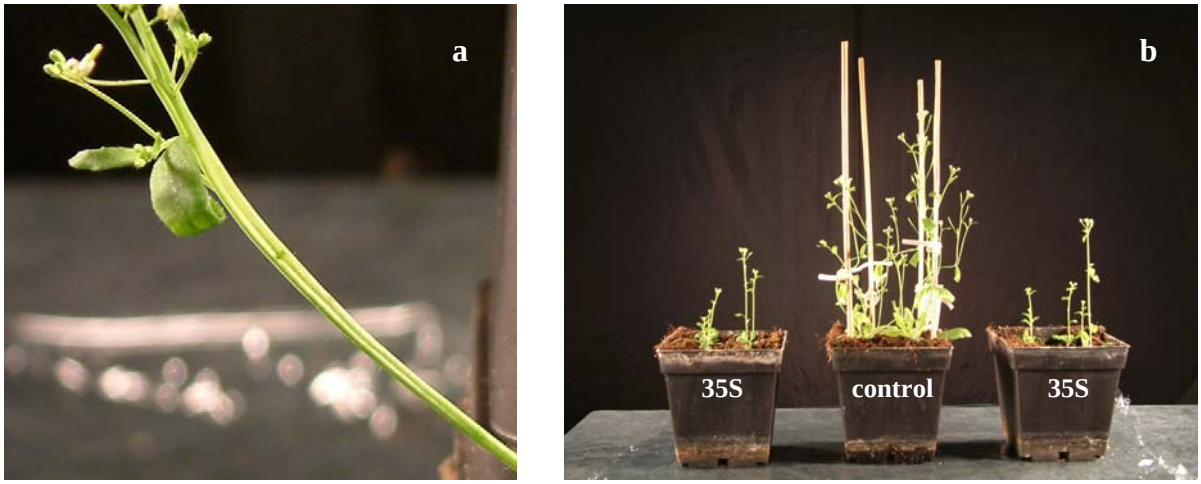


Figure 2-14

Pictures of 1 month-old 35S::*AtPrx33* and control plants.

(a) Example of fasciated stems in a 35S::*AtPrx33* plant

(b) Comparison of 35S::*AtPrx33* (35S) and control (Ws with empty pCHF3 plasmid) plants

were sometimes combined. RT-PCR reactions were performed on the aerial parts of each plant. Results obtained with 5'CA/3'CA primers show that in control plants (C) as well as in two (out of three) 35S::*AtPrx33* plants with a wild type phenotype (N), either *AtPrx33* is not expressed or levels of *AtPrx33* are too low to be detected (in line with the figure 4 from the article: **"Two cell wall associated peroxidases from *Arabidopsis* influence root**

elongation"). In contrast, fasciated (F) and small (S) plants had variable but detectable levels of *AtPrx33*, implying that *AtPrx33* is over-expressed in those plants (Fig. 2-15). Absence of *AtPrx33* expression in "N" plants could be explained by a silencing of the transgene. Plants were considered as "small" (dwarfism) when their size was lower than 50% of the mean height of wild type plants. Importantly, a few 35S::*AtPrx33* seedlings were not able to generate stems (only the rosette formed), and some of them died before observation (Table 2-1).

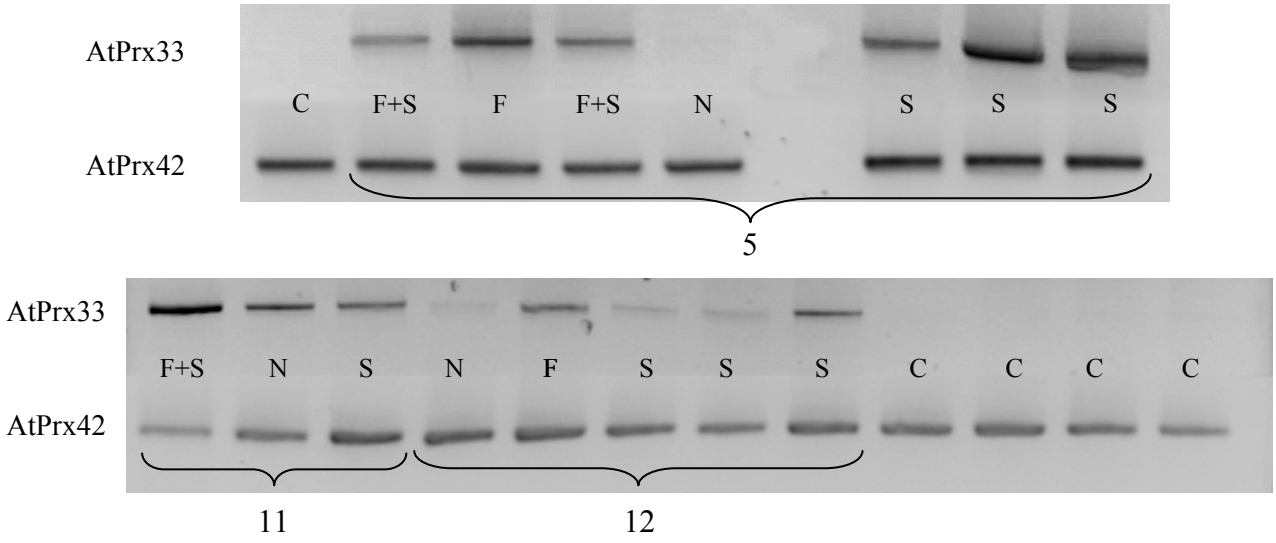


Figure 2-15

RT-PCR reactions with primers specific for *AtPrx33* and *AtPrx42* on 35S::*AtPrx33* lines and Ws pCHF3 control lines (1 month-old).

The two experiments represent two independent series with three different homozygous 35S::*AtPrx33* lines (noted "5", "11" and "12").

C Ws pCHF3; **N** 35S::*AtPrx33* with wild type phenotype; **F** 35S::*AtPrx33* showing fasciation; **S** 35S::*AtPrx33* showing dwarfism; **F+S** combined fasciation and dwarfism.

The main problem of this series of experiments was reproducibility. The three independent lines "5", "11" and "12" were studied. Phenotypes were clearly observed on one occasion for each line and during two separate experiments (Fig. 2-15 and Table 2-1). For lines "11" and "12", a significant proportion of plants were not forming any stem, and others even died before the observation day. The same study was repeated (24.10.2005) with lines "5", but no phenotype was observed. This result was rather unexpected, since the conditions

used for all experiments were the same (T2 generations, same lamps, same shelf, same soil). More surprisingly, in the experiment with lines "12" started on October 31st, the "12.1" T2 seedlings looked quite normal except for a high rate of fasciation: in striking contrast, "sub-lines" "12.5", "12.6" and "12.10" (all four sub-lines are derived from the same T0 transformant) were generally small, many of them did not form a stem or even died before the day of observation, and fasciation was never observed. "12.1" seedlings were grown at the same time and in the same conditions as "12.5", "12.6" and "12.10".

Following these contrasting results, I then tried to find whether an external or internal factor could have an influence on the triggering of the dwarf and fasciate phenotypes. All three lines were sown again (8.12.2005), and 7 day-old seedlings were transferred onto soil. For each sub-line, seedlings were split into two pots (4 seedlings per pot), one pot going into a container that would then be watered with 1600ml twice a week ("1/1 water"), and the other into a container watered with 800ml ("1/2 water") twice a week, a volume just sufficient to cover the surface of the container. Each container included all "sub-lines" and Ws pCHF3 controls. 25 days after transfer, controls and 35S::*AtPrx33* plants looked very similar in both containers: watering cannot explain the differences observed in previous experiments. More extreme watering conditions were not tested, as the goal was to repeat the first successful experiments, in which growth conditions, though not precisely monitored, were quite "standard". During this same experiment, light intensity was measured above each pot, and showed to vary between pots from 190 and 280μE/m^2/sec. Differences in light intensity cannot hence explain the triggering of the small/fasciate phenotype. Finally, within each container, pots were placed onto three rows: the first row was much more exposed to wind (cooling system of the growth chamber), however no difference was observed between the first and the last row. The last external parameter studied was treatment with the insecticide Marshal 25 EC® (Carbosulfan 250g/l-Maag, Switzerland) at a concentration of 0.15%. This product is indeed one of the main insecticides used by our gardeners. Either 50ml of Marshal 0.15% or water were sprayed weekly onto Ws pCHF3 controls and 35S::*AtPrx33* lines, but no significant effect was noted on their growth: adult plants looked quite similar to a third batch of non-sprayed plants (not shown in Table 2-1).

In a further attempt to try and reproduce the small/fasciate phenotype of adult 35S::*AtPrx33* plants, seeds of the two successful experiments (28.09 and 31.10) were collected separately for small (S), fasciated (F) and small/fasciated (F+S) plants. Seeds were pooled for plants without phenotype ("pool"). After a 7 day growth on ½ MS medium and transfer onto soil, plants were let grow and were observed after 25 days: Although we cannot

exclude appearance of the "small" phenotypes in T3 lines "5.3" and "12.1" descending from "F+S" T2 plants, results were not very convincing: the phenotype does not seem to be transmitted to the next generation.

Lines "5"			5.3	5.5	5.8	5.10	Total
28.09.2005	T2	(n)	(11)	(10)	(10)	(10)	(41)
		F	27%	---	---	20%	12%
		S	82%	80%	90%	100%	87%
		F+S	18%	---	---	20%	10%
24.10.2005	T2	(n)		(12)		(11)	(23)
		F		8%		---	4%
		S		---		---	---
08.12.2005	T2, 1/2 water	(n)	(4)		(4)		(8)
		F	---		---		---
		S	---		---		---
	T2, 1/1 water	(n)	(4)		(4)		(8)
		F	---		---		---
		S	---		---		---
07.02.2006	T3 from T2 F+S	(n)	(14)				(14)
		F	---				---
		S	21%				21%
	T3 pool	(n)	(20)				(20)
		F	---				---
		S	---				---
23.10.2006	Kana T3 from T2 F+S	(n)	(12)			(12)	(24)
		F	---			8%	4%
		S	16%			25%	21%
		F+S	---			8%	4%
		no stem	8%			8%	8%
		dead	---			33%	17%

Table 2-1

Observation of 32day-old *Arabidopsis* 35S ::*AtPrx33* plants.

Dates noted correspond to transfer of MS-sown seeds to the phytotron (after 2 days of vernalisation).

(n) total number of plants observed

N.B.: none of the phenotypes described here was ever reported for Ws pCHF3 controls.

Lines "11"			11.1	11.7	Total
31.10.2005	T2	(n)	(12)	(12)	(24)
		F	8%	8%	8%
		S	16%	50%	33%
		F+S	8%	---	4%
		no stem	8%	16%	12.5%
08.12.2005	T2, 1/2 water	(n)	(4)	(4)	(8)
		F	---	---	---
		S	---	---	---
	T2, 1/1 water	(n)	(4)	(4)	(8)
		F	---	---	---
		S	---	---	---
07.02.2006	T3 from T2 F	(n)		(20)	
		F		---	---
		S		5%	5%
	T3 pool	(n)		(20)	(20)
		F		---	---
		S		15%	15%
23.10.2006	Kana T3 from T2 F+S	(n)	(12)		(12)
		F	16%		16%
		S	25%		25%
		F+S	8%		8%
		no stem	8%		8%
		dead	---		---

Lines "12"			12.1
31.10.2005	T2	(n)	(12)
		F	16%
		S	8%
		F+S	8%
		no stem	8%
		dead	8%
08.12.2005	T2, 1/2 water	(n)	(4)
		F	---
		S	---
	T2, 1/1 water	(n)	(4)
		F	---
		S	---
07.02.2006	T3 from T2 F+S	(n)	(20)
		F	---
		S	25%
	T3 pool	(n)	(20)
		F	---
		S	---
23.10.2006	Kana T3 from T2 F+S	(n)	(12)
		F	---
		S	8%
		F+S	---
		no stem	33%
		dead	8%

35S::*AtPrx34* lines (Columbia ecotype, previously obtained in the laboratory) were also grown and observed in the experiment of 8.12.2005 (two watering conditions), but did not show any phenotypic difference compared to Ws pCHF3 controls (data not shown).

Width of central stems (under lower cauline leaves) was measured on 35S::*AtPrx33* lines exhibiting a normal size (experiment 24.10.2005), and ratio height/width was calculated, in order to check whether, although of a normal size, stems of *AtPrx33* over-expressor plants were not thinner: no difference was however observed (data not shown).

Finally (23.10.2006), I sowed seedlings on kanamycin (50µg/ml): I noticed that in both successful (appearance of phenotype) experiments, seedlings had been selected for homozygosity on kanamycin, whereas in all other experiments, the antibiotic had not been used. After one week, all the seedlings were less developed than on MS; 35S::*AtPrx33* lines were even smaller and with shorter roots than controls (Fig. 2-16). Moreover, several seedlings (up to 70%) were showing a mosaic resistance pattern, suggesting the presence of a partial silencing of the transgene. A significant proportion of adult plants from the four 35S::*AtPrx33* lines exhibited a small and/or fasciated phenotype, in contrast to controls, which all looked normal. However, some transgenic lines also had a "wild type" phenotype ("N" plants"). RT-PCR reactions were performed, and gave similar results as in figure 2-15: "N" plants had a partial silencing, whereas "F+S" or "S" plants were over-expressing *AtPrx33*. Control plants had almost no detectable levels of *AtPrx33* (Fig. 2-17). Interestingly, 35S::*AtPrx33* plants that were "S" or "F+S" also had strong levels of the *PR-1* gene, in contrast to "N" plants and controls, where its expression was barely detectable. The *PR-1* gene is known to be up-regulated during systemic acquired resistance (Uknes et al., 1992; Ryals et al., 1996):35S::*AtPrx33* plants may hence be experiencing a stress throughout their whole lifespan, resulting in a severely reduced size. This experiment also suggests that 35S::*AtPrx33* seedlings are more likely to develop a fasciated and/or dwarf phenotype when they are stressed in their early age. It finally explains why previous treatments on adult plants did not have any consequence on growth and organ formation.

In order to verify if stress, and not the kanamycin itself, is responsible for appearance of the phenotype, seedlings were sown on various chemicals (IAA, sirtinol, brassinolide, TIBA and KI) at concentrations known to disturb plant growth (same concentrations as in the article project: "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**"). Additionally, 5 day-old seedlings were incubated during 2 days in the cold room before transfer on soil. One month-old plants were observed, and confirmed the phenotype observed in kanamycin-selected 35S::*AtPrx33* seedlings. However, in all other treatments, transgenic plants did not have a small phenotype. Fasciation was observed in 35S::*AtPrx33* lines at a rate of around 3 to 10%, independently from the treatment; wild type controls did not show any sign of fasciation. Curiously, fasciation was also observed in non-treated plants,

with a similar penetrance (data not shown). This experiment hence suggests that stresses in general do not affect growth of adult 35S::*AtPrx33* plants compared to wild type controls, except for kanamycin. Appearance of fasciation in non-treated plants is quite unexpected, and indicates that another still undefined actor has to be discovered.

The RT-PCR reactions performed on various organs of adult (5 week-old) *Arabidopsis* plants have shown that expression of *AtPrx33* is the strongest in roots (Passardi et al., 2006). Moreover, there is significant evidence that *Atprx33* can be involved in tissue differentiation

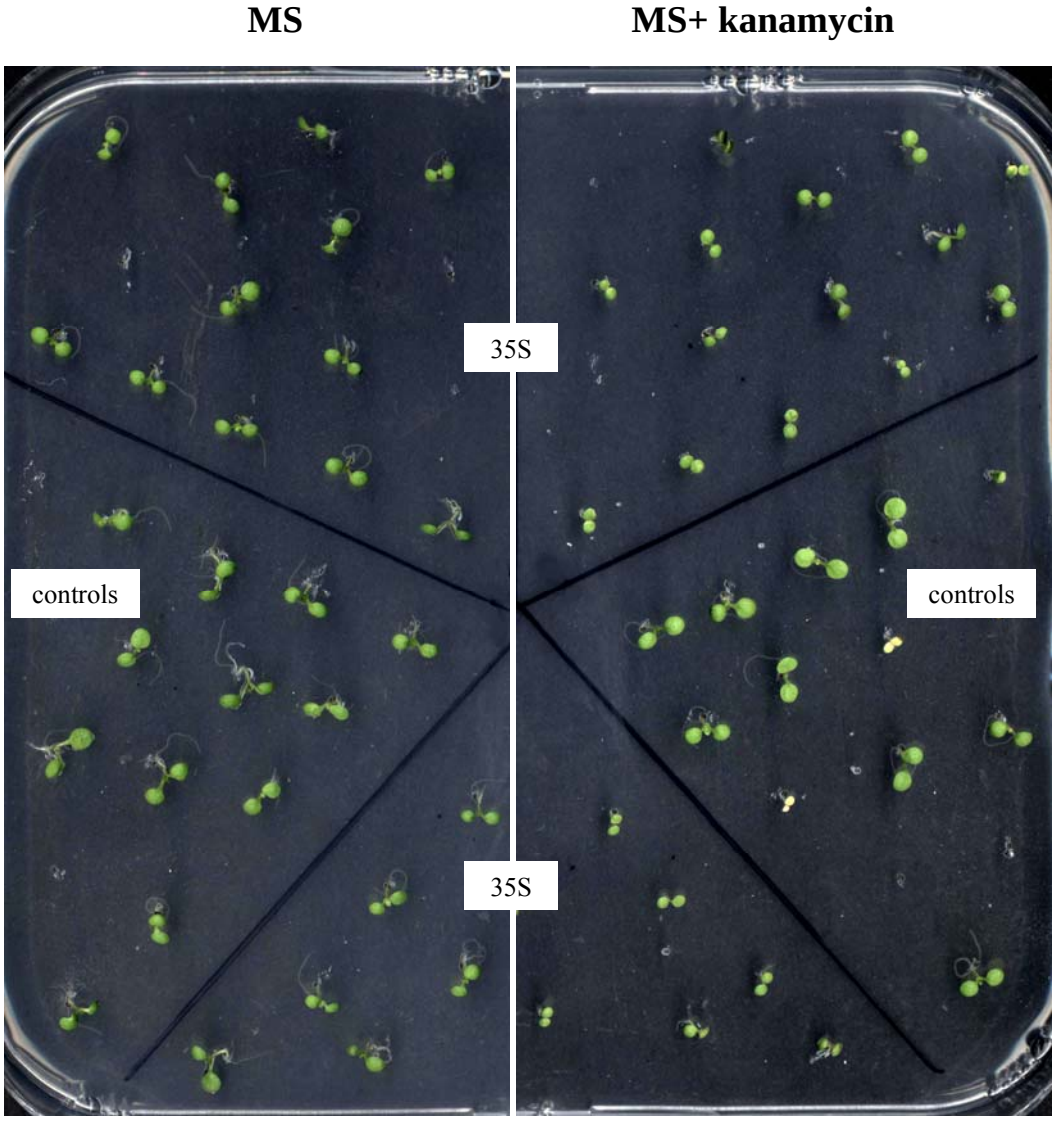


Figure 2-16

One week old seedlings grown on MS agar or MS agar with kanamycin (50µg/ml).

Some seedlings are kanamycin-sensitive in controls (heterozygous line).

(article project "***Arabidopsis* peroxidase AtPrx33: a new actor on the scene of cell differentiation**"). If this peroxidase is over-expressed in the aerial parts of *Arabidopsis*, we could hence have expected an effect on the morphology of the different plant organs. The observation of a fasciation phenotype was then not so surprising. Concerning the small size, absence of stem and higher mortality rate, all these effects are more difficult to explain, but show that ectopic expression of *AtPrx33* can have dramatic consequences. They could be due to a possible function of *AtPrx33* in auxin homeostasis, as suggested in the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**". The relationship with kanamycin is however still obscure. Phenotypic differences were observed in three different 35S::*AtPrx33* lines, hence discarding the possibility of a phenotype due to transgene insertion into a coding region or a regulatory sequence of *Arabidopsis* genome. A hypothesis not tested could be presence of specific bacterium species at a certain stage of plant growth. Several different bacteria are known indeed to induce fasciation both in mono- and dicotyledonous plants by modifying the plant hormone metabolism (Thimann and Sachs, 1966; Nilsson et al., 1996; Bertaccini et al., 2005). In our case, if a similar bacterium was present, the over-expressed *AtPrx33* may have already modified auxin content, resulting in increased susceptibility of transgenic plants to form fasciated stems. Alternatively, presence of a bacterium, or even an insect, may have triggered a defence reaction that led to fasciation in transgenic plants due to interference of the over-expressed *AtPrx33* with the whole defence process.

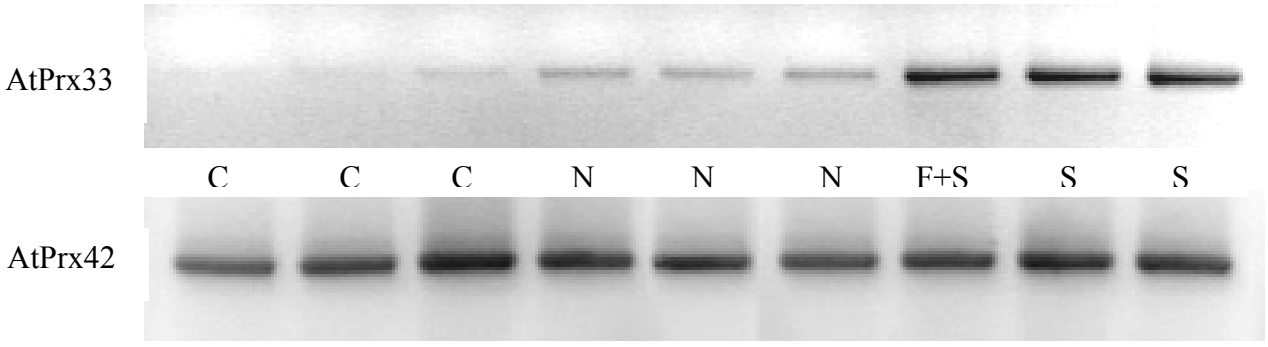


Figure 2-17

RT-PCR reactions with primers specific for AtPrx33 and AtPrx42 on 35S ::AtPrx33 line "5.10" and a Ws pCHF3 control line (experiment 23.10.2006).

Plants were one month-old.

C Ws pCHF3; **N** 35S::*AtPrx33* with wild type phenotype; **S** 35S::*AtPrx33* showing dwarfism; **F+S** combined fasciation and dwarfism.

D. Importance of ecotypes in the study of peroxidases

Since the first days of my PhD thesis, I have been working with two *Arabidopsis thaliana* ecotypes: Columbia (Col) and Wassilewskija (Ws). Indeed, the two peroxidase T-DNA mutants isolated by Michael and Christophe, *atprx33* and *atprx34*, had been obtained in a Ws and Col background, respectively. In the beginning, nobody considered this "detail" as relevant. However, it began soon to gain a growing importance, ending with being an essential parameter to be taken into account. With Christophe and other colleagues, we noticed that ecotypes could be responsible for several phenotypes first attributed to the loss of a peroxidase. For instance, *atprx33* mutant was often compared to a Columbia wild type, leading to erroneous conclusions in callus formation and differentiation studies, as well as root and cell length studies. In order to better understand the many facets characterising each mutant, we decided to screen all phenotypic differences that could be significant for future studies in the laboratory. The Landsberg *erecta* (La*er*) strain was added, as it was the ecotype of knock-out *atprx42* lines studied by Jan Dobias.

D.1 Article: "Morphological and physiological traits of three major Arabidopsis thaliana *accessions"*

The article "**Morphological and physiological traits of three major *Arabidopsis thaliana* accessions**" describes the comparison of the *Arabidopsis thaliana* Columbia, Wassilewskija and Landsberg *erecta* ecotypes. Although similar and more exhaustive studies have already been published (Larkin et al., 1996; Yanovsky et al., 1997; Debeaujon et al., 2000; Boyes et al., 2001; El-Lithy et al., 2004), our work focuses on some aspects that have so far not been thoroughly analysed, particularly on the Ws ecotype (Alonso-Blanco and Koornneef, 2000). Besides introduction and conclusion, only the experimental part concerning callus differentiation is shown hereafter, as it is the one I contributed the most.

This article is cited as follows:

Passardi F, Dobias J, Valerio L, Guimil S, Penel C, Dunand C.
Morphological and physiological traits of three major *Arabidopsis thaliana* accessions.
J Plant Physiol. 2006 Aug; in press
DOI: http://dx.doi.org/10.1016/j.jplph.2006.06.008

Morphological and physiological traits of three major *Arabidopsis thaliana* accessions

ARTICLE IN PRESS

Filippo Passardi[1], Jan Dobias[1], Luisa Valério, Sonia Guimil, Claude Penel and Christophe Dunand[*]

Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland

[1] These authors contributed equally to this work

[*] Corresponding author

Introduction

Native *Arabidopsis thaliana* is widely distributed across Europe, continental Asia and Northern Africa (Ratcliffe, 1965; van der Zwan et al., 2000; Al-Shehbaz and O'Kane, 2002; Hoffmann et al., 2003). The hypothesis of a central Asiatic, west Himalayan origin is supported by the high number of species distributed in these regions (Röbbelen, 1965; Price et al., 1994). Its presence on American and Australian territories is due to human migrations. Its limited geographical distribution indicates its preference for temperate climate with low precipitation (Landolt, 1977). It is a pioneer plant that does not support competition in dense vegetation. *Arabidopsis* was discovered in the Harz mountains in the sixteenth century by Johannes Thal (hence, thaliana), who called it *Pilosella siliquos*. The current *Arabidopsis* name was proposed by de Candolle (de Candolle, 1821). The very first *A. thaliana* ecotypes appeared 5 MYA ago (Koch et al., 2000). This gave sufficient time for the creation of current ecotypes through natural mechanisms.

Over 750 natural ecotypes or accessions of *A. thaliana* have been collected from around the world and are available from two major seed stock centers, the Arabidopsis Biological Resource Center (ABRC) and the Nottingham Arabidopsis Stock Centre (NASC). Researchers are using the differences within this natural variety (Quiroga et al., 2000) to uncover complex genetic interactions such as those underlying plant responses to

environment and evolution of morphological traits. These different «ecotypes» or accessions are quite variable in terms of morphology and development (e.g. leaf shape, trichoblast formation) and physiology (e.g. flowering time, disease resistance).

The most popular *Arabidopsis* accessions are Columbia (Col), Landsberg *erecta* (La*er*) and Wassilewskija (Ws). These three ecotypes are widely used for both molecular and genetic studies and are the chosen genetic background for the majority of *Arabidopsis* T-DNA insertion mutant collections. The Col and La*er* ecotypes both originate from a wild type Landsberg (La) strain selected by Professor George Rédei, whereas the Ws (N1602) ecotype originally comes from Vasljevici, in Belarus (http://seeds.nottingham.ac.uk). In contrast to Col, the La*er* line was obtained by X-ray induced mutagenesis (Torii et al., 1996). Col ecotype was proposed to be the molecular model for *Arabidopsis* by E. Meyerowitz and since then its genome has been entirely sequenced (*Arabidopsis* Genome Initiative, 2000). La*er*, is currently being sequenced with the purpose of obtaining new molecular markers for chromosome walking (http://www.tigr.org/tdb/e2k1/ath1/atgenome/Ler.shtml). As *Laer* sequence data becomes available, its comparison with Col will facilitate the discovery of gene sequence variation between the two ecotypes.

To illustrate genetic differences between the three ecotypes, we focused on the class III plant peroxidases (EC 1.11.1.7), a reactive multigenic family that is easy to study at the biochemical level. These enzymes reduce hydrogen peroxide (H_2O_2) by transferring electrons from various donor molecules, such as phenolic compounds, lignin monomers or auxin. Due to this substrate diversity, the class III peroxidases are implicated in a large range of biological processes such as cell elongation, cell wall differentiation, auxin metabolism and defense against pathogens (Heloir et al., 1996; Hiraga et al., 2001; Passardi et al., 2004; Passardi et al., 2005). Class III form multigenic families in all green plants (Bakalovic et al., 2006).

Because of their high reactivity and implication in a large range of physiological processes, peroxidase proteins are considered to be good environmental and molecular markers. Many members of this large gene family are strongly expressed in different organs and are regulated by numerous stimuli (Valério et al., 2004). Furthermore, peroxidases have experienced a high rate of duplication and hence serve as interesting markers for plant evolution (Passardi et al., 2004).

The major *Arabidopsis* ecotypes, Col, La*er* and Ws, are well known for their morphological differences and for their genetic variability (Alonso-Blanco and Koornneef, 2000). Numerous papers have been published concerning either the exhaustive study of

growth and development of one ecotype (Boyes et al., 2001) or specific trait analyses between Col and La*er* ecotypes, such as trichome number (Larkin et al., 1996), seed weight, anthocyanin content (Debeaujon et al., 2000; El-Lithy et al., 2004), and growth responses to very low light conditions (Yanovsky et al., 1997). As yet, a comprehensive study that characterizes the major traits of each ecotype and compares their differences has not been published. Recently, a few studies have shown a heterosis effect observed in the F1 generation of a cross between different *Arabidopsis* accessions (Meyer et al., 2004; Rohde et al., 2004). Identification of new phenotypic differences between these ecotypes would further characterize the heterosis effect in *Arabidopsis.* Additionally, identification of genes differentially expressed between ecotypes could provide a putative molecular genetic explanation for the heterotic observation. For this reason, we have analyzed the phenotypes of the three main *A. thaliana* ecotypes from germination to silique formation, and subsequently examined cellular and molecular aspects linked to peroxidase gene expression.

Callus formation and shoot regeneration

Besides the anatomical differences described above, the three *A. thaliana* ecotypes can be distinguished at the cellular level. It is well known that plants can generate various kinds of tissues from undifferentiated cell clumps called calli (Skoog and Miller, 1957). The fate of a callus depends on the balance of hormones (mainly auxins and cytokinins) present in their culture medium. Excised roots from Col and Ws plants formed a similar number of calli at a similar rate when transferred to an appropriate medium. La*er* formed as many calli as Col and Ws, but at a slower rate (data not shown). On this same medium, however, Ws calli started to generate shoots 10 days after transfer and, had over 80% differentiation 22 days after transfer (Fig. 7B). La*er* calli

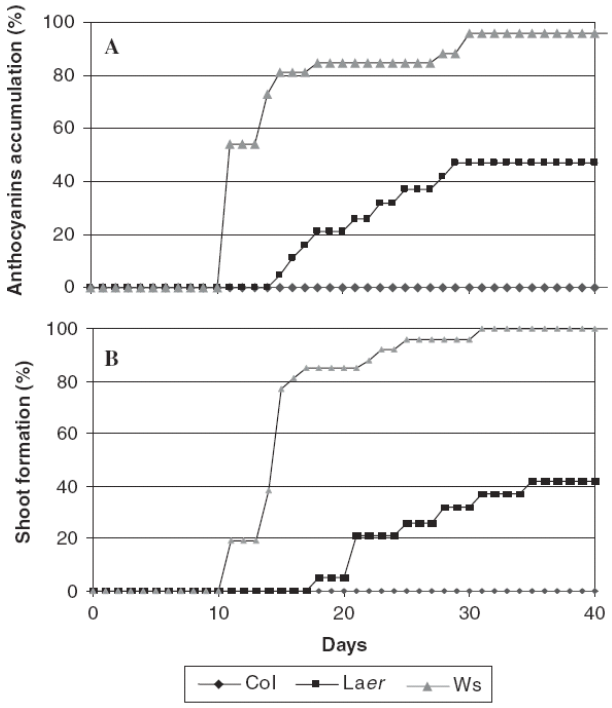


Figure 7. Anthocyanins accumulation expressed in percentage of roots producing anthocyanins (A) and shoot regeneration (percentage of differentiated roots) (B) on root segments from the three ecotypes. The 100% is based on the total number of roots. Difference between Ws and La*er* is statistically significant (*P*-values < 0.001 using Log Rank, Breslow and Tarone–Ware pairwise comparisons).

are also able to generate shoots, some days after Ws. Anthocyanin accumulation is a marker for the onset of differentiation and appeared just before the onset of shoot (Fig. 7A). In striking contrast, no organogenesis is observed for 40 days in Col root calli. Col calli only start forming shoots when the concentration of auxin is increased ten-fold (data not shown). All shoots from the different ecotypes eventually lead to fully differentiated plants (although much smaller than plants coming from seeds), with leaves, flowers and siliques. When transferred to soil, however, plants were not able to grow further, mainly due to the lack of developed roots.

Although there is still no direct experimental proof, one may speculate that this discrepancy of organogenesis between the three ecotypes could be due to a change in the balance of hormones in relation to variation in the expression of certain class III peroxidases. Indeed, as peroxidases are able to degrade auxin *in vitro* (Gaspar et al., 1982), some of them may act *in vivo* and hence create a particular hormone environment around the roots of each ecotype. The change of auxin balance would then influence the rate of tissue formation.

Conclusions

The phenotypic differences observed amongst the three ecotypes can be explained by genetic variability. The compact appearance of La*er* plants is due to a mutation in the *erecta* gene which determined the shape of organs originating from the shoot apical meristem (Torii et al., 1996). The faster growth rate of Ws is almost equally related to higher mature cortical cell length and to faster rate of cell production probably correlated with cyclin-dependent kinase activity (Beemster et al., 2002). Curiously, in some phenotypic aspects such as seed shape, root structure on horizontal plates or rosette morphology, La*er* was more similar to Ws than Col. Although La*er* has been X-ray irradiated, both Col and La*er* come from the same seed stock and would be expected to be more similar to one another than to Ws. This discrepancy is either due to a broad effect of X-rays or to an initial heterogeneity between the subpopulations of original Landsberg stock.

The divergence between *Arabidopsis* ecotypes observed in this work can be related to several factors. Since their appearance 5 million years ago in central Asia (Koch et al., 2000), *Arabidopsis* plants have spread across Asia and Europe and have been facing an extensive variety of biotopes, with radical changes in temperature, light, and exposure to pathogen. Even in a given region, climatic changes, such as during the glaciation periods, have had a strong influence on *Arabidopsis* selective evolution (Sharbel et al., 2000). As class III

peroxidases are very sensitive to a wide range of stress factors (Hiraga et al., 2001), the cumulative changes in biotopes and climate affected the evolution of *Arabidopsis* genes. Some of these changes gave rise to novel phenotypes which in turn gave rise to novel ecotypes. Moreover, the evolution of peroxidases may have had a more dramatic impact on the appearance of ecotypes and on environmental adaptation than other protein families. Indeed, peroxidases have a particularly higher gene duplication rate probably related to external stress conditions (Passardi et al., 2004).

This article does not intend to show an exhaustive list of all differences existing between Col, La*er* and Ws. However, we would like to draw attention to the fact that there can be considerable discrepancies between ecotypes in numerous traits, which should not be mistaken as a novel mutation during plant screenings. This study differs from previous publications in that we have analyzed various morphological, cellular and molecular traits in three major ecotypes instead of one specific trait deeply analyzed in many ecotypes or only in one ecotype versus a mutant. It provides a global overview of principal differences. The variability of features among ecotypes has developed over a relatively short evolutionary time frame (5 MYA). Since *Arabidopsis* is used as a scientific model, it is probable that its evolution has been accelerated through human impact (i.e. crossing between ecotypes). Ideally, a more complete phenotypic analysis of several ecotypes could be the starting point for future heterosis experiments.

E. Discussion and prospects

Following three years of work on functional characterisation of AtPrx33 and AtPrx34, several questions have been answered, both on localisation and role of these two peroxidases in *Arabidopsis thaliana*. Both peroxidases seem to share similar functions, and are expressed mainly in roots. When aligning the two peroxidases genes, the contrast is striking: a very high identity between the coding sequence, and almost none in the promoter region. In order to differentiate both peroxidases, AtPrx34::*GFP* transgenic plants are definitely needed. A transgenic AtPrx34::*AtPrx34*::*GFP* line had been already obtained by Christophe. I observed it in parallel with AtPrx33::*AtPrx33*::*GFP* seedlings, but was confronted to the same problem: absence of signal.

Despite being almost identical, slight variations in amino acid sequences could be responsible for differences in substrate oxidising activities, substrate affinities, or other physico-chemical properties. One hint already came with the study of baculovirus-produced AtPrx33 and AtPrx34, where pectin-binding activity was stronger for AtPrx34. Unfortunately, Mireille could only produce little amount of both peroxidases, just enough in fact to perform the pectin-binding assay. It could be worthwhile, in a near future, to try and produce more peroxidase through the baculovirus-insect system, or even to start again bacterial production with lower growth temperatures or with more aggressive techniques to break inclusion bodies and solubilise proteins. Obtaining higher amounts of peroxidases will allow testing a wide variety of substrates, and bacterial-produced peroxidases could, ideally, be crystallised and bring useful information on the importance of the few amino acids that differ between AtPrx33 and AtPrx34.

Regarding the interaction between auxin and AtPrx33 *in planta*, working with the GH3::*GUS* reporter proved to be rather tedious, as this marker is not always representative of auxin levels *in vivo*. Using DR5::*GUS* could help confirming some results, but it could also bring more confusion (Bierfreund et al., 2003). Whereas auxin-responsive promoters can still be used in order to rapidly obtain hints on auxin distribution, two other techniques should be more often employed. Auxin dosage by chromatography is probably the most accurate method, although only for relatively large portions of plants. For seedlings already, one measurement necessitated hundreds of plants. This approach could prove particularly useful with other peroxidase mutant lines available in the laboratory, as well as with the *AtPrx33* over-expressor line, which, at the time of performing auxin dosage experiments, was not yet

available. It is indeed quite intriguing that shoots of all transgenic lines assayed in the article project "**The intricate relationship of auxin and AtPrx33 peroxidase: a dance or a glance?**" (Fig.6) were showing a tremendous drop in their auxin levels. The second technique to consider is the use of anti-auxin antibodies (Leverone et al., 1991; Benkova et al., 2003; Hou and Huang, 2005). Anti-IAA monoclonal antibody is commercially available (www.agdia.com) and was shown to be highly specific for free auxin (Weiler, 1981; Mertens et al., 1985). The drawback is that, until now, no antibody is available against the many conjugated forms of auxin. Nevertheless, if we want to monitor auxin distribution in precise parts of the plant, including at the cellular level, this is certainly the most straightforward method.

Studies on callus differentiation demonstrated a role for AtPrx33 in organ formation, but also showed that many other peroxidases are playing a similar function. The most surprising conclusion came out from the double-mutant *atprx33*/*atprx34*, which showed a differentiation rate half-way between wild type and *atprx33*. In the article project, I proposed that this phenotype was probably caused by RNAi down-regulation of additional peroxidases, due to the generally high identity among *Arabidopsis* peroxidases. If this hypothesis is true, then at least one peroxidase must exist that should prevent organ differentiation. In a screening of 12 mutants, 4 were found with a delayed and reduced differentiation rate, and none with an opposite phenotype: putative "organogenesis-inhibiting" peroxidases would then be less numerous that the "organogenesis-promoting" ones. The peroxidase AtPrx36 may be one of these putative "organogenesis-inhibiting" peroxidases, as its overexpressor mutant provoked a delay in callus differentiation. However, this is not sure, since the same phenotype was observed with calli of both *atprx33* and 35S::*AtPrx33* lines. If new peroxidase mutants are isolated in the laboratory in the future, it may be interesting to monitor their effect on callus differentiation, especially if we can isolate knock-out and overexpressor lines of a same peroxidase.

The series of experiments conducted on calli with different coloured filters are also a source of many future exciting projects. By using the standardised protocol that I set up in the laboratory, photoreceptor knock-out mutants could be assayed to define which light-sensing proteins are the most involved in callus differentiation, and then try and establish a regulation mechanism between photoreceptors and peroxidases. Many transcription factors have already been identified that are directly related to activation of a photoreceptor (Wang and Deng, 2003). Performing band-shift assays with selected peroxidase promoters could hence link all players together, for the first time.

3

Class III peroxidases in other plants

A. Introduction

Arabidopsis thaliana is the first plant to have its whole genome fully sequenced. As soon as this data was available, Michael Tognolli, from our laboratory, searched for all *Arabidopsis* peroxidase genes and found as many as 73, in addition to 6 pseudogenes. In parallel, EST data was already quite considerable, and he could see that 70% of the genes found were expressed (Tognolli et al., 2002). Further work by Luisa Valério and others (Welinder et al., 2002; Valério et al., 2004) allowed increasing this rate to almost 100%. At the start of my thesis, sequencing of the rice genome was almost completely achieved: it was a unique opportunity to search for all peroxidase genes present in another organism. This would bring information about total number of peroxidases per organism, presence of organism-specific peroxidases and conservation of peroxidases throughout evolution. This last aspect further led us to screen many other plants for presence of peroxidase genes, and gain a better understanding of peroxidase appearance and evolution. Although the final proof has not been obtained yet, there is strong evidence for an appearance of the first peroxidases in evolved algae, and hence a key role of peroxidases for colonisation of land by plants, about 500 million years ago (Kenrick and Crane, 1997; Strother, 2000; Sanderson, 2003; Wellman et al., 2003).

B. Article: “The class III peroxidase multigenic family in rice and its evolution in land plants”

As mentioned in the introductory part of this chapter, the beginning of my thesis coincided with the achievement of the rice genome sequencing. With Christophe, we hence decided to screen rice genome for all peroxidase genes. The screening procedure and the results obtained are described in the article: “**The class III peroxidase multigenic family in rice and its evolution in land plants**”. The main conclusions are that:

- rice possesses almost twice the number of peroxidase genes than *Arabidopsis thaliana*.
- rice peroxidase genes accept unusually long introns (up to 10kb) and splice them efficiently.
- tandem and interchromosomal duplications are very frequent (more than in *Arabidopsis*).
- peroxidase genes in rice are evolving rapidly.

This last point suggests a strong impact of human selection pressure on evolution of rice (and probably of other cereals) genome. Due to their putative function in the defence of the plant against different stress factors, peroxidase genes probably evolve more rapidly than other genes (new pathogens, biotopes, chemicals, etc…).

This article is cited as follows:

Passardi F, Longet D, Penel C, Dunand C.
The class III peroxidase multigenic family in rice and its evolution in land plants.
Phytochemistry. 2004 Jul;65(13):1879-93.
DOI: http://dx.doi.org/10.1016/j.phytochem.2004.06.023

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 65 (2004) 1879–1893
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

The class III peroxidase multigenic family in rice and its evolution in land plants ☆,☆☆

Filippo Passardi [a], David Longet [b], Claude Penel [a], Christophe Dunand [a,*]

[a] Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland
[b] Department of Zoology and Animal Biology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland

Received 16 February 2004; received in revised form 11 June 2004
Available online 20 July 2004

Abstract

Plant peroxidases (class III peroxidases, E.C. 1.11.1.7) are secreted glycoproteins known to be involved in the mechanism of cell elongation, in cell wall construction and differentiation, and in the defense against pathogens. They usually form large multigenic families in angiosperms. The recent completion of rice (*Oryza sativa japonica* c.v. Nipponbare) genome sequencing allowed drawing up the full inventory of the genes encoding class III peroxidases in this plant. We found 138 peroxidase genes distributed among the 12 rice chromosomes. In contrast to several other gene families studied so far, peroxidase genes are twice as numerous in rice as in *Arabidopsis*. This large number of genes results from various duplication events that were tentatively traced back using a phylogenetic tree based on the alignment of conserved amino acid sequences. We also searched for peroxidase encoding genes in the major phyla of plant kingdom. In addition to gymnosperms and angiosperms, sequences were found in liverworts, mosses and ferns, but not in unicellular green algae. Two rice and one *Arabidopsis* peroxidase genes appeared to be rather close to the only known sequence from the liverwort *Marchantia polymorpha*. The possible relationship of these peroxidases with the putative ancestor of peroxidase genes is discussed, as well as the connection between the development of the class III peroxidase multigenic family and the emergence of the first land plants.

Keywords: *Arabidopsis thaliana*; Rice; Multigenic family; Evolution; Duplication; Laccase; Phytochrome; Peroxidase ancestor

1. Introduction

Plant peroxidases (EC 1.11.1.7), often designated as class III peroxidases (Welinder, 1992a,b), are heme-containing proteins generally encoded by large numbers of paralogous genes. *Arabidopsis* genome, known to harbor a high number of multigenic families (Arabidopsis Genome Initiative, 2000), contains 73 peroxidase genes (Tognolli et al., 2002; Welinder et al., 2002). Almost every living organism contains peroxidases (donor: hydrogen peroxide oxidoreductase) that catalyze the reduction of H_2O_2 by taking electrons to various donor molecules. In the case of class III plant peroxidases, the donor molecules can be phenolics, lignin precursors, or secondary metabolites. Plant peroxidases can also oxidize the growth hormone auxin, as well as other substrates (Gaspar et al., 1982) and produce H_2O_2 (Blee et al., 2001) and hydroxyl radicals (Chen and Schopfer, 1999), two activated oxygen species involved in oxidative burst and in cell elongation (Bolwell et al., 1998; Joo et al., 2001; Rodriguez et al., 2002; Liszkay et al., 2003). The diversity of the reactions catalyzed by plant peroxidases explains the implication of these proteins in a broad range of physiological processes, such as auxin metabolism, lignin and suberin formation, cross linking of cell wall components, defense against pathogens or cell elongation (Penel et al., 1992; Hiraga et al., 2001). The plant peroxidase protein sequence is characterized by the presence of highly conserved amino acids, such as

☆ Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.phytochem.2004.06.023.
☆☆ All rice peroxidase sequences are being submitted to the EMBL database and will be released in case of publication.
* Corresponding author. Tel.: +41-22-379-3012; fax: +41-22-379-3017.
E-mail address: christophe.dunand@bota.unige.ch (C. Dunand).

0031-9422/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2004.06.023

two histidine residues interacting with the heme (distal and proximal histidines) and eight cysteine residues forming disulfide bridges (Fig. 1). The distal histidine is necessary for the catalytic activity. These histidine residues are present in all known heme-containing peroxidase sequences. Plant peroxidase proteins exhibit a constant molecular organization based on the presence of 13 α-helices, 10 of these helices being already present in prokaryotic peroxidases (Welinder, 1992a,b; Gajhede et al., 1997).

Multigenic families originate from gene duplications resulting from different mechanisms (Zhang, 2003): unequal crossing-over, various transposition events, duplication of large chromosome segments or polyploïdization. One of the transposition events, the retrotransposition of cDNA, is characterized by the loss of all introns and regulatory sequences and by a random insertion within the genome. In this case, duplicated genes are usually not transcribed and rapidly degenerate into pseudogenes (Casacuberta and Santiago, 2003). On the other hand, direct transposition of genomic sequences (without an RNA intermediate) creates new genes with higher chances to retain their functionality. Miniature inverted-repeat transposable elements (MITE) belong to this category. These sequences are not able to transpose autonomously. Their replication is probably mediated by a *trans*-acting transposase that recognizes specific sequences present on every MITE called terminal inverted repeats (Casacuberta and Santiago, 2003). MITEs are extremely frequent in rice genome and are generally found in gene-rich regions (Mao et al., 2000). Finally, segmental, chromosomal or whole genome duplications are relatively frequent in plants and are not a source of tandem repeats (Skrabanek and Wolfe, 1998; Blanc et al., 2000; Gebhardt et al., 2003; Zhang, 2003).

It seems difficult to estimate the rate and the conservation of gene duplication without considering the impact of external factors such as the natural or human selections, the chromosomal location and the gene function. Lynch and Conery (2000), have estimated that the rate of duplication is one per gene and per 100 million years (MY) in eukaryotes. A duplicated gene, as every gene, is exposed to nucleotide substitutions at an estimated rate of 0.1 to 0.5 substitution per site and per 100 MY (Graur and Li, 1999). Due to this substitution rate, the usual fate for duplicated genes is to become pseudogenes and to be finally deleted after numerous mutations, insertions and deletions, unless it leads to a selective advantage. In that case, plant will preserve it from degenerative mutations. The retention of a duplicated gene occurs for example when its presence is beneficial for the plant because an extra amount of the encoded protein is profitable for a given biological mechanism. Such cases are known as functional redundancy and are exemplified by the multiple copies of rRNA genes or chloroplast and mitochondria entire genomes (Gillham, 1994). A modification of the duplicated gene expression (subfunctionalization) can also explain its persistence. Lastly, a duplicated gene can acquire a new function (neofunctionalization): this is the most important outcome of gene duplication and can explain the presence of large multigenic families, each paralog becoming specialized for a specific task (Zhang, 2003). The number of retained duplications and of pseudogenes gives information on the gene stability, the gene family evolution, and the importance of the protein function in a given organism.

In September 2003, the Gramene website (Ware et al., 2002), which collects all data concerning grass genomes and particularly rice data from the International Rice Genome Sequencing Project (IRGSP), announced the completion of the whole rice genome sequencing. Consequently, the entire genomic sequences of two Angiosperms, *Arabidopsis* (a Eudicotyledon) and rice (a Monocotyledon) became available for comparison. Both genomes differ in several aspects: *Arabidopsis* genome is composed of five chromosomes, 125 Mbp and more than 26,000 genes, whereas rice genome is nearly four times larger with 12 chromosomes, 420–466 Mbp and probably twice more putative genes (Delseny, 2003; Schoof and Karlowski, 2003). Comparison of the large peroxidase multigenic family in the two plants provided the opportunity to better understand the evolution of these two genomes that diverged from a common ancestor about 150 million years ago, according to one molecular clock approach (Wikstrom et al., 2001). In addition, the completed genome sequencing of rice and

L----Y---**C**P--E-[-V---V-----]------[--A-L-**R**L-F**HDC**]FV-[**GCDAS**VLL]----(--)------[EK]-(-)A-PN--

-S[LRG](--)[F---D--K---E]--(--)**C**P-[**VSCADI**]-[A-AARD-V-]--GGP---V-[**G**R-**DG**-]----

------[LP-P----](-)[--L---F]---GL(-)[---D-V-**LS**GA**H**T-**G**---]-[C--]F--R-----(+1 to 13 a.a.)-----A--

L---C----(+1 to 15 a.a.)----D--TP--[FDN-YY-------GL--S**D**--][-L]---------V---A-----[F---F---]

[MVKMG]-----TG-[-GEI**R**-](-)-[**C**--]-N

Fig. 1. Rice peroxidase consensus amino acid sequence. The 60% consensus sequence has been obtained from the alignment of the 138 OsPrx protein sequences. The boxes represent the residues used for the phylogenetic analysis. The bolded amino acids correspond to a consensus of 100%. Hyphens stand for variable amino acids. Bracketed hyphens represent additional amino acids present exceptionally in a few peroxidases. Bracketed numbers indicate the size range of variable regions.

Arabidopsis and the existence of many EST projects allowed performing an exhaustive analysis of peroxidase genes in *Oryza sativa* ssp. *japonica* and a comparative study of the class III peroxidases among green plants. We found that the rice peroxidase family is composed of 138 genes and 14 pseudogenes, dispersed within the genome. Due to this dispersion, the members of the peroxidase family can be used as milestones to study gene duplications. The intron/exon structure shows different patterns than in *Arabidopsis* with a high intron size variability. The presence of genes encoding class III peroxidases has been observed in every branch of the green plants, except algae, suggesting that ancestor of class III peroxidases appeared around 470 MY ago together with the emergence of the liverworts or other primitive land plants (Kenrick and Crane, 1997).

2. Results

2.1. Rice peroxidase genes and their orthologs in Arabidopsis

The complete genome sequence obtained from the Rice Genome Project (RGP) allowed drawing up the exhaustive list of the peroxidase genes in rice. We identified 138 genes and 14 pseudogenes in the annotated and unannotated BAC sequences (table available online). Rice peroxidases are named hereafter OsPrx, followed by a number. *OsPrx1* is the first peroxidase gene on chromosome I, and so on following the order of the BACs on chromosomes, until *OsPrx138*, the last peroxidase gene on chromosome XII. Automated annotation of rice genome was mainly based on sequence comparisons with the *Arabidopsis* genes (Schoof and Karlowski, 2003). Despite this large and detailed information, numerous BAC sequences were not well annotated by automated programs such as FGenesh, Genscan, RiceHMM, NetGene2, GlimmerR and GeneMark.hmm (Schoof and Karlowski, 2003). In a few cases, these gene-predicting programs did not detect peroxidase sequences. We looked for the presence of peroxidase genes in the unannotated BACs with the programs FGenesh and Genscan, and checked by visual analysis the putative sequences found in order to correct possible prediction mistakes. We also verified every annotated BAC, which allowed correcting predicted peroxidase sequences as well as finding sequences that had been missed in the BAC by automated annotation methods. EST sequences were also searched for in the RGP website (http://rgp.dna.affrc.go.jp). Surprisingly, only about one third of the inventoried peroxidase genes showed registered cDNA sequences in the NCBI EST database. However, when searching in another EST database (Hiraga et al., 2000), we found a few more peroxidases, thus increasing the total expressed peroxidase number to 42%. This expression level remains low compared to *Arabidopsis* (82%; Tognolli et al., 2002). It seems very unlikely that so many functional genes have been preserved but not expressed. The apparent low level of peroxidase gene expression is probably due to the absence of cDNA libraries produced from plants grown in conditions inducing peroxidase gene expression (pathogen interaction, biotic and abiotic stresses). On the other hand, the amount of EST sequences (266,000, http://www.ncbi.nlm.nih.gov) is certainly still too low for the number of genes, considering that many EST sequences are overlapping and hence represent only one gene.

The predicted protein sequences encoded by the peroxidase genes identified in rice genome contain the two histidine and eight cysteine residues characteristic of class III peroxidases (Fig. 1). A 60% consensus protein sequence obtained after the alignment of the 138 predicted OsPrx shows conserved residues and domains also found in *Arabidopsis* peroxidases (AtPrxs), such as FHDC, VSCAD, GAHT or GEIR, the first and the third sequence, respectively, containing the two essential histidines. Similar regions are also typically found with a high conservation in peroxidases from many species, for instance horseradish, barley, tomato, wheat and cucumber (Welinder, 1992a,b). Because our aim was to describe the different peroxidase sequences and families found in rice and also to deduce their possible phylogenetic relationships, only homologous amino acids positions had to be conserved for phylogenetic inference. Therefore, only boxed sequences and positions shown in Fig. 1 were used to construct the OsPrx distance tree (Fig. 2), thus discarding variable regions that were not clearly aligned among all identified peroxidase sequences.

The distance tree of rice peroxidases (Fig. 2) provides a general overview of this large family. We divided it into distinct groups and subgroups based on observed genetic distances and bootstrap supports. To the exclusion of pseudogenes, groups II and VI are strongly supported by bootstrap values. Larger groups, I, IV and V, were divided into generally well-supported subgroups. In the group V, only one subgroup is well defined. We used the peroxidase sequence of *Marchantia polymorpha* to root the tree. Indeed, it provides the maximal sequence length when compared to other outgroups and it is the only known peroxidase sequence in liverworts. Interestingly, OsPrx73 and, to a lesser extent OsPrx116, appeared as the most basal sequences and did not present a phylogenetic association with any other peroxidases in 95% and 51% of bootstrap replicates, respectively. Further analysis with Maximum-Likelihood (ML) tree (data not shown) supported the basal sequence of OsPrx73, Os Prx116 and most of the subgroups.

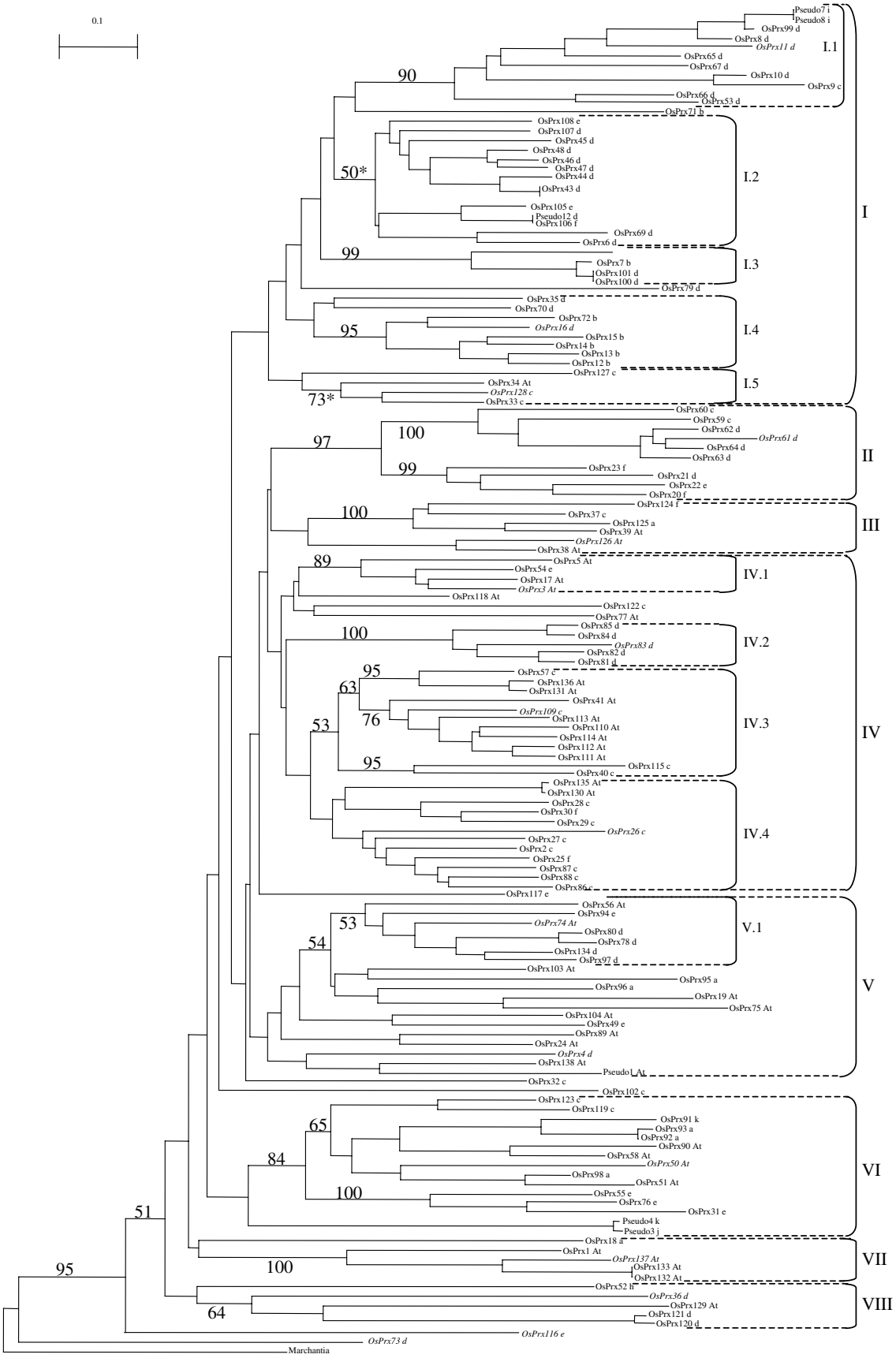

0.1
Pseudo7 i
Pseudo8 i
OsPrx99 d
OsPrx8 d
OsPrx11 d
I.1
OsPrx65 d
OsPrx67 d
OsPrx10 d
OsPrx9 c
OsPrx66 d
OsPrx53 d
OsPrx71 b
90
OsPrx108 e
OsPrx107 d
OsPrx45 d
OsPrx48 d
OsPrx46 d
OsPrx47 d
OsPrx44 d
OsPrx43 d
50*
I.2
OsPrx105 e
Pseudo12 d
OsPrx106 f
OsPrx69 d
OsPrx6 d
I
99
OsPrx7 b
OsPrx101 d
OsPrx100 d
I.3
OsPrx79 d
OsPrx35 d
OsPrx70 d
OsPrx72 b
OsPrx16 d
95
OsPrx15 b
OsPrx14 b
OsPrx13 b
OsPrx12 b
I.4
OsPrx127 c
OsPrx34 At
OsPrx128 c
OsPrx33 c
73*
I.5
OsPrx60 c
OsPrx59 c
OsPrx62 d
OsPrx61 d
100
97
OsPrx64 d
OsPrx63 d
II
99
OsPrx23 f
OsPrx21 d
OsPrx22 e
OsPrx20 f
OsPrx124 f
100
OsPrx37 c
OsPrx125 a
OsPrx39 At
III
OsPrx126 At
OsPrx38 At
89
OsPrx5 At
OsPrx54 e
OsPrx17 At
OsPrx3 At
IV.1
OsPrx118 At
OsPrx122 c
OsPrx77 At
OsPrx85 d
100
OsPrx84 d
OsPrx83 d
OsPrx82 d
OsPrx81 d
IV.2
95
OsPrx57 c
63
OsPrx136 At
OsPrx131 At
OsPrx41 At
OsPrx109 c
76
OsPrx113 At
53
OsPrx110 At
OsPrx114 At
OsPrx112 At
OsPrx111 At
IV.3
95
OsPrx115 c
OsPrx40 c
IV
OsPrx135 At
OsPrx130 At
OsPrx28 c
OsPrx30 f
OsPrx29 c
OsPrx26 c
OsPrx27 c
OsPrx2 c
OsPrx25 f
OsPrx87 c
OsPrx88 c
OsPrx86 c
IV.4
OsPrx117 e
OsPrx56 At
OsPrx94 e
53
OsPrx74 At
OsPrx80 d
OsPrx78 d
54
OsPrx134 d
OsPrx97 d
V.1
OsPrx103 At
OsPrx95 a
OsPrx96 a
OsPrx19 At
OsPrx75 At
V
OsPrx104 At
OsPrx49 e
OsPrx89 At
OsPrx24 At
OsPrx4 d
OsPrx138 At
Pseudo1 At
OsPrx32 c
OsPrx102 c
OsPrx123 c
OsPrx119 c
65
OsPrx91 k
OsPrx93 a
OsPrx92 a
OsPrx90 At
84
OsPrx58 At
OsPrx50 At
OsPrx98 a
OsPrx51 At
100
OsPrx55 e
OsPrx76 e
OsPrx31 e
VI
51
Pseudo4 k
Pseudo3 j
OsPrx18 a
OsPrx1 At
OsPrx137 At
95
100
OsPrx133 At
OsPrx132 At
VII
OsPrx52 h
OsPrx36 d
OsPrx129 At
64
OsPrx121 d
OsPrx120 d
VIII
OsPrx116 e
OsPrx73 d
Marchantia

Fig. 2. Phylogenetic tree of the rice peroxidases based on predicted protein sequences. The tree was constructed by the neighbour-joining method and rooted with the peroxidase sequence of *Marchantia polymorpha*. Values at nodes indicate bootstrap supports greater than 50%. All bootstraps were supported by an ML tree except those marked with an asterisk. All branches are drawn to scale and the scale bar represents 0.1 substitution per site. The letter following each peroxidase stands for its intron pattern (Fig. 7). Italicized peroxidases were used to build the phylogenetic tree shown in Fig. 3.

To provide another evidence of the independent and ancestral origin of OsPrx73 and OsPrx116, a ML tree was inferred from peroxidase cDNA sequences instead of protein sequences. The input data consisted of *OsPrx73*, *OsPrx 116*, and one expressed *OsPrx* (EST RGP database; Hiraga et al., 2000) of each subgroup, as well as expressed peroxidases from the liverwort *M. polymorpha*, the moss *Physcomitrella patens* (PPPrx) and the fern *Ceratopteris richardii* (CrPrx). The ML tree thus obtained (Fig. 3), again suggests that the closest rice homolog to *Marchantia* peroxidase is *OsPrx73*, although the bootstrap support for this relationship is not high. *Ceratopteris* and *Physcomitrella* did not have any close rice homolog, and the branching position of *Osprx116* was not supported enough by a strong bootstrap to be considered as a close rice peroxidase homolog to *Marchantia* peroxidase.

We also checked the codon usage for peroxidases in rice, *Marchantia*, *Physcomitrella* and *Ceratopteris*, by computing their cDNA sequences into the GCUA program (McInerney, 1998). Rice came out with a different codon usage as compared to the other plants (data not shown). Neither *OsPrx73* nor *OsPrx116* showed any significant codon usage deviation from the other *OsPrx*s. A closer look to the cDNA alignment used to build the tree of Fig. 3 showed that variation of sequence between *OsPrx73* and *Marchantia* peroxidase mainly affected the third position of codons.

We further combined OsPrx and AtPrx protein sequences in another phylogenetic tree (data not shown) in order to study the relations between the various peroxidases groups in the two plants. It appeared that rice subgroups I.1, I.2, IV.2 and V.1 had no closely related AtPrx proteins. Therefore, they may correspond to rice or Monocotyledon specific groups. Moreover, when we aligned all OsPrx groups with a peroxidase sequence of *Spirodela polyrrhiza* (NCBI Accession No. Z22920, Chaloupkova and Smart, 1994), a basal Monocotyledon (Judd et al., 2002), group V.1 came out as the one with the highest homology to this protein. In order to look

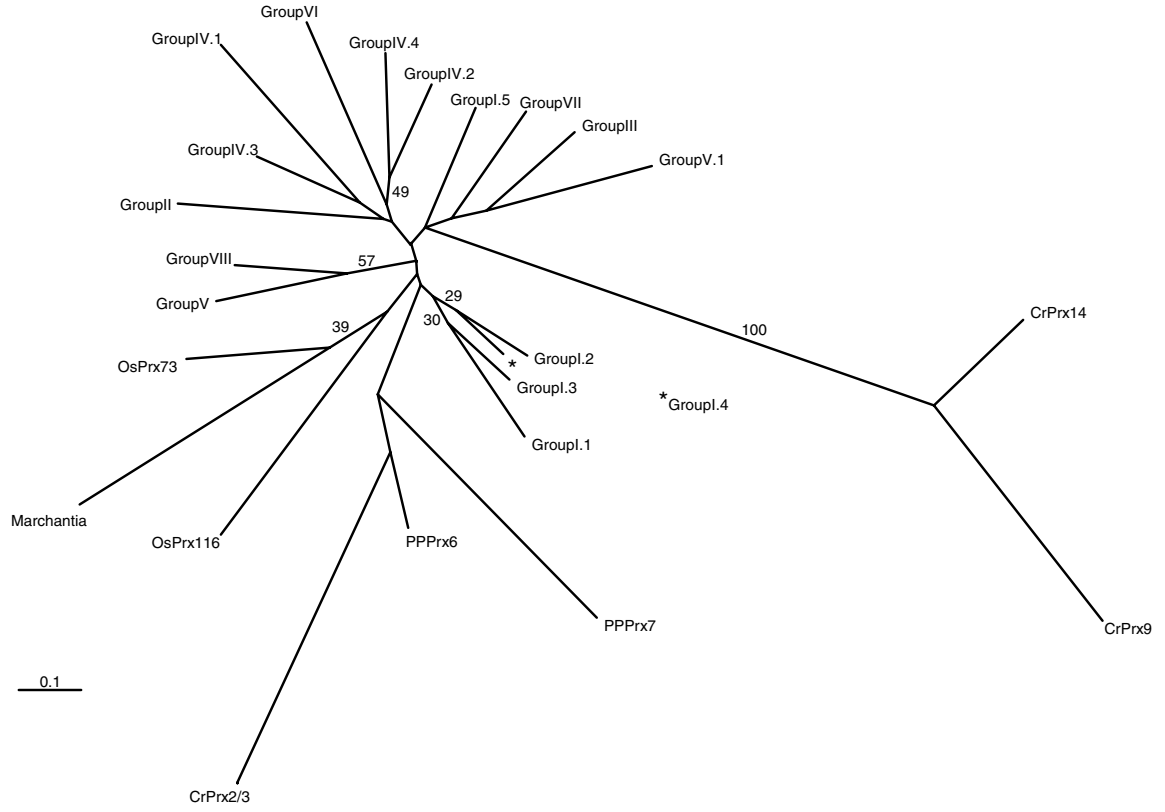


Fig. 3. Unrooted phylogenetic tree of one expressed member of each *Osprx* group and complete peroxidase sequences from the fern *C. richardii*, the liverwort *M. polymorpha* and the moss *P. patens*. The tree is based on cDNA sequences and was constructed by the maximum-likelihood method (see text). Values at nodes indicate bootstrap supports greater than 20%. All branches are drawn to scale and the scale bar represents 0.1 substitution per site.

for a possible rice or Monocotyledon specific motif in peroxidases, we performed a ClustalW alignment including four OsPrx protein sequences from group V.1, four other OsPrx sequences from various groups having AtPrx homologs, as well as four randomly chosen AtPrx and sequences from *Spirodela* and *Hordeum vulgare* (barley). We found a 16-amino acid sequence common to members of group V.I, situated between two highly conserved domains near the signal peptide (Fig. 4). Though less conserved, *Spirodela* sequence obtained a higher homology score with each group V.I peroxidase than with any other OsPrx or AtPrx sequence. The barley peroxidase included in the alignment did not show significant homology with group V.I peroxidases. Barley genome probably contains a large number of peroxidase genes. This means that more peroxidase genes must be identified in this species as well as in other Monocotyledons to confirm the existence of a Monocotyledon-specific motif in class III peroxidases. Finally, it should be mentioned that OsPrx73 and OsPrx116 form a small group of sequences, containing also AtPrx42 known to be constitutively expressed and to have no close paralog in *Arabidopsis* (Tognolli et al., 2002).

2.2. Duplication events among rice peroxidases

Peroxidase genes are not homogeneously distributed among the 12 chromosomes. There are several large clusters of closely homologous genes, most likely resulting from various duplication events (Fig. 5(a)). Six simple tandem duplications can be identified in rice. The most obvious case is *OsPrx100* and *OsPrx101* that are completely identical and probably result from a very recent tandem duplication. Rice genome also shows larger duplications involving 3–6 peroxidase genes. For example, the relationships linking the members of the cluster *OsPrx59* to *OsPrx64* (Fig. 2) strongly suggest that they result from five successive tandem duplications. Furthermore, the presence of introns in all the genes of the cluster excludes retrotransposition events, hence suggesting several rounds of crossing-overs in this region as source of the cluster. Interestingly, the same region of chromosome IV containing this cluster has also been sequenced in *O. sativa* ssp. *indica* c.v. Guangluai 4 (GLA4) in order to contribute to the ab initio annotation of *O. sativa* ssp. *japonica* c.v. Nipponbare (Feng et al., 2002). The BACs corresponding to this region contain *OsPrx59* to *OsPrx64* disposed in a colinear arrangement in GLA4 (respectively, in BACs OSJN00076 and AJ245900), excepted for *OsPrx62*, which was absent. Moreover, the distances between *OsPrx60* and *OsPrx61*, *OsPrx63* and *OsPrx64* as well as *OsPrx61* and *OsPrx63* were much larger in Nipponbare than in GLA4 cultivars. The two first discrepancies are due to large gag-pol polyprotein encoding sequence insertions in Nipponbare cultivar. The OsPrx tree (Fig. 2) reveals that *OsPrx62* is clearly older than *OsPrx61* and *OsPrx64*. These observations show that cluster *OsPrx59* to *OsPrx64* existed before the separation of *indica* and *japonica* subspecies and that, consequently, the difference concerning *OsPrx62* was due to a deletion of this peroxidase in GLA4, rather than to the formation of a new gene in Nipponbare. Another large cluster, *OsPrx42* to *OsPrx48*, is associated with many MITE ele-

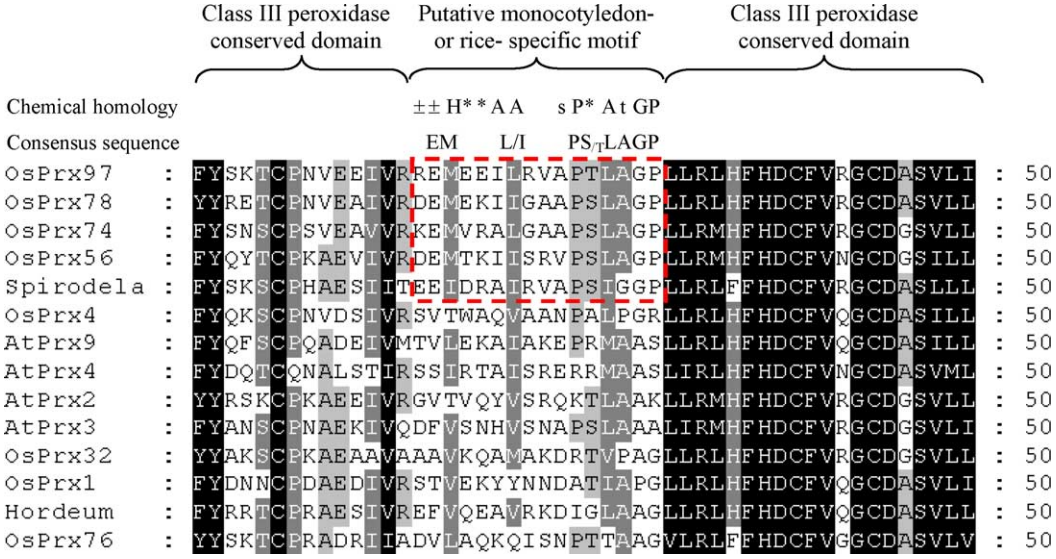


Fig. 4. ClustalW alignment of a peroxidase 50 amino acid sequences located close to the signal peptide sequence. OsPrx 56, 74, 78 and 97 belong to rice-specific subgroup V.1. AtPrx and the remaining OsPrx sequences were chosen randomly. Two typical class III peroxidase conserved domains are shown. The putative rice specific motif is surrounded by a dashed line. Consensus sequences are indicated for this latter motif. One is a chemical homology sequence, whose symbols represent the following categories of amino acids: ±, charged; ∗, polar; H, hydrophobic; A, aliphatic; s, small, t, tiny; P, proline; G, glycine. The second is an amino acid sequence consensus.

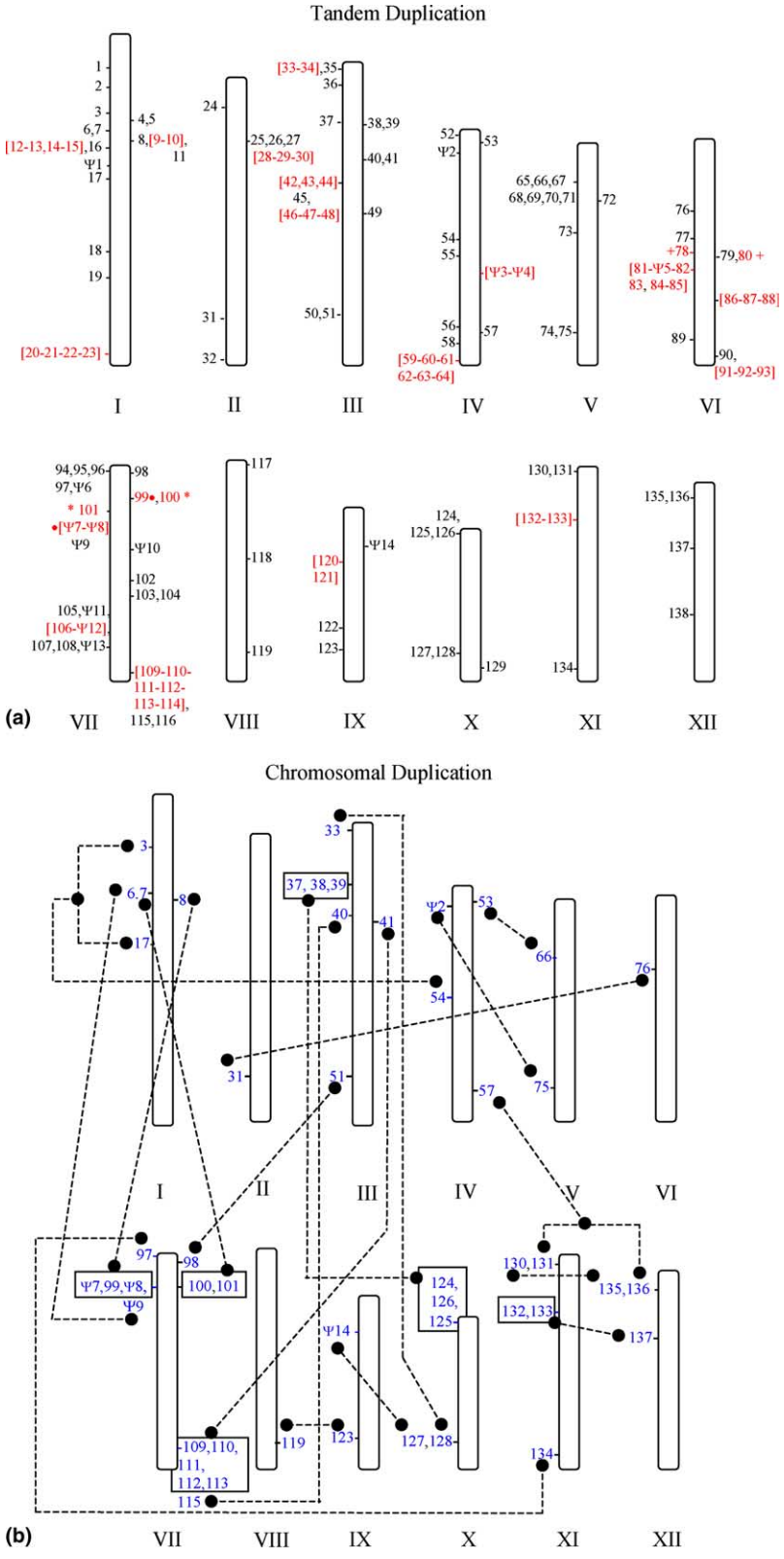


Fig. 5. Distribution of the peroxidase loci on the 12 rice chromosomes. (a) Position of the 138 genes and 14 pseudogenes. Bracketed peroxidase numbers in red, as well as similar symbols (+, ∗ or •) correspond to duplicated genes. (b) Position of the peroxidase genes involved in segmental duplications. Duplicated peroxidases or peroxidase clusters are linked by a dotted line.

ments, located between peroxidase sequences, but none of them was found within a coding region.

Fig. 6 shows a case of double segmental duplication in chromosome VII (Fig. 6). Three large repeated fragments (two of 44,000 bp and one of 32,000 bp) present in BACs AP005409 and AP005454 contain three peroxidase genes (*OsPrx99*, *OsPrx100* and *OsPrx101*) and two pseudogenes (*Ψ7* and *Ψ8*). Other undefined homologous sequences have been used to position correctly the repeated regions. These three segments are surrounded by several *gag-pol* sequences and *RIRE* orf elements.

Gene duplications often result in plants from the duplication of chromosomal segments. Seven chromosomal duplications involving peroxidase genes were found in *Arabidopsis* (Tognolli et al., 2002). The homologies observed in the OsPrx phylogenetic tree (Fig. 2) suggest that more than twenty inter-chromosomal duplications occurred in rice (Fig. 5(b)). Some chromosomes, such as II, VI and VIII, have only one gene exhibiting a close homology with a gene on another chromosome. This may indicate that they are more recent, and therefore have had fewer possibilities to share duplications with older chromosomes. Chromosomes XI and XII look closely related due to two duplication events involving *OsPrx130–OsPrx131* and *OsPrx135–OsPrx136*, as well as *OsPrx132–OsPrx133* and *OsPrx137*. Homologies were also found between chromosomes III and X. The comparison of entire BAC sequences in a "BLAST2 sequences" alignment (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html) revealed that the regions encompassing *Osprx33* to *OsPrx34* and *OsPrx37* to *OsPrx39* exhibited a very strong homology with two regions on chromosome X, respectively, around *OsPrx127* to *OsPrx128* and *OsPrx124* to *OsPrx126*. The phylogenetic tree shown in Fig. 2 does not allow to determine if chromosome X is younger than chromosome III. Similarities in segmental regions between chromosomes III and VII may further support the idea that chromosomes VII and X originated from chromosome III. Confirming this assertion of an ancestral chromosome containing the initial information, phytochrome family genes, which play a key role in plant photomorphogenesis (Fankhauser, 2001), were uniquely found on chromosome III.

2.3. Intronic structure of rice and Arabidopsis peroxidase genes

Forty eight of the 72 peroxidase encoding genes in *Arabidopsis* consist of 3 introns and 4 exons (Tognolli et al., 2002). This pattern, named hereafter *At*, is also well represented in rice (Fig. 7). *O. sativa* ssp. *japonica* has three main intronic patterns: *At* (27%), *c* (17%) and *d* (31%). Intron models *a* to *f* either lack one, two, or all three introns present in *At* model. The last five intron

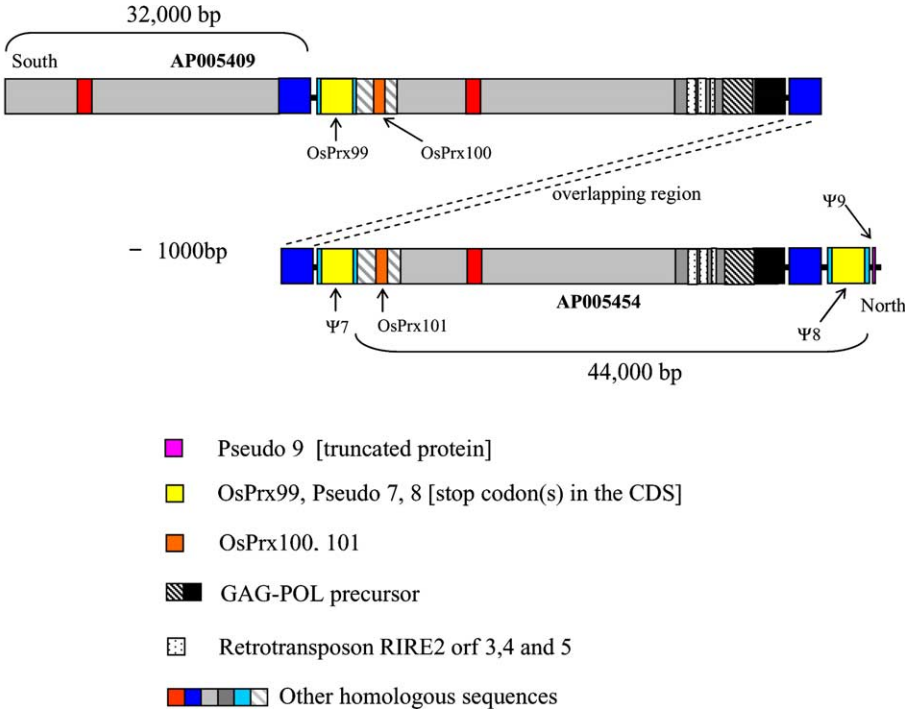


Fig. 6. Structure of AP005454 and AP005409, two contiguous BACs on chromosome VII with an overlapping region of about 2800 bp. Similar patterns denote homologous sequences. Lines represent unique sequences.

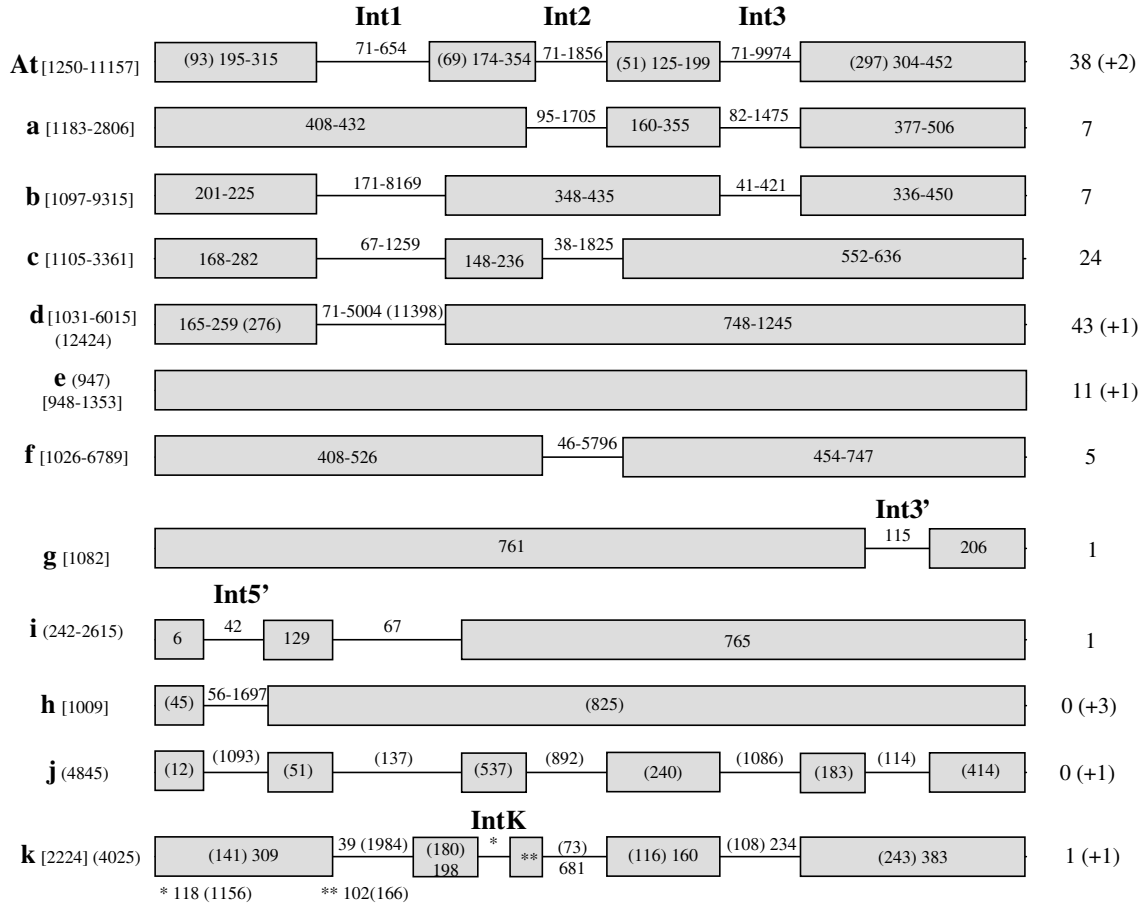


Fig. 7. Structure of rice peroxidase genes. Intron patterns have been identified according to the nomenclature used for *Arabidopsis* peroxidase genes (Tognolli et al., 2002). Grey boxes represent exons, and lines introns. Numbers above each intron/exon region or in front of each gene type indicate the range of the base pairs length. When out of range, values belonging to pseudogenes are written into brackets. Numbers of peroxidase genes in each family are indicated on the right, as well as numbers of pseudogenes in brackets.

patterns (*g* to *k*) are found only in few genes or in pseudogene sequences. They are characterized by different intron positions, that were named *Int5'*, *Int3'*s and *IntK*. In contrast to intron patterns *At* and *a* to *e*, patterns *f* to *k* are not present in *Arabidopsis thaliana*. Rice introns (38–9974 bp) exhibited a broader size range than *Arabidopsis* (71–2850 bp). Surprisingly, several genes were expressed, despite their long intronic sequences or their unusual intronic pattern (Table 1, RGP EST database). *OsPrx89*, *OsPrx124* genes and *Ψ12* were characterized by very large introns containing retrovirus-related *gag-pol* sequences. An EST was detected for *OsPrx89*, but not for *OsPrx124*.

The intron patterns are often conserved within duplicated genes, without conservation of the intronic sequence. The distribution of the intron patterns in the phylogenetic tree (Fig. 2) shows that gene type *e* (no introns) can generate only other genes without intron (type *e*), in line with the hypothesis of a natural intron deletion process described for catalase genes evolution (Frugoli et al., 1998; Iwamoto et al., 1998). Moreover, model *At* and *c* are frequently conserved after duplication, for example in group IV.

2.4. Origin and expansion of class III peroxidases

Class III peroxidases are present in all plants and form large multigenic families (Gaspar et al., 1982; Penel et al., 1992; Welinder, 1992a,b). The development of several genome and EST projects is continuously bringing new elements to confirm this fact. Typical peroxidase sequences were found in numerous Mono- and Eudicotyledonous genomes and in Gymnosperms (Fig. 8), with at least 60% of homology to *Arabidopsis* peroxidases. For example, an EST search in NCBI shows that many peroxidases are present, among others, in tomato (*Lycopersicon esculentum*), spinach (*Spinacia oleracea*), tobacco (*Nicotiana tabacum*) and two species of pine trees (*Pinus sylvestris*, *Pinus pinaster*). But the presence of peroxidases is not restricted to Gymno- and Angiosperms. They have been found as EST sequences in other groups. We identified eight peroxidase

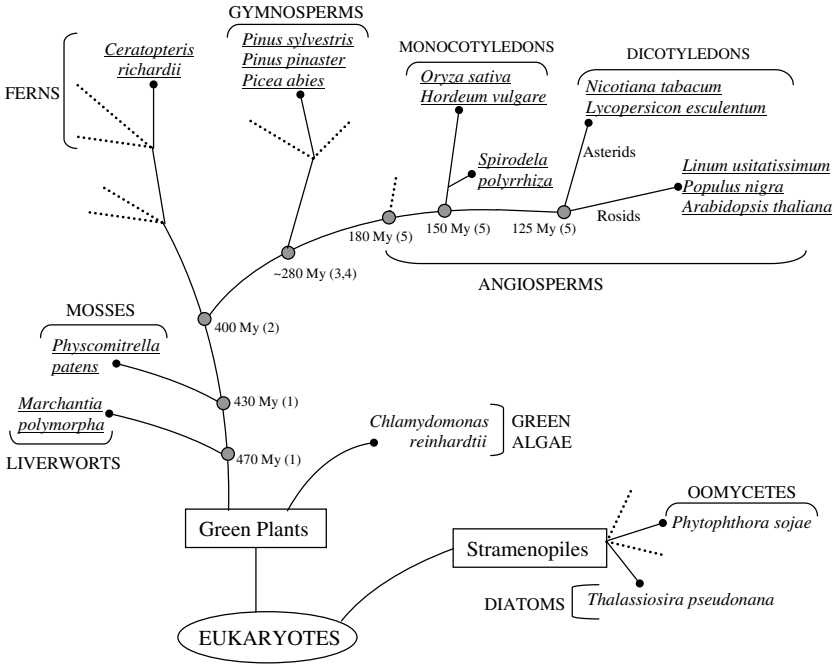


Fig. 8. Plant taxonomic tree. The branches are not drawn to scale. The underlined species contain at least one sequence coding for a class III peroxidase. Many more species were found in databases within the groups listed on the taxonomic tree, excepted for groups with one species shown. Dotted lines are used for groups in which no peroxidase was found, probably due to lack of data (at least in terrestrial green plants). (1) (Kenrick and Crane, 1997); (2) (Pryer et al., 2001); (3) (Savard et al., 1994); (4) (Schmidt and Schneider-Poetsch, 2002); (5) (Wikstrom et al., 2001).

sequences in *C. richardii* (fern) and one in *M. polymorpha* (liverwort) within small EST databases (3600 and 1400 EST, respectively, http://www.plantgdb.org). In *P. patens* (moss), we found 11 to 14 different putative non-overlapping peroxidase sequences in the PEP EST database (www.moss.leeds.ac.uk). Importantly, this EST database has been obtained from two different culture conditions, one of them being an auxin-rich environment. It is known that auxin increases the expression of peroxidase genes (Lee et al., 2002). Since several *Arabidopsis* peroxidase promoters contain auxin-inducible elements, as found through a PLACE search (http://www.dna.affrc.go.jp/htdocs/PLACE/; Penel et al., 2002), it can be hypothesized that the majority of *Physcomitrella* peroxidase transcripts were present in the cDNA library. On the other hand, the entirely sequenced genome of the green alga *Chlamydomonas reinhardtii* and the draft genome sequence of the stramenopiles *Thalassiosira pseudonana* and *Phytophthora sojae* do not contain class III peroxidase encoding genes nor any laccase homolog (secreted proteins), though both algae and stramenopiles genomes contain catalase (cytosolic localization) and cytochrome *C* peroxidase (mitochondrial) sequences. A further search within a cyanobacterium genome database containing sequence information on seven different cyanobacterium species (www.kazusa.or.jp/cyano/) did not retrieve any class III peroxidase sequence. None were found either in other prokaryotic organism through a TBLASTN search in the NCBI website (http://www.ncbi.nlm.nih.gov). Guaiacol oxidation in the presence of H_2O_2, which is a common method used to detect class III peroxidase activity (Greppin et al., 1986), was tested on *C. reinhardtii* and four other algae: *Zygnema* sp., *Staurastrum lacustris*, *Pediastrum boryanum* and *Chloromonas* sp. *Zygnema* and *S. lacustris* belong to *Streptophyta*, which is the Division that contains all the terrestrial plants. The two last ones are members of the *Chlorophyceae* class, like *Chlamydomonas*, which is issued from the Division of *Chlorophyta*. None of these algae showed any guaiacol oxidation in the presence of H_2O_2 (data not shown).

3. Discussion

3.1. Evolution and duplication events

The main goal of this work was to compare the peroxidase multigenic family in rice and in *Arabidopsis*. It appeared that rice genome contains twice as many peroxidase genes as *Arabidopsis*. Other reports demonstrated so far a one to one ratio for several multigenic families (Iwamoto et al., 1998; Baumberger et al., 2003; Baxter et al., 2003; Lijavetzky et al., 2003). In parallel to our search for peroxidases, we also looked for the phytochrome kinase substrate (PKS) family (data not shown). It was found to be composed of 4 genes in

Arabidopsis and only 3 paralogs in *O. sativa* ssp. *japonica*. However, a recent report revealed that expansins are also encoded by twice the number of genes in rice than in *Arabidopsis* (Li et al., 2003). The total number of genes in rice is still hypothetical. Predictions range from 32,000 to 62,000, that is from a 1:1 to a 2:1 ratio compared to *Arabidopsis* (Delseny, 2003; Schoof and Karlowski, 2003). However, the current opinion tends to bet on a number of genes closer to 32,000 than to 62,000 (Delseny, 2003), because the latter number was based on chromosome 1 gene density (Sasaki et al., 2002), that turned out to be higher than the average of the other chromosomes (Chen et al., 2002). A 1:1 ratio in the number of genes within a family would then be more likely expected between rice and *Arabidopsis*. Interestingly, expansin and peroxidase gene families that show both a 2:1 ratio encode cell wall associated proteins. They may have followed a similar gene expansion during evolution in connection with their role in cell elongation and cell wall mechanical properties.

Monocotyledons (rice) branched off from Eudicotyledons (*Arabidopsis*) some 150 MY ago, based on a NPRS molecular clock approach (Wikstrom et al., 2001), although fossil-based methods tend to place this divergence more recently, around 120 MY ago (Magallon et al., 1999). It can be assumed that, at their origin, the first Monocotyledons and Eudicotyledons had a similar number of peroxidase genes. The evolution leading to today's *Arabidopsis* and rice yielded a quite different number of peroxidase genes in the two plants. This difference could result from a better conservation of duplicated genes in rice, maybe because they met some needs in herbaceous Monocotyledonous plants, while they were less useful in Eudicotyledons. Additionally, a polyploidy event in rice may explain the twofold increase in the number of peroxidase genes. This difference also indicates that the estimation of a duplication rate of one gene per 100 million years in eukaryotes proposed by Lynch and Conery (2000) seems a little low in view of our findings.

Rice is under mankind selection since 9000 years (Khush, 1997). This could have favored the conservation of some peroxidase encoding genes. Human influence may also be responsible for differences in peroxidases numbers between the two subspecies of rice, *O. sativa indica* (GLA4 c.v.) and *japonica* (Nipponbare c.v.), that diverged from a common ancestor only one MY ago (Khush, 1997). In contrast to Angiosperms, mosses, exemplified by *P. patens*, contain apparently a rather small number of peroxidase genes. This difference could be explained by the greater level of organ and tissue differentiation in Angiosperms. The presence in these plants of molecules polymerized by peroxidases such as lignin and suberin and the fact that mosses generally live near the soil, in a rather humid environment, while Angiosperms became adapted to a wider range of conditions and reached larger sizes could also play a role. All these considerations are based on the assumption that the rate of gene duplication was the same in the various phyla. Other events such as MITE-mediated duplication or still undefined mechanisms could have also played some role in the differential enlargement of peroxidase multigenic family among the various phyla.

3.2. Implication of retrotransposons in peroxidase gene duplication

MITE elements are supposed to have been generated from autonomous MITE precursors through a two-step process (Feschotte et al., 2002): every time that a precursor excises itself, it creates double-stranded breaks that are repaired by an internally-deleted copy of the transposon, which becomes non-autonomous due to loss of its transposase-encoding region. This deleterious mechanism is named abortive repair. This non-autonomous transposon generates a large number of the so-called MITE elements, probably through a "cut and paste" mechanism, using the same transposase that recognizes specific repeats on the MITE sequence. The "cut and paste" process generates double-stranded breaks, which are believed to create recombination regions on chromosomes (Shalev and Levy, 1997). This series of events may be responsible for the formation of MITE-containing gene clusters such as *OsPrx42* to *OsPrx48* through numerous recombination rounds. Such multiple gene duplications may then confer selective advantages to the plant and the resulting clusters would hence be retained. For instance, peroxidases play a role in plant resistance against pathogens (Moerschbacher, 1992). This could be a possible function for the cluster *OsPrx42* to *OsPrx48*, in which recombinations helped plant to widen its pathogen resistance range. Interestingly, the members of this cluster do not have any close homologs in *Arabidopsis*. They may have evolved to adapt the defense of rice against new pathogens different from the *Arabidopsis* pathogens. On the other hand, the presence of retrotransposable elements such as *RIRE* and *gag-pol* sequences in a cluster of peroxidases could argue in favor of duplication processes correlated with retrotransposition events.

3.3. OsPrx introns

Introns in rice peroxidase genes show large size variability (Fig. 7) and, in a few cases, they can reach a considerable length. Rice splicing machinery has the capacity to deal with these large introns, some of them containing even internal retroviral-related polyprotein coding sequences. We may then assume that rice introns can play a role of viral genomic insertions "buffers", protecting the plant against gene inactivation pheno

mena. However, this defense system against insertions is apparently not always efficient, since truncated proteins may be produced or tissue-specific expression may be altered by the presence of very large introns (Casacuberta and Santiago, 2003).

Intron distribution within rice peroxidase genes is quite variable. The evolution does not seem to have preserved a major intron organization pattern, even between paralogous genes. As the *At* intronic model is found in both rice and *Arabidopsis* in a significant proportion, it may well reflect the ancestral intronic model. Loss of introns would then have occurred in rice to create most of the other intronic models observed, excepted in rare cases such as models *g* to *k* (Fig. 7), in which new introns appeared at unusual positions. The insertion of new introns in models *i* and *j* seems to have been deleterious, since these models are found only in pseudogenes. This hypothesis of the loss of intron during the evolution would be in line with previous studies on catalase genes in *Poaceae* (Iwamoto et al., 1998), but definitely needs more sequencing data of primitive organisms to be confirmed. The putative ancient origin of model *At* and the hypothesis of an evolutionary intron loss converge to the conclusion that *At* would be the ancestor model in Angiosperms. In this respect, a particularly interesting plant would be *Amborella trichopoda*, supposedly the sister group of all Angiosperms (Qiu et al., 1999), whose lineage appeared around 180 MY ago (Wikstrom et al., 2001) and whose EST sequences listed in NCBI do not contain any peroxidase so far. Several key questions arise from these observations, such as why not only one intronic pattern has been preserved in rice during evolution and how organisms managed to reorganize their gene sequences. One justification could be the transcriptional control and the RNA stabilization by intronic sequences (Gonzalez et al., 2002).

3.4. Searching for the ancestral peroxidase

Despite a high duplication rate of the encoding genes, the crucial peroxidase specific motifs and residues were conserved throughout the evolution from liverworts and mosses to rice and *Arabidopsis*. This is the case of two conserved domains, $LxRLxxHDC_2xxxGC_3DxS$ and $LxxxHxxGxxxC_6$ (Fig. 1), necessary for the interaction with the heme moiety through histidine residues. Disulfide bridges involving cysteines C_2, C_3 and C_6 have also important functions in catalytic cleavage of hydrogen peroxide (Poulos and Kraut, 1980). The two histidine residues are the signature of all the hemo-peroxidases and are conserved in every organism (Zamocky et al., 2000). A third conserved domain, C_4xxVSC_5xD, has an unknown function. The other cysteine residues involved in disulfide bridges (C_1, C_7 and C_8) are also preserved. These conserved motifs were very useful not only to search for class III peroxidases, but also to correct rice predicted OsPrx sequences. Variable regions of the protein can also be very informative. A 16-amino acid domain, found in rice subgroup V.1, showed significant homology with a domain of *Spirodela* peroxidase. As *Spirodela* is considered as a basal Monocotyledon, it may have preserved ancient sequences that should be found in a major number of Monocotyledons. Therefore, this motif might also be considered as Monocotyledon-specific. The importance of this putative sequence would be either structural or functional.

The fact that class III peroxidases have been found in all the major divisions of land plants but not in green algae denotes the functional importance of these proteins and points to the existence of a common peroxidase ancestor 470 million years ago, before or during the emergence of liverworts (Kenrick and Crane, 1997), the very first land plants (Wellman et al., 2003). Nevertheless, it will be necessary to wait for sequencing of other green algae and stramenopiles species to definitively confirm the exclusive presence of peroxidases in terrestrial green plants.

Distance (Fig. 2) and maximum-likelihood (Fig. 3) phylogenetic analyses converge towards an interesting point: OsPrx73 is the closest peroxidase rice homolog to *Marchantia* peroxidase. Although this conclusion is supported by a relatively low bootstrap value in Fig. 3, the congruence observed between both analyses is strengthened by the fact that two very different data sources, such as protein (Fig. 2) and cDNA (Fig. 3), tend to the same conclusion. Moreover, codon usage analysis showed that *OsPrx73* has not preserved the codon usage of *Marchantia* peroxidase, but has instead followed the general OsPrx usage. Therefore, evolution has successfully maintained OsPrx73 amino acid sequence despite the modification of its corresponding DNA sequence, thanks to codon third position flexibility. The role of OsPrx73 in plant is unknown, but it could be assumed that it has been preserved because of an important function possibly existing also in *Marchantia*. Why is OsPrx73 so different from the other OsPrx, and why is it so similar to a *Marchantia* peroxidase? Convergent evolution may be an explanation, though very improbable, as rice and liverworts do not live in the same biotopes, and are morphologically very dissimilar. A second possibility is that OsPrx73 represents the closest rice sequence to the class III peroxidase ancestor, which probably appeared around 470 MY ago. Interestingly, *AtPrx42* is also not a member of any *Arabidopsis* peroxidase subgroup (Tognolli et al., 2002) and is homolog to *OsPrx73* and *OsPrx116*. Moreover, it is strongly expressed in all organs of *Arabidopsis*, which may indicate an ancient regulation system.

The apparition and the diversification of class III peroxidases may be related to the adaptation of plant to the life out of water. Two major challenges may have

played a role in the birth of the first class III peroxidase. A radical change in oxygen environment has probably had a strong impact in chemical reaction rates involving oxygen, and thereby in the various concentrations of oxygen containing molecules, such as the peroxidase substrate H_2O_2. As well, there may be a correlation between the appearance of class III peroxidases and the emergence of cell wall structures adapted to terrestrial life. In line with our results on peroxidases, we also searched for laccase homologs in plants and other kingdoms, and only found these enzymes in land plants. Interestingly, laccases are also involved in cell wall construction (O'Malley et al., 1993; Ranocha et al., 1999). Their absence in other organisms than terrestrial plants suggests that both proteins may have appeared at the same time during evolution in order to build novel walls of terrestrial plants.

Class III peroxidases are probably absent from prokaryotes, which contain class I peroxidases. As both classes have strikingly conserved a general similar structure as well as two critical histidine residues involved in the interaction with heme (Welinder, 1992a,b; Zamocky et al., 2000), class III peroxidases may well have originated from class I peroxidases. The first land plants in possession of modified class I peroxidases able to deal with changes such as increased H_2O_2 availability and cell wall structure constraints were then probably favored in their progress on land colonization.

Present-day Angiosperm peroxidases catalyze many essential reactions in cell wall, including cross-linking between structural polysaccharides (Fry, 1986), oxidative lignin or suberin polymerization (Ros Barceló, 1997), or formation of activated oxygen species that are involved in the modulation of the mechanical properties of cell wall (Rodriguez et al., 2002; Liszkay et al., 2003). The appearance of these diversified functions during land plant evolution most likely came from neofunctionalization of duplicated peroxidase genes. The formation of new organs, such as stems, leaves and flowers, during green plants evolution also favored the diversification of new peroxidases necessary for the specific needs of these new organs. Many plant peroxidases have been shown to be involved in the responses of plants to a wide range of biotic and abiotic stresses (Castillo, 1992). In this respect, stress factors also have varied during evolution with the colonization of new habitats and the constant emergence of novel plant pathogens. The appearance of insects about 420 MY ago, for instance, has been a major threat to plants, which, as suggested by Gaunt and Miles (Gaunt and Miles, 2002), induced ferns to build higher stems to escape herbivory. Ultimately, the development of agriculture subjected plants, like rice, to a strong selective pressure. In order to face these challenges, subfunctionalization and neofunctionalization were critical survival events that allowed plants to evolve.

4. Experimental

4.1. Rice data mining

Arabidopsis class III peroxidases protein sequences have been used as a starting point for the rice data mining. Each of the 73 different amino acid sequences of *Arabidopsis* class III peroxidases was submitted against the whole genomic rice database with a tblastn search on the Rice Genome Project (RGP) website (http://rgp.dna.affrc.go.jp/), that reports genomic sequence information on the Nipponbare cultivar of *O. sativa japonica*. When the BAC sequence was not annotated or contained putative peroxidase sequences, BACs and sequences were analyzed for gene presence with different programs such as FGenesh (http://www.softberry.com/berry.phtml) and GenScan (http://genes.mit.edu/GENSCAN.html). The corresponding coding sequence (CDS) was translated with the "translate" tool on Expasy (http://us.expasy.org/tools/dna.html) and controlled for specific peroxidases motifs. BACs were positioned on the chromosomes following the indications of the Gramene website (http://www.gramene.org/).

In order to find peroxidases in other organisms, various OsPrx protein sequences were used as input sequences in TBLASTN searches within different databases. Peroxidases were sought in the NCBI website (http://www.ncbi.nih.gov/BLAST) for *M. polymorpha*, in the PEP EST database (www.moss.leeds.ac.uk) for *P. patens*, in the PlantGDB database (http://www.plantgdb.org/) for *C. richardii* and in the DOE Joint Genome Institute (JGI) website for *C. reinhardtii*, *T. pseudonana* and *P. sojae* (http://genome.jgi-psf.org/). A similar search was performed for cytochrome *C* peroxidases laccases, catalases, phytochromes, and phytochrome kinase substrates (PKS). *Arabidopsis* sequences drawn from NCBI were used as input sequences.

4.2. Comprehensive phylogenetic analysis of peroxidase sequences identified in O. sativa

All 138 peroxidase sequences present in rice and the single peroxidase sequence of *M. polymorpha*, used as an outgroup, were aligned using Clustal W (Thompson et al., 1994). The alignment was further inspected and visually adjusted. Only unambiguously aligned amino acid positions were retained for analyses, leading to an alignment consisting of 145 sequences and 156 sites. Distance and maximum likelihood analyses were performed with the help of the PHYLIP 3.6a3 package (Felsenstein, 1993). The distance tree was constructed with the NEIGHBOR option under the JTT substitution frequency matrix, and 1000 bootstrap replicates were carried out for this method with the SEQBOOT option. The maximum likelihood tree was inferred with the ProML option, under the JTT substitution

frequency matrix, by using the global rearrangement option. Maximum likelihood bootstrap analysis was not performed due to computational limitations. The Njplot software was used to visualize phylogenetic trees.

Putative Monocotyledon-specific domains were located visually and then realigned with ClustalW. One 16 aa putative sequence gave a significant similarity score (computed with GeneDoc software) for OsPrx group V.I and *Spirodela*. On a 23 aa sequence encompassing the putative 16 amino acids (between highly conserved amino acid sequences IVR and LLRL), all comparisons within group V.I and *Spirodela* were scored below 290 (sum of pairs method) and with a minimal homology of 69%, whereas other OsPrx groups and AtPrx compared with group V.1 and *Spirodela* were scored above and with lower homologies.

4.3. Phylogenetic analysis of expressed peroxidase sequences present in different plant lineages

We also examined the phylogenetic relationships among expressed peroxidases present in rice and in several distantly related plants. We analyzed a restricted data set consisting of 18 expressed (RGP website EST search) *O. sativa* peroxidase sequences representing the global phylogeny, along with 2 peroxidase sequences of the moss *P. patens*, 3 of the fern *C. richardii* and the single sequence of *M. polymorpha*. OsPrx peroxidases chosen for this alignment are shown in italic in the phylogenetic tree of Fig. 2. All 24 amino acid sequences were aligned as explained above, resulting in 80 unambiguously aligned amino acids positions. Because of this limited sequence length and consequently limited phylogenetic signal, we used nucleotide sequences for phylogenetic reconstruction. All 24 cDNA sequences were aligned manually against the 80 unambiguously aligned amino acids, resulting in 240 aligned nucleotides. The limited number of analyzed sequences allowed us to perform maximum likelihood (ML) analyses using PAUP* (Swofford, 1998), and to carry out a bootstrap analysis using 100 replicates. ML trees were inferred from all codon positions. Because preliminary distance analyses suggested that some expressed peroxidases (e.g., CrPrx 14 and 9) are fast evolving sequences, we chose to minimize possible long branch attraction artifacts by taking into account rate variations among sites. A gamma distribution with 8 categories plus invariable sites was therefore added to the model of sequence evolution. All necessary parameters (substitution matrix, nucleotide frequencies, among-site rate variations) were estimated via ML from a starting distance tree. Heuristic ML tree searches were then performed starting from this distance tree under the General-Time-Reversible model of sequence evolution since it allows all six pairs of substitutions to have different rates.

4.4. Codon usage analysis

Codon usage was studied by inserting cDNA sequences of *M. polymorpha*, *P. patens* and *C. richardii* peroxidases, as well as all rice peroxidases into the GCUA program (McInerney, 1998). Codon usage was then calculated for each amino acid and visually analysed.

Acknowledgements

We thank Dr. P. Cuenoud for helpful discussion and critical reading of the manuscript. F.P. is supported by the Swiss National Science Foundation (31-068003.02).

References

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815.

Baumberger, N., Doesseger, B., et al., 2003. Whole-genome comparison of leucine-rich repeat extensins in Arabidopsis and rice. A conserved family of cell wall proteins form a vegetative and a reproductive clade. Plant Physiol. 131, 1313–1326.

Baxter, I., Tchieu, J., et al., 2003. Genomic comparison of P-type ATPase ion pumps in Arabidopsis and rice. Plant Physiol. 132, 618–628.

Blanc, G., Barakat, A., et al., 2000. Extensive duplication and reshuffling in the Arabidopsis genome. Plant Cell 12, 1093–1101.

Blee, K.A., Jupe, S.C., et al., 2001. Molecular identification and expression of the peroxidase responsible for the oxidative burst in French bean (*Phaseolus vulgaris* L.) and related members of the gene family. Plant Mol. Biol. 47, 607–620.

Bolwell, G.P., Davies, D.R., et al., 1998. Comparative biochemistry of the oxidative burst produced by rose and french bean cells reveals two distinct mechanisms. Plant Physiol. 116, 1379–1385.

Casacuberta, J.M., Santiago, N., 2003. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311, 1–11.

Castillo, F.J., 1992. Peroxidases and stress. In: Penel, C., Gaspar, T., Greppin, H. (Eds.), Plant Peroxidases. University of Geneva, Switzerland.

Chaloupkova, K., Smart, C.C., 1994. The abscisic acid induction of a novel peroxidase is antagonized by cytokinin in *Spirodela polyrrhiza* L. Plant Physiol. 105, 497–507.

Chen, M., Presting, G., et al., 2002. An integrated physical and genetic map of the rice genome. Plant Cell 14, 537–545.

Chen, S.X., Schopfer, P., 1999. Hydroxyl-radical production in physiological reactions. A novel function of peroxidase. Eur. J. Biochem. 260, 726–735.

Delseny, M., 2003. Towards an accurate sequence of the rice genome. Curr. Opin. Plant Biol. 6, 101–105.

Fankhauser, C., 2001. The phytochromes, a family of red/far-red absorbing photoreceptors. J. Biol. Chem. 276, 11453–11456.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package). Seattle, University of Washington.

Feng, Q., Zhang, Y., et al., 2002. Sequence and analysis of rice chromosome 4. Nature 420, 316–320.

Feschotte, C., Jiang, N., et al., 2002. Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. 3, 329–341.

Frugoli, J.A., McPeek, M.A., et al., 1998. Intron loss and gain during evolution of the catalase gene family in angiosperms. Genetics 149, 355–365.

Fry, S.C., 1986. Polymer-bound phenols as natural substrates of peroxidases. In: Greppin, H., Penel, C., Gaspar, T. (Eds.), Molecular and Physiological Aspects of Plant Peroxidases. University of Geneva, Switzerland, pp. 169–182.

Gajhede, M., Schuller, D.J., et al., 1997. Crystal structure of horseradish peroxidase C at 2.15 Å resolution. Nat. Struct. Biol. 4, 1032–1038.

Gaspar, T., Penel, C., et al., 1982. Peroxidases 1970–1980. University of Geneva, Switzerland.

Gaunt, M.W., Miles, M.A., 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol. Biol. Evol. 19, 748–761.

Gebhardt, C., Walkemeier, B., et al., 2003. Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. Plant J. 34, 529–541.

Gillham, N.W., 1994. Organelle Genes and Genomes. Oxford University Press, Oxford, GB.

Gonzalez, J.M., Penzes, Z., et al., 2002. Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by insertion of an intron. J. Virol. 76, 4655–4661.

Graur, D., Li, W.H., 1999. Fundamentals of Molecular Evolution. Sinauer Associates, Sunderland, MA.

Greppin, H., Penel, C., et al., 1986. Molecular and Physiological Aspects of Plant Peroxidases. University of Geneva, Switzerland.

Hiraga, S., Sasaki, K., et al., 2001. A large family of class III plant peroxidases. Plant Cell. Physiol. 42, 462–468.

Hiraga, S., Yamamoto, K., et al., 2000. Diverse expression profiles of 21 rice peroxidase genes. FEBS Lett. 471, 245–250.

Iwamoto, M., Maekawa, M., et al., 1998. Evolutionary relationship of plant catalase genes inferred from exon-intron structures: isozyme divergence after the separation of monocots and dicots. Theor. Appl. Genet. 97, 9–19.

Joo, J.H., Bae, Y.S., et al., 2001. Role of auxin-induced reactive oxygen species in root gravitropism. Plant Physiol. 126, 1055–1060.

Judd, W.S., Campbell, C.S., et al., 2002. Plant Systematics: A Phylogenetic Approach. Sinauer Associates, Sunderland, MA.

Kenrick, P., Crane, P.R., 1997. The origin and early evolution of plants on land. Nature 389, 33–39.

Khush, G.S., 1997. Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. 35, 25–34.

Lee, D.J., Kim, S.S., et al., 2002. The regulation of Korean radish cationic peroxidase promoter by a low ratio of cytokinin to auxin. Plant Sci. 162, 345–353.

Li, Y., Jones, L., et al., 2003. Expansins and cell growth. Curr. Opin. Plant Biol. 6, 603–610.

Lijavetzky, D., Carbonero, P., et al., 2003. Genome-wide comparative phylogenetic analysis of the rice and *Arabidopsis* Dof gene families. BMC Evol. Biol. 3, 17.

Liszkay, A., Kenk, B., et al., 2003. Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth. Planta 217, 658–667.

Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.

Magallon, S., Crane, P.R., et al., 1999. Phylogenetic pattern, diversity, and diversification of eudicots. Ann. Missouri Bot. Garden 86, 297–372.

Mao, L., Wood, T.C., et al., 2000. Rice transposable elements: a survey of 73,000 sequence-tagged-connectors. Genome Res. 10, 982–990.

McInerney, J.O., 1998. GCUA: general codon usage analysis. Bioinformatics 14, 372–373.

Moerschbacher, B.M., 1992. Plant peroxidases: involvement in response to pathogens. In: Penel, C., Gaspar, T., Greppin, H. (Eds.), Plant Peroxidases. University of Geneva, Switzerland.

O'Malley, D.M., Whetten, R., et al., 1993. The role of laccase in lignification. Plant J. 4, 751–757.

Penel, C., Dunand, C., et al., 2002. In search of function for 73 *Arabidopsis* peroxidases genes. In: Acosta, M., Rodriguez-Lopez, J.N., Pedreno, M.A. (Eds.), Plant Peroxidase Biochemistry and Physiology. University of Murcia, pp. 135–141.

Penel, C., Gaspar, T., et al., 1992. Plant Peroxidases. University of Geneva, Switzerland.

Poulos, T.L., Kraut, J., 1980. The stereochemistry of peroxidase catalysis. J. Biol. Chem. 255, 8199–8205.

Pryer, K.M., Schneider, H., et al., 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409, 618–622.

Qiu, Y.L., Lee, J., et al., 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402, 404–407.

Ranocha, P., McDougall, G., et al., 1999. Biochemical characterization, molecular cloning and expression of laccases – a divergent gene family – in poplar. Eur. J. Biochem. 259, 485–495.

Rodriguez, A.A., Grunberg, K.A., et al., 2002. Reactive oxygen species in the elongation zone of maize leaves are necessary for leaf extension. Plant Physiol. 129, 1627–1632.

Ros Barceló, A., 1997. Lignification in plant cell walls. Intern. Rev. Cytol. 176, 87–132.

Sasaki, T., Matsumoto, T., et al., 2002. The genome sequence and structure of rice chromosome 1. Nature 420, 312–316.

Savard, L., Li, P., et al., 1994. Chloroplast and nuclear gene sequences indicate late Pennsylvanian time for the last common ancestor of extant seed plants. Proc. Natl. Acad. Sci. USA 91, 5163–5167.

Schmidt, M., Schneider-Poetsch, H.A., 2002. The evolution of gymnosperms redrawn by phytochrome genes: the Gnetatae appear at the base of the gymnosperms. J. Mol. Evol. 54, 715–724.

Schoof, H., Karlowski, W.M., 2003. Comparison of rice and *Arabidopsis* annotation. Curr. Opin. Plant Biol. 6, 106–112.

Shalev, G., Levy, A.A., 1997. The maize transposable element Ac induces recombination between the donor site and an homologous ectopic sequence. Genetics 146, 1143–1151.

Skrabanek, L., Wolfe, K.H., 1998. Eukaryote genome duplication – where's the evidence? Curr. Opin. Genet. Dev. 8, 694–700.

Swofford, D.L., 1998. PAUP*: phylogenetic analyses under parsimony (*and other methods). Sinauer Associates, Sunderland, MA.

Thompson, J.D., Higgins, D.G., et al., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Tognolli, M., Penel, C., et al., 2002. Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288, 129–138.

Ware, D., Jaiswal, P., et al., 2002. Gramene: a resource for comparative grass genomics. Nucleic Acids Res. 30, 103–105.

Welinder, K.G., 1992a. Plant peroxidases: structure–function relationships. In: Penel, C., Gaspar, T., Greppin, H. (Eds.), Plant Peroxidases. University of Geneva, Switzerland.

Welinder, K.G., 1992b. Superfamily of plant, fungal and bacterial peroxidases. Curr. Opin. Struct. Biol. 2, 388–393.

Welinder, K.G., Justesen, A.F., et al., 2002. Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. Eur. J. Biochem. 269, 6063–6081.

Wellman, C.H., Osterloff, P.L., et al., 2003. Fragments of the earliest land plants. Nature 425, 282–285.

Wikstrom, N., Savolainen, V., et al., 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond. B: Biol. Sci. 268, 2211–2220.

Zamocky, M., Janecek, S., et al., 2000. Common phylogeny of catalase-peroxidases and ascorbate peroxidases. Gene 256, 169–182.

Zhang, J., 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18, 292–298.

Table 1 (supplementary material) Rice class III peroxidases.

The **accession numbers** refer to the genomic sequence (BAC). **Chr:** chromosome on which the BAC is located; **Int:** type of intron pattern (see Figure 7, **nd**: not classified as the sequence is incomplete); **cM:** position of the BAC on the chromosome in centimorgans. **Arabidopsis homolog**: closest related *Arabidopsis* peroxidase (AtPrx#) orthologous gene, **Homol:** percentage of homology with the Arabidopsis ortholog. **EST:** expression of the *OsPrxs* according to the last indexed RGP EST database from RGP [(-) means no expression, (+) 1 to 3 clones found, (++) 4 to 10 clones found]. **Gene position:** relative gene position in bp into the BACs. The peroxidase number follows the chromosomal position from OsPrx1 at the top of chromosome I to OsPrx138 at the bottom of chromosome XII.

	Accession number	Chr	Int	cM	Arabidopsis homolog	Homol.	EST	Gene position		Accession number	Chr	Int	cM	Arabidopsis homolog	Homol.	EST	Gene position
OsPrx1	AP003215	I	At	16.4-19.9	At2g41480-AtPrx25	78%	-	4692-6472	OsPrx73	AC145396	V	d	49.4	At3g17070-AtPrx29	52%	-	90230-91662
OsPrx2	AP001551	I	c	28.9	At5g05340-AtPrx52	63%	-	113253-114357	OsPrx74	AC113332	V	At	103	At1g05260-AtPrx3	67%	+	39801-41111
OsPrx3	AP001366	I	At	45.4-46.3	At5g66390-AtPrx72	81%	++	138626-140416	OsPrx75	AC113332	V	At	~~	At1g05260-AtPrx3	63%	+	43873-45206
OsPrx4	AP001073	I	d	49.3	At2g39040-AtPrx24	56%	+	115763-117548	OsPrx76	AP004995	VI	e	40.2	At5g47000-AtPrx65	61%	-	95557-98540
OsPrx5	AP001081	I	At	49.6	At1g44970-AtPrx9	81%	+	138626-140416	OsPrx77	AP003044	VI	At	54.1	At1g68850-AtPrx11	78%	++	65481-67381
OsPrx6	AP002482	I	d	55.7	At4g11290-AtPrx39	57%	-	174174-180188	OsPrx78	AP006164	VI	d	62.9	At1g05260-AtPrx3	67%	-	30481-31741
OsPrx7	AP002482	I	b	~~	At1g05260-AtPrx3	48%	-	184097-185193	OsPrx79	AP003528	VI	d	65.8	At5g15180-AtPrx56	54%	-	111127-112371
OsPrx8	AP002820	I	d	58.1	At4g36430-AtPrx49	54%	-	22516-24020	OsPrx80	AP003577	VI	d	65.8	At1g05260-AtPrx3	66%	-	156912-158061
OsPrx9	AP002820	I	c	~~	At4g33420-AtPrx47	51%	-	34029-35137	OsPrx81	AP003684	VI	d	67.4	At3g50990-AtPrx36	60%	-	59244-60351
OsPrx10	AP002820	I	d	~~	At4g33420-AtPrx47	52%	-	41309-42513	Pseudo5	AP003684	VI	e	67.4	At1g44970-AtPrx9	52%	-	66074-67021
OsPrx11	AP002820	I	d	~~	At5g14130-AtPrx55	48%	+	69782-71036	OsPrx82	AP003684	VI	d	~~	At5g58390-AtPrx67	58%	+	73082-74191
OsPrx12	AP004358	I	b	66.2	At1g05260-AtPrx3	55%	-	21626-25553	OsPrx83	AP003684	VI	d	67.4	At3g50990-AtPrx36	58%	+	126904-128079
OsPrx13	AP004358	I	b	~~	At5g64120-AtPrx71	55%	-	31220-40534	OsPrx84	AP003684	VI	d	~~	At5g05340-AtPrx52	62%	-	133154-134240
OsPrx14	AP004358	I	b	~~	At1g05260-AtPrx3	60%	+	72991-78610	OsPrx85	AP003684	VI	d	~~	At5g05340-AtPrx52	62%	-	138360-139411
OsPrx15	AP004358	I	b	~~	At5g66390-AtPrx72	54%	+	85953-91274	OsPrx86	AP004731	VI	c	72.6	At5g05340-AtPrx52	79%	+	76343-77530
OsPrx16	AP004358	I	d	~~	At1g05260-AtPrx3	62%	++	99827-103601	OsPrx87	AP004731	VI	c	~~	At5g05340-AtPrx52	73%	-	91435-92604
Pseudo1	AP002968	I	At	73.4	At2g39040-AtPrx24	61%	-	92693-96379	OsPrx88	AP004731	VI	c	72.6	At5g05340-AtPrx52	73%	-	105615-106753
OsPrx17	AP004233	I	At	78.8-83.7	At5g66390-AtPrx72	83%	+	56112-60226	OsPrx89	AP004797	VI	At	110.6	At1g05260-AtPrx3	69%	+	54224-65381
OsPrx18	AP003410	I	a	123.2	At4g33870-AtPrx48	60%	-	74021-75309	OsPrx90	AP003771	VI	At	113.4-115.6	At2g18980-AtPrx16	67%	+	101729-104419
OsPrx19	AP003345	I	At	136	At5g15180-AtPrx56	66%	-	52178-53608	OsPrx91	AP003771	VI	k	~~	At4g30170-AtPrx45	51%	-	119108-121331
OsPrx20	AP003263	I	f	176.3	At1g71695-AtPrx12	70%	+	81318-84064	OsPrx92	AP003771	VI	a	~~	At2g18980-AtPrx16	67%	-	128995-130077
OsPrx21	AP003263	I	d	~~	At1g71695-AtPrx12	69%	-	88285-89516	OsPrx93	AP003771	VI	a	~~	At2g18980-AtPrx16	66%	-	134058-135546
OsPrx22	AP003263	I	e	~~	At1g71695-AtPrx12	66%	-	93420-94490	OsPrx94	AP004673	VII	e	0.8	At1g05260-AtPrx3	64%	-	3817-4866
OsPrx23	AP003263	I	f	~~	At1g71695-AtPrx12	73%	-	97697-99155	OsPrx95	AP004673	VII	a	~~	At5g15180-AtPrx56	49%	+	10787-13442
OsPrx24	AP004126	II	At	15.0	At4g11290-AtPrx39	60%	-	96867-99427	OsPrx96	AP004673	VII	a	~~	At1g05260-AtPrx3	63%	-	15467-16649
OsPrx25	AP003991	II	f	36.8	At5g05340-AtPrx52	54%	-	42481-43506	OsPrx97	AP005187	VII	d	0.8	At1g05260-AtPrx3	67%	-	69862-70968
OsPrx26	AP003991	II	c	~~	At5g05340-AtPrx52	68%	+	51299-52558	Pseudo6	AP005187	VII	At	~~	At1g05240-At1g05250A	71%	-	73379-74809
OsPrx27	AP003991	II	c	36.8	At5g05340-AtPrx52	70%	+	60426-62246	OsPrx98	AP005486	VII	a	3.9-4.4	At4g37530-AtPrx51	78%	-	36004-38644
OsPrx28	AP005613	II	c	37.0-37.4	At5g05340-AtPrx52	72%	-	76998-78201	OsPrx99	AP005409	VII	d	11.0-24.2	At4g36430-AtPrx49	56%	-	41564-44285
OsPrx29	AP005613	II	c	~~	At5g05340-AtPrx52	66%	-	88090-89275	OsPrx100	AP005409	VII	d	~~	At1g05260-AtPrx3	52%	-	38142-39272
OsPrx30	AP005613	II	f	~~	At5g05340-AtPrx52	73%	-	98378-99436	Pseudo7	AP005454	VII	i	24.2	At4g33420-AtPrx47	51%	-	3935-6548

OsPrx31	AP004815	II	e	131.0-134.5	At5g47000-AtPrx65	70%	-	55923-57275
OsPrx32	AP005290	II	c	157.9	At5g42180-AtPrx64	67%	-	21421-22570
OsPrx33	AC105363	III	c	3.5	At1g05260-AtPrx3	61%	+	55986-59346
OsPrx34	AC105363	III	At	~~	At1g05260-AtPrx3	67%	-	62012-63472
OsPrx35	AC105363	III	d	~~	At1g05260-AtPrx3	60%	-	74187-75217
OsPrx36	AC137635	III	d	14.4	At4g26010-AtPrx44	61%	+	132485-134048
OsPrx37	AC103891	III	c	34.8-36.1	At5g06730-AtPrx54	67%	-	4328-5515
OsPrx38	AC137696	III	At	36.1	At5g19890-AtPrx59	73%	-	94465-96145
OsPrx39	AC137696	III	At	~~	At5g06720-AtPrx53	76%	-	97723-99727
OsPrx40	AC131175	III	c	57-59	At5g05340-AtPrx52	72%	-	83034-84215
OsPrx41	AC131175	III	At	56-57	At5g05340-AtPrx52	74%	+	85606-90621
OsPrx42	AC082644	III	g	65.4-66.2	At4g11290-AtPrx39	54%	++	23726-24807
OsPrx43	AC082644	III	d	~~	At4g11290-AtPrx39	54%	++	32979-34060
OsPrx44	AC082644	III	d	~~	At4g11290-AtPrx39	56%	+	42143-43903
OsPrx45	AC082644	III	d	65.4-66.2	At5g58390-AtPrx67	57%	-	57230-58320
OsPrx46	AC082644	III	d	~~	At4g11290-AtPrx39	57%	+	68094-69660
OsPrx47	AC082644	III	d	~~	At4g33420-AtPrx47	58%	-	82402-83777
OsPrx48	AC082644	III	d	~~	At4g11290-AtPrx39	60%	-	92252-93387
OsPrx49	AC084767	III	At	86	At5g15180-AtPrx56	66%	+	55156-60291
OsPrx50	AC090713	III	At	141.6	At4g37530-AtPrx51	70%	+	30073-31324
OsPrx51	AC090713	III	At	~~	At3g49960-AtPrx35	78%	+	34240-36539
OsPrx52	OSJN00193	IV	h	4	At4g26010-AtPrx44	56%	-	50608-51616
OsPrx53	OSJN00254	IV	d	7.9	At5g05340-AtPrx52	43%	-	94017-97134
Pseudo2	OSJN00103	IV	i	10.1	At2g41480-AtPrx25	64%	-	60756-60997
OsPrx54	OSJN00233	IV	e	61	At5g66390-AtPrx72	79%	-	8768-9805
OsPrx55	OSJN00104	IV	e	70.5	At5g47000-AtPrx65	69%	-	48885-49892
Pseudo3	OSJN00025	IV	j	76.8	At2g24800-AtPrx18	61%	-	16508-19504
Pseudo4	H0711G06	IV	k	76.8	At2g24800-AtPrx18	60%	-	129823-131864
OsPrx56	OSJN00186	IV	At	109.9	At3g01190-AtPrx27	66%	-	25530-27447
OsPrx57	OSJN00105	IV	c	114	At5g05340-AtPrx52	70%	-	96501-97867
OsPrx58	OSJN00090	IV	At	114.3-120.3	At2g18980-AtPrx16	68%	-	110310-116099
OsPrx59	OSJN00076	IV	c	128.3	At1g71695-AtPrx12	57%	-	18662-19951
OsPrx60	OSJN00076	IV	c	~~	At1g71695-AtPrx12	65%	+	23426-24646
OsPrx61	OSJN00076	IV	d	~~	At1g71695-AtPrx12	59%	++	38313-39490
OsPrx62	OSJN00076	IV	d	~~	At1g71695-AtPrx12	58%	-	46409-47594
OsPrx63	OSJN00076	IV	d	~~	At1g71695-AtPrx12	53%	-	53845-54942
OsPrx64	OSJN00076	IV	d	~~	At1g71695-AtPrx12	54%	+	76550-77640
OsPrx65	AC087552	V	d	20.5-21.5	At4g33420-AtPrx47	54%	+	41108-45334
OsPrx66	AC087552	V	d	~~	At5g05340-AtPrx52	49%	+	58689-60613
OsPrx67	AC087552	V	d	~~	At4g33420-AtPrx47	52%	+	75884-77070
OsPrx68	AC087552	V	d	~~	At5g05340-AtPrx52	57%	+	81398-82744
OsPrx69	AC087552	V	d	~~	At5g05340-AtPrx52	60%	-	93962-95069
OsPrx70	AC087552	V	d	~~	At1g05260-AtPrx3	62%	-	105807-107162
OsPrx71	AC087552	V	b	~~	At1g05260-AtPrx3	58%	+	112778-114513
OsPrx72	AC137614	V	b	30.7	At1g05260-AtPrx3	52%	+	3016-4831
OsPrx101	AP005454	VII	d	~~	At1g05260-AtPrx3	52%	-	8841-9971
Pseudo8	AP005454	VII	i	~~	At4g33420-AtPrx47	51%	-	48774-51383
Pseudo9	AP005454	VII	nd	~~	At1g77100-AtPrx13	73%	-	52719-52958
Pseudo10	AP005774	VII	nd	45.5	At3g21770-AtPrx30	64%	-	88570-88916
OsPrx102	AP005177	VII	c	62.4-67.0	At1g30870-AtPrx7	61%	-	44433-45684
OsPrx103	AP004674	VII	At	73.2	At1g05260-AtPrx3	63%	-	82202-83887
OsPrx104	AP004674	VII	At	~~	At4g11290-AtPrx39	62%	-	102564-104416
OsPrx105	AP005292	VII	e	99.3	At4g11290-AtPrx39	57%	-	49003-50025
Pseudo11	AP005292	VII	nd	~~	At5g17820-AtPrx57	53%	-	52098-52370
OsPrx106	AP005292	VII	f	~~	At4g11290-AtPrx39	59%	-	53406-54979
Pseudo12	AP005292	VII	d	~~	At4g11290-AtPrx39	59%	-	59128-71551
OsPrx107	AP005292	VII	d	~~	At4g11290-AtPrx39	58%	+	95375-96441
OsPrx108	AP005292	VII	e	99.3	At4g11290-AtPrx39	52%	-	107256-108290
Pseudo13	AP005292	VII	nd	~~	At4g26010-AtPrx44	46%	-	108926-109683
OsPrx109	AP003817	VII	c	115.5	At5g05340-AtPrx52	74%	++	23516-24848
OsPrx110	AP003817	VII	At	~~	At5g05340-AtPrx52	72%	-	32866-34446
OsPrx111	AP003817	VII	At	~~	At5g05340-AtPrx52	72%	-	37104-38353
OsPrx112	AP003757	VII	At	115.8	At5g05340-AtPrx52	74%	-	69-1322
OsPrx113	AP003757	VII	At	~~	At5g05340-AtPrx52	70%	-	5792-7105
OsPrx114	AP003757	VII	At	~~	At5g05340-AtPrx52	71%	-	10105-11651
OsPrx115	AP003757	VII	c	~~	At5g05340-AtPrx52	64%	-	12412-14162
OsPrx116	AP005199	VII	e	118.6	At2g37130-AtPrx21	74%	+	21048-22040
OsPrx117	AP003910	VIII	e	3.6	At4g33420-AtPrx47	64%	-	142342-143289
OsPrx118	AP004041	VIII	At	54.3	At4g16270-AtPrx40	56%	-	63992-65385
OsPrx119	AP005529	VIII	c	109.3	At4g37520-AtPrx50	71%	+	52975-55451
Pseudo14	AP006235	IX	nd	20.7	At5g66390-AtPrx72	67%	-	107271-107411
OsPrx120	AP005879	IX	d	26.7-30.6	At5g17820-AtPrx57	60%	-	4508-5584
OsPrx121	AP005879	IX	d	~~	At5g17820-AtPrx57	61%	-	9979-11048
OsPrx122	AP005555	IX	c	65.1	At2g22420-AtPrx17	83%	+	74845-77589
OsPrx123	AP005706	IX	c	77-77.2	At5g14130-AtPrx55	67%	+	168538-170797
OsPrx124	AC037197	X	f	1.1	At5g06730-AtPrx54	60%	-	65866-72654
OsPrx125	AC068654	X	a	1.1-1.9	At5g06720-AtPrx53	71%	++	102618-105423
OsPrx126	AC068654	X	At	~~	At5g19890-AtPrx59	67%	-	123965-125457
OsPrx127	AC074196	X	c	68-69	At3g21770-AtPrx30	60%	-	16680-17806
OsPrx128	AC074196	X	c	~~	At5g05340-AtPrx52	56%	+	21265-23058
OsPrx129	AC087599	X	At	73.7-83	At5g17820-AtPrx57	58%	-	67502-69808
OsPrx130	CNS08CE2	XI	At	2.5-2.8	At5g05340-AtPrx52	81%	+	158568-160762
OsPrx131	CNS08CE2	XI	At	~~	At5g05340-AtPrx52	76%	-	171100-172627
OsPrx132	AF172282	XI	At	27.8-30.8	At4g25980-AtPrx43	70%	-	51800-53237
OsPrx133	AC123522	XI	At	30	At4g25980-AtPrx43	70%	-	167846-169283
OsPrx134	AC125783	XI	d	111.5-111.8	At1g05260-AtPrx3	66%	+	16476-17999
OsPrx135	CNS08CE0	XII	At	7.4-9.4	At5g05340-AtPrx52	80%	+	23942-26139
OsPrx136	CNS08CE0	XII	At	~~	At5g05340-AtPrx52	79%	-	33998-35527
OsPrx137	CNS08CBN	XII	At	37	At4g25980-AtPrx43	64%	+	98562-99834
OsPrx138	CNS07YPC	XII	At	71-72	At2g39040-AtPrx24	55%	+	11669-20808

C. Back to the roots: a search for the first class III peroxidases

In the previous article, we mentioned that class III peroxidases might have originated in the very first land plants, a conclusion shared by another independent study on peroxidase evolution (Duroux and Welinder, 2003). This hypothesis was supported by two observations. First, BLAST searches in the NCBI dbEST database were unable to find any protein similar to a class III peroxidase in algae. Secondly, I had performed guaiacol oxidase assays on extracts of three green algae: *Chlamydomonas reinhardtii*, *Chloromonas sp*. and *Pediastrum boryanum*. No activity was detected for any of the three organisms (Fig. 3-1). Further guaiacol experiments were performed on additional species: the results obtained supported the idea that the first peroxidases appeared in the order of Charales. I hence undertook a search for class III peroxidase in two species of *Chara*, as described hereafter.

C.1 Preliminary work: guaiacol oxidase activities of algae

The guaiacol activity test is used as a standard and specific way of detecting class III peroxidases. Other heme-containing peroxidases are able to oxidise guaiacol, though with a much lower efficiency. Generally, class III peroxidases in plants are more numerous and mask the activity of the other peroxidases: in presence of H_2O_2, activity is detected within seconds. Non-specific activity can be detected by omitting H_2O_2 in the reaction medium: it is mainly due to laccases and other polyphenol oxidases (Badiani et al., 1983; Shin and Lee, 2000; Beloqui et al., 2006). Caution must also be taken when no guaiacol activity is detected: phenolic substances and ascorbate can indeed mask peroxidase activity (Castillo et al., 1984). Ascorbate reduces guaiacol radicals as soon as they are formed and, until ascorbate is consumed, a “lag phase” is established where tetraguaiacol (470nm) is not formed. Concerning phenolic compounds, the mechanism is more complex: each phenolic compound has a different influence on peroxidase activity, leading mainly to kinetic changes rather than complete inhibition of guaiacol oxidase activity (Pickering et al., 1973; Penel, 1976; Volpert et al., 1995). In the case of the three algae tested, absence of activity could mean total absence of heme peroxidases. However, EST libraries of green algae (*Chlamydomonas reinhardtii, Dunaliella salina, Acetabularia acetabulum, Scenedesmus*

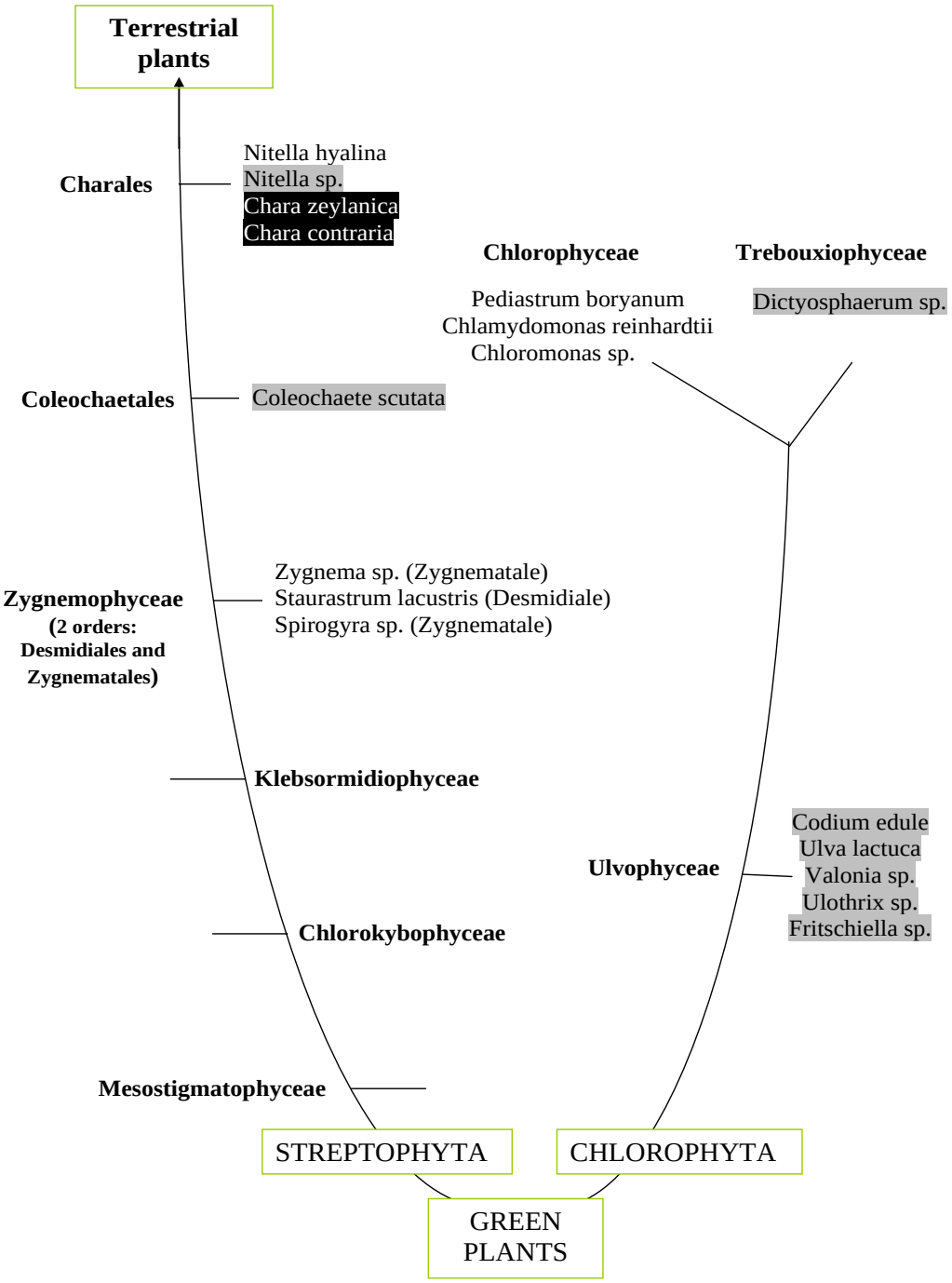


Figure 3-1

Taxonomical distribution of green algae (McCourt, 1995; McCourt et al., 2004)

Guaiacol oxidase activity was detected in the species shaded in grey (Siegel and Siegel, 1970; Siegel and Siegel, 1986) or in black (our laboratory).

All the species where guaiacol oxidase was not detected (no shading) were analysed in our laboratory.

obliquus, Prototheca wickerhamii) show that they encode ascorbate peroxidase proteins (APX). Guaiacol oxidation may have been prevented, at least in the algal species studied, by contaminant molecules, as mentioned previously.

In contrast to my observations, earlier studies using various known class III peroxidase substrates, including guaiacol, have concluded to presence of class III peroxidases in five Ulvophyceae (Fig. 3-1) and one Trebouxiophycea (Siegel and Siegel, 1970, 1986). As there are probably no class III peroxidases in Chlorophyta, according to EST libraries in NCBI, the activity observed was due to another heme-containing peroxidase (e.g. ascorbate peroxidase). Why did I not observe guaiacol activity in the algae I analysed? The difference between my study and Siegel's was that I tested Chlorophyceae, whereas Siegel and Siegel studied Trebouxiophyceae and Ulvophyceae. Buffers used and extraction procedure were very similar between my studies and Siegel and Siegel's: some substance contained in Chlorophyceae may have hence prevented APX to oxidise guaiacol.

Despite Siegel and Siegel's observation, it seems that class III peroxidases are likely to be absent from the Chlorophyta branch of green algae. I hence concentrated my research in identifying whether class III peroxidases appeared in algal forms of Streptophyta or in the first land plants. Again, I performed guaiacol activity assays on available algae:

- three Zygnemophyceae: *Zygnema sp., Staurastrum lacustris, Spirogyra sp.*
- one Coleochaetale: *Coleochaete scutata*
- two Charales: *Chara zeylanica*, *Nitella hyalina*

Whereas no activity could be detected in Zygnemophyceae, *Coleochaete* and *Chara* oxidised guaiacol in a H_2O_2-dependent fashion. *Coleochaete* activity was very low (Table 3-1) and even disappeared when the *Coleochaete* extract was re-assayed after being stored overnight at 4°C. *Chara* activity was much higher, but still rather moderate compared to the land plants *Conocephalum sp.* (liverwort), *Selaginella sp.* (Lycophyta), *Arabidopsis* or the aquatic plant *Elodea densa* (monocotyledon). Following these experiments, it seems more appropriate to report activities per mass of total protein rather than fresh plant weight. Two independent *Elodea* extracts gave indeed very different values when they were normalised to fresh weight. These results suggest that the total number of peroxidases in algae is lower than in land plants. It is also possible that primitive forms of peroxidases, such as the one(s) in *Coleochaete*, have a less efficient activity. Surprisingly, *Nitella hyalina* showed no activity, although another undefined species (*Nitella sp.*) has already been described to oxidise guaiacol and other known peroxidase substrates (Siegel and Siegel, 1970). Siegel and Siegel

further found (Siegel and Siegel, 1986) that guaiacol oxidation activities (per mg fresh weigth) of *Elodea densa* and *Chara sp.* were quite similar, and that *Chara* activity was ten to twenty-fold higher than *Marchantia polymorpha* and *Selaginella rupestris*, in contrast with my findings. Their experiment was conducted in a very similar way as mine: the major difference was a 20-fold higher concentration of H_2O_2, which alone could not explain this discrepancy. Culture conditions, organs (Siegel's *Elodea*: leafy shoots and branches), and development stage are possible explanations to this striking difference. This series of experiments demonstrates that guaiacol activity assays should be taken as indicative.

Despite being contradictory to Siegel and Siegel's observations, it is tempting to speculate that class III peroxidases may be present in the evolved algal forms that gave rise to land plants, and that the number of peroxidase encoding genes has then increased in land plants, explaining the higher guaiacol activity observed. My next goal was hence to isolate an algal peroxidase sequence, either genomic or peptidic.

Species	OD_{470}/min	OD_{470}/min*mg
Coleochaete scutata	0.004	0.001
Chara zeylanica	0.097	0.014
Conocephalum sp.	3.750	0.120
Selaginella sp.	6.383	0.030
Arabidopsis thaliana	0.554	0.082
Elodea densa	1.643 (1.758)	0.120 (0.032)

Table 3-1

Guaiacol oxidase activity results (average of three independent measurements)

First column: values are normalised to protein concentration (20µg of total protein)

Second column: values per mg of fresh weight

Extraction material: whole organism (algae, *Conocephalum* and *Selaginella*) or leaves (*Arabidopsis* and *Elodea*)

Elodea: the values in parentheses refer to a second independent extraction.

C.1.1 Materials and methods

C.1.1.a Plant material

Coleochaete scutata, Nitella hyalina and *Chara zeylanica* were purchased from Carolina Biological Supply Company (Burlington, NC, USA). *Chara contraria*, *Zygnema sp.*, *Staurastrum lacustris*, *Spirogyra sp.* and *Selaginella sp.* were from our laboratory (Jordan Voluntaru and Yann Emmenegger). *Chara contraria* was collected in the Lake of Geneva (Port Choiseul, Versoix) and the species was identified by Dr. Dominique Auderset-Joye (Laboratoire d'écologie et biologie aquatique, Geneva). *Conocephalum sp.* was obtained from Dr. Michelle Price (Conservatoire et Jardin Botaniques, Geneva); *Elodea densa* was a gift from Marion Minouflet (Insitut Forel, Versoix, Switzerland).

C.1.1.b Peroxidase extracts and guaiacol assays

Fresh material was collected and dounce homogenised in HEPES buffer 20mM (pH 7.0, EGTA 2mM), with a ratio of 2µl solution per mg fresh weight. After centrifuging at 12'000g for 20 minutes and at 4°C (Biofuge "Pico", Heraeus), the supernatant was collected and directly tested for guaiacol oxidase activity, as well as for protein quantification.

Peroxidase extract was assayed at room temperature in 1ml of "guaiacol mix" (50mM phosphate buffer pH 6.1, guaiacol 0.125%, i.e. ~11mM) supplemented with 125µl H_2O_2 11mM. Optical density (470nm) was measured every minute for at least 3 time points.

Protein concentration was measured at 595nm with Bio-Rad reagent (ref. 500-0205, Bio-Rad Laboratories GmbH, Germany) and calculated by comparison with a bovine serum albumin standard titration.

C.2 Isolation of a peroxidase from Chara zeylanica *and* Chara contraria

C.2.1 DNA approach

The strategy used for cloning a peroxidase gene from *Chara* was, in a first attempt, to extract genomic DNA and amplify it with primers "Cons" and "Sca". These two primers hybridise quite well, *in silico*, with DNA sequences of all plant peroxidases, on two highly conserved regions. The amplified bands were then directly ligated into the vector pGEM T, cloned and sequenced. This first approach gave rather disappointing results: usually, no

sequence could be amplified by PCR, and if any band appeared on agarose gels, it subsequently proved to be a non-specific amplification product.

In order to avoid non-specific amplification and to prevent competition of Cons and Sca primers with other *Chara* genes, the strategy illustrated in figure 3-2 was applied. Although digestion with HaeIII restriction enzyme was readily achieved, PCR reactions always generated smears of DNA on agarose gels, independently of the PCR conditions. Parallel experiments were performed with *Coleochaete* DNA, without any better result. Why did this approach fail? The most probable answer is that primers Cons and Sca are too unspecific. A second hypothesis is that the sequence of *Chara* may be quite different from land plants. Codon usage is probably not a major problem, since the sequence targeted by Cons and Sca is conserved among all land plants, which all also have different codon usages. Alternatively, the DNA extraction procedure could have been revised. In this respect, Marion Koch, from our laboratory, tried a DNA extraction with a plant-specific DNA purification kit (QiaGen), but she was not able to obtain a satisfactory DNA yield, nor any signal following PCR amplification with Cons and Sca primers.

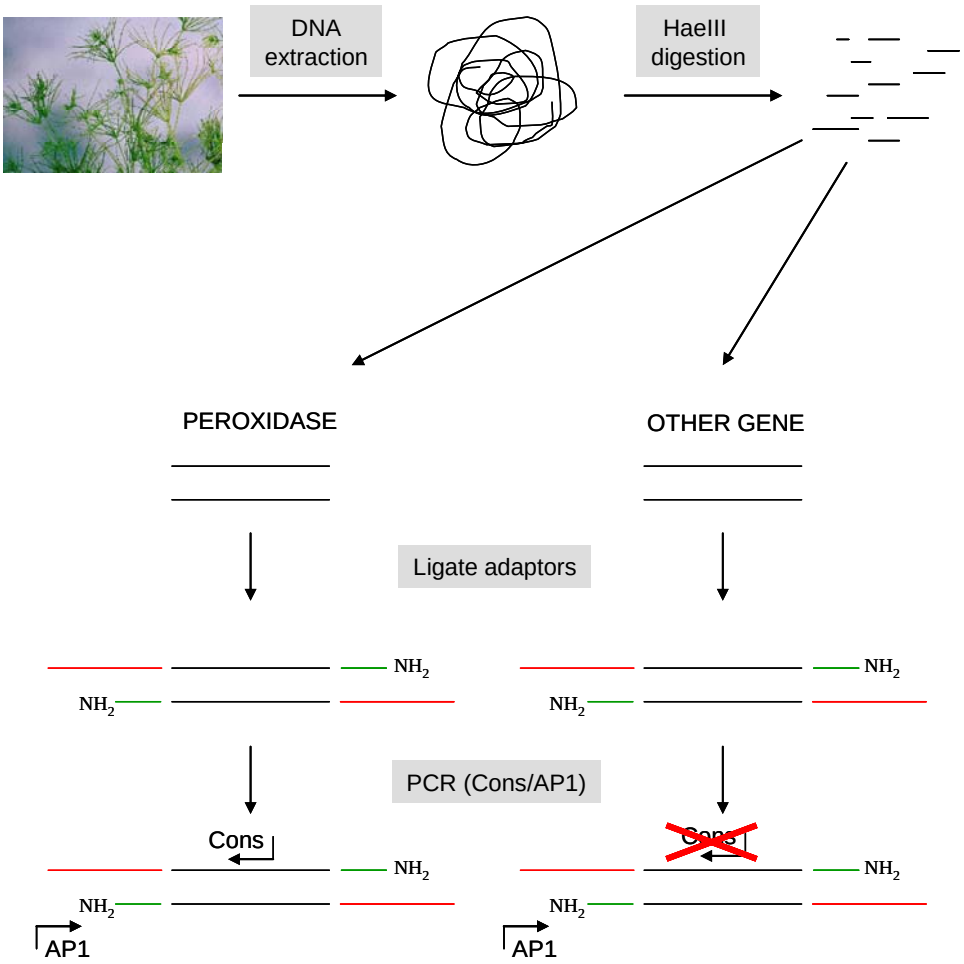


Figure 3-2 (previous page)

Strategy employed to specifically amplify peroxidase genes (modified from Marathon™ cDNA amplification protocol, Clontech laboratories Inc.).

Genomic DNA is first extracted and purified. Secondly, HaeIII digestion (blunt ends) is performed overnight. Adaptor DNA molecules are then ligated to the sheared DNA. The adaptor short strand last nucleotide is covalently linked to an amino group that prevents DNA polymerase filling the short strand using the long strand as a template. The primer AP1 is specific for the complementary strand of the long (red line) adapter. If a peroxidase gene is present, then Cons or Sca primers will amplify the upper strand, thus generating the lower strand that will be hybridised by primer AP1. Adaptor molecules increase the chances that at least one primer (AP1) will not randomly hybridise to other genes than peroxidases.

C.2.2 Protein approach

Since I was not able to amplify a peroxidase gene in *Chara* (and *Coleochaete*), I chose to verify the presence of peroxidase proteins and eventually purify, for the first time, an algal class III peroxidase. This approach requires a fair amount of fresh material, which was not commercially available for *Coleochaete* or *Chara*. Concerning this last alga, however, several species are growing in the Lake of Geneva (Auderset Joye et al., 2002). A colony of *Chara contraria* was found close to the Versoix harbour Port-Choiseul (Fig.3-3), at 50-100cm depth, thus allowing collecting enough material.

Figure 3-3

Localisation of the *Chara* collection site (provided by Google Earth v 3.0)

Chara contraria was collected in the circled area. Scale bar: 50m.

Preliminary work with *Chara zeylanica* already indicated that several isoforms (both acidic and basic), and not one, as initially thought, are present in this alga (Fig. 3-4).

Figure 3-4

IEF gel of *Chara zeylanica* stained with *o*-dianisidine.

pI 6 (red) to 9 (black). Guaiacol activity loaded: 0.34 OD_{470}/min.

The first tests performed with Claude (*Chara contraria*) showed that if protein extracts were loaded on a negatively charged sepharose column (heparin) at slightly alkaline pH, peroxidase activity was not retained on the column. With a positively charged column (diethylaminoethyl, DEAE) and the same pH however, peroxidase activity was retained on the column and could not be eluted, even with NaCl concentrations of 2M. Another positively charged column (Q-Sepharose) was then successfully used, with alkaline (pH 8.6) glycyl-glycine buffer. The protocol is summarised in figure 3-5. As shown in the figure, a large part of total peroxidase activity was lost in the column flow-through, confirming that most of *Chara* peroxidases are basic isoforms (Fig.3-4).

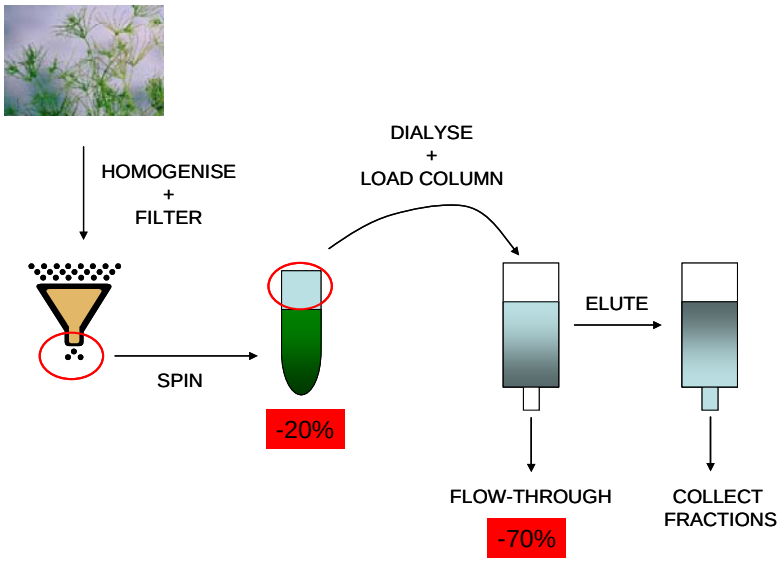


Figure 3-5

Schematical description of peroxidase purification in *Chara contraria*

Guaiacol activity comparisons indicate a 20% loss in the pellet obtained from centrifugation of the filtered crude extract, and 70% loss in the flow-through of the Q-Sepharose column.

After a first passage through the column and subsequent elution in a NaCl gradient, four samples were collected. Fractions were first loaded on a IEF gel, together with the flow-through: as expected, they consisted of acidic isoforms, whereas the flow-through was composed of several basic isoforms (data not shown). Aliquots of the fractions were then migrated on a SDS-PAGE gel (Fig. 3-6). They all appeared to contain a large amount

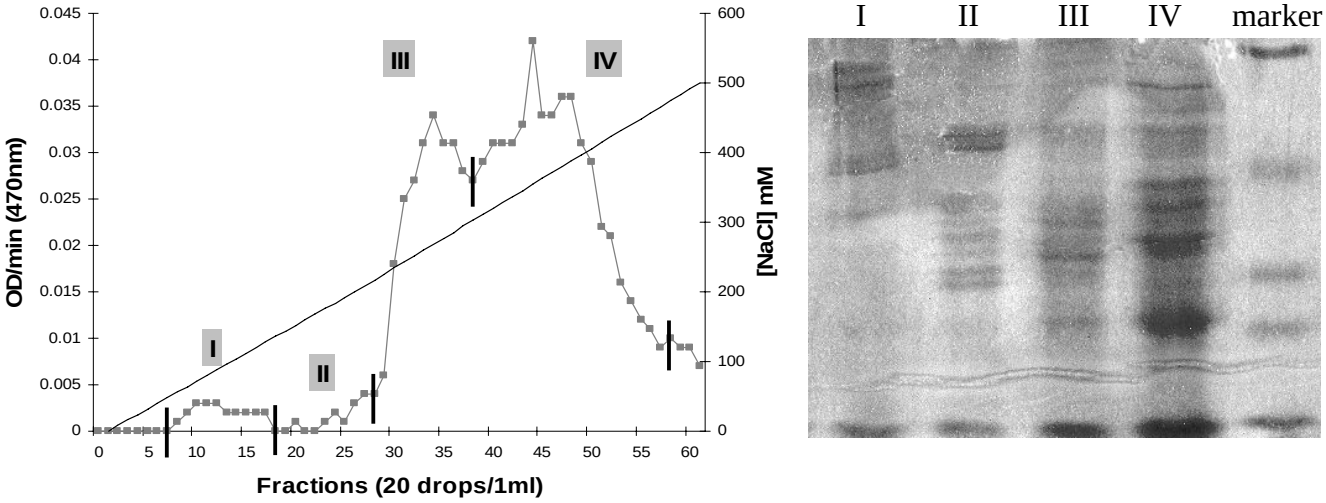


Figure 3-6

Elution of peroxidases from a Q-Sepharose column

Left: elution profile following a NaCl gradient of 0 to 500mM. Four different pooled fractions were selected (I to IV). Left axis: guaiacol oxidase activity. Right axis: NaCl concentration of the elution gradient.

Right: SDS-PAGE gel stained with Coomassie blue with fractions I to IV

marker (from upper to lower band): 77-50-34-29-21 kDa (New England Biolabs P7708S).

of different proteins. Fraction III was chosen for further purification because of its high activity (easily detectable) compare to fractions 1 and 2 and the lower number of proteins (easier purification) compared to fraction IV (SDS-PAGE gel). I loaded again fraction III on a Q-Sepharose column, but with a NaCl gradient close to the concentration needed for elution (250 to 350mM NaCl). After dialysis and concentration, the sample was run in duplicate on a SDS-PAGE gel (non-denaturing conditions) and, although peroxidase activity was rather faint, a band co-localised on Coomassie and *o*-dianisidine staining, at 25kDa (Fig. 3-7). The size of the protein is smaller than what usually expected (36-44 kDa) for peroxidases (Shannon et al., 1966; Duarte-Vazquez et al., 2001; Shah et al., 2004). According to the SDS-PAGE in figure 3-6, the protein could have a molecular weight of 30kDa (different markers

were used). Proteins from the Coomassie-stained gel were further transferred onto a PVDF membrane for protein sequencing (Edman degradation method). However, the staining procedure involved use of acetic acid, which cross-links amino-terminal groups and prevents further protein sequencing. Moreover, the acetic acid contained in the Coomassie solution fixes proteins in the gel and decreases efficiency (or even prevents) transfer on PVDF membranes. A new SDS-PAGE gel was hence run, but without Coomassie staining. Unfortunately, the "fraction III" sample did not behave at all like in the first run: peroxidase activity was concentrated in high molecular weights. Coomassie staining had then to be performed, and confirmed presence of high molecular weight proteins: as the sample was frozen at -20°C between the two gels, it is probable that during thawing, protein aggregates formed. Subsequent tries to break aggregates failed. The first gel obtained (fig. 3-7b, *o*-dianisidine stained gel) was hence rehydrated and the sample was sent for mass spectrometry analysis. The predominant protein detected was not a peroxidase, but a NAD-dependent malate dehydrogenase from *Chara vulgaris* (TrEMBL accession Q8GTZ6). This enzyme

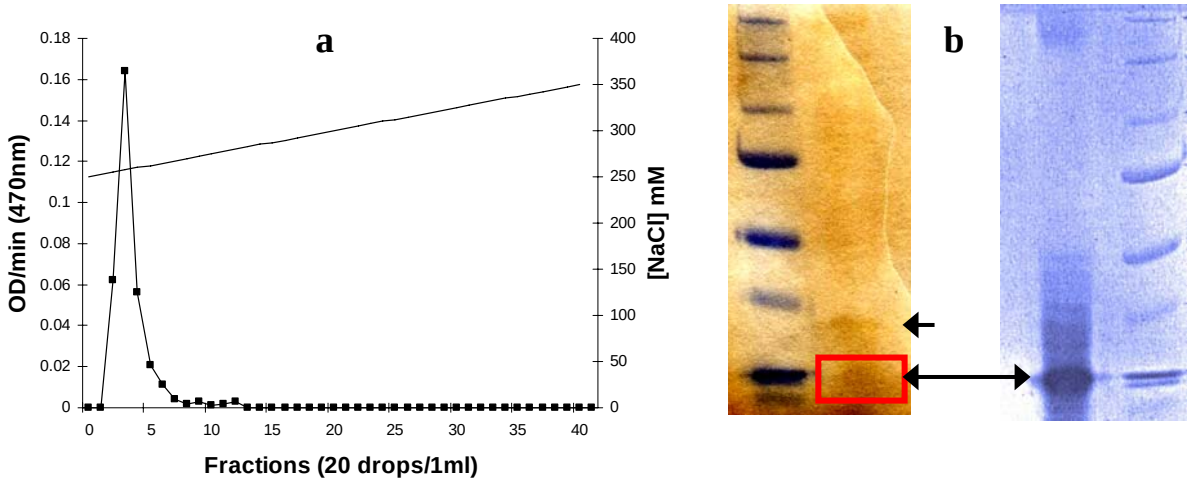


Figure 3-7

Purification of fraction III

(a) Elution profile on a NaCl gradient of 250 to 350mM.

(b) Dried SDS-PAGE gel stained with *o*-dianisidine (left) or Coomassie blue (right). Molecular weight marker (from top to bottom): 250-150-100-75-50-37-[25+20] kDa (BioRad Precision Plus 161-0373).

Red rectangle: region of the gel excised for mass spectrometry analysis

N.B.: Intensity of the two bands stained with *o*-dianisidine was much more pronounced in the gel before drying.

possesses a similar mass as peroxidases (36kDa) and a slightly acidic isoelectric point (6.12, predicted on the website http://www.expasy.ch/tools/pi_tool.html), in line with the Q-Sepharose separation procedure (binds negatively charged proteins). Malate dehydrogenases are not known to oxidise *o*-dianisidine: a peroxidase protein may be contained in the same band as malate dehydrogenase, however the latter protein is probably much more abundant than the peroxidase, therefore masking its presence. In future experiments, a further purification step should be performed in order to separate both proteins, and finally isolate a *Chara* peroxidase protein.

The results obtained so far indicate that *Chara contraria* possesses various peroxidases, both acidic and basic. A protocol has been set up for purification of acidic peroxidases, and shows that one peroxidase (fraction III) can give a quite clean band, almost exempt of other proteins. However, since homogenisation of the alga, samples should not be frozen and procedure should be carried out in a continuous fashion. If, in the future, a suitable protocol can be set up and lead to purification of the peroxidase protein, and Edman sequencing could give us the first amino acids of a *Chara* peroxidase. Appropriate primers would then have to be designed for amplification of genomic DNA (or, if possible, cDNA) and obtain the full peroxidase sequence.

While I was not able to conduct this experiment to its end, I recently (September 2006) found in the GenBank database a class III peroxidase sequence in the zygnemophycean alga *Closterium peracerosum-strigosum-littorale* (CpslPrx01). This last alga belongs to a group that appeared much earlier than Charales during evolution, and its representatives are much simpler organisms (McCourt et al., 2004). Knowing the sequence of one among many *Chara* class III peroxidases is still interesting for phylogenetic studies, although it would be more appropriate now to search in earlier algae, or to sequence several *Chara* peroxidases in order to have a more global view on the different forms of peroxidases in the ancestors of land plants. In this respect, a project headed by professor Charles Delwiche (University of Wisconsin, USA) is currently under way for creating EST databanks of several streptophytean algae, including *Chara*: this project has started, and should bring very soon information on several sequences of algal class III peroxidases.

Did class III peroxidases still play a role in colonisation of land by algae? The discovery of peroxidases in *Closterium* refutes the hypothesis for a key role of peroxidases in the coming out of Charales from a life under water to a life on land. Moreover, life under water is not incompatible with presence of class III peroxidases. In this respect, IEF experiments performed on the aquatic plant *Elodea densa* clearly show that there is no

particular loss in the number and activity of peroxidases when compared to *Arabidopsis thaliana* (Fig. 3-8). Nevertheless, it cannot be excluded that appearance of one peroxidase and subsequent duplication in later algae contributed to establish a genetic background for the creation of the first land plant. Importantly, many Zygnemophyceae (as well as the earlier groups Klebsormidiophyceae and Chlorokybophyceae) are able to live out of water (Cameron, 1960; Lange et al., 1992; Lewis and Lewis, 2005), suggesting that the transition from water to land occurred several times (Lewis and McCourt, 2004), and that appearance of embryophytes took place by small and successive tries and errors. Among these evolutionary steps, specific genes such as class III peroxidases successively appeared. Importantly, several species of Chlorophyceae can live out of water as well (Lange et al., 1992; Lewis and Flechtner, 2002), but none of them achieved sufficient evolutionary changes to give rise to a second lineage of "large-sized" land plants. Some genes are certainly missing in Chlorophyceae: peroxidases are probable candidates, as they are completely absent from this group of algae.

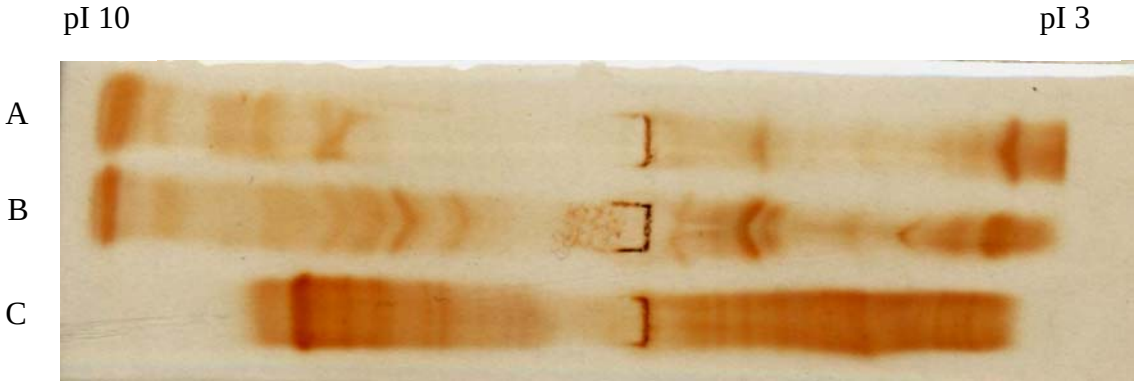


Figure 3-8

IEF gel of *Elodea densa* and *Arabidopsis thaliana.*

(A-B) Adult flowers of *Arabidopsis thaliana* (provided by Claudia Cosio).

(C) *Elodea densa* (thallus and leaves).

C.2.3 Materials and methods

C.2.3.a Amplification of peroxidase sequences from genomic DNA

DNA was extracted from fresh or frozen (-70°C) material. About 2cm of whole plant were taken for *Chara*, and 150µl wet pellet for *Coleochaete*. Algae were ground in eppendorf tubes with a plastic pestle in 100µl (*Chara*) or 300µl (*Coleochaete*) of cold EB buffer (Tris 0.2M, EDTA 25mM, NaCl 250mM, SDS 0.5%, pH 7.5). After a 10-minute centrifugation at 12'000g (4°C), supernatant was supplemented with 1 volume of phenol-chloroform (1:1, equilibrated in Tris 100mM pH 8.0), vortexed vigorously and centrifuged (12'000g). The aqueous phase (upper phase) was washed with 1 volume of chloroform. After ethanol precipitation, purified DNA was resuspended in TE buffer 1x (10mM Tris-HCl, 1mM EDTA, pH 8) and stored at -20°C.

a) First amplification strategy

PCR was performed on total genomic DNA with Cons (5'-GNCTNCWYTTCCACGAYTGYTTYGT-3') and Sca (5'-ANRATRTCRGCRCAWGANAC-3') primers (N = A/T/C/G; W = A/T; Y = T/C; R = G/A). 13 out of 20µl of the resulting DNA were loaded on 1.5% agarose gels. If amplified bands were present, an overnight ligation was directly performed (room temperature) to vector pGEM T (Promega) with the following volumes:

- 1µl pGEM T (50ng/µl)
- 1µl T4 DNA ligase buffer 10x
- 1µl T4 DNA ligase (3U/µl)
- 7µl PCR

3µl of ligation were used to transform *E.Coli* DH5α bacteria by the thermal shock procedure. After an overnight incubation at 37°C, colonies were screened with pGEM T-specific primers T7 (5'-TAATACGACTCACTATAGG-3') and SP6 (5'-GATTTAGGTGACACTATAG-3') for presence of an insert. Positive colonies were cultured overnight in 5ml LB medium; plasmid was purified in a final volume of 50µl (Sigma GenElute^{TM} Plasmid Miniprep Kit) and sent for sequencing with T7 and SP6 primers (Fasteris Life Sciences, Geneva or Synergene Biotech, Zurich).

b) Second amplification strategy

Whole genomic DNA was extracted from plant material as described above. HaeIII digestion (30 units per reaction, QBioGene) was performed overnight at 37°C, in a 30µl reaction volume on 3 to 10µg genomic DNA, and compared the next day to undigested DNA on a 1% agarose gel. Adaptor molecules were obtained from Microsynth (Balgach, Switzerland), with the following sequences.

- Long strand: 5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3'
- Short strand: 5'-ACCTGCCC-NH_2-3'

Adaptor molecules were first incubated together (25µM each) during 5' at 95°C and then slowly cooled down until reaching room temperature. Ligation was performed in a 10µl reaction volume (room temperature, 4 hours) with 2 to 5µg HaeIII-digested DNA, 3 units of T4 DNA ligase and 5µM adaptors. DNA ligase was inactivated 10 minutes at 70°C, and PCR reaction was performed with primers Cons (or Sca) and AP1 (5'-CCATCCTAATACGACTCACTATAGGGC-3') or AP2 (5'-ACTCACTATAGGGCTCGAGCGGC-3').

C.2.3.b Isolation of Chara contraria *peroxidases by chromatography*

All the solutions used in protein extraction were bubbled with argon prior to use, in order to remove as much oxygen as possible and hence minimise oxidative damage on peroxidases. The experiment was entirely performed at 4°C.

50g of frozen material were homogenised (TurMix blender) in 80ml glycyl-glycine 100mM, NaCl 200mM pH 8.6 buffer and 5g polyvinylpyrrolidone (Polyclar AT, removes phenolic compounds). After filtration on a nylon cloth, the solution was centrifuged 10 min. at 10'000 rpm on (ALC 4239R centrifuge, fixed angle rotor A-18C, i.e. 8'380g) and supernatant was dialysed twice (2-3 hours, then overnight; MW cut-off 6-8'000 Dalton, SpectraPor membranes) in 2-3 litres glycyl-glycine 1mM (pH 8.6, no NaCl).

The dialysed solution was passed through a Q-sepharose column (Pharmacia, fresh, 2cm diameter, 3cm high, equilibrated with glycyl-glycine buffer 1mM). After a wash with 200ml glycyl-glycine 1mM, elution was performed in two steps (due to volume limitations of the gradient mixer): a first gradient with 40ml glycyl-glycine 0 to 200mM NaCl, and a second

gradient with 40ml glycyl-glycine 200 to 500mM NaCl. Fractions containing 20 drops each were collected and assayed (20µl) for guaiacol oxidation activity.

Fractions of interest were further dialysed by two rounds of dialysis against glycyl-glycine 1mM and passed again through a Q-Sepharose column (freshly made), but eluted with tighter ranges of NaCl (closer to the concentration used to elute the fraction). Desalting was then achieved through one overnight dialysis against glycyl-glycine 1mM, and samples were finally concentrated by spinning in Centricon YM-10 filters (4'700g) until a volume of ~200µl was reached. Concentrated samples were recovered by gently vortexing Centricon filters, inverting them and spinning them one minute at 750g.

C.2.3.c SDS-PAGE gels

Two protocols were used for SDS-PAGE sample preparation, depending on whether samples had to be stained for peroxidase activity on the gel (non-denaturing conditions, figure 3-7) or stained with Coomassie blue (denaturing conditions, figure 3-6).

For non-denaturing conditions, 2x15µl of sample were mixed with loading dye (50mM TrisCl pH6.8, 0.1% bromophenol blue and 5% glycerol). SDS-PAGE (12.5%) was run at 120V. After migration, one sample lane was stained in Coomassie (0.3% Coomassie brilliant blue R250, 50% MeOH, 10% acetic acid) during 30 minutes and non-specific staining was washed out during several hours with a destaining solution (5% acetic acid, 5% methanol). The second sample lane was stained for peroxidase activity with ortho-dianisidine (100mM, acetate buffer pH 4.5).

For denaturing conditions, 15µl of sample were mixed with 5µl urea 8M and protein sample buffer (1x final concentration: 50mM TrisCl pH6.8, 1% sodium dodecyl sulphate [SDS], 5% β-mercaptoethanol [βme]). After boiling during 5 minutes, loading dye was added to each sample. Gels were run and stained with Coomassie blue as described in the previous paragraph.

After staining, gels were vacuum-dried at 80°C (Gel drier model 583, BioRad) on a gel blotting paper and stored at room temperature. For rehydration, the region of the gel shown in figure 3-7 was cut and soaked overnight at 4°C in $(NH_4)HCO_3$ 50mM solution. Blotting paper was then carefully removed; the gel was transferred to a new vial with 50µl EtOH 10% and sent for mass spectrometry analysis (Proteomics Core Facility, CMU, Geneva). Prior to analysis, the solubilised gel was purified by microcapillary liquid chromatography (C18 column) in the same facility. This procedure was also performed with a

negative control (fragment of gel without protein) and a positive control (BSA from the molecular weight marker). Fragment ion spectra were compared to the SwissProt database with the Mascot^{TM} algorithm.

C.2.3.d Transfer on PVDF membranes

The following steps were used for transfer of proteins from a SDS-PAGE gel to a PVDF membrane:

- rehydrate a gel-sized PVDF membrane (BioRad 162-0184) in MeOH and then in water
- soak the PVDF membrane and four pieces of gel blotting paper (0.8 mm thick, Schleicher & Schuell SS-10426892, gel-sized) in protein transfer buffer (25mM Tris, 200mM glycine, 20% MeOH)
- soak fibre pads of transfer cassette (Hoefer TE22 Mighty Small Transphor apparatus) in transfer buffer
- starting from the black side of the cassette, put in this order:
    * fibre pad
    * two gel blotting papers
    * SDS-PAGE gel
    * PVDF membrane
    * two gel blotting papers
    * fibre pad
- transfer 1-2 hours at 80V or overnight at 30V.

After transfer, membrane was briefly washed in water and stained with Ponceau S solution (0.2% Ponceau, 3% trichloroacetic acid, 3% sulfosalycilic acid). Protein bands were revealed by washing with water.

C.2.3.e Isoelectric focusing gels

Servalyt Precotes (Serva, Germany; Cat.No. 42866.02 and 42878.02) polyacrylamide gels were run on a Multiphor 2117 (LKB Co., Sweden) electrophoresis apparatus, on a thin layer of decan. The starting voltage was 300V (constant power at ~4W for a full 12x12cm gel). Migration was stopped until a plateau of voltage was reached (usually 3 hrs and 1300V).

The IEF gel was then immersed in ca. 200ml acetic acid buffer (100mM AcOH, ~45mM NaOH, pH4.5) and coloured with *o*-dianisidine (spatula tip dissolved in an eppendorf of warm EtOH) and ~200µl H_2O_2 30%. Reaction was stopped by rinsing with water. Gels were air dried for storage.

D. PeroxiBase: a comprehensive peroxidase database

D.1 Article: "PeroxiBase: a class III plant peroxidase database"

Following the extensive search for rice peroxidases, it appeared that rice-specific and *Arabidopsis*-specific groups of peroxidases could be distinguished. This observation raised a potential interest for practical application on rice culture: some peroxidases could have borne, as mentioned in the article "**The class III peroxidase multigenic family in rice and its evolution in land plants**", in response to the diversification of pathogenic attacks. Furthermore, since functional studies on peroxidases were far more advanced in *Arabidopsis* than in rice, identifying roles or expression profiles of *Arabidopsis* peroxidases could be very informative for rice orthologs. This reasoning led Christophe and I to create a database for rice and *Arabidopsis* peroxidase protein sequences, including fields with additional information, such as "inducers", "repressors" and "tissue type". As we noticed that a large amount of peroxidase sequences was already known in other plants, we extended our database to the whole plant kingdom. While constantly adding more sequences, with the help of Nenad Bakalovic, we decided to make this repository publicly available: we hence started a collaboration with Vassilios Ioannidis and Laurent Falquet (Swiss Institute of Bioinformatics) and created the website "PeroxiBase": http://peroxidase.isb-sib.ch. In order to validate the database and announce to peroxidase researchers its creation, we published the article "**PeroxiBase: a class III plant peroxidase database**".

This article is cited as follows:

Bakalovic N, Passardi F, Ioannidis V, Cosio C, Penel C, Falquet L, Dunand C.
PeroxiBase: a class III plant peroxidase database.
Phytochemistry. 2006 Mar;67(6):534-9.
DOI: http://dx.doi.org/10.1016/j.phytochem.2005.12.020

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 67 (2006) 534–539
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Update in Bioinformatics

PeroxiBase: A class III plant peroxidase database

Nenad Bakalovic [a,1], Filippo Passardi [a,1], Vassilios Ioannidis [b], Claudia Cosio [a], Claude Penel [a], Laurent Falquet [b], Christophe Dunand [a,*]

[a] *Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland*
[b] *Swiss Institute of Bioinformatics, CH-1066 Epalinges/Lausanne, Switzerland*

Received 13 September 2005; received in revised form 9 December 2005
Available online 26 January 2006

Abstract

Class III plant peroxidases (EC 1.11.1.7), which are encoded by multigenic families in land plants, are involved in several important physiological and developmental processes. Their varied functions are not yet clearly determined, but their characterization will certainly lead to a better understanding of plant growth, differentiation and interaction with the environment, and hence to many exciting applications. Since there is currently no central database for plant peroxidase sequences and many plant sequences are not deposited in the EMBL/GenBank/DDBJ repository or the UniProt KnowledgeBase, this prevents researchers from easily accessing all peroxidase sequences. Furthermore, gene expression data are poorly covered and annotations are inconsistent. In this rapidly moving field, there is a need for continual updating and correction of the peroxidase superfamily in plants. Moreover, consolidating information about peroxidases will allow for comparison of peroxidases between species and thus significantly help making correlations of function, structure or phylogeny. We report a new database (PeroxiBase) accessible through a web server (http://peroxidase.isb-sib.ch) with specific tools dedicated to facilitate query, classification and submission of peroxidase sequences. Recent developments in the field of plant peroxidase are also mentioned.

Keywords: Database; Multigenic family; Evolution; Phylogeny; Peroxidases

1. Introduction

Class III plant peroxidases (EC 1.11.1.7, donor:hydrogen-peroxide oxidoreductase) are present in all land plants (Table 1). Genes encoding this enzyme family are particularly numerous in Angiosperms. The high number of isoenzymes and their remarkable catalytic versatility allow them to be involved in a broad range of physiological and developmental processes all along the plant life cycle (Passardi et al., 2005). Plant peroxidases have been shown to be involved in the cross-linking of cell wall constituents, lignin polymerization, the catabolism of auxin – a hormone having a critical role in plant growth and development – and the formation of reactive oxygen species (superoxide, hydroxyl radical). They also play a prominent role in defence reactions against many pathogenic organisms. However, until now the in vivo functions of a particular peroxidase have not been reported. This knowledge is however crucial to understand the evolution, the roles and the regulations of this key multifunctional enzyme. Plant peroxidases are an example of a multigenic family whose number of members increased since the conquering of land by plants due to constant evolution. The *Arabidopsis* genome contains 73 genes encoding a peroxidase (Tognolli et al., 2002) and rice contains 138 (Passardi et al., 2004). The homology between paralogs in a plant ranges from 30% to 100%, but very close orthologs exist, even between evolutionarily distant plants. All plant peroxidases contain invariant amino acids essential for their catalytic properties and for their proper folding (Welinder et al., 2002). They are structurally related to other heme-containing

* Corresponding author. Tel.: +41 223793012; fax: +41 223793017.
E-mail address: christophe.dunand@bota.unige.ch (C. Dunand).
[1] The two first authors have contributed equally to this work.

0031-9422/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.12.020

Table 1
Representation of the major plant lineages found in the PeroxiBase

Order (number of species/number of sequences): Genus
Angiosperms
Dicotyledons
Rosids
Cucurbitales (3/7): Cucumis (2x), Cucurbita
Fabales (14/314): Arachis, Cicer, Glycine, Lotus, Lupinus (2x), Medicago (2x), Phaseolus (2x), Pisum, Stylosanthes, Trifolium, Vigna
Rosales (3/8): Ficus, Malus, Urtica
Fagales (1/2): Quercus
Malpighiales (10/92): Euphorbia, Linum, Manihot, Mercurialis, Populus (6x)
Malvales (3/58): Gossypium, Theobroma
Sapindales (3/20): Citrus (2x), Poncirus
Brassicales (7/103): Arabidopsis, Brassica (2x), Armoracia, Raphanus, Thellungiella (2x)
Vitaceae (1/21): Vitis
Asterids
Gentianales (3/10): Coffea, Hedyotis (2x)
Lamiales (6/10): Avicennia, Eucommia, Orobanche (2x), Scutellaria, Striga
Solanales (11/199): Capsicum (2x), Ipomoea (2x), Lycopersicon (2x), Nicotiana (3x), Petunia, Solanum
Asterales (8/73): Artemisia, Cichorium, Helianthus (2x), Gerbera, Lactuca, Stevia, Zinnia
Apiales (1/1): Petroselinum
Ericales (1/1): Vaccinium
Saxifragales (1/6): Ribes
Caryophyllales (4/56): Beta, Mesembryanthemum, Mirabilis, Spinacia
Ranunculales (2/37): Eschscholzia, Aquilegia
Monocotyledons
Magnolids
Laurales (1/10): Persea
Magnoliales (1/11): Liriodendron
Piperales (1/0): Saruma
Liliopsida
Poales (15/655): Aegilops, Ananas, Avena, Cenchrus, Hordeum, Lolium, Oryza, Saccharum (2x), Secale, Setoria, Sorghum, Triticum (2x), Zea
Zingiberales (0/0) EST project in run
Liliales (2/2): Lilium, Alstroemeria
Acorales (1/5): Acorus
Alismatales (1/1): Spirodela
Arecales (1/1): Elaeis
Asparagales (3/36): Allium, Asparagus, Hyacinthus
Basal magnoliophyta
Austrobaileyales (1/0): Illicium
Nymphaeales (1/3): Nuphar
Amborellaceae (1/13): Amborella
Gymnosperms
Gnetopsida
Gnetales
Welwitschiales (1/1): Welwitschia
Coniferales (6/24): Cryptomeria, Pinus (3x), Picea (2x)
Ginkgoales (1/8): Ginkgo
Cycadales (3/12): Cycas, Zamia (2x)
Cryptogames (4/21)
Marchantia, Physcomitrella, Selaginella, Ceratopteris
Algae
Charales (0/0) no sequence available/activity found
Chlamydomonadales (0/0) no sequence available/no activity found

Number of species and of peroxidases found in each lineage are represented in brackets.

proteins, like peroxidases from prokaryotes, fungi. The key amino acids that interact with heme are also found in hemoglobins and cytochromes. The broad molecular diversification of plant peroxidases mainly results from gene duplication events. Newly duplicated genes were likely conserved because they acquired new modes of expression, regulation (subfunctionalization) or novel functions (neofunctionalization).

The automated annotation of the whole genomes of *Arabidopsis* (Arabidopsis Genome Initiative, 2000) and *Oryza sativa* (Goff et al., 2002; Yu et al., 2002), the automated clustering and assembling of EST sequences, and numerous EST projects led to the identification of a large number of sequences coding for class III plant peroxidases. We decided to construct a database devoted to this large, multigenic family because in our experience automated processing sometimes yields sequences of poor quality. Specificity is compromised and BLAST searching often requires manual sorting. Using the highly conserved motifs of the class III peroxidases (Welinder, 1992), manual annotation and editing can retrieve whole peroxidase sequences, that are unrecognized in automation due to poor quality sequences. *Arabidopsis* and rice are completely sequenced and are considered to be plant models, but they are not representative of plant diversity. The large number of EST projects developed with more diverse plants will provide a better overview of peroxidase evolution throughout green plants. The first goal of the PeroxiBase is to centralize most of the annotated and non-annotated class III peroxidase-encoding sequences and to make them publicly available, so that the research community has a unique tool for discovery, comparison, and exchange of peroxidase sequences. The second goal is to compile information concerning putative function and transcription regulation in order to facilitate cross-checking between close paralogs and orthologs. The final goal is to confirm the hypothesis that the number of class III isoforms increased after the emergence of the land plants.

2. Construction of the database

The database was constructed following two parallel procedures: one exhaustive and another more specific (Fig. 1). Firstly, the plant/fungal/bacterial heme peroxidase proteins are characterized by the motif PEROXIDASE_4 (http://ca.expasy.org/cgi-bin/nicesite.pl? PS50873). Using this signature, systematic data mining with MyHits (Pagni et al., 2004) from different predicted protein databases (TrEST, TrGEN) (Pagni et al., 2001) provides a global view of the peroxidase encoding sequences. The resulting hits are already treated data (assembled and translated sequences) with the risk of automatic compilation and translation.

We have then used AtPrx42 and OsPrx73, two sequences potentially related to an ancestral sequence (Passardi et al., 2004), for a second, more specific approach by scanning numerous public sources of plant ESTs and genome sequences in order to obtain a large collection of peroxidase-encoding sequences. For rare species, a tBLASTn was first performed against the NCBI whole database using limited queries for the date and for the organism. The Plant Genome Project (http://www.pgn.cornell.edu, 2004), Plant GDB (Dong et al., 2004) and Sputnik (Rudd, 2005) were used to complete the short sequences obtained from TIGR and to find new ones.

Each sequence obtained (assembly or singleton sequences) was individually translated and the presence of characteristic peroxidase motifs was verified using FingerPRINTScan and InterPro Scan softwares. Low quality

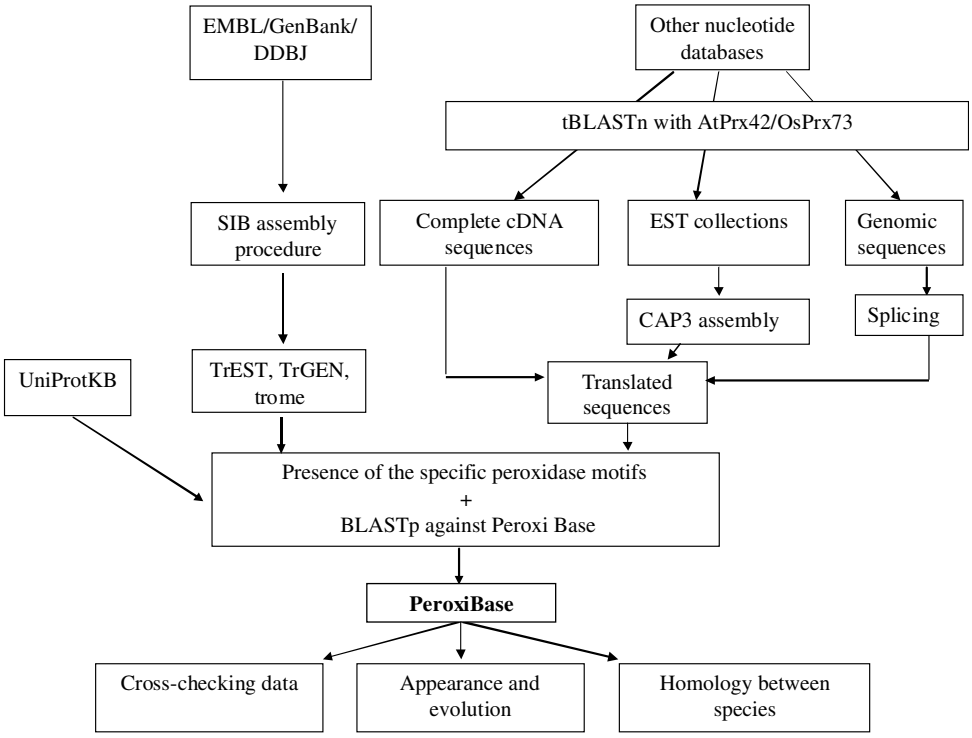


Fig. 1. Procedure of data analysis for generation of the PeroxiBase. Various EST and genomic databases have been used as sources of peroxidase encoding sequence.

sequences are not included in TIGR consensus sequences due to the high sequence stringency TIGR uses. Using peroxidase motifs as a guide, manual inspection of these poor quality sequences allows for increasing their length and assembling new sequences. All distinct sequences, even short ones, are kept in the database. We have also compared the UTR regions to confirm that sequences with high homology are truly distinct.

In addition, as in numerous other genera, ESTs from *Gossypium* (cotton), *Picea* (spruce), *Populus* (poplar) have been assembled by TIGR. However, the sequences used for the construction of the contigs are treated as if derived from a single species although they originate from different species. For example, cotton is derived from *Gossypium arboreum* and *hirsutum*, spruce from a collection of *Picea abies*, *Picea glauca* and *Picea sitchensis* and poplar from a mix of *Populus alba*, *Populus balsamifera*, *Populus euphratica*, *Populus kitakamiensis*, *Populus nigra*, *Populus tremula*, *Populus tremuloides*, and *Populus trichocarpa*. The Sputnik database (http://sputnik.btk.fi/) helped us unscramble this mixture of sequences for *Gossypium* and *Populus* and *Picea* species. The other species have been assembled directly from the NCBI entries.

After analysis of sequence alignment within each species (ClustalW and BioEdit), the protein sequences were individually entered in the database with their corresponding accession numbers as well as various information concerning the putative functions and transcription regulation (localization, induction and repression).

3. Web interface

The PeroxiBase web interface includes four main modules. (i) *Search*: this module enables a text query from the entire dataset with keywords such as tissue type, accession number, inducer/repressor, and name of sequences and organisms. (ii) *Organism*: each organism possesses a link with its taxonomic identity. Information for the peroxidases present in each organism can be viewed independently. Each file contains a direct link to the corresponding database (NCBI, TIGR, PGN, Sputnik) and to Swiss-Prot and DNA sequences when these entries exist. In addition, numbers of ESTs, cellular localization and tissue type are all included. The three closest homologous sequences, with their corresponding score and *E*-value are also described in the files. (iii) *BLAST*: two BLAST searches can be performed against the whole peroxidase database (Altschul et al., 1997), BLASTp for protein sequence and BLASTx for nucleotide sequences. The alignments are visualized and linked to the entry of each peroxidase. (iv) *FingerPrintscan*: this tool helps to find out which peroxidase family the sequence belongs to.

Three minor modules *Tissue type*, *Inducers/repressors* and *Cellular localization* are used as alternative ways of viewing sequences.

4. Modelization of the number of class III isoforms evolution

Class III peroxidase encoding sequences and peroxidase activity are both absent from the green alga *Chlamydomonas reinhardtii* (Passardi et al., 2004). On the contrary, in various *Chara* species, guaiacol oxidation in the presence of H_2O_2, specific to class III peroxidase activity (Greppin et al., 1986), can be detected (data not shown). The exhaustive data mining performed for the setup of the PeroxiBase confirms that the class III peroxidases are present in all land plants. For some species the total EST count is low (less than 1000), yet several independent isoforms were identified. In the case of large EST libraries (over 10,000), multigenic families with numerous putative class III peroxidases sequences can be found confirming previous results obtained with *Arabidopsis* and *Oryza* (Tognolli et al., 2002; Duroux and Welinder, 2003). In addition, the case of *Physcomitrella patens* seems to further validate the

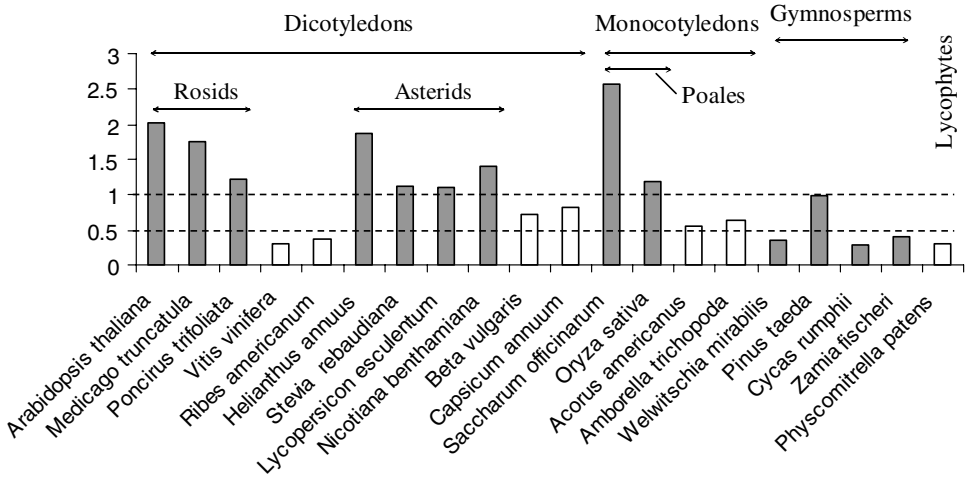


Fig. 2. Evolution of the peroxidase encoding genes versus the total number of genes in various key organisms. The *y*-axis represents a ratio obtained as following: (number of peroxidase encoding EST/number of independent peroxidase encoding genes)/(number of total EST/number of unigenes).

hypothesis that the number of peroxidase encoding genes increases over evolution. With more than 100,000 EST and only 12 peroxidase encoding sequences, the number of peroxidases sequences for this moss is far below those from other organisms such as *Arabidopsis*.

A potential evolution of peroxidase gene numbers can be drawn up based on information concerning each species such as total EST count and number of unigenes with the following formula: (number of peroxidase encoding EST/number of independent peroxidase encoding genes)/(number of total EST/number of unigenes). The value is proportional to the number of peroxidase genes in each species and gives information regarding the putative evolution of the peroxidase isoforms. Other completed genome sequencing and increasing EST sequences should confirm this hypothesis. Independently of the EST number, the size of the family seems also to follow a gradual increase from *Charales* (few isoforms) to the higher plants (numerous isoforms) confirming the previous hypothesis. The species can be classified in two major groups following the value of this ratio (Fig. 2). Rosids, Asterids and Poales considered to be higher plants, show a high diversification rate value over 1. On the other hand, species issued from basal Gymnosperms, and from small Mono- and Dicotyledons orders such as Vitaceae, Saxifragales, Caryophyllales, and Ranunculales have values around 0.5 or smaller.

5. Current status and future developments

The first goal of the PeroxiBase was to develop an efficient tool for the study of the evolution of a plant multigenic family. We tried not to be exclusive and to include as many sequences as possible from different organisms. The base currently consists of a core dataset containing over 2000 complete or partial peroxidase-encoding sequences from 125 organisms (Table 1), and it is still in constant evolution.

New peroxidase encoding sequences can be easily and directly added to the database by external people with individual user name and password. Continuous data mining will be performed until a putative complete analysis of the available sequences is achieved (EST and genomic sequences). At this point, a semi-automatic update will be set up to collect the peroxidase encoding sequences newly submitted to general databases (NCBI, Swiss-Prot). Information concerning the expression profile will also be updated and new features such as results of knock-out, knock-down or overexpression studies will be added when available.

The superfamily of plant, fungal and bacterial heme peroxidases contains *class I* (Cytochrome *C* peroxidase (EC 1.11.1.5), catalase peroxidase (EC 1.11.1.6) and ascorbate peroxidase (EC 1.11.1.11)), *class II* (lignin peroxidases (EC 1.11.1.14) and manganese peroxidases (EC 1.11.1.13)) and *class III*. The next update will include class I and class II peroxidases in the database. To our knowledge, the PeroxiBase will become the first repository devoted exclusively to a superfamily composed of multigenic families. The database could help to confirm the hypothesis that the three classes evolved from a single ancestral sequence (Zamocky, 2004). Another major addition will be to relate the major lineages containing peroxidase-encoding sequences to a schematic evolutionary tree. Peroxidases may become key markers for the evolution of plants, from as early as the first moments of land colonization to the human impact on genetics of cultivated plants today. The varied functions of peroxidases will be characterized and lead to a better understanding of plant growth, differentiation and interaction with the environment, and eventually to many exciting applications.

6. Useful web links

BioEdit: http://www.mbio.ncsu.edu/BioEdit/bioedit.html
ClustalW: http://www.ebi.ac.uk/clustalw/
Expasy translate: http://us.expasy.org/tools/dna.html
FingerPRINTScan: http://www.bioinf.man.ac.uk/fingerPRINTScan/
InterPro Scan: http://www.ebi.ac.uk/InterProScan/
MyHits: http://myhits.isb-sib.ch/cgi-bin/motif_query
NCBI: http://www.ncbi.nlm.nih.gov/
PlantGDB: http://zmdb.iastate.edu/PlantGDB/
Plant Genome Network: http://pgn.cornell.edu/
Reverse-Complement: http://bioinformatics.org/sms/rev_comp.html
SoftBerry- FGENESH: http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind
Sputnik: http://sputnik.btk.fi/ests
TIGR: http://www.tigr.org/tdb/tgi/plant.shtml

Acknowledgments

We thank Sonia Guimil for her critical reading. The financial support of the Swiss National Science Foundation (Grant 31-068003.02) to C.P. and C.D. is gratefully acknowledged, N.B. is paid by the Office Cantonal de l'Emploi.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815.

Dong, Q., Schlueter, S.D., Brendel, V., 2004. PlantGDB, plant genome database and analysis tools. Nucleic Acids Res. 32, D354–D359.

Duroux, L., Welinder, K.G., 2003. The peroxidase gene family in plants: a phylogenetic overview. J. Mol. Evol. 57, 397–407.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., Briggs, S., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). Science 296, 92–100.

Greppin, H., Penel, C., Gaspar, T., 1986. Molecular and Physiological Aspects of Plant Peroxidases. University of Geneva, Switzerland.

Pagni, M., Iseli, C., Junier, T., Falquet, L., Jongeneel, V., Bucher, P., 2001. trEST, trGEN and Hits: access to databases of predicted protein sequences. Nucleic Acids Res. 29, 148–151.

Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C.V., Falquet, L., 2004. MyHits: a new interactive resource for protein annotation and domain identification. Nucleic Acids Res. 32, W332–W335.

Passardi, F., Longet, D., Penel, C., Dunand, C., 2004. The class III peroxidase multigenic family in rice and its evolution in land plants. Phytochemistry 65, 1879–1893.

Passardi, F., Cosio, C., Penel, C., Dunand, C., 2005. Peroxidases have more functions than a Swiss army knife. Plant Cell Rep. 24, 255–265.

Rudd, S., 2005. openSputnik: a database to establish comparative plant genomics using unsaturated sequence collections. Nucleic Acids Res. 33, D622–D627.

Tognolli, M., Penel, C., Greppin, H., Simon, P., 2002. Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288, 129–138.

Welinder, K.G., 1992. Plant peroxidases: structure–function relationships. In: Penel, C., Gaspar, T., Greppin, H. (Eds.), Plant Peroxidases. University of Geneva, Switzerland.

Welinder, K.G., Justesen, A.F., Kjaersgard, I.V., Jensen, R.B., Rasmussen, S.K., Jespersen, H.M., Duroux, L., 2002. Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. Eur. J. Biochem. 269, 6063–6081.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Li, J., Liu, Z., Qi, Q., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Zhao, W., Li, P., Chen, W., Zhang, Y., Hu, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Tao, M., Zhu, L., Yuan, L., Yang, H., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). Science 296, 79–92.

Zamocky, M., 2004. Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. Eur. J. Biochem. 271, 3297–3309.

D.1.1 Supplementary comments

The article published on the PeroxiBase describes the generation and development of a database destined to class III peroxidases. Since then, several improvements have been performed, and the database is constantly evolving. Among the major changes implemented, there have been:

- addition of class I and II peroxidases from the superfamily
- BLAST tool for searches within the PeroxiBase
- Finger Print Scan tool to recognise specific peroxidase motifs
- Standardisation of fields “inducers”, “repressors” and “tissue type” with the use of pre-defined key words
- Classification of organisms according to NCBI TaxId
- Multi-criteria searches
- Addition of the “cellular localisation” field (particularly useful for ascorbate peroxidases)
- Obligatory double-check of peroxidase sequences by the PeroxiBase administrators
- cross-references with the SwissProt and TrEMBL repositories

Future changes will mainly include addition of all haem and non haem peroxidase sequences, BLAST searches limited to selected groups of organisms or peroxidase types and searches by single peroxidase type. External contributors have also joined the PeroxiBase team: Felipe Teixeira and Marcia Margis-Pinheiro (Rio de Janeiro, Brazil), Marcel Zamocky (Bratislava, Slovakia), Jean-Pierre Jacquot and Nicolas Rouhier (Nancy, France). The database “webmaster” is now Gregory Theiler, from the Plant Biology department.

D.2 A last search for the first sequences…

One approach that could be used to identify the remnants of ancestral peroxidase sequences in evolved land plants is to compare whole multigenic families with peroxidases of earlier organisms, such as liverworts, mosses and ferns. For these organisms however, the amount of available data is still rather poor: among the cryptogams listed in the PeroxiBase, complete peroxidase sequences are almost only present for the moss *Physcomitrella patens* (PpaPrx). Nevertheless, I performed a phylogenetic analysis comprising all complete PpaPrx sequences, a few complete sequences from other cryptogams (*Ceratopteris richardii*,

Selaginella moellendorffii, *Marchantia polymorpha*) as well as peroxidases from one well-represented gymnosperm (*Pinus taeda*) and *Arabidopsis thaliana*. The results were quite interesting (Fig. 3-9): AtPrx29 came out in a group comprising one fern peroxidase (CrPrx03), the peroxidase of *Marchantia* and two *Physcomitrella* peroxidases. The cluster is strongly supported and well separated from other groups. The remaining cryptogam peroxidase sequences (six PpaPrx and three SmPrx) formed separate groups or small clusters with a few PtaPrx or AtPrx. The closest sister group to the AtPrx29 cluster comprises AtPrx42 and AtPrx21. However, when performing a similar analysis with OsPrx and AtPrx, Nenad and Christophe found that AtPrx21/42 are well separated from AtPrx29; in Nenad and Christophe's study, AtPrx29 came out exclusively with OsPrx73. The association AtPrx29-OsPrx73 was also found in further studies by Nenad and Christophe, where they compared sequences from rice, *Arabidopsis*, *Medicago*, tomato, soybean and wheat: AtPrx29 and OsPrx73 came out together alone, without any orthologous sequence from the other four organisms. Considering that the PeroxiBase contains almost all peroxidase sequences from the six organisms (very large EST database and/or genomic sequences), wheat, tomato, soybean and *Medicago* do not encode any AtPrx29/OsPrx73 orthologous gene. The close association of AtPrx29 and cryptogamous sequences suggests that AtPrx29/OsPrx73 are likely remnants of an "ancestral" peroxidase sequence. Interestingly, the ancient origin of OsPrx73 was already hinted in the article "**The class III peroxidase multigenic family in rice and its evolution in land plants**" (Fig. 2). Loss of this peroxidase in other Tracheophyta indicates that throughout evolution, it was no longer required. The hypothesis that ancestral peroxidases fulfil essential functions is hence not valid.

Figure 3-9 (next page)

Phylogenetic tree (Neighbor-Joining method) of AtPrx (*Arabidopsis thaliana*), PtaPrx (*Pinus taeda*) and cryptogam protein peroxidase sequences.

PpaPrx: *Physcomitrella patens* (moss)

CrPrx: *Ceratopteris richardii* (fern)

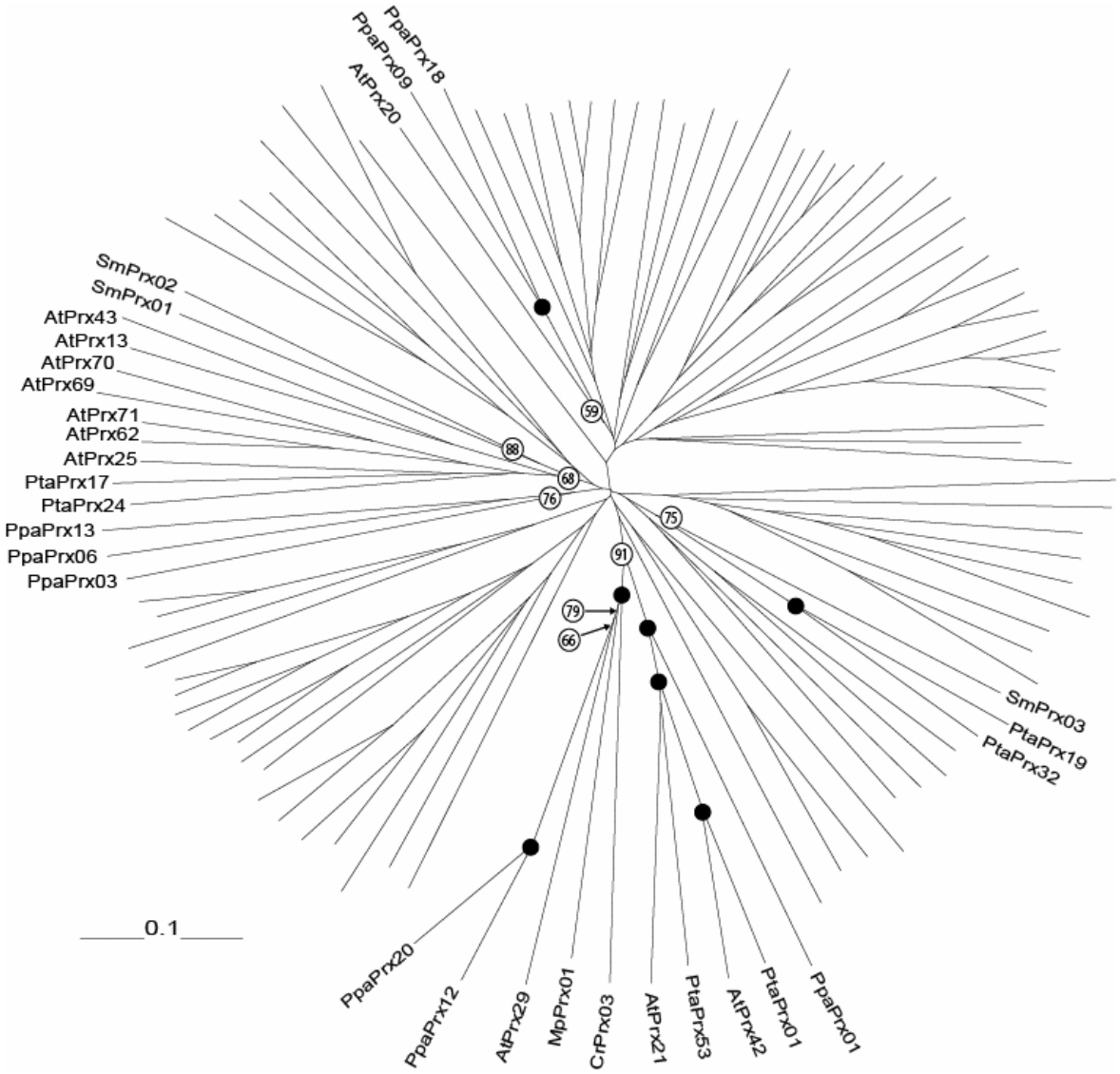
SmPrx : *Selaginella moellendorffii* (spikemoss)

MpPrx : *Marchantia polymorpha* (liverwort)

Only groups containing cryptogams are shown.

Bootstraps are indicated in white circles when their value is below 95. Black dots stand for values >95.

Scale bar represents 0.1 substitution per site.


PpaPrx18
PpaPrx09
AtPrx20
SmPrx02
SmPrx01
AtPrx43
AtPrx13
AtPrx70
AtPrx69
AtPrx71
AtPrx62
AtPrx25
PtaPrx17
PtaPrx24
PpaPrx13
PpaPrx06
PpaPrx03
59
88
68
76
75
91
79
66
SmPrx03
PtaPrx19
PtaPrx32
0.1
PpaPrx20
PpaPrx12
AtPrx29
MpPrx01
CrPrx03
AtPrx21
PtaPrx53
AtPrx42
PtaPrx01
PpaPrx01

4

AtPrx42 phylogeny and promoter studies

A. Introduction

Among the 73 peroxidase genes identified in *Arabidopsis thaliana*, *AtPrx42* is probably the most peculiar. First, it does not have any particularly high identity with other *AtPrx* genes: its closest paralog is *AtPrx21*, with 57% identity. This value is rather low when, for example, *AtPrx33* and *AtPrx34* show 87% identity. For this reason, AtPrx42 appears isolated in a phylogenetic tree (Tognolli et al., 2002). Secondly, *AtPrx42* is by far the most expressed gene: an exhaustive study performed by Luisa Valério, from our laboratory, reported 340 EST sequences for *AtPrx42*, while the following most expressed genes were *AtPrx32* and *AtPrx34* with, respectively, only 135 and 93 genes (Valério et al., 2004). Moreover, the EST libraries used were issued from many diverse organs, at various developmental stages. The last particular feature of AtPrx42 is its ubiquitous presence and the low variability of its expression, as demonstrated in many studies (Kjaersgard et al., 1997; Welinder et al., 2002; Valério et al., 2004; Passardi et al., 2006) and during my thesis work.

When Michael had been searching for the exhaustive list of peroxidases in *Arabidopsis*, he had conjectured that "the P42 group was present in an ancestral plant species" (Tognolli et al., 2002), following the finding that AtPrx42 shares high identity (~80%) with other peroxidases from distantly related plants. Given the exponential accumulation of data in NCBI and other databases, and the rapid increase in the sequences present in the PeroxiBase, I performed a search of all organisms available in order to better understand the distribution and evolution of *AtPrx42*. In a second part of the study, different deletion mutants of the *AtPrx42* promoter were fused to the glucuronidase reporter gene for transformation into *Arabidopsis thaliana*.

B. Phylogeny of AtPrx42 in plants

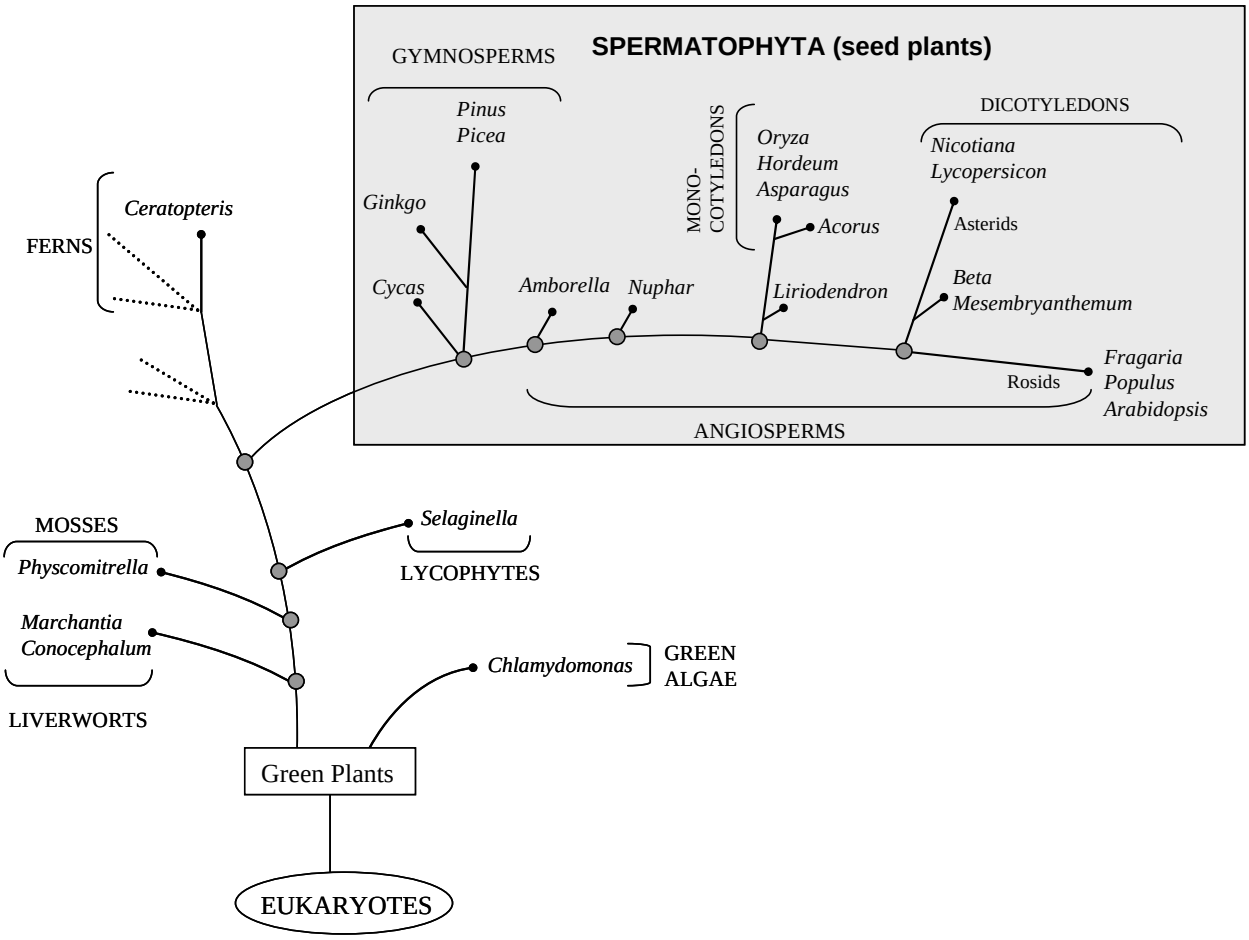


Figure 4-1

Simplified taxonomical distribution of green plants

(Savard et al., 1994; Kenrick and Crane, 1997; Soltis et al., 2000; Pryer et al., 2001; Wikstrom et al., 2001; Schmidt and Schneider-Poetsch, 2002).

Representative genera of the species described in this thesis are reported on the taxonomic tree.

Atprx42 gene is present in all the seed plants, with some exceptions in monocotyledons (see text).

B.1 Study of taxonomical distribution of AtPrx42 *orthologs*

The results obtained through BLAST searches on the PeroxiBase and in NCBI are summarised in Table 4-1. The count of EST sequences was reported for each new gene found and compared with the total number of EST genes available in NCBI (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html, last update: September 8th, 2006), TIGR or SPUTNIK. The ratio # $EST_{peroxidase}$/#EST_{total} was set to "1" for *AtPrx42* for more clarity. It rapidly appeared that *AtPrx42* orthologs were present in most of the plant organisms screened. The ratio # $EST_{peroxidase}$/#EST_{total} was rather heterogeneous and not specific to definite groups of plants. On the other hand, *AtPrx42* orthologs usually had two typical features: highest identity ("rank in species") to *AtPrx42* within the organism and higest EST count ("rank in EST"). Exceptions to this last rule were found in:

- *Cichorium intybus* (Asterales): CiPrx01
- *Liriodendron tulipifera* (Magnoliales): LtPrx01
- *Lycopersicon pennellii* (Solanales): LperPrx01
- *Mesembryanthemum crystallinum* (Caryophyllales): McPrx10
- *Physcomitrella patens* (mosses): PpaPrx01
- *Triphysaria versicolor* (Lamiales): TvePrx01

These organisms all had a "rank in EST" lower than 3. Whereas this discrepancy may be explained by the small EST libraries available (less than 10'000 ESTs in total), it is not the case for *Mesembryanthemum crystallinum* (27'348 EST) and *Physcomitrella patens* (174'908 EST). Concerning the latter organism, it is probable that it does not encode any ortholog to *AtPrx42*: the percentage of identity is very low (42%). Further search in primitive organisms (earlier than Spermatophyta, i.e. hornworts, liverworts, mosses and ferns) in NCBI and the PeroxiBase confirmed absence of *AtPrx42* orthologs (no identity higher than 46%, data not shown). This absence cannot be only explained by small EST databases, considering that two primitive organisms possess large EST librairies: *Physcomitrella patens* and *Marchantia polymorpha* (33'692 EST). The ancestral form of *AtPrx42* hence probably appeared together with the first seed plants (Spermatophyta).

Species	Peroxidase	# EST	total EST	ratio * 1000	%identity	score	e value	rank in species	rank in EST	total perox	Database
Acorus americanus	AaPrx01	7	9'695	0.63	79.00	531	e-152	1	1	6	NCBI
Allium cepa	AcPrx13	13	19'582	0.58	78.00	534	e-153	1	2	28	TIGR
Amborella trichopoda	AtrPrx09	2	8'706	0.20	79.00	214	e-16	1	3	13	NCBI
Antirrhinum majus	AmaPrx01	2	25'310	0.07	74.00	347	e-97	1	3=	13	NCBI
Apium graveolens var. dulce	AgPrx02	8	1'218	5.76	78.00	136	e-30	1	1	4	NCBI
Aquilegia formosa x pubescens	AfpPrx01	28	85'039	0.29	81.00	530	e-152	1	3	28	NCBI
Arabidopsis thaliana	AtPrx42	710	622'972	1.00		675	0		1	73	NCBI
Arachis hypogaea	AhPrx02	9	7'538	1.05	83.00	546	e-157	1	1	5	NCBI
Asparagus officinalis	AoPrx08	2	8'422	0.21	85.00	534	e-153	1	1	9	NCBI
Avicennia marina	AmPrx01	na	1'893	na	78.00	530	e-152	1	na	2	NCBI
Beta vulgaris	BvPrx03	21	22'347	0.82	77.00	535	e-153	1	1	29	NCBI
Brassica napus	BnPrx01	77	71'720	0.94	96.00	654	0	1	1	24	TIGR
Brassica oleracea	BoPrx42	5	13'425	0.33	95.00	538	e-155	1	3	47	NCBI
Brassica rapa	BrPrx42	22	8'639	2.23	96.00	654	0	1	1	43	NCBI
Capsicum annuum	CanPrx17	24	29'894	0.70	84.00	541	e-155	1	1	17	TIGR
Catharanthus roseus	CroPrx02	na	179	na	78.00	210	e-55	1	na	2	NCBI
Cicer arietinum	CarPrx01	na	724	na	82.00	542	e-155	1	na	2	NCBI
Cichorium intybus	CiPrx01	0	3'424	0.00	81.00	271	e-74	1	<3	5	NCBI
Citrus clementina	CclPrx01	32	61'404	0.46	81.00	560	e-161	1	1	8	NCBI
Citrus sinensis	CsPrx01	152	93'926	1.42	81.00	561	e-162	1	1	17	NCBI
Citrus x paradisi x poncirus	CppPrx01	49	7'954	5.41	82.00	561	e-161	1	1	14	NCBI
Cycas rumphii	CruPrx01	9	8'061	0.98	73.00	498	e-142	1	1	7	NCBI
Descurainia sophia	DsPrx01	4	1'024	3.43	96.00	649	0	na	na	1	NCBI
Elaeis guineensis	EgPrx01	3	2'411	1.09	71.00	225	e-60	na	na	1	NCBI
Eschscholzia californica	EcaPrx01	3	9'083	0.29	77.00	245	e-66	1	2	9	NCBI
Fragaria vesca	FvPrx02	13	10945	1.04	80.00	550	e-158	1	1	2	NCBI
Gerbera hybrida cv Terra Regina	GhyPrx01	18	15'851	1.00	79.00	553	e-159	1	1	5	SPUTNIK
Ginkgo biloba	GbPrx01	18	6'250	2.53	75.00	476	e-136	1	1	8	NCBI
Glycine max	GmPrx02	142	330'436	0.38	80.00	553	e-159	1	2	85	TIGR
Glycine max	GmPrx35	145	330'436	0.39	78.00	548	e-157	2	1	85	TIGR
Gossypium arboreum	GaPrx01	44	39'230	0.98	81.00	560	e-161	2	1	15	SPUTNIK
Gossypium hirsutum	GhPrx36	543	177'030	2.69	81.00	553	e-159	4	1	35	NCBI
Gossypium raimondii	GrPrx01	89	63'577	1.23	81.00	557	e-160	na	na	1	NCBI
Hedyotis centranthoides	HcPrx01	3	5'416	0.49	77.00	531	e-153	1	1	6	NCBI
Hedyotis terminalis	HtPrx01	2	4'875	0.36	73.00	387	e-109	1	1	3	NCBI
Helianthus annuus	HaPrx06	22	94'110	0.21	80.00	545	e-157	1	1	23	NCBI
Helianthus argophyllus	HarPrx01	58	12'787	3.98	80.00	546	e-157	1	1	16	SPUTNIK
Ipomoea nil	InPrx01	43	62'282	0.61	81.00	554	e-160	1	1	8	NCBI
Lactuca perennis	LperPrx05	37	29'125	1.11	78.00	546	e-157	1	1	29	NCBI
Lactuca saligna	LsPrx05	23	30'696	0.66	78.00	548	e-157	1	1	36	NCBI
Lactuca sativa	LsaPrx05	47	80'763	0.51	78.00	548	e-157	1	1	54	NCBI
Lactuca serriola	LsePrx05	31	55'490	0.49	78.00	548	e-157	1	1	24	NCBI
Lactuca virosa	LvPrx05	22	30'068	0.64	78.00	548	e-157	1	1	30	NCBI
Liriodendron tulipifera	LtPrx01	1	9'531	0.09	85.00	288	e-79	1	<3	10	NCBI
Lotus corniculatus var. japonica	LjPrx01	398	109'618	3.19	79.00	549	e-158	1	1	70	TIGR
Lupinus albus	LaPrx11	3	2'128	1.24	74.00	447	e-127	1	1	17	NCBI
Lycopersicon esculentum	LePrx17	93	213'947	0.38	80.00	542	e-156	1	1	65	TIGR
Lycopersicon pennellii	LpPrx01	1	8'346	0.11	76.00	284	e-78	1	<3	8	NCBI
Malus domestica	MdPrx02	151	253'992	0.52	81.00	546	e-157	1	1	6	NCBI
Manihot esculenta	MePrx05	4	17'954	0.20	83.00	327	e-91	1	2	10	NCBI
Medicago sativa	MsPrx10	27	6'534	3.63	82.00	528	e-151	1	1	15	NCBI
Medicago truncatula	MtPrx47	375	226'923	1.45	84.00	545	e-156	1	1	94	TIGR
Mesembryanthemum crystallinum	McPrx10	2	27'348	0.06	75.00	330	e-92	1	<3	13	NCBI
Nicotiana tabacum	NtPrx60	50	73'847	0.59	79.00	539	e-155	1	1	11	NCBI
Nuphar advena	NaPrx02	17	10'208	1.46	84.00	543	e-156	1	1=	6	NCBI
Ocimum basilicum	ObPrx02	7	23'232	0.26	75.00	363	e-102	1	1	2	NCBI
Phaseolus coccineus	PcoPrx03	118	20'120	5.15	79.00	550	e-158	1	1	8	SPUTNIK
Phaseolus vulgaris	PvPrx11	18	21'346	0.74	78.00	547	e-157	1	1	11	NCBI
Physcomitrella patens	PpaPrx01	29	174'908	0.15	42.00	270	e-74	1	<3	33	NCBI
Picea sitchensis	PsiPrx01	2	80'789	0.02	77.00	501	e-144	1	na	1	NCBI
Pinus pinaster	PpPrx01	29	27'283	0.93	75.00	504	e-144	1	1	28	SPUTNIK
Pinus taeda	PtaPrx01	155	329'469	0.41	75.00	504	e-144	1	1	52	NCBI
Poncirus trifoliata	PtriPrx01	21	28'737	0.64	82.00	561	e-162	na	na	1	SPUTNIK
Populus alba	PalPrx01	15	10'446	1.26	81.00	559	e-161	1	1	5	SPUTNIK
Populus balsamifera	PbPrx01	29	30'296	0.84	80.00	550	e-158	2	1	23	SPUTNIK
Populus euphratica	PePrx01	19	13'905	1.20	79.00	551	e-158	1	1	15	SPUTNIK
Populus tremula	PtrePrx01	46	37'313	1.08	80.00	558	e-161	1	1	23	SPUTNIK
Populus tremuloides	PtremPrx01	39	12'813	2.67	81.00	531	e-152	1	1	15	NCBI
Populus trichocarpa	PtPrx02	na	na	na	80.00	556	e-160	1	na	116	JGI
Prunus armeniaca (apricot)	ParPrx01	3	15'105	0.17	83.00	291	e-81	na	na	1	NCBI
Prunus persica (peach)	PpePrx03	1	66'249	0.01	79.00	151	e-37	1	na	10	NCBI

Species	Peroxidase	# EST	total EST	ratio	%identity	score	e value	rank in species	rank in EST	total perox	Database
Ribes americanum	RaPrx01	4	2'238	1.57	77.00	385	e-108	1	1	6	NCBI
Saruma henryi	ShePrx01	na	10'274	na	84.00	327	e-91	1	na	5	NCBI
Solanum tuberosum	StPrx01	81	219'917	0.32	80.00	543	e-156	2	1	71	NCBI
Stevia rebaudiana	SrPrx01	11	5'548	1.74	80.00	546	e-157	1	1	4	NCBI
Thellungiella halophila	ThPrx01	2	1'739	1.01	96.00	505	e-144	1	1	2	NCBI
Thellungiella salsuginea	TsPrx01	24	6'537	3.22	97.00	656	0	na	na	1	NCBI
Theobroma cacao	TcPrx06	1	6'581	0.13	77.00	232	e-62	1	na	7	TIGR
Thlaspi caerulescens	TcaPrx01	1	4'289	0.20	97.00	384	e-108	na	na	1	NCBI
Triphysaria versicolor	TvePrx01	2	9'305	0.19	83.00	269	e-73	1	<3	7	TIGR
Vigna unguiculata	VuPrx01	na	336	na	85.00	236	e-63	na	na	1	NCBI
Vitis vinifera	VvPrx08	33	191'616	0.15	81.00	546	e-157	1	1	22	TIGR
Yucca filamentosa	YfiPrx01	2	1'743	1.01	79.00	430	e-122	na	na	1	NCBI
Zantedeschia aethiopica	ZaPrx02	1	4'283	0.20	84.00	179	e-43	1	na	3	NCBI
Poaceae (cereals)											
Aegilops speltoides (wild wheat)	AspPrx01	1	4'315	0.20	32	97	e-19	1		1	NCBI
Hordeum vulgare (barley)	HvPrx14	19	370'546	0.04	38	202	e-53	1		104	TIGR
Oryza sativa ssp. japonica (rice)	OsPrx116	7	1'163'134	0.01	51	342	e-95	1		138	TIGR
Oryza sativa ssp. japonica (rice)	OsPrx61	158	1'163'134	0.12	31	130	e-28		1	138	TIGR
Saccharum officinarum (sugar cane)	SofPrx10	1	246'301	0.00	54	168	e-40	1		10	NCBI
Sorghum bicolor	SbPrx116	5	208'197	0.02	51	331	e-92	1		133	TIGR
Triticum aestivum (wheat)	TaPrx67	15	855'066	0.02	51	341	e-95	1		105	TIGR
Triticum monococcum (wild wheat)	TmPOX7	1	10'139	0.09	34	187	e-49	1		11	NCBI
Zea Mays (maize)	ZmPrx108	3	695'811	0.00	37	210	e-55	1		129	TIGR
Other monocots (no AtPrx42)											
Ananas comosus	AncPrx01	3	5'649	0.47	34	173	e-41	1		1	NCBI
Curcuma longa	CloPrx01	2	12'593	0.14	55	381	e-107	1		1	NCBI
Zingiber officinale	ZoPrx01	5	38'083	0.12	55	373	e-105	1		1	NCBI

Table 4-1

***AtPrx42* orthologous genes.**

EST: number of ESTs found for each peroxidase listed

total EST: total number of ESTs in each organism (last update: September 8th, 2006)

ratio: ratio total EST/# EST (normalised to 1 for AtPrx42)

rank in species: position of the peroxidase compared to the other peroxidases within the same organism, according to the percentage of identity

rank in EST: position of the peroxidase compared to the other peroxidases within the same organism, according to the number of ESTs

total perox: total number of peroxidases of each organism loaded in the PeroxiBase database

Database: database used to collect information about EST counts

na: not applicable (not enough available data)

1=, 3=: 1st or 3rd rank *ex aequo*

* % identity, score, e value and rank in species all refer to the comparison of the peroxidase protein with AtPrx42.

* Populus trichocarpa: genomic DNA data (http://www.jgi.doe.gov).

* Monocotyledonous species are shaded in gray.

* For cereals and other monocotyledons, “rank in EST” field was not analysed (except for rice).

AtPrx42 orthologs are also absent from a particular group of seed plants: the cereals (Poaceae). The highest identity found within complete sequences of rice, sorghum and wheat was 51%, and 54% for a partial sequence in sugar cane (Table 4-1). In maize and barley, identities even fell down to, respectively, 37 and 38%. When the peroxidase with the highest EST count was searched in rice (OsPrx61), the level of identity also fell to a very low percentage (31%). Considering that other monocotyledons (Liliopsida) possess *AtPrx42* paralogs (*Acorus*, *Allium*, *Asparagus*, *Yucca*, *Zantedeschia*, *Elaeis*), Poaceae have lost this peroxidase during evolution. In order to better identify the plants that do not encode *AtPrx42*, I further screened EST libraries of other groups taxonomically close to Poaceae (Fig. 4-2). I found that, like Poaceae, a member of the family Bromeliaceae (*Ananas comosus*) and two Zingiberales (*Curcuma longa* and *Zingiber officinale*) had no *AtPrx42* ortholog: the closest peroxidases had only 55% of identity and not a particularly high EST count (see "ratio" in table 4-1). In Arecales, *Elaeis guineensis* probably possesses an *AtPrx42* ortholog: *EgPrx01*. This peroxidase has quite a high identity (71%), although it is the lowest identity found among higher plants. Even gymnosperms (*Picea*, *Pinus*, *Ginkgo* and *Cycas*), which are the most distantly related organisms from *Arabidopsis thaliana*, show between 73% and 77% identity with *AtPrx42*. The EST database of *Elaeis guineensis* is however still too low to confirm or not a high expression level. The relatively low level of identity of *EgPrx01* suggests that in Arecales, *AtPrx42* orthologous genes started undergoing through a series of mutations resulting in a degradation of the original "AtPrx42-like" sequence. This would mean that *AtPrx42* orthologs in Zingiberales and Commelinales probably disappeared through degradation of the "AtPrx42-like" sequence rather than through loss of the gene following, for instance, genome rearrangements or transposon insertions.

A last peculiarity observed was presence of two *AtPrx42* orthologous genes in soybean (Glycine max: GmPrx02 and GmPrx35). Both genes were highly expressed; they differ only in their peptide signal and their C-terminal sequence. It is the first case reported of a duplication of an *AtPrx42* ortholog.

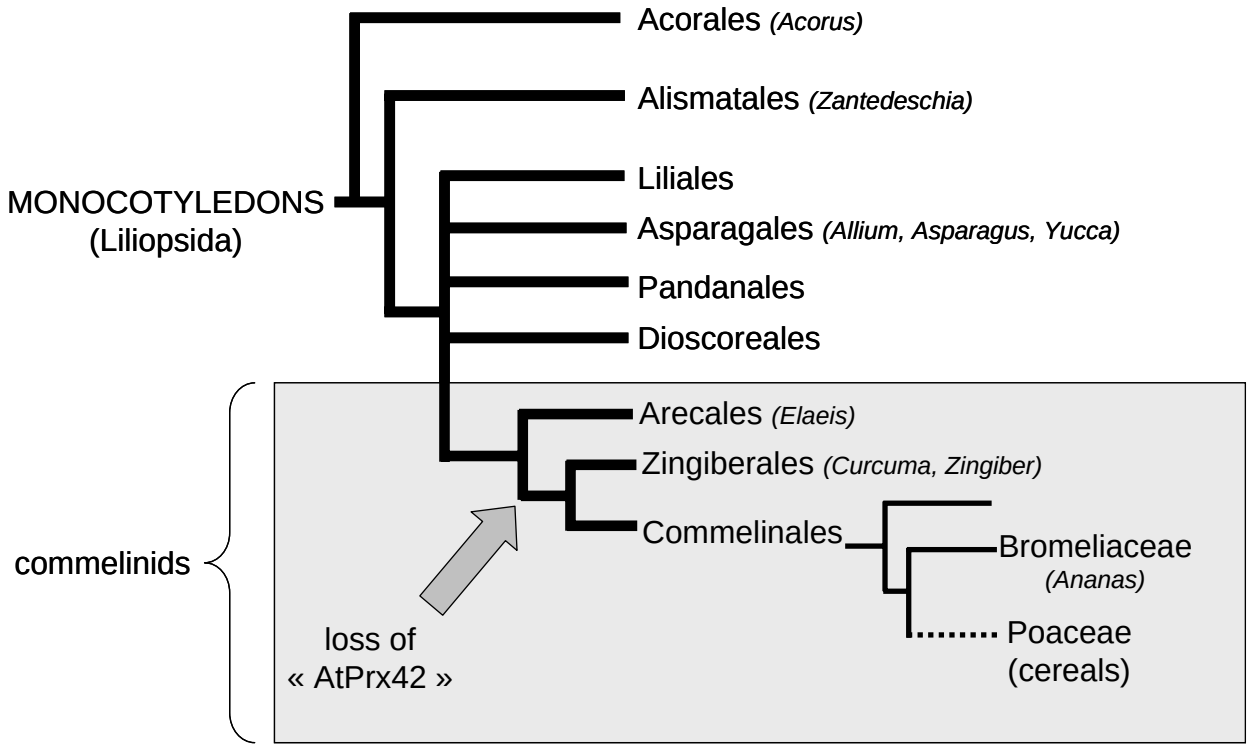


Figure 4-2

Taxonomical classification of Monocotyledons (Liliopsida).

From http://www.tolweb.org (tree of life project) and Soltis *et al.*, 2000.

B.2 Phylogenetic relationships among AtPrx42 *orthologs*

Neighbor-Joining (NJ), maximum-parsimony (MP) and maximum-likelihood analysis of AtPrx42 orthologous proteins gave quite similar results. The trees are rather devoid of evolutionary peculiarities: generally, AtPrx42 orthologs follow taxonomical distribution of species (Fig.4-3). Peroxidases belonging to organisms of the same orders or families (e.g. Fabaceae, Brassicaceae, Asparagales) cluster together. Isolated peroxidases (e.g. EcaPrx01, CroPrx02, ObPrx02) are the only representatives of other specific families. Among average to weakly supported nodes:

- node 1: *Nuphar advena* (NaPrx02) is an ancestor of all flowering plants, whereas the family of Brassicaceae is a more recent group among the subclass of eurosids. This node is only supported by MP analysis. Despite extremely weak bootstrap values for the two other computing methods, NaPrx02 was still appearing at the base of the

Brassicaceae in the NJ and ML trees obtained. This relationship is a unique proof of convergent evolution among AtPrx42 orthologs.

- nodes 2 to 7: closely related organisms. Weakness of the node is mainly due to presence of partial sequences (4, 6, 7); different families (node 2: InPrx01 belongs to Convolvulaceae, and not to Solanaceae); different subfamilies (node 5); or simply divergent evolution (node 3).
- node 8: putative AtPrx21 orthologs.

The peroxidase PpaPrx01, from the moss *Physcomitrella patens*, clearly does not appear as an AtPrx42 ortholog, and it does not seem either to be a basal (ancestral) AtPrx42 peroxidase. Cereals, *Ananas* and Zingiberales also form outgroups, as expected, as well as AtPrx21.

The main conclusion issued from this study is the striking degree of conservation of AtPrx42 orthologous proteins, even between extremely distant species. Within each organism, such a high conservation is rarely observed between paralogs: the closest AtPrx42 paralog, AtPrx21, already comes out of the tree as an outgroup. Evolution has retained during millions of years the intrinsic sequence of AtPr42 orthologs, and probably also some structural features, for a still unknown reason. In our laboratory, two collaborators (Jan Dobias, Marion Koch) have tried to isolate *atprx42* knock-out mutants from *Arabidopsis* T-DNA libraries, but the seeds obtained gave plants that were still expressing *AtPrx42* at a high level. T-DNA was either inserted always in the same intron (third intron), but the plant was still able to normally splice out the intron, or T-DNA was inserted in the promoter region, which was not sufficient to completely turn off *AtPrx42* expression. AtPrx42 seems then to be an essential peroxidase during germination, and knocking it out is probably lethal. Considering that *AtPrx42* probably appeared with the first seed plants (Spermatophyta), we could speculate that as long as early plants did not have an ancestral sequence of *AtPrx42*, they could not form viable seeds. A question however still remains: why have cereals, *Ananas* and Zingiberales lost *AtPrx42*? How could they form viable seeds? Is this simply due to the fact that all organisms belonging to these groups are cultivated plants and that *AtPrx42* was lost through artificial selection? Although this possibility cannot be discarded, a BLAST search through EST libraries of the wild wheat species *Triticum monococcum* (10’139 EST) and *Aegilops speltoides* (4’315 EST) showed that the closest *AtPrx42* orthologous EST sequences had, like all cereals, a very low identity level (Table 4-1). In order to better understand the role of AtPrx42, Marion has started making a transgenic construct bearing the *AtPrx42-linker-24xrPtA* RNAi sequence under the control of an estradiol-inducible promoter (Zuo et al., 2000). Hopefully this strategy will generate for the first time an *atprx42* line.

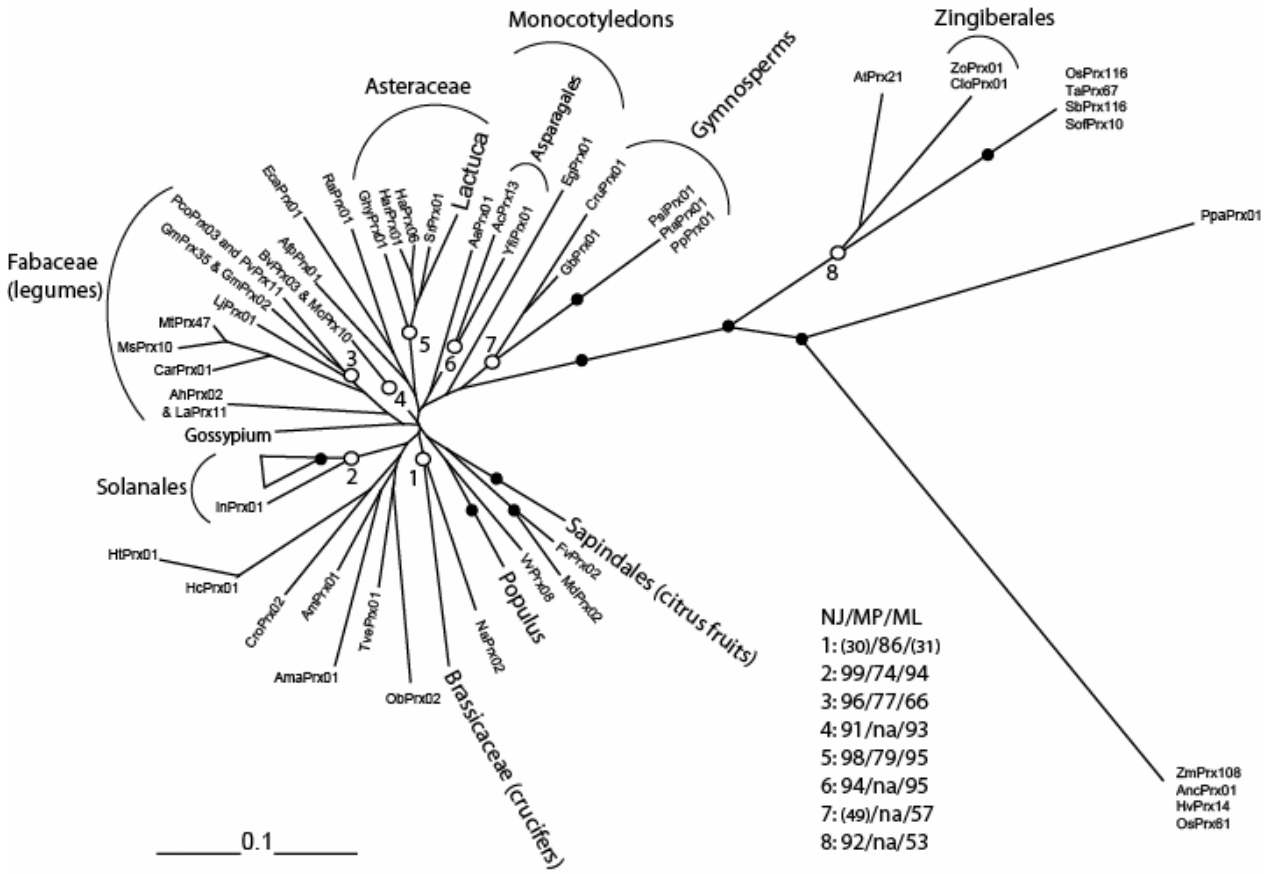


Figure 4-3

Simplified view of a NJ tree comprising AtPrx42 orthologs.

The closest AtPrx42 orthologous peroxidases from cereals, *Ananas*, Zingiberales and AtPrx21 were added as outgroups.

Black dots on nodes represent bootstrap support higher than 95 with distance ("NJ"), MP and ML methods.

White circles on nodes have lower bootstrap values, which are listed next to the tree. MP bootstraps were obtained on complete sequences only: when the cluster contains partial sequences, bootstrap calculation is then not applicable ("na").

Unsupported nodes do not bear any label.

Sapindales: PtriPrx01, CppPrx01, CsPrx01, CclPrx01

Populus: PtPrx02, PalPrx01, PtrePrx01, PbPrx01, PePrx01

Brassicaceae : AtPrx42, DsPrx01, TsPrx01, BoPrx42, BrPrx42, BnPrx42

Solanales : InPrx01, NtPrx60, CanPrx17, StPrx01, LePrx17

Gossypium : GaPrx01, GhPrx36, GrPrx01

Lactuca : LperPrx05, LsPrx05, LsaPrx05, LvPrx05

The scale bar represents 0.1 substitution per site.

B.3 Identification of a motif specific to the AtPrx42 *promoter*

The protein sequence of AtPrx42 orthologs has been well preserved during evolution. The next step was to screen for the promoter sequence of the peroxidase-encoding genes. As most of sequences were obtained through EST libraries, very few contained a significant amount of 5'UTR sequence. Accessions containing at last 75bp of promoter sequence are listed in Table 4-2. Concerning other peroxidases, 5'UTR sequences were found within databases, however they were rather short (less than 30 bp). For these peroxidases, I ordered the clones containing the 5'UTR, in order to hopefully obtain a longer sequence. I received only two clones. They were cultured and subsequently sequenced with appropriate primers on the vector plasmid sequence (Table 4-3). Whereas clone CA411339 (LaPrx11) gave no additional information, I was able to add 32 new base pairs to clone CV034771 (PtaPrx01), which was, unfortunately, not sufficient for promoter analysis.

Table 4-2
GenBank accessions of ESTs containing promoter sequences (>75bp).

genomic DNA : obtained from genome sequencing projects.

AtPrx42	AfpPrx01	BvPrx03	CanPrx17	GhPrx36
genomic DNA	DR945226	CK136611	CA523818	CD485878
GmPrx35	GrPrx01	InPrx01	LePrx17	McPrx10
AW706845	CO096540	BJ562022	BG628012	BE037041
MtPrx47	ObPrx02	PcoPrx03	PtPrx02	PtrePrx01
CA918349	DY342870	CA909365	genomic DNA	BU888470
PvPrx11	StPrx01	VvPrx08		
CV538186	BG597413	CB974099		
AtPrx21	PpaPrx01	OsPrx61	OsPrx116	
genomic DNA	BJ191820	genomic DNA	genomic DNA	

Table 4-3
Details concerning the two clones obtained containing 5' UTR sequences of *AtPrx42* orthologous genes.

Peroxidase	Clone (GenBank accession)	Vector	Antibiotic resistance	Primers used for sequencing	Provider	Additional sequence obtained (5' - 3')	Final length of promoter obtained
LaPrx11	CA411339	pBluSK(+)	Ampicillin	T3/T7	Dr C. Uhde-Stone, Minnesota, USA	none	17bp (=)
PtaPrx01	CV034771	pSL1180	Ampicillin	M13-rev(-48) M13-fwd(-20)	Dr M.Cordonnier-Pratt, Georgia, USA	ACGAGGCCTCGTGC CGAATTCGGCACGAG G	32bp (+30)

Alignment of promoter sequences was then performed, in parallel with a search of highly conserved regions with the MEME software (http://meme.sdsc.edu). Both methods allowed detecting a specific motif, located between 31 (InPrx01) and 92 (AtPrx42) base pairs from the START codon of the peroxidase gene (Fig. 4-4): TTATTTAAGxAG. When searching in the PLACE (http://www.dna.affrc.go.jp/PLACE/) web signal scan software (Prestridge, 1991), the closest sequences found were two TATA boxes: "TATABOX5" (TTATTT) and "TATABOXOSPAL" (TATTTAA). The first motif (Tjaden et al., 1995) was found in *Pisum sativum* (pea), very close (-27 to -21) to the START codon; deletion mutant studies determined that this element was essential for gene expression, as expected from a TATA box. Regarding the second motif (Zhu et al., 1995), it belongs to the phenylalanine ammonia-lyase (*pal*) gene of rice, is located from -34 to -28, and is also essential for translation initiation. We can hence suppose that the "AtPrx42 motif" TTATTTAAGxAG is a particular form of TATA box. In addition, no other TATA box was found closer to the START codon than this motif in all promoters available for *AtPrx42* orthologs. Conservation of the sequence "AAGxAG" is quite remarkable, as these nucleotides are not known to be necessary in TATA boxes (Smale and Kadonaga, 2003; Grace et al., 2004). When adding promoters from *AtPrx42* non-orthologous genes, this motif was not found. A screening of all *Arabidopsis thaliana* promoters (available in our laboratory) revealed that a similar (TTATTTAAGgtt) motif was only present in *AtPrx08*, but at a far longer distance from the START codon (813 bp). The sequence of this motif and its position close to the START codon seem hence to be unique to promoters of *AtPrx42* orthologous genes.

One exception was however found for two *AtPrx42* orthologs: *BvPrx03* and *McPrx10*. These peroxidases are encoded, respectively, by *Beta vulgaris* and *Mesembryanthemum crystallinum*: both organisms are members of the order Caryophyllales, a sister group to asterids (Figure 4-1) (Soltis et al., 2000). Whereas *BvPrx03* still bears some degree of conservation, *McPrx10* has lost the motif. A screening of both promoters revealed that no other TATA box was present up to 150 bp from the START codon. When looking at the EST count of each peroxidase, we can see that *BvPrx03* is still the most expressed peroxidase in *Beta vulgaris*; in contrast, the peroxidase *McPrx10* is almost not expressed at all. On the other hand, the protein encoded has preserved a high degree of identity (75%) with the other AtPrx42 orthologs. This observation indicates a possible role of the "AtPrx42 motif" for the high and constitutive expression of *AtPrx42* orthologous genes. Complete loss of TATA box in *McPrx10* promoter probably explains its low expression level. *Beta* and *Mesembryanthemum* are the only two representatives of Caryophyllales having enough EST

Figure 4-4 (previous page)

Alignment of promoters from *AtPrx42* orthologous genes.

The promoter sequences from the non-orthologs *AtPrx21*, *OsPrx61*, *OsPrx116* and *PpaPrx01* have been added as outgroups.

Dark grey: 80% conservation; light grey: 60% conservation.

sequences available. In this respect, it would be extremely interesting screening for more ESTs from this family, in order to check whether these plants are undergoing a loss of *AtPrx42*, or at least of its constitutive expression, and hence gain information on the function and regulation of this peroxidase in plants.

Another interesting sequence to obtain would be the promoter sequence of *GmPrx02*. As previously mentioned, *Glycine max* is the only organism found that shows a duplication of an *AtPrx42* ortholog (*GmPrx02* and *GmPrx35*). Unfortunately, only the promoter sequence of *GmPrx35* was long enough to detect the "AtPrx42 motif". Alignment of the 21 available nucleotides of *GmPrx02* promoters with *GmPrx35* is shown hereafter:

```
GmPrx02 :  CCCACGCGTCCGGCGAAGAGA
GmPrx35 :  GAAGAGCGAAGAGTGAAGAGA
```

This alignment indicates that, despite a recent duplication event, the two promoters have already diverged quite importantly. Genome sequencing of soybean is already under way (http://soybase.agron.iastate.edu/): obtaining a longer sequence of *GmPrx02* promoter will allow verifying if, in a rapidly diverging promoter, the putative "AtPrx42 motif" is still conserved. Expression of both peroxidases is indeed still very high, and presence of the sequence TTATTTAAG in *GmPrx02* would add substantial evidence for a role of this new motif in strong and constitutive expression.

B.4 Materials and methods

B.4.1 Database mining

AtPrx42 orthologous genes were first searched in the PeroxiBase database (http://peroxidase.isb-sib.ch). The protein sequence of AtPrx42 was used as a template for a BLASTp search in the database. For every plant organism, the first peroxidase to appear in the resulting window was selected. A parallel BLASTp search was performed with the closest *AtPrx42* paralog: *AtPrx21*. If the candidate sequence found showed higher identity with AtPrx21, then it was discarded. Additionally, the NCBI database was screened for more orthologs through BLASTp and tBLASTn searches (with the "database" options: nr or est_others), and any new candidate found was entered into the PeroxiBase. Information was also collected on 5' and 3' UTR sequences. The number of EST sequences was determined through a tBLASTn on NCBI, TIGR (www.tigr.org) or SPUTNIK (http://sputnik.btk.fi).

For *Populus trichocarpa*, the total number of EST sequences is not available: the NCBI dbEST database reports 89'943 EST sequences, however they consist in a mixture of *Populus trichocarpa*, *Populus trichocarpa x deltoides* and *Populus trichocarpa x nigra*. The sequence of PtPrx02 comes from genomic DNA sequencing (JGI website: http://www.jgi.doe.gov/) and cannot be found on NCBI with a tBLASTn search.

B.4.2 Identification of a common motif for the AtPrx42 promoter

5' UTR sequences of *AtPrx42* and its putative orthologous genes were submitted to the motif detection software MEME (http://meme.sdsc.edu) for prediction of a common motif, with options "zero or one motif per sequence" and "search given strand only". Alignments of promoters were performed with Clustal W (http://www.ebi.ac.uk/clustalw), and manually corrected with BioEdit software v 7.0.4.1 (Hall, 1999). Promoters of *BvPrx03* and *McPrx10* were further re-aligned side by side through Clustal W for more accuracy. Graphical representation was finally improved with GeneDoc software v.2.6.002 (Nicholas and Nicholas Jr., 1997).

B.4.3 Phylogenetic trees

When computing phylogenetic trees, some sequences shown in Table 4-1 had to be removed, mainly because of the too short fragment available. Only partial sequences containing a significant proportion of the peroxidase 5' part were considered: they had no

overlap with short 3' sequences, hence the latter ones were discarded. A first alignment was performed with ClustalW and manually checked. Distance (Neighbor-Joining option, JTT substitution frequency matrix) and maximum-parsimony (MP) trees were obtained through the Phylip package v.3.65 (Felsenstein, 2004), with 100 bootstrap replicates. Concerning MP trees, only complete peroxidase sequences were analysed (if partial sequences are entered, they tend to cluster together). Maximum-likelihood trees were computed with the PhyML software v.2.4.4 (Guindon and Gascuel, 2003) and the following parameters: 100 bootstrap replicates, 4 categories of substitution rates and estimated gamma parameter.

B.4.4 Sequencing of peroxidase-containing clones

The clones were received in solid LB medium. They were cultured overnight on Petri dishes with antibiotic selection according to the provider's indications. The day after, colonies were screened by PCR reactions with the appropriate primers, and positive clones were cultured overnight under agitation in 5ml liquid LB medium. Purified plasmids were sequenced with primers T3, T7, M13 rev(-48) or M13 fwd(-20). Once sequencing results were obtained (Fasteris Life Sciences, Geneva), vector contaminating sequences were trimmed by VecScreen analysis (http://www.ncbi.nlm.nih.gov/VecScreen/). Results were then communicated to each provider.

C. Deletion and mutation variants of the *AtPrx42* promoter

C.1 Strategy

In order to better characterise the function of the putative "AtPrx42 motif", two deletion mutants of the *AtPrx42* promoter were designed for cloning into the binary vector pCAMBIA 1281Z close to the glucuronidase (GUS) gene. In addition, an *AtPrx42* promoter bearing four point mutations in the "AtPrx42 motif" was designed for cloning into the same binary vector (Fig. 4-5).

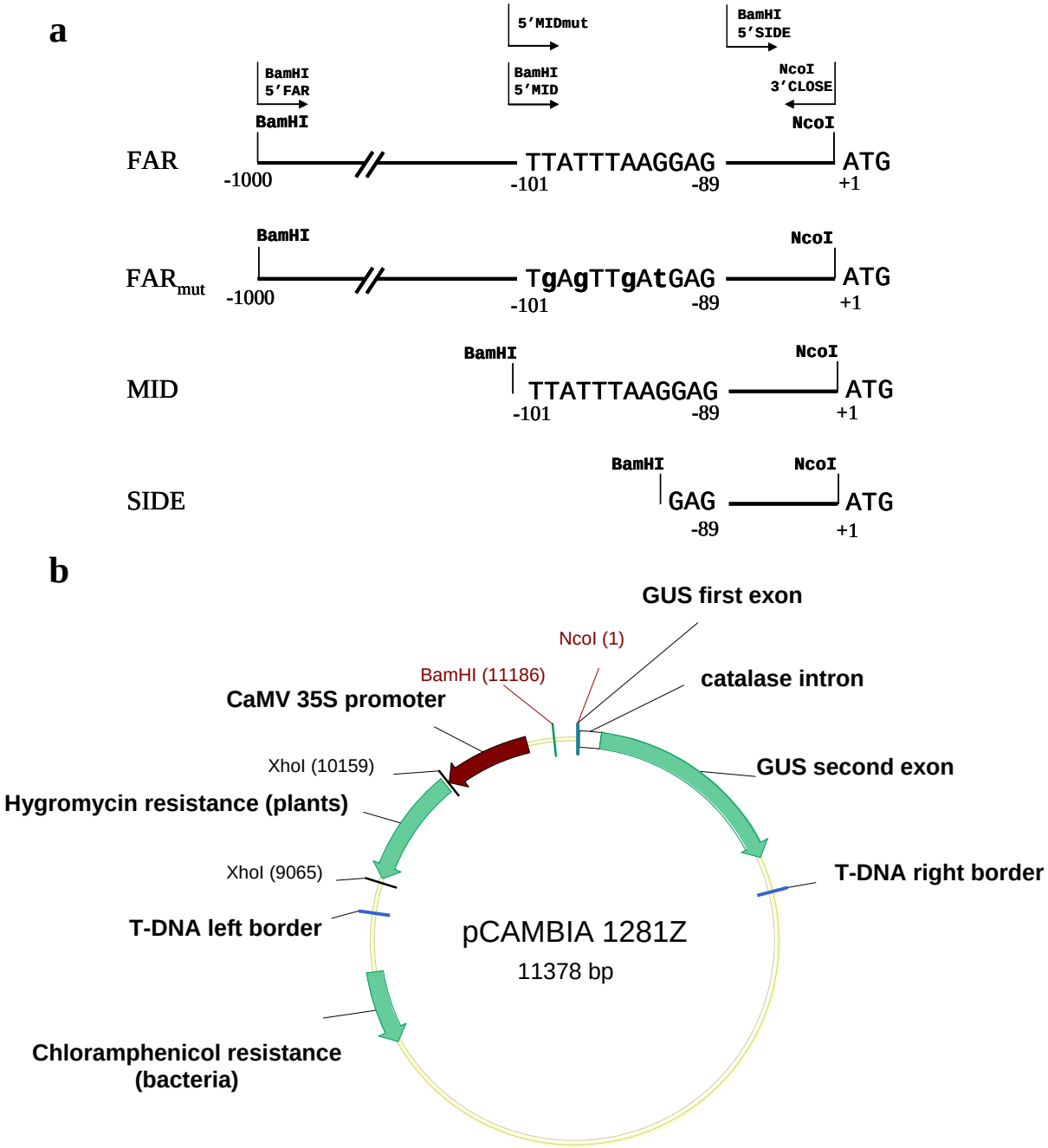


Figure 4-5 (previous page)

Strategy designed for the characterisation of the putative "AtPrx42 motif".

(a) Schematical representation of the *AtPrx42* promoter ("FAR"), two deletion mutants ("MID", "SIDE") and one full promoter bearing four point mutations ("FAR_{mut}"). Position of the primers used is shown above the constructs.

(b) Simplified map of the binary vector pCAMBIA 1281Z (NCBI accession AF234294).

pCAMBIA 1281Z possesses a promoterless *GUS* gene interrupted by a catalase intron. Presence of the intron ensures selective expression by eukaryotic organisms. The hygromycin resistance gene is under the control of a CaMV 35S promoter for expression in plants. T-DNA left and right border repeats delimit the region to be transferred into the plant genome by *Agrobacterium*-mediated infection. This plasmid can also be used for blue/white colony selection (see Appendix).

In order to target mutations to four specific nucleotides in the AtPrx42 promoter (FAR_{mut}), the "megaprimer" approach was used (Sarkar and Sommer, 1990): first, a PCR primer (5'MID_{mut}) was designed that hybridises to the "AtPrx42 motif", but with the sequence T**g**A**g**TT**g**A**t**GAG instead of the wild type sequence. This primer was used for amplifying the fragment comprised between the "AtPrx42 motif" and the peroxidase START codon. Once obtained, a second round of PCR amplification was performed by using this time the megaprimer as a classical primer to amplify the full 1000bp of the *AtPrx42* promoter.

The goal of the above described strategy is to assess the importance of the "AtPrx42 motif" in gene expression. The MID construct will allow knowing if presence of the "AtPrx42 motif" alone is sufficient for constitutive expression. Moreover, it will give information on presence of regulator sequences upstream from the motif. The FAR_{mut} construct should confirm the essential role of this motif as a TATA box. If GUS expression is abolished in FAR_{mut}, then it will be interesting to specifically mutate the residues AAGGAG, downstream from the classical TATA motif. Conservation of these residues is indeed quite intriguing. The SIDE construct is rather a control, which should yield no, or a low basal level, of GUS expression.

Importantly, it will be necessary to obtain several independent *Arabidopsis* transformant lines for each construct in order to compare them. In previous experiments indeed (chapter 2, figure 2-4), I have noticed that GUS expression is rather variable between different transformants. Another caveat should be considered: studies on pCAMBIA 1281Z

performed by the pCAMBIA manufacturer (http://www.bios.net/daisy/cambia) revealed that presence of the 35S promoter (driving the plant selection gene) provokes a significant level of GUS activity, even though it is not oriented towards the GUS gene. A control transformation with empty pCAMBIA 1281Z should hence be performed.

C.2 Results

The details of the cloning procedure are described in the "Materials and methods section". Despite generation of the different constructs and subsequent cloning into pCAMBIA 1281Z were successfully achieved, I was not able to obtain any *Arabidopsis* transformant. Two different lines of *Agrobacterium* were used: GV3101/2 (gentamicin^R) and LBA4404 (streptomycin^R). Both lines were efficiently transformed by all four pCAMBIA constructs. The only concern was that growth of GV3101/2 on Petri dishes was rather "smeary"-like, with few isolated colonies. Selection of *Arabidopsis* seeds was done on an optimised concentration of hygromycin B (25 μg/ml), following preliminary tests with wild types and hygro^R positive controls (gift of Maria Perez-Gonzalez). Transformation with the LBA4404 *Agrobacterium* line resulted in hygromycin-resistant seedlings (~25 per construct), which were analysed by Marion for GUS activity. However, Marion has not been able to detect any activity for these seedlings with the MUG or the X-Gluc substrates.

The failure to obtain pCAMBIA-transformed *Arabidopsis* plants is rather puzzling. A possible explanation could be a transgene-induced silencing of endogenous *AtPrx42* expression. If we assume that *AtPrx42* has an essential role in seed formation and/or seed germination, then such a silencing would be lethal. When searching on the TIGR database for *AtPrx42* EST sequences, it appears that at least 50 nucleotides from the promoter are transcribed into RNA. These nucleotides could be the source of double-stranded RNA molecules, synthesised by *Arabidopsis* RNA-dependent RNA polymerase, and known to trigger silencing (Napoli et al., 1990; Fire et al., 1998; Dalmay et al., 2000). As previously mentioned, even if the *AtPrx42* promoter is inactive, there is always some level of expression of the GUS gene due to vicinity of the strong 35S promoter: *AtPrx42* promoter RNA molecules can hence be synthesised in all four constructs used. So far, however, no study has examined if silencing could be triggered since the very first moments of pollinisation, during seed formation or just before germination. In line with my results, Jennifer Minardi, from our laboratory, has made an RNAi construct destined to silence as many peroxidases as possible

in *Arabidopsis*. Despite a considerable work and several rounds of selection, she has not been able to obtain *Arabidopsis* transformants: she might as well have generated *AtPrx42*-silenced plants.

C.3 Materials and methods

C.3.1 Generation of point mutations by the megaprimer technique

In order to obtain targeted mutations in the putative "AtPrx42 motif", the AtPrx42 promoter was amplified from *Arabidopsis* genomic DNA (40 cycles, T_m = 55°C) with the promoter 5'MID_{mut} p42 (5'-CATTTCCCTT**g**A**g**TT**g**A**t**GAGGCTGCA-3') and NcoI 3'CLOSE p42 (5'-CATGCCATGGCATCACACCTTTGCCTC-3'). The fragment hence obtained ("megaprimer" or M_m) was gel-purified and precipitated in 2.5 volumes EtOH 100%, 0.1 volume NaAc 3M, glycerol 2% and $MgCl_2$ 0.01M: this last substance improves precipitation yield of small DNA fragments (Sambrook et al., 1989). The megaprimer was then used in a second PCR reaction on genomic DNA with the primer BamHI 5'FAR p42 (5'-CGGGATCCATATTATTTGTTATAGCC-3') in order to obtain the final FAR_{mut} construct. No amplification was however achieved, and the following improvements had to be applied in order to obtain the FAR_{mut} construct:

- the megaprimer was amplified from a pGEM-T::M_m template (instead of genomic DNA)
- the PCR product obtained was cleaned with a GenElute® PCR Clean-up Kit (Sigma)
- the megaprimer fragment was gel purified with a GenElute® Gel Extraction Kit (Sigma)
- the PCR reaction BamHI 5'FAR p42-M_m was performed with a miniprep pCAMBIA-FAR as a template (200ng per reaction) and approximately 50ng of megaprimer. The PCR mixture was incubated at 95°C during 5 minutes before starting amplification (40 cycles, T_m = 56°C)

If any one of these steps was bypassed or modified, PCR reactions with primers BamHI 5'FAR p42 and the megaprimer were not giving any product.

C.3.2 Generation of inserts and cloning into pCAMBIA 1281 Z

Amplification of genomic DNA or pGEM-M_m with primers BamHI 5'FAR p42, BamHI 5'MID p42 (5'-CGGGATCCCTTTATTTAAGGAGGCTG-3'), BamHI 5'SIDE p42 (5'-CGGGATCCGAGGCTGCATTGCTTGCT-3') and NcoI 3'CLOSE p42 generated the designed constructs that were directly inserted into pCAMBIA 1281Z through BamHI and NcoI digestion and subsequent T4 DNA ligase-mediated ligation (Promega). The obtained pCAMBIA-F, -M, S and F_m plasmids were then cloned in *E.coli* DH5α strain.

C.3.3 *Agrobacterium*-mediated transformation of *Arabidopsis*

pCAMBIA-F, -M, S and F_m plasmids were cloned in *Agrobacterium tumefaciens* strains GV3101/2 or LBA4404. Plants were subsequently transformed by the spraying technique (see appendix). Selection of transformants was made on MS agar with hygromycin B (Sigma) at 25μg/ml.

5

Beyond classes : the peroxidase superfamily

A. Introduction

The PeroxiBase database has first been intended to be a repository of all existing class III peroxidases. After most of green plant genomes and EST data had been searched for class III peroxidases, the decision was taken to collect all the other peroxidases belonging to the "plant, fungal and bacterial" peroxidase superfamily (Welinder, 1992). This new task would hopefully increase the number of researchers interested in using the PeroxiBase. More importantly, as for class III peroxidases, the two other classes of peroxidases are spread among a wide variety of species: the PeroxiBase would then be a powerful tool for screening hundreds of class I and II peroxidases and performing many exciting studies. In this respect, I have been involved in the two following projects, whose results have been recently submitted to scientific journals:

1) to screen all available bacterial genomes for catalase-peroxidases

2) to study the transmission mechanisms of the various classes of peroxidases among living organisms

B. Article project: “Phylogenetic distribution of catalase-peroxidases in bacteria: are there patches of order in chaos?”

The first project included an exhaustive screen of bacterial genomes for presence but also absence of catalase-peroxidases. Various surprising observations came out from this study, such as the almost random distribution of catalase-peroxidases among bacterial species, the presence of catalase-peroxidases in other reigns (protists, green plants), and the absence of non-duplicated forms of catalase-peroxidase sequences. This work is detailed in the article project “**Phylogenetic distribution of catalase-peroxidases in bacteria: are there patches of order in chaos?**”: it especially highlights the mode of transmission of catalase-peroxidases and the rationale of their presence or absence in bacteria. It finally formulates a hypothesis about the appearance of catalase-peroxidases (also based on the results of the article project described in section C), and hence of the whole peroxidase superfamily.

Phylogenetic distribution of catalase-peroxidases: are there patches of order in chaos?

Filippo Passardi[a], Jocelyne Favet[b], Marcel Zamocky[c], Christa Jakopitsch[d], Claude Penel[a], Christian Obinger[d] and Christophe Dunand[a]*

[a]Laboratoire de Physiologie Végétale, Université de Genève, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland.
[b]Laboratoire de Biologie Moléculaire des Plantes Supérieures, Université de Genève, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland.
[c] Institute of Molecular Biology, Slovak Academy of Sciences c/o Department of Chemistry, BOKU-University of Natural Resources and Applied Life Sciences, Muthgasse 18, A-1190 Vienna, Austria. [d]Department of Chemistry, Division of Biochemistry, BOKU-University of Natural Resources and Applied Life Sciences, Muthgasse 18, A-1190 Vienna, Austria.

*Corresponding author. Tel.: +41223793012; Fax +41223793017; e-mail: christophe.dunand@bioveg.unige.ch

Keywords: Catalase-peroxidase; catalatic activity; hydrogen peroxide; phylogenetic analysis, lateral gene transfer.

Abbreviations: CP, catalase-peroxidase; APX, ascorbate peroxidase; C*c*P, cytochrome *c* peroxidase; PX, peroxiredoxin; SOD, superoxide dismutase; ROS, radical oxygen species; LGT, lateral gene transfer.

Abstract

Hydrogen peroxide features in many biological oxidative processes and must be continuously degraded enzymatically either via a catalatic or a peroxidatic mechanism. For this purpose ancestral bacteria designed a battery of different heme and non-heme enzymes, among which heme-containing catalase-peroxidases (CP) are one of the most widespread representatives. They are unique since they can follow both H_2O_2-degrading mechanisms, the catalase activity being clearly dominant. With the fast increasing amount of genomic data available, we were able to perform an extensive search for CP and found more than 300 sequences covering a large range of microorganisms. Most of them were encoded by bacterial genomes, but we could also find some in eukaryotic organisms other than fungi, which has never been shown until now. Our screen also reveals that more than 60% of the bacteria do not possess CP genes. Chaotic distribution among species and incongruous phylogenetic construction indicated existence of numerous lateral gene transfers in addition to duplication events and regular speciation. The results obtained show an impressively complex gene transmission pattern, and give some new insights about the role of CP and the origin of life on earth. Finally, we propose for the first time bacterial candidates that may have participated in the transfer of CP from bacteria to eukaryotes.

1-Introduction

Both aerobic and anaerobic organisms are exposed to external and internal reactive radical or non-radical oxygen species (ROS). Atmospheric oxygen ($^3O_2$) contains two unpaired electrons and is hence a free radical. Other biologically important free radicals related to O_2 are the superoxide anion ($O_2^{\bullet -}$) and the hydroxyl radical ($HO^{\bullet}$). The non-radical species, which feature also in many biological oxidative processes, are hydrogen peroxide (H_2O_2) and singlet oxygen ($^1O_2$). Since activated oxygen is toxic, it has to be continuously under strict control of integral detoxification processes, mediated either by detoxifying enzymes or by organic antioxidants.

Hydrogen peroxide can be degraded by catalatic or peroxidatic pathways. In the catalatic pathway (Reaction 1), two molecules of H_2O_2 are degraded to water and oxygen. There is a considerable diversity of enzymes that are catalatically active: (i) the “classical” heme-containing monofunctional catalases, (ii) heme containing bifunctional catalase-peroxidases (CPs) and (iii) the nonheme manganese-containing enzymes [1].

$$2\ H_2O_2 \rightarrow 2\ H_2O + O_2 \qquad \text{Reaction 1}$$

Hydrogen peroxide degradation via the peroxidatic pathway needs two one-electron donors to reduce hydrogen peroxide (Reaction 2). Examples are the heme enzyme ascorbate peroxidase (AH_2 = ascorbate) or the selenoenzyme glutathione peroxidase (AH_2 = glutathione).

$$H_2O_2 + 2\ AH_2 \rightarrow 2\ H_2O + 2\ {}^{\bullet}AH \qquad \text{Reaction 2}$$

Bifunctional catalase-peroxidase (CP or KatG) is unique since it is the only member of the plant, fungal and bacterial heme peroxidase superfamily [2] that exhibits an overwhelming catalatic activity and a peroxidase activity. The latter becomes significant in the presence of a suitable electron donor and low levels of H_2O_2. The *in vivo* peroxidatic substrate has not been identified, leaving the actual role of the peroxidatic reaction undefined [3]. Together with ascorbate peroxidases (APX), present in chloroplastic organisms, and cytochrome *c* peroxidase (C*c*P), found mainly in mitochondrial organisms, CPs constitute class I of the plant, fungal, and bacterial heme peroxidase superfamily [2-4]. They have been shown to occur in prokaryotes and also in some Ascomycetes [5,6]. The predominant form of CPs in solution is a dimer or tetramer [3] and each subunit is composed of two distinct sequence-

related N- and C-terminal domains, which led to the proposal that the large gene size of CP had arisen through a gene duplication and fusion event of a primordial peroxidase gene after which the C-terminal domain likely lost its functionality [7].

Today the crystal structures of the CPs from *Haloarcula marismortui* (1ITK) [8], *Burkholderia pseudomallei* (1MWV) [9], *Mycobacterium tuberculosis* (1SJ2) [10] and *Synechococcus* 7942 (1UB2) [11] have been defined. Structural analysis revealed that the proximal and distal heme pockets in the N-terminal domains contain conserved amino acids at

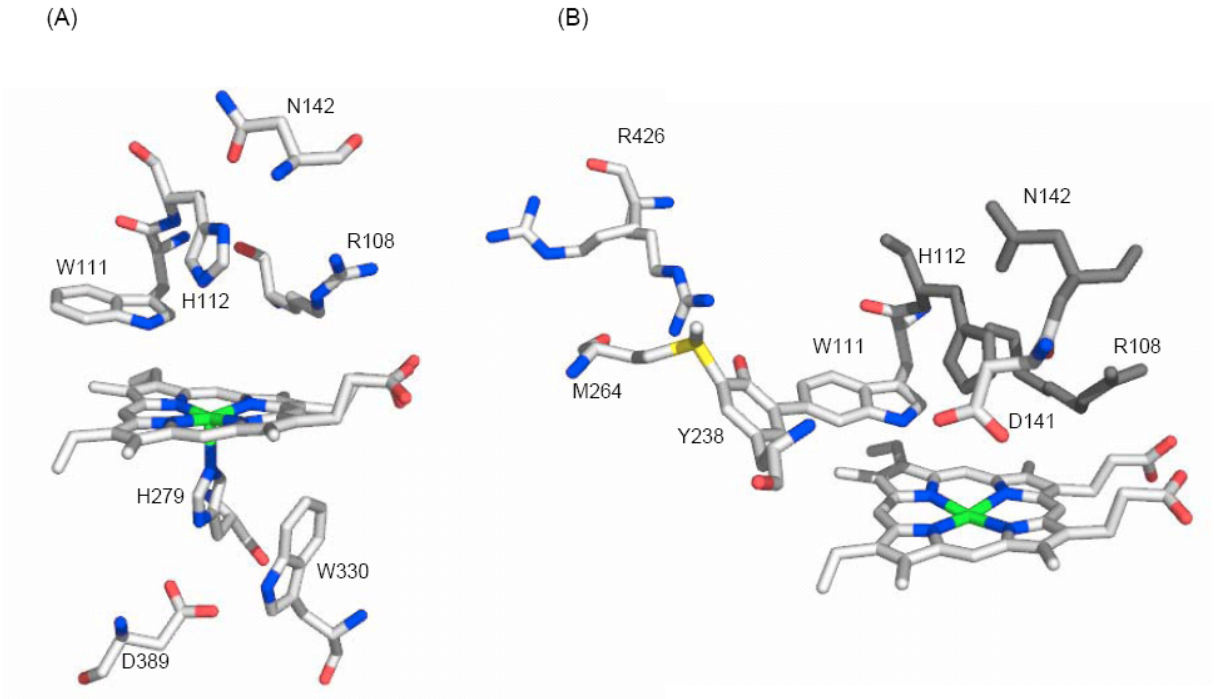


Figure 1

(A) Active site of CP from *Burkholderia pseudomallei* [9] showing the proximal triad His-Trp-Asp and the distal triad His-Arg-Trp. In addition, an Asn hydrogen-bonded to distal His is shown. These residues are highly conserved and at almost identical positions in class I peroxidases (catalase-peroxidase, ascorbate peroxidase and cytochrome c peroxidase). (B) Distal heme pocket of CP from *Burkholderia pseudomallei* showing the CP-specific structural features that are known to be essential for the catalatic activity. In particular the covalent adduct Trp-Tyr-Met and its association with an arginine, which can adopt two conformations [9], is depicted. Additionally, all catalatically active CPs contain an aspartate in the distal cavity. The figures are constructed using PyMol [43].

almost identical positions as in the other class I peroxidases APX and C*c*P. In particular, both triads His-Trp-Asp (His279, Trp330 and Asp389 in *Burkholderia pseudomallei* numbering, see sequence alignment in supplemental Fig. 1) and His-Arg-Trp (His112, Arg108 and Trp111), are conserved (Fig. 1A). Additionally, an asparagine (Asn142) hydrogen-bonded to the distal histidine (His112) is found in all class I peroxidases (Fig. 1A). However the X-ray structures and recent mutational and kinetic analyses [3] also revealed unique structural features essential for the catalatic activity of CP. The distal site Trp111 is part of a covalent adduct with Tyr238 and Met264 [8-11], which can be associated with a highly conserved arginine (Arg426) [9]. In addition the distal site contains a conserved aspartate (Asp141) (Fig. 1B), which participates in the CP-typical hydrogen bond network that is necessary for H_2O_2 oxidation [3]. Whenever these residues were manipulated by site-directed mutagenesis, the catalatic activity significantly decreased or disappeared, whereas the peroxidase activity remained unchanged or even increased [3].

The first review of CP phylogeny appeared in 2000 and included 19 sequences [12] and the following report from 2003 has included 58 sequences, the majority from bacteria but with five from archaebacteria and fungi [5]. The latter report underlined the frequency of lateral exchange of CP genes among prokaryotes, but the obtained phylogenetic tree was incongruent with a lateral gene transfer [LGT] from Archaebacteria to pathogenic Eubacteria proposed in [12]. Generally, LGT represents one of the frequently used events for the exchange of genes between two organisms that can be phylogenetically related or totally different. This process is widespread among prokaryotes and less frequent between prokaryotes and eukaryotes. Genes can be transmitted through three different modes among bacteria, namely transformation, conjugation and transduction [13,14]. For prokaryotic organisms, LGT is a rapid way of introducing novel genes and thus helps the organism adapting to environmental stress. Several features have been used to detect LGT, such as variation in codon usage or nucleotide composition. However, the most straightforward approach to highlight LGT events is probably incongruous phylogeny [15]. LGT can be controlled by genetic elements such as transposase, integrase and inverted repeats sequences. The precise mechanisms of LGT are not well understood, but different factors influence the transfer. The physical proximity and taxonomical relatedness are major points, but environmental parameters such as growing temperature and pH can favour the transfer. The G/C content compatibility between genes and the recipient organism genome, as well as the gain of new functions have an important role in LGT [16].

In this study, we performed an extensive data mining in order to find all complete CP encoding sequences in living organisms. The number of CP gene and protein sequences has more than quintupled since the last report. We found over 300 sequences, mainly in bacteria but also in fungi (Ascomycota and Basidiomycota), protists and a green alga. A global phylogenetic analysis comprising 152 representative CP sequences revealed that at first sight the distribution of bacterial CP is rather random, indicating that CP may have an important yet not essential function for bacterial viability. Additionally, phylogenetic relationships of CPs seemed to be rather chaotic. However, after careful analysis, some extent of order appeared that gave valuable information about the modes of gene transmission among prokaryotes and between bacteria and fungi, chlorophyceae and various protistean organisms, as well as about appearance and evolution of catalase-peroxidases.

2. Materials and Methods

2.1 Data mining

Catalase-peroxidase sequences were obtained either from the literature [6,17] or by the means of tBLASTN searches in public databases, mainly NCBI and the Gordon and Betty Moore Foundation Marine Microbial Genome Sequencing Project (https://research.venterinstitute.org/moore/). *Chlamydomonas reinhardtii* CP was found in JGI genome database (http://www.jgi.doe.gov/). All bacterial genomes listed in the NCBI webpage (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) were screened for CP sequences. Each CP containing sequence has been verified for the presence of specific residues that have been shown to be essential for the catalatic activity of these H_2O_2 degrading enzymes, namely the proximal triad His-Trp-Asp, the distal triad His-Arg-Trp including an associated asparagine, the covalent adduct Trp-Tyr-Met with an associated Arg as well as the conserved distal side Asp (see Figure 1 and sequence alignment in supplemental Fig. 1). The last search was run on December 4th, 2006 and 299 CP encoding sequences were found during our searches. Updates will be performed until submission of this article. All CP sequences found were deposited in the PeroxiBase at http://peroxidase.isb-sib.ch [18]. Any CP abbreviation cited in this article follows the nomenclature used in the PeroxiBase for more clarity. 16S RNA data comes from the NCBI repository and has been used to perform a comparative evolutionary analysis focused on the taxonomy.

2.2 Phylogenetic trees

152 protein sequences of catalase-peroxidases chosen proportionally from all taxonomic groups were aligned with ClustalW [19]. The total length of the multiple alignment used for further analysis was 720 amino acids. Distance (Neighbor-Joining option, JTT substitution frequency matrix) and maximum parsimony (MP) trees were reconstructed with the Phylip package v.3.65 [20], using 100 bootstrap replicates. Maximum likelihood trees were inferred with the PhyML software v.2.4.4 [21] under the JTT substitution frequency matrix and the following parameters: 100 bootstrap replicates, 4 categories of substitution rates and estimated gamma parameter. Specific subgroups were separately realigned with ClustalW and visually examined. NJ, ML and MP trees were constructed from these alignments. The same procedure was performed in parallel with 16S RNA. BioEdit software v.7.0.4.1 [22] was used to inspect the various sequence alignments and NJplot software to visualize phylogenetic trees.

2.3 Analysis of bacteria characteristics

Bacterial characteristics (GC content, respiration mode, growth temperature) were searched for in NCBI (Genome projects) and in the Bergey's manual of systematic bacteriology [23]. GC content of peroxidases was calculated with BioEdit. Search for inverted repeats was performed with the "einverted" program from the Institut Pasteur (http://bioweb.pasteur.fr/seqanal/interfaces/einverted.html).

3. Results and discussion

3.1 Distribution among organisms

Catalase-peroxidases can be found in eubacteria, archaebacteria (Supplemental Table 1), Ascomycota and Basidiomycota (Supplemental Table 2), as well as in few protists (*Heterocapsa triquetra*, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, various *Phytophthora* species, *Karenia brevis*) and Viridiplantae (*Chlamydomonas reinhardtii*) (Supplemental Table 3). In eubacteria and archaebacteria, the distribution of CPs is not homogenous and it is quite intriguing to note that sequences are absent from a large number of bacterial species: only 47% of the 620 screened genomes (excluding unfinished genomes) encoded a CP sequence (Supplemental Table 1). In phyla such as Chlamydiae or Spirochaetes, where CP is absent from all the organisms, other enzymes have been recruited

Kingdom-Phylum	Class-Order	Genus	Species	DCCP	SOD	catalase	GPX	AH	PX	CP	Metabolism
Chlamydiae (10)					X				X		INTRA
Spirochaetes		Borrelia (2)			X						MICRO
		Leptospira (2)		X		X	X		X		OX
		Treponema	Treponema denticola ATCC 35405				X		X		MICRO
			Treponema pallidum subsp. pallidum str. nichols						X		MICRO
Firmicutes	Lactobacillales	Enterococcus	Enterococcus faecalis V583		X	X			X		FERM
			Enterococcus faecium DO(p)				X		X		FERM
		Lactobacillus	Lactobacillus acidophilus NCFM						X		FERM
			Lactobacillus casei ATCC 334(p)		X		X		X		FERM
			Lactobacillus johnsonii NCC 533						X		FERM
			Lactobacillus plantarum WCFS1		X	X	X		X		FERM
			Lactobacillus sakei subsp. sakei 23K		X	X			X		FERM
		Lactococcus	Lactococcus lactis subsp. Lactis II1403		X		X		X		FERM
		Streptococcus	Streptococcus agalactiae (3)		X				X		FERM
			Streptococcus mutans UA159		X				X		FERM
			Streptococcus pneumoniae (2)		X		X		X		FERM
			Streptococcus pyogenes (7)		X		X		X		FERM
			Streptococcus thermophilus (2)		X				X		FERM
	Clostridia	Clostridium	Clostridium acetobutylicum ATCC 824		X		X				FERM
			Clostridium perfringens str. 13		X		X		X		FERM
			Clostridium tetani E88		X				X		FERM
Alpha-proteobacteria	Rhizobiales	Bartonella (2)			X				X		INTRA
		Brucella (4)			X	X			X		INTRA
		Agrobacterium	Agrobacterium tumefaciens str. C58	X	X	X		X	X	X	OX
		Bradyrhizobium	Bradyrhizobium japonicum USDA 110	X	X		X	X	X	X	OX
		Mesorhizobium	Mesorhizobium loti MAFF303099	X	X	X		X	X	X	OX
		Rhizobium	Rhizobium etli		X		X	X	X	X	OX
		Rhodopseudomonas	Rhodopseudomonas palustris CGA009	X	X	X	X	X	X	X	OX
		Sinorhizobiun	Sinorhizobium meliloti 1021		X	X		X	X	X	OX
Gamma-proteobacteria	Enterobacteriales	Buchnera (3)			X				X		INTRA
		Blochmannia (2)			X				X		INTRA
		Photorhabdus	Photorhabdus luminescens subsp. Laumondii TTO1		X	X			X		INTRA
		Wigglesworthia	Wigglesworthia glossinidia		X				X		INTRA
		Erwinia	Erwinia carotovora subsp. Atroseptica SCRI1043		X		X	X	X	X	OX+FERM
		Escherichia	Escherichia coli (5)	X	X	X	X	(x)	X	X	OX+FERM
		Shigella (5)		X	X	X	X	(x)*	X	X	OX+FERM
		Salmonella (4)		X	X	X	X	(x)	X	X	OX+FERM
		Yersinia (3)		X	X	X	X	X	X	X	OX+FERM
	Alteromonadales	Pseudoalteromonas	Pseudoalteromonas haloplanktis TAC125	X	X	X			X		OX
			Pseudoalteromonas atlantica T6c(p)		X	X		X	X	X	OX
		Colwellia	Colwellia psychrerythraea 34H	X	X	X		X	X	X	OX
		Idiomarina	Idiomarina loihiensis			X			X	X	OX
		Shewanella	Shewanella oneidensis MR-1	X	X	X			X	X	OX

Table 1

Distribution of hydrogen peroxide degrading enzymes and SOD among some prokaryotic groups. Except for a few cases (p), only organisms with complete genomes were searched for various antioxidative enzymes,. Grey shaded boxes represent organisms lacking CP. Presence of of proteins is noted with an X in the corresponding box. (x) stands for protein sequence showing low homology. The numbers in brackets correspond to the number of species for one genus or the number of strains for one species. DCCP: di-haem cytochrome c peroxidase; SOD: superoxide dismutase; GPX: glutathione peroxidase; AH: alkylhydroperoxidase; PX: peroxiredoxin or undefined thiol peroxidase; FERM: fermentative metabolism; OX: oxidative metabolism; INTRA: intracellular bacterium; MICRO: microaereobes. (x)* absent from *Shigella sonnei* Ss046.

in order to degrade hydrogen peroxide (Table 1). It also happens that particular species within a genus lack CP genes (e.g. *Streptomyces*, *Rhodobacter*, *Geobacter*, *Mycobacterium*). Surprisingly, within one species, CP can even be present in one strain and absent from

another one (*Synechococcus sp.*). In this case, the explanation becomes much less evident. One hypothesis could be that a non-duplicated form of CP is present in organisms that do not contain CP, but that it has become so distantly related that it is now considered as another peroxidase. However, our search for degenerated sequences did not yield any interesting candidate (data not shown).

Due to their function as a H_2O_2 degrading enzyme [24,25], absence of CP should bring a severe disadvantage to the bacterial cell. Thus, we have checked in organisms lacking CP whether they encoded other kinds of enzymes that remove hydrogen peroxide. In particular the genomes were screened for monofunctional catalase (EC 1.11.1.6), glutathione peroxidase (EC 1.11.1.9), di-haem cytochrome *c* peroxidase (EC 1.11.1.5), thioredoxin peroxide reductase (peroxiredoxins, PX) or alkyl hydroperoxide reductase (EC 1.11.1.15). In addition the occurrence of superoxide dismutase (EC 1.15.1.1) was investigated.

Representative and distinct groups of bacteria were chosen for this analysis, as shown in Table 1. CP distribution is rather heterogeneous and difficult to predict, with important variations within one phylum or order (Spirochaetes, Lactobacillales, Enterobacteriales) but also within one genus (*Lactobacillus, Clostridium, Streptococcus*). There is nevertheless some degree of order in specific groups, such as intracellular bacteria, represented by Chlamydiae and a subgroup of Enterobacteriales (*Buchnera, Blochmannia, Photorhabdus, Wigglesworthia*) and Rhizobiales (*Bartonella, Brucella*), where the enzymes detected are almost exclusively peroxiredoxin/thiol reductase and superoxide dismutase. Considering that other Enterobacteriales and Rhizobiales possess the majority of the enzymes screened, and that intracellular bacteria tend to reduce their genome when starting their intracellular life [26,27], presence of PX only is certainly due to loss of the other enzymes. It has to be mentioned that PXs have a more than 100-fold higher specificity for H_2O_2 than CPs or monofunctional catalases [28]. Thus, PXs are efficient H_2O_2-degrading enzymes at low (micromolar) peroxide concentrations.

For the other bacterial groups screened in Table 1, the general rule is that microaerobes (*Borrelia, Treponema*) and fermentative bacteria (Firmicutes) possess less ROS-detoxifying enzymes than oxidative/mixed oxidative-fermentative bacteria (*Leptospira*, other Rhizobiales, other Entrobacteriales and Alteromonadales). For a particular species however, distribution of these enzymes remains unpredictable, as exemplified by the various *Lactobacillus* species. Interestingly, among the bacteria screened for catalases and peroxidases, we never found any without a hydrogen peroxide scavenging enzyme, with the exception of the genus *Borrelia* (microaerobes), which possesses only SOD. It appears hence

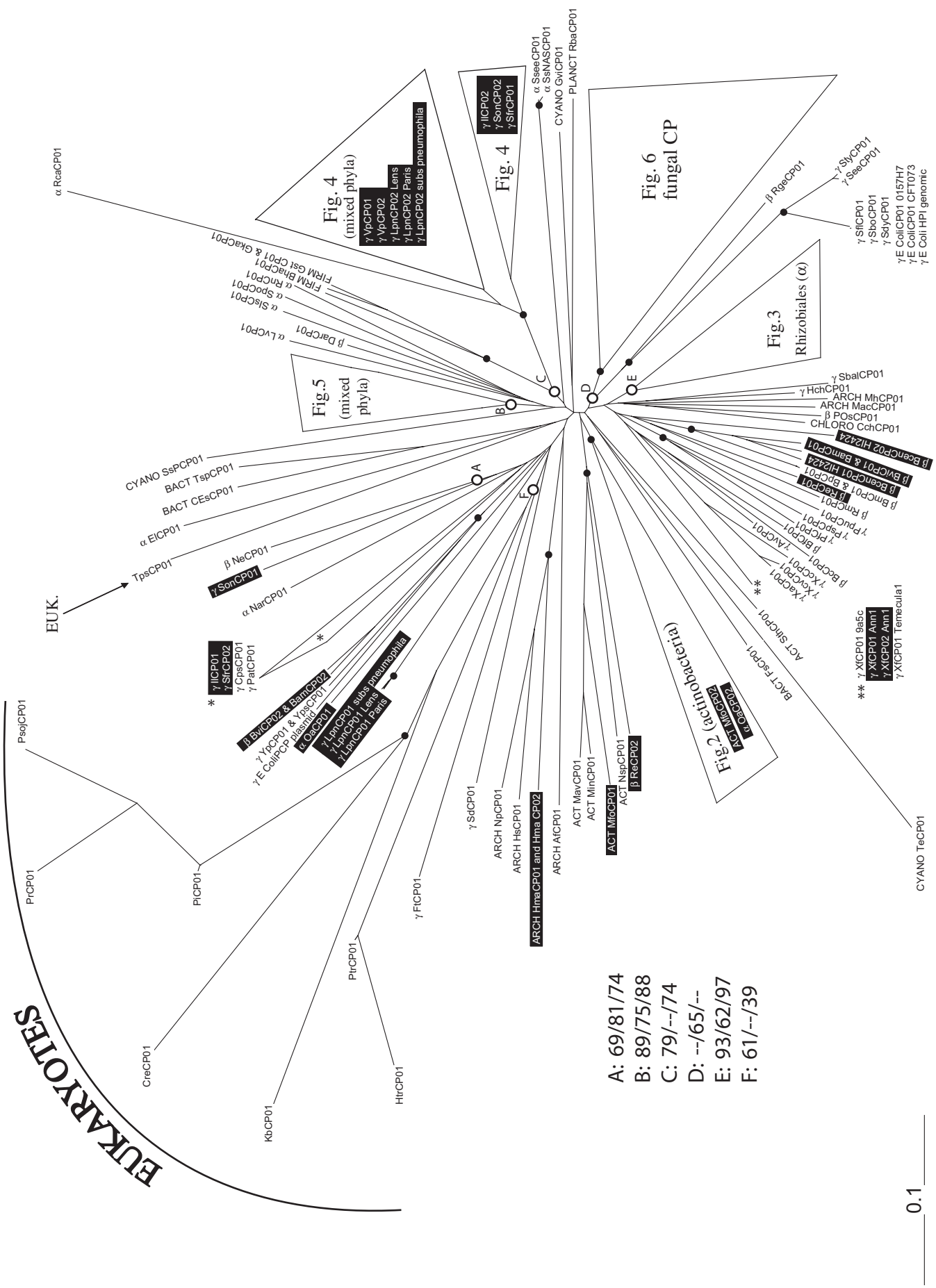

EUKARYOTES
EUK.
PsojCP01
PrCP01
PiCP01
CreCP01
KbCP01
PtrCP01
HtrCP01
γ FtCP01
γ SdCP01
ARCH NpCP01
ARCH HsCP01
ARCH HmaCP01 and Hma CP02
ARCH AfCP01
ACT MavCP01
ACT MinCP01
ACT MfoCP01
ACT NspCP01
β ReCP02
Fig.2 (actinobacteria)
ACT MfoCP02
α OscCP01
BACT FscP01
ACT SlnCP01
γ XfCP01 9a5c
γ XfCP01 Ann1
γ XfCP02 Ann1
γ XfCP01 Temecula1
CYANO TeCP01
γ XacP01
γ XcvCP01
γ XccCP01
γ AvCP01
β BlCP01
γ PicCP01
γ PspCP01
γ PnuCP01
β RmCP01
β BmCP01 & RpCP01
β BvICP01 & Bpe CP01
β BcenCP01 HI2424
β BcenCP02 HI2424
β BamCP01
ARCH MacCP01
ARCH MhCP01
γ HchCP01
γ SbalCP01
β POsCP01
CHLORO CchCP01
Fig.3 Rhizobiales (α)
β RgeCP01
γ StyCP01
γ SeeCP01
γ SflCP01
γ SboCP01
γ SdyCP01
γ E ColiCP01 0157H7
γ E ColiCP01 CFT073
γ E Coli HPI genomic
Fig. 6 fungal CP
α SseeCP01
α SsNASCP01
CYANO GviCP01
PLANCT RbaCP01
γ IICP02
γ SonCP02
γ SfrCP01
Fig. 4
Fig. 4 (mixed phyla)
γ VpCP01
γ VpCP02
γ LpnCP02 Lens
γ LpnCP02 Paris
γ LpnCP02 subs pneumophila
α RcaCP01
FIRM GstCP01 & GkaCP01
FIRM BhaCP01
α RmCP01
α SpocCP01
α SlsCP01
β DarCP01
α LvCP01
Fig.5 (mixed phyla)
CYANO SsPCP01
BACT TspCP01
BACT CEsCP01
α ElCP01
TpsCP01
β NeCP01
γ SomCP01
α NarCP01
γ IICP01
γ SfrCP02
γ CpsCP01
γ PatCP01
β BvrCP02 & BamCP02
γ YpCP01 & YpsCP01
γ E ColiPCP plasmid
α OscCP01
γ LpnCP01 subs pneumophila
γ LpnCP01 Lens
γ LpnCP01 Paris
A
B
C
D
E
F
A: 69/81/74
B: 89/75/88
C: 79/--/74
D: --/65/--
E: 93/62/97
F: 61/--/39
0.1

Figure 2 (previous page)
Global phylogenetic tree of catalase-peroxidases. Schematic representation has been used to group well-supported clusters of sequences. Triangles were formed by linking the longer branch of each group to the shortest branch. Phylogenetic analysis of these clusters is developed in Figs. 3 to 7. Peroxidases in black boxes belong to organisms possessing two CP. Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap values of more than 50% in at least two of the three computational methods (except for nodes D and F). Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bar represents 0.1 substitution per residue. ACT: actinobacteria; ARCH: archaebacteria; BACT: bacteroidetes; CHLORO: chlorobia; CYANO: cyanobacteria; FIRM: firmicutes; PLANCT: planctomycetes; α: α-proteobacteria; β: β-proteobacteria; γ: γ-proteobacteria. Abbreviations of CP names refer to accessions in the PeroxiBase repository (www.peroxidase.isb-sib.ch).

that the general rule among bacteria is to have at least one H_2O_2-degrading enzyme, but not necessarily the same peroxidatic or catalatic enzyme. By this way, bacteria protect themselves against peroxides, even if they are not supposed to meet any oxygenic condition such as *Clostridium*, various Chlorobiales, Delta-proteobacteria and Firmicutes (Fig. 6). It is known for instance that several *Clostridium* species are able to grow on microoxic environments [29,30].

Usually, CP-containing organisms possess one single CP gene, but we have found several exceptions with two genes (Fig. 2 and Supplemental Table 1). Within one genus (*Burkholderia, Ralstonia, Vibrio, Xylella, Shewanella*) or even one species (*Burkholderia cenocepacia*), we can find bacteria with one or two CPs. In *Vibrio*, depending on the species, genomes encode zero, one or two CPs. So far, none of the organisms screened shows more than two CP genes. In fungi, only one CP was found in each organism, with the exception of two Ascomycetes (*Magnaporthe grisea* and *Gibberella zeae*) which contain two isoforms. One of the Ascomycete sequences, GzCP01 (*Gibberella zeae*), was exceptionally long, and further analysis revealed a possible fusion with a cytochrome P450 protein. Cooperation between CP and cytochrome P450 proteins in protecting the organism against oxidative damage may be a reason for this fusion. A similar case has been proposed for *Chlamydomonas reinhardtii*, with CrePrx01 and 02 [4]. These two proteins show a first

region similar to the classical peroxidase motif (like the N-terminal region of CP) and a second region very similar to a multicopper oxidase. As for GzCP01 the proximity of these two different proteins participating in redox reactions may bring some advantage in reactions involving various oxygen species. To our knowledge, this is the first time that such a fusion event with important functional consequences has been reported.

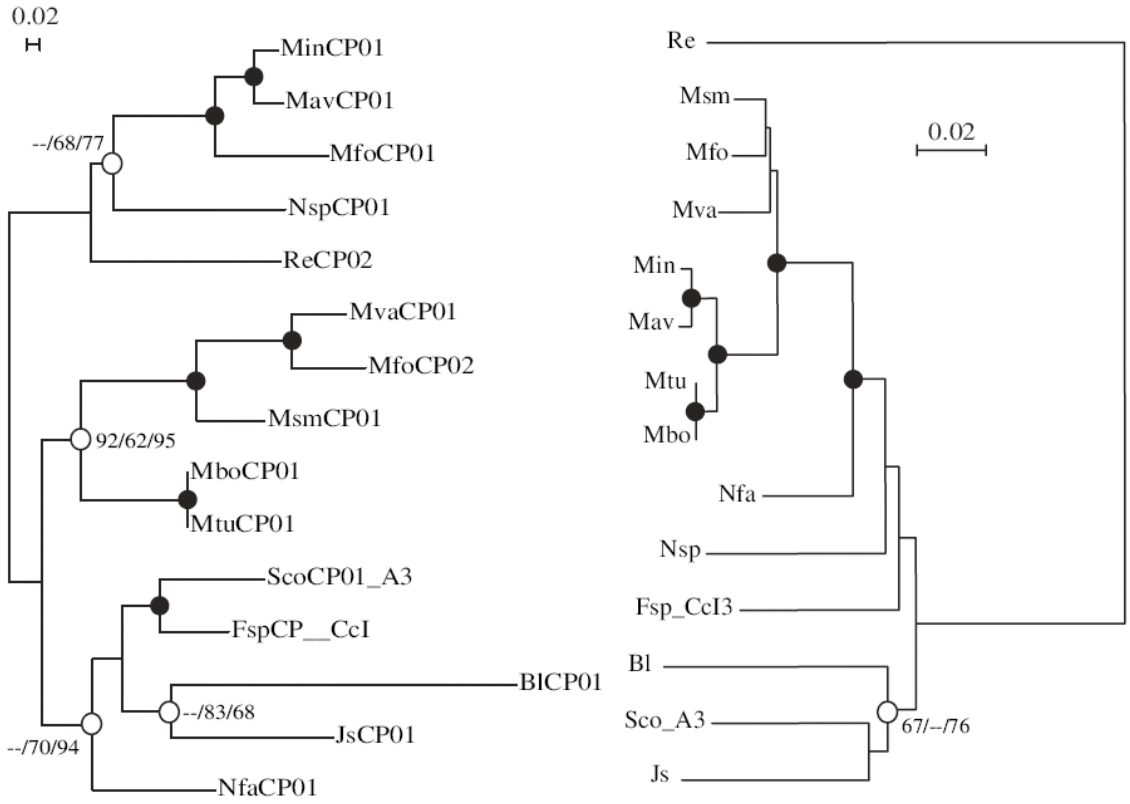


Figure 3

Comparative phylogentic analysis of CP sequences (left tree) and 16S RNA genes (right tree) of the actinobacterial cluster (Fig. 2). Five other peroxidases sequences closely related to the cluster have been added to root the tree (four actinobacterial peroxidases (NspCP01, MfoCP01, MavCP01 and MinCP01) and one alpha-proteobacterium sequence (ReCP02). Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap supports of more than 50% in at least two of the three computational methods. Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bars in each tree represents 0.02 substitution per residue.

Following data collection of CP sequences, we performed a phylogenetic analysis (using distance, MP and ML methods) with 152 representative (all phyla) and complete (except for some eukaryotes) CP protein sequences. CP sequences were not included when too similar and hence not informative, particularly *Burkholderia* species sequences. Several subgroups came out well-supported, though many CPs could not be assigned to a particular subgroup (Fig. 2).

Protistean/algal CPs form a distinct outgroup, with the exception of *Thalassiosira pseudonana* (TpsCP01). The same observation applies for fungi, which are, however, included within the bacterial/archaeal tree. Addition or not of fungal CP in the alignment only slightly affects bootstrap values and does not modify the presence of well-supported groups.

Among bacteria, some clades form strongly supported groups (actinobacteria, Rhizobiales), but other members of the same clade are scattered across the tree (Rhodobacterales, Alteromonadales). Among other peculiarities of the alignment of CP, it appears that archaebacterial sequences are clustered together, with the exceptions of MacCP01 (*Methanosarcina acetivorans*) and MhCP01 (*Methanospirillum hungatei*). This discrepancy strongly suggests lateral gene transfer with other bacteria, as discussed further in the text.

3.2 Transmission of catalase-peroxidases

3.2.1 Lateral gene transfer among prokaryotes

According to the phylogenetic trees obtained, catalase-peroxidases have been transmitted among species by several manners. Although the well-supported and separated actinobacterial and rhizobial groups (Figs. 3 and 4) suggest transmission of CP through speciation, it seems that LGT is the most frequent mechanism. In the case of sequences resulting from LGT events, when aligning regions ±10’000 bp up- and downstream from the CP sequence, the only homologies found were for the CP alone. Catalase-peroxidase is therefore not transferred within the large DNA fragments commonly referred to genomic islands [31,32], but more likely alone. However, it cannot be excluded that speciation occurred and was followed by genomic rearrangements, especially within a same genus (*Mycobacterium*). Similar conclusions were found in the Rhizobiales subgroup.

CP distribution does not exactly parallel the taxonomical distribution of the corresponding organisms, as demonstrated in actinobacteria when comparing the trees obtained with CP and 16S RNA sequences (Fig. 3). This discrepancy reflects convergent and

divergent evolution patterns, which were responsible for separation of *Mycobacterium* CP into two well-defined groups. The difference can be also related to the differential evolution rate of an essential gene (16S RNA) *versus* an important but non essential gene (CP). Similarly, the relationships between Rhizobiales obtained with 16S RNA alignments show some discrepancy with the CP phylogenetic tree (Fig. 4).

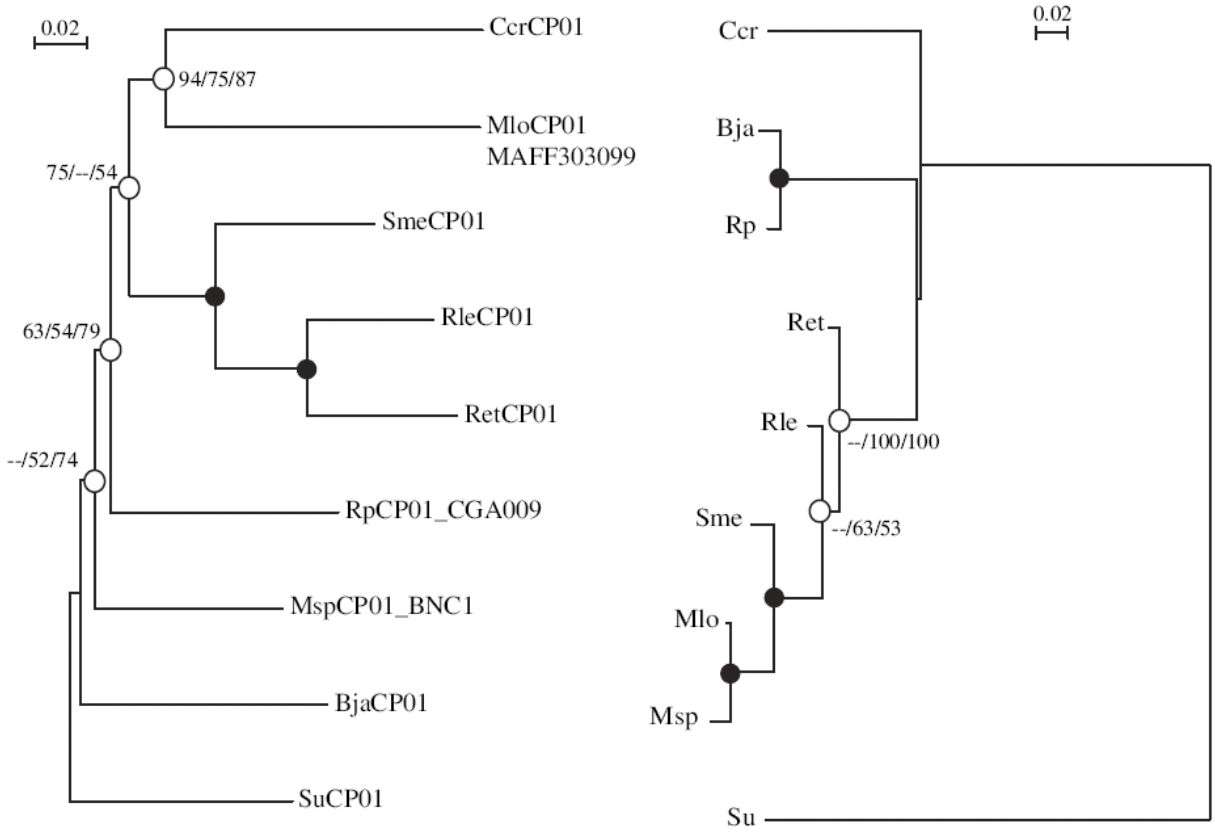


Figure 4

Comparative phylogenetic analysis of CP sequences (left tree) and 16S RNA genes (right tree) of the rhizobial cluster (Fig. 2). Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap supports of more than 50% in at least two of the three computational methods. Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bars in each tree represents 0.02 substitution per residue. Note that *Caulobacter crescentus* (CcrCP) and *Solibacter usitatus* (SuCP) are not Rhizobiales, but, respectively, another class of alpha-proteobacterium and an acidobacterium.

The most evident proof of LGT comes from the well-supported subgroup detailed in Fig. 5, where cyanobacteria and alpha/gammaproteobacteria are mixed together. LGT and speciation can also be combined. When aligning a region ± 20'000 bp up- and downstream from LpnCP02 (Paris and subs_pneumophila), we noticed a homology over 4'000 bp upstream and 19'000 bp downstream. Strong synteny was also observed with LpnCP02 from the strain Lens. Speciation is hence responsible for LpnCP02 distribution. However, for all the other CP of the subgroup, the same analysis showed no synteny and hence indicates probable LGT events.

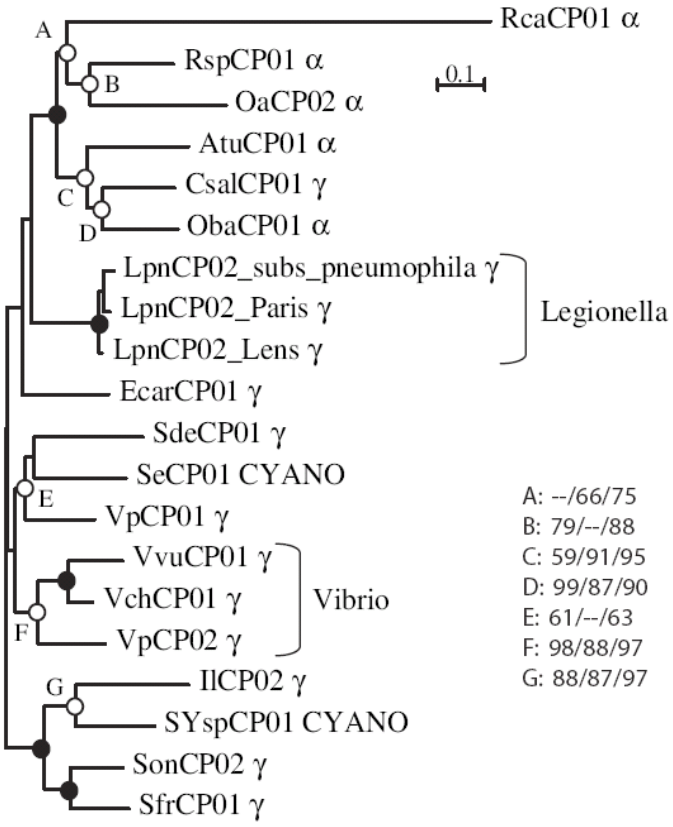


Figure 5

Phylogenetic tree of protein CP sequences of a cluster including various bacterial phyla (Fig. 2). RcaCP01, which is also related to this group (Fig. 1), has been added. Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap supports of more than 50% in at least two of the three computational methods. Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bar represents 0.1 substitution per residue.

Following our results, it clearly appears that LGT is quite a common way for dissemination of CP sequences among bacteria. We then tried to look for factors that could favour the transfer of CP between bacteria. Among various parameters, it has been proposed that the G/C content can have a positive associative effect on LGT [16]. We found that this rule is valid for the subgroup of Fig. 5, where RcaCP01, RspCP01, AtuCP01, ObaCP01, OaCP02 and CsalCP01 form a distinct cluster and the corresponding sequences have a high G/C content (60-67%), in line with the high G/C content of the respective species genome (59-67%). All other CP in this group (and their respective species genome) are separated from Rca, Rsp, Csal, Oba, Oa and Atu and have an average G/C content of 40-50% (Table 2). This search also showed that the G/C content of CP usually follows the G/C content of the whole genome of the corresponding organism.

The “G/C content rule” does not apply in other cases, as for instance in the group of Fig. 6, where the G/C contents are rather heterogeneous (Table 2). In this group, respiration mode and growing temperature are the same, and could hence explain the lateral transmission of CP. All organisms are indeed anaerobic and mesophilic. Similarly, BviCP02, BamCP02, E_Coli pCP, YpsCP01 and YpCP01 (Fig. 2) have different G/C contents but are all facultative aerobic and mesophilic. Furthermore, most organisms have a genomic G/C content very different from the corresponding CP, with the exception of DaCP01 (Table 2). Acquisition of CP by these organisms is thus quite recent, or this CP sequence is resistant to adaptive mutations. When manually looking at the G/C content up to 10’000 bp upstream and downstream from each CP, we observed that it is gradually increasing towards the CP sequence, with a maximum on the CP sequence itself. Alignment of sequences 10’000 bp upstream and downstream from each CP shows that within a same cluster from the Fig. 6 subgroup (same bacterial phylum), the only homology within this region was found for the CP sequence itself, similarly to the actinobacterial and rhizobial groups (Figs. 3 and 4). Despite the lack of synteny, there seems to be a preferential transfer between bacteria belonging to the same phylum. This transfer may be favoured by proximity of closely related bacteria, or through an ancient speciation followed by genomic rearrangements.

Furthermore, in order to better understand the mechanisms of LGT, we searched for transposon typical signatures in the heterogenous subgroup represented in Fig. 6. Inverted repeats were found very close to CP sequences in DaCP, GsuCP and PprCP, as well as in DhaCP and AmeCP. However, we could not find any close homology (sequence or position) between these inverted sequences. Moreover, none was found in PcaCP, PaeCP and CphCP. When looking at integrase and transposase encoding sequences, we found that DaCP

				GC content (%)			
	Name	Species	Phylum/subdivision	CP	genome	aerobic/anaerobic	Optimal temperature
Fig.5	RcaCP01	Rhodobacter capsulatus	alpha	67	67	facultative anaerobic	Meso
	RspCP01	Roseovarius sp.	alpha	63	60	aerobic	Meso
	OaCP02	Oceanicaulis alexandrii	alpha	63	64	aerobic	Meso
	AtuCP01	Agrobacterium tumefaciens	alpha	62	59	aerobic	Meso
	CsalCP01	Chromohalobacter salexigens	gamma	60	63	aerobic	Meso
	ObaCP01	Oceanicola batsensis	alpha	65	66	aerobic	Meso
	LpnCP02_subs_pneumophila	Legionella pneumophila subs pneumophila	gamma	46	38	aerobic	Meso
	LpnCP02_Paris	Legionella pneumophila Paris	gamma	46	38	aerobic	Meso
	LpnCP02_Lens	Legionella pneumophila Lens	gamma	46	38	aerobic	Meso
	EcarCP01	Erwinia carotovora	gamma	55	52	facultative anaerobic	Meso
	SdeCP01	Shewanella denitrificans	gamma	47	45	facultative anaerobic	Meso
	SeCP01	Synechococcus elongatus	cyano	54	55	aerobic	Meso
	VpCP01	Vibrio parahaemolyticus	gamma	48	46	facultative anaerobic	Meso
	VvuCP01	Vibrio vulnificus	gamma	48	46	facultative anaerobic	Meso
	VchCP01	Vibrio cholerae	gamma	48	48	facultative anaerobic	Meso
	VpCP02	Vibrio parahaemolyticus	gamma	47	46	facultative anaerobic	Meso
	IICP02	Idiomarina loihiensis	gamma	50	49	aerobic	Psychro
	SSPPCCP01	Synechocystis sp.	cyano	49	47	aerobic	Meso
	SonCP02	Shewanella oneidensis	gamma	48	45	facultative anaerobic	Meso
	SfrCP01	Shewanella frigidimarina	gamma	42	41	facultative anaerobic	Psychro
Fig.6	PaeCP01	Prosthecochloris aestuarii	Chlorobia	60	50	anaerobic	Meso
	CphCP01	Chlorobium phaeobacteroides	Chlorobia	61	48	anaerobic	Meso
	PphCP01	Pelodictyon phaeoclathratiforme	Chlorobia	58	48	anaerobic	Meso
	DaCP01	Desulfuromonas acetoxidans	Delta	53	51	anaerobic	Meso
	GsuCP01	Geobacter sulfurreducens	Delta	67	60	anaerobic	Meso
	PcaCP01	Pelobacter carbinolicus	Delta	61	55	anaerobic	Meso
	PprCP01	Pelobacter propionicus	Delta	63	58	anaerobic	Meso
	DhaCP01	Desulfitobacterium hafniense	Firmicutes	57	53	anaerobic	Meso
	AmeCP01	Alkaliphilus metalliredigenes	Firmicutes	48	36	anaerobic	Meso
	BviCP02	Burkholderia vietnamensis	beta	63	65	facultative anaerobic	Meso
	BamCP02	Burkholderia ambifaria	beta	63	66	facultative anaerobic	Meso
	E_Coli pCP plasmid	Escherichia Coli (plasmid p0157)	gamma	46	47	facultative anaerobic	Meso
	YpsCP01	Yersinia pseudotuberculosis	gamma	46	47	facultative anaerobic	Meso
	YpCP01	Yersinia pestis	gamma	46	47	facultative anaerobic	Meso

Table 2

Catalase-peroxidase subgroups with specific features of the CP and the organism. Grey shaded boxes represent differences >5% in the GC content between CP and genome. Meso and Psychro: Mesophilic (25-40°C) and Psychrophilic (≤25°C).

sequence is very close (<1000bp) to a transposase and an integrase, and is preceded by a tRNA synthetase, commonly found next to transposable elements [33,34]. Inverted repeats were also detected, suggesting that CP can be transmitted also through transposition. This observation was confirmed by the finding of transposases and inverted repeats close to several other CP not belonging to subgroup of Fig. 6, such as GkaCP01, FspCP01 or ScoCP01. In most of cases though, none of these elements was found.

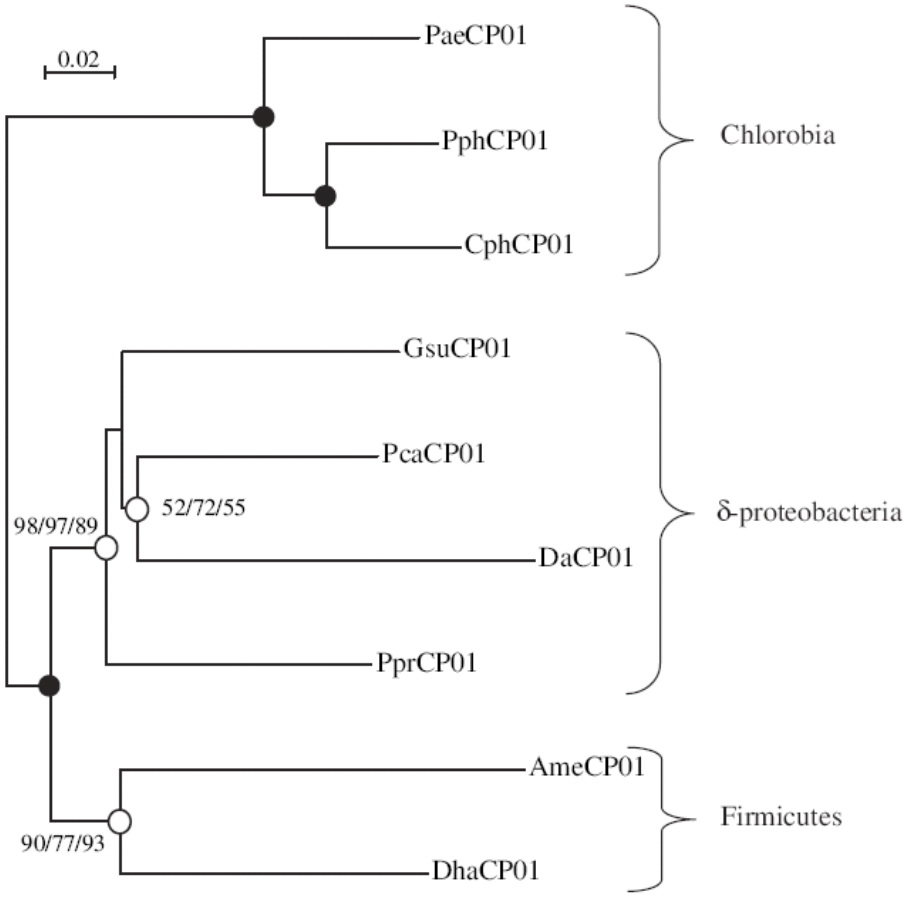


Figure 6

Phylogenetic analysis of protein CP sequences of a cluster including various bacterial phyla (Fig. 2). Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap supports of more than 50% in at least two of the three computational methods. Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bar represents 0.02 substitution per residue.

Marginally, we found CP sequences in plasmids (E_ColiPCP, SmeCP01), which could favour a more rapid transmission of CP. If transmission of CP is still very active between bacteria, some of them can also lose this gene. For example, in the *Mycobacterium* genus (Supplemental Table 1) only *M. leprae* is lacking a CP sequence. This obligate parasitic bacterium is known to have experienced a reductive evolution, in which it probably lost, among many others, its CP gene [26]. After screening *M. leprae* genome, we indeed found a CP pseudogene (locus ML2009). It does not seem, however, that CP is always lost during reductive evolution, as exemplified by presence of CP in *Salmonella typhi* and *Shigella flexneri*, two bacteria also known to have lost a significant part of their genome [35].

3.2.2 Speciation

Acquisition of a CP with a different G/C content is possible, but the CP DNA will then be modified in order to adapt its G/C content to the host genome G/C content. For a more recent transfer, the current G/C content is still different from the whole genomic one, thus creating the so-called "genomic islands" [31,32]. By comparing the G/C content of CP sequences and genomes in the various well-supported subgroups of our NJtree (Table 2), we searched for recent transfers. LpnCP02_Paris and LpnCP02_subs_pneumophila show a significant higher G/C content than *Legionella pneumophila* genomes, and are hence good candidates for genomic islands. The donor bacterium must have had a higher G/C content than *Legionella*, which is the case for most of the organisms in this subgroup. The donor bacterium cannot be identified, nevertheless we can conclude that acquisition of CP sequences by *Legionella pneumophila* occurred quite recently. When screening the region surrounding both LpnCP02_Paris and LpnCP02_subs_pneumophila, we interestingly saw that G/C content very rapidly decreased from 46% for the CP sequence to 40% and 38% for 500 bp and 2000 bp surrounding sequences respectively. Therefore, these CP do not belong to a genomic island, according to the G/C content. When aligning regions encompassing both CP sequences, results show that they have more than 90% homology on a large region. This was not the case for other organisms belonging to this subgroup. These LpnCP02 hence come from a *Legionella* ancestor that acquired the CP alone (probably not clustered in a genomic island) and later on transferred its CP to the current *Legionella* through speciation. When performing the same analysis with LpnCP02 from strain "Lens", the same conclusions apply.

Other clear cases of speciation (synteny ±20'000 bp up- and downstream from the CP gene) were found for the LpnCP01 cluster, as well as for MboCP01 and MtuCP01, and exceptionally between two different species of a same genus (BcenCP01 and BamCP01). We

were not able to detect speciation between different genera of a same order, probably due to genomic rearrangements.

3.2.3 Organisms bearing two catalase-peroxidases

A few bacterial organisms possess two CP sequences. Phylogenetic distribution of these CP varies according to the organism. In the subgroup of Fig. 5, the three independent LpnCP02 clearly appeared through speciation, as mentioned before. Their counterparts (LpnCP01) are also clustered together in another well defined group (Fig. 2). The subgroup of Fig. 5 contains other organisms with two CP sequences. VpCP01 and 02 are clear duplicates of an ancestral CP. Regarding IlCP02, SonCP02 and SfrCP01, their duplicated counterparts (respectively IlCP01, SonCP01 and SfrCP02) are not regrouped in a cluster. This observation indicates independent duplication events, and hence suggests that bacteria from the Alteromonadales order have a high tendency to undergo katG catalase-peroxidase gene duplication (supplemental Table 1).

Independent CP sequence acquisition probably also occurred in *Ralstonia eutropha* (ReCP), *Oceanicaulis alexandrii* (OaCP) and two *Burkholderia* species (BamCP and BviCP). *Burkholderia cenocepacia* (BcenCP) and *Mycobacterium fortuitum* (MfoCP) more likely duplicated their CP, although phylogenetic relationship is weakly supported for both cases. Concerning BcenCP01 and BamCP01, a strong homology was observed in a region ± 10’000 up- and downstream of the CP sequences, thus indicating transmission through speciation. Recent acquisition by duplication is also observed for XfCP01/02 and HmaCP01/02.

These observations show that duplicated forms of CP come from two different events, namely independent acquisitions by LGT and speciation. Most probably CP duplication also happened in the two fungal species *Magnaporthe grisea* and *Gibberella zeae* (MagCP and GzCP). The two isoforms appeared in two independent branches (Fig. 7). Regarding GzCP01, the duplication event was followed by a fusion with a cytochrome P450 oxidase gene, a unique event in catalase-peroxidase evolution. It is very unlikely that MagCP02 and GzCp01 were acquired by a second independent bacterial gene transfer, as they would not appear in the fungal group with a high bootstrap support.

3.2.4 Eukaryotic catalase-peroxidases

Transfer of genetic material from bacteria to fungi has been demonstrated in various studies [36,37]. Presence of CP in fungi is hence certainly due to a bacterial gene transfer. All fungal CP are well separated from bacterial CP, indicating a single and early LGT between a

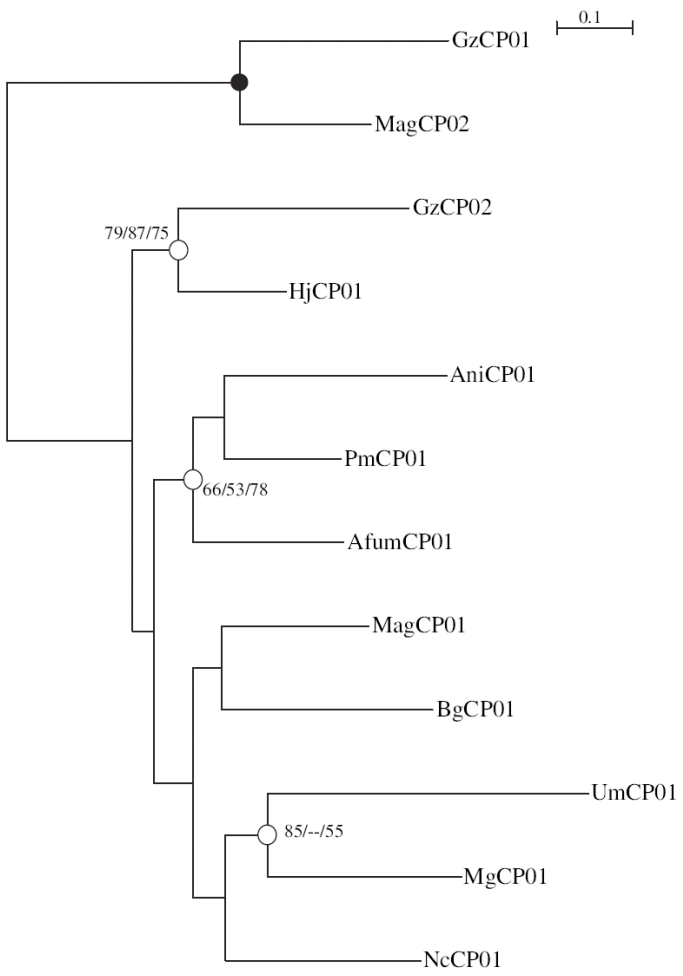


Figure 7

Phylogenetic analysis of protein CP sequences of a cluster including only fungal CP (Fig. 2). Black circles stand for nodes supported by more than 95% with distance, MP and ML methods. White circles refer to nodes with bootstrap supports of more than 50% in at least two of the three computational methods. Bootstrap values are written in the following order: NJ (distance), MP and ML. All branches are drawn to scale and the scale bar represents 0.1 substitution per residue.

bacterium and a fungal ancestor, in agreement with a previous study [5]. This bacterium may well be a close relative of the Enterobacteriaceae family (*Escherichia*, *Shigella*, *Salmonella*), or the Burkholderiale *Rubrivivax gelatinosus* (see node "D" in figure 2). Both types of organisms are however not known to have particular interrelationships with fungi. The presence of a CP in *Ustilago maydis* (UmCP01) is quite intriguing, considering that this organism is the only basidiomycete possessing a CP, in the middle of ascomycetous fungi (Fig. 7). If UmCP01 has derived from a second independent bacterial gene transfer, its

sequence should be separated from the other Ascomycete sequences. The fact that UmCP01 appeared in the same phylogenetic branch may simply be due to the lack of sequences currently available for the Ustilaginomycete branch. It is probable that more basidiomycete catalase-peroxidase genes can occur with ongoing sequencing projects (basidiomycete genomes are less covered in databases at the moment in comparison to ascomycetes).

When analysing regions ± 10'000 bp up- and downstream from fungal CP sequences, we only found homology on the CP sequence alone, similarly to the bacteria. This is certainly due to genome recombinations rather than transposon-mediated events. Even between closely related fungi (*Aspergillus*: AniCP and AfumCP), it is known that the extent of recombination within the genome is already quite important [38]. Some fungal CP sequences (e.g. MagCP02, GenBank accession no. MG09834) have intronic regions, whereas others (AfumCP, AniCP) have preserved their primitive intronless form.

Transmission of genes from bacteria to chlorophyceae and to various protistean organisms (Diatoms, Oomycetes and Dinophyceae) was also found on several occasions. On the phylogenetic tree (Fig. 2), all of these CPs are separated from the bacterial group, excepted for TpsCP01 (*Thalassiosira pseudonana*). Concerning TpsCP, it is probably the result of a recent LGT, following an interaction with a yet unidentified bacterium (no close homolog on the NJ and ML trees). The grouping of different protists together with the green alga *Chlamydomonas* is certainly due to lack of sequences from related organisms. We can expect to discover many more CPs in protists together with the building up of new genomic and EST databases. Concerning the original transfer of a CP from bacteria to eukaryotes, an interesting candidate came out from our NJ tree (Fig. 2), namely FtCP (*Francisella tularensis*), which has a fairly well supported basal position on the eukaryotic CP branch. Although bootstrap support was rather low in the ML tree (39%), both methods were positioning this CP as basal to the eukaryotic group. This bacterium is an intracellular host of protists [39], and it is hence possible that its ancestral relative is the source of eukaryotic CP. Marine Protists and Bacteria are known to cohabitate through endosymbiosis or in closely aggregated communities [40-42].

3.3 Conclusions and prospects

Our extensive search for catalase-peroxidase genes showed that a significant percentage of the so far sequenced 620 bacterial genomes do not possess this enzyme,

including aerobic organisms. Despite this chaotic distribution, a high degree of conservation was generally observed, and all predicted proteins resulted from the fusion of a duplicated gene, as already recently reported [5,8]. Based on this observation, the duplication must have been extremely old. In this respect, a study by Klotz and Lowen (2003) proposed that CP are at the origin of APX and C*c*P. According to the two authors, these two peroxidases would have lost the "duplicated pattern" by intron insertion. Our study confirms a very ancient origin of the catalase-peroxidase sequence, as we never found any non-duplicated CP form in the screened bacterial genomes.

However, recent investigations by our group on the whole peroxidase superfamily [4] suggest that the non-duplicated form of the ancestral sequence was still present in the bacteria that were engulfed by other organisms to become the mitochondrion and the chloroplast. This assumption is mainly based on the finding that APX and C*c*P were acquired in two distinct organelles (mitochondrion and chloroplast), and at two different times. Insertion of an intron in the same position for both peroxidases is hence very unlikely. After these primitive endosymbiotic events, a duplication of this ancestral sequence occurred in prokaryotes, and spread across bacteria and archaebacteria through speciation and lateral gene transfer. We cannot exclude that in the future the ancestral CP can be located in some yet unsequenced or even undiscovered bacteria. Finding out this ancestral form of peroxidase is a very exciting prospect. It would be the link that is still missing between the class I peroxidases (CP-APX-C*c*P) of this heme peroxidase superfamily, and it would bring precious information about the core sequence that gave rise to several thousands of peroxidases from the whole superfamily, leading to a "unified theory" of the peroxidase superfamily evolutionaty history.

Acknowledgments

The financial support of the Swiss National Science Foundation (grant 31-068003.02) to Claude Penel and Christophe Dunand is gratefully acknowledged, as well as the Austrian Science Fund (FWF project P18751) to Marcel Zamocky, Christa Jakopitsch and Christian Obinger. The authors also thank Robert Peck (University of Geneva) for useful discussion on the protistean reign.

Kingdom-Phylum	Class-Order	Genus	Species
Archaea-Crenarcheotes			*Aeropyrum pernix (c)* *Pyrobaculum aerophilum str.2 (c)* *Sulfolobus acidocaldarius DSM639 (c)* *Sulfolobus solfataricus P2 (c)* *Sulfolobus tokodaii str.7 (c)*
Archaea-Nanoarcheotes			*Nanoarchaeum equitans Kin4-M (c)*
Archaea-Euryarcheotes	Thermococci		*Pyrococcus abyssi GE5 (c)* *Pyrococcus furiosus DSM3638 (c)* *Pyrococcus horikoshii OT3 (c)* *Thermococcus kodakarensis KOD1 (c)*
	Methanococci		*Methanocaldococcus jannaschii DSM 2661 (c)* *Methanococcus maripaludis S2 (c)*
	Thermoplasmata		*Ferroplasma acidarmanus Fer1(p)* *Picrophilus torridus DSM9790 (c)* *Thermoplasma acidophilum DSM1728 (c)* *Thermoplasma volcanium GSS1 (c)*
	Methanopyrales		*Methanopyrus kandleri AV19 (c)*
	Methanobacteriales		*Methanosphaera stadtmanae DSM 3091 (c)* *Methanothermobacter thermautotrophicus Delta H (c)*
	Archaeoglobales		*Archaeoglobus fulgidus DSM 4304 (c)*
	Halobacteriales		*Haloarcula marismortui ATCC 43049 (c)* *Halobacterium sp. NRC-1 (c)* *Natronomonas pharaonis DSM 2160*
	Methanomicrobiales		*Methanospirillum hungatei JF-1 (p)*
	Methanosarcinales		*Methanosarcina acetivorans C2A (c)*
Aquificae			*Aquifex aeolicus VF5 (c)*
Chlamydiae			*Chlamydia muridarum Nigg (c)* *Chlamydia trachomatis A/HAR-13 and D/UW-3/CX (c)* *Chlamydophila abortus S26/3 (c)* *Chlamydophila caviae GPIC (c)* *Chlamydophila pneumoniae AR39, CWL029, J138, TW-183 (c)* *Parachlamydia sp. UWE25 (c)*
Chloroflexi			*Chloroflexus aurantiacus J-10-fl (p)* *Dehalococcoides ethenogenes 195 (c)* *Dehalococcoides sp. CBDB1 (c)*
Deinococcus-thermus			*Deinococcus geothermalis DSM 11300 (p)* *Deinococcus radiodurans R1 (c)* *Thermus thermophilus HB27 (c)* *Thermus thermophilus HB8 (c)*

Fusobacteria			*Fusobacterium nucleatum subsp. nucleatum ATCC 25586; subsp. vincentii ATCC 49256 (c)*
Spirochaetes			*Borrelia burgdorferi B31 (c)* *Borrelia garinii PBi (c)* *Leptospira interrogans serovar (c) Copenhageni str. Fiocruz L1-130 (c), serovar Lai str. 56601 (c)* *Treponema denticola ATCC 35405 (c)* *Treponema pallidum subsp. pallidum str. Nichols (c)*
Thermotogae			*Thermotoga maritima MSB8 (c)*
Actinobacteria	Rubrobacteridae	*Rubrobacter*	*Rubrobacter xylanophilus DSM 9941 (p)*
	Actinobacteridae	*Bifidobacterium*	*Bifidobacterium longum NCC2705 (c)*
		Corynebacterium	*Corynebacterium diphtheriae NCTC 13129 (c)* *Corynebacterium efficiens YS-314 (c)* *Corynebacterium glutamicum ATCC 13032 (c)* *Corynebacterium jeikeium K411 (c)*
		Kineococcus	*Kineococcus radiotolerans SRS30216 (p)*
		Leifsonia	*Leifsonia xyli subsp. xyli str. CTCB07 (c)*
		Propionibacterium	*Propionibacterium acnes KPA171202 (c)*
		Thermobifida	*Thermobifida fusca YX (c)*
		Tropheryma	*Tropheryma whipplei TW08/27, Twist (c)*
		Mycobacterium	*Mycobacterium leprae TN (c)* *Mycobacterium tuberculosis H37Rv, CDC1551 (c)* *Mycobacterium bovis AF2122/97 (c)* *Mycobacterium avium subsp. Paratuberculosis K-10 (c)* *Mycobacterium fortuitum ATCC 6841 (p)* *Mycobacterium intracellulare (p)*
		Streptomyces	*Streptomyces avermitilis MA-4680 (c)* *Streptomyces coelicolor A3(2) (c)* *Streptomyces reticuli (p)*
		Brevibacterium	*Brevibacterium linens BL2 (p)*
		Frankia	*Frankia sp. Ccl3, EAN1pec, R43 (p)*
		Janibacter	*Janibacter sp. HTCC 2649 (c)*
		Nocardia	*Nocardia farcinica IFM 10152 (c)*
		Nocardioides	*Nocardioides sp. JS614 (c)*
	unclassified	*Symbiobacterium*	*Symbiobacterium thermophilum IAM 14863 (c)*

Firmicutes	Lactobacillales		*Enterococcus faecalis V583 (c)* *Enterococcus faecium DO (p)* *Lactobacillus acidophilus NCFM (c)* *Lactobacillus casei ATCC 334 (p)* *Lactobacillus delbrueckii subsp. bulgaricus ATCC BAA-365 (p)* *Lactobacillus johnsonii NCC 533 (c)* *Lactobacillus plantarum WCFS1 (c)* *Lactobacillus sakei subsp. sakei 23K (c)* *Lactococcus lactis subsp. cremoris SK11 (p)* *Lactococcus lactis subsp. lactis Il1403 (c)* *Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293 (p)* *Oenococcus oeni PSU-1 (p)* *Pediococcus pentosaceus ATCC 25745 (p)* *Streptococcus agalactiae 2603V/R, A909, NEM316 (c)* *Streptococcus mutans UA159 (c)* *Streptococcus pneumoniae R6 (c)* *Streptococcus pneumoniae TIGR4 (c)* *Streptococcus pyogenes M1 GAS, MGAS10394, MGAS315, MGAS5005, MGAS6180, MGAS8232, SSI-1 (c) ; M49591(p)* *Streptococcus suis 89/1591 (p)* *Streptococcus thermophilus CNRZ1066, LMG18311 (c), LMD-9 (p)*
	Mollicutes		*Mesoplasma florum L1 (c)* *Mycoplasma capricolum subsp. capricolum ATCC 27343 (c)* *Mycoplasma gallisepticum R (c)* *Mycoplasma genitalium G37 (c)* *Mycoplasma hyopneumoniae 232 (c)* *Mycoplasma hyopneumoniae 7448 (c)* *Mycoplasma hyopneumoniae J (c)* *Mycoplasma mobile 163K (c)* *Mycoplasma mycoides subsp. mycoides SC str. PG1 (c)* *Mycoplasma penetrans HF-2 (c)* *Mycoplasma pneumoniae M129 (c)* *Mycoplasma pulmonis UAB CTIP (c)* *Mycoplasma synoviae 53 (c)* *Onion yellows phytoplasma OY-M (c)* *Ureaplasma parvum serovar 3 str. ATCC 700970 (c)*

Firmicutes (continued)	Clostridia	*Carboxydothermus*	*Carboxydothermus hydrogenoformans Z-2901 (c)*
		Clostridium	*Clostridium acetobutylicum ATCC 824 (c)* *Clostridium perfringens str. 13 (c)* *Clostridium tetani E88 (c)* *Clostridium thermocellum ATCC 27405 (p)*
		Moorella	*Moorella thermoacetica ATCC 39073 (c)*
		Syntrophomonas	*Syntrophomonas wolfei str. Goettingen (p)*
		Thermoanaerobacter	*Thermoanaerobacter tengcongensis MB4 (c)*
		Alkaliphilus	*Alkaliphilus metalliredigenes QYMF (p)*
		Desulfitobacterium	*Desulfitobacterium hafniense DCB-2 (p)*
	Bacillales	*Exiguobacterium*	*Exiguobacterium sp. 255-15 (p)*
		Listeria	*Listeria innocua Clip11262 (c)* *Listeria monocytogenes EGD-e (c); 1/2a F6854 (p)* *Listeria monocytogenes str. 4b F2365 (c) and 4b H7858 (p)*
		Oceanobacillus	*Oceanobacillus iheyensis HTE831 (c)*
		Staphylococcus	*Staphylococcus aureus RF122 (c)* *Staphylococcus aureus subsp. Aureus COL MSSA252, MSS0476, MW2, Mu50, N315 (c)* *Staphylococcus epidermidis ATCC 12228, RP62A (c)* *Staphylococcus haemolyticus JCSC1435 (c)* *Staphylococcus saprophyticus sub. saprophyticus ATCC 15305 (c)*
		Bacillus	*Bacillus anthracis str. 'Ames Ancestor' (c)* *Bacillus anthracis A1055, A2012, Australia 94, CNEVA9066, Kruger B, Vollum, Western North America USA6153 (p); Ames, Sterne (c)* *Bacillus cereus G9241 (p) ; ATCC 10987, ATCC14579, E33L, (c)* *Bacillus clausii KSM-K16 (c)* *Bacillus licheniformis ATCC 14580 (c)* *Bacillus sp. B14911 (c)* *Bacillus subtilis subsp. subtilis str. 168 (c)* *Bacillus thuringiensis serovar konkukian str. 97-27 (c)* *Bacillus halodurans C-125 (c)*
		Geobacillus	*Geobacillus kaustophilus HTA426 (c)* *Geobacillus stearothermophilus (p)*

Proteobacteria	Alpha subdivision-Rickettsiales		*Anaplasma marginale str. St. Maries (c)* *Candidatus Pelagibacter ubique HTCC1062 (c)* *Ehrlichia canis str. Jake (c)* *Ehrlichia ruminantium Gardel (c)* *Ehrlichia ruminantium str. Welgevonden (c)* *Rickettsia akari str. Hartford (p)* *Rickettsia conorii str. Malish 7 (c)* *Rickettsia felis URRWXCal2 (c)* *Rickettsia prowazekii Madrid E (c)* *Rickettsia rickettsii (p)* *Rickettsia sibirica 246 (p)* *Rickettsia typhi str. Wilmington (c)* *Wolbachia endosymbiont of Drosophila ananassae (p)* *Wolbachia endosymbiont of Drosophila melanogaster (c)* *Wolbachia endosymbiont of Drosophila simulans (p)* *Wolbachia endosymbiont strain TRS of Brugia (c)*
	Alpha subdivision-Rhodospirillales		*Gluconobacter oxydans 621H (c)* *Magnetospirillum magneticum AMB-1 (c)* *Magnetospirillum magnetotacticum MS-1 (p)* *Rhodospirillum rubrum ATCC 11170 (c)*
	Alpha subdivision-Rhodobacterales	*Rhodobacter*	*Rhodobacter sphaeroides 2.4.1 (c)* *Rhodobacter capsulatus B10 (p)*
		Loktanella	*Loktanella vestfoldensis SKA53 (c)*
		Oceanicaulis	*Oceanicaulis alexandrii HTCC2633 (c)*
		Oceanicola	*Oceanicola batsensis HTCC2649 (c)* *Oceanicola granulosus HTCC2516 (c)*
		Roseovarius	*Roseovarius sp. 217 (c)* *Roseovarius nubinhibens ISM (c)*
		Silicibacter	*Silicibacter pomeroyi DSS-3 (c)* *Silicibacter sp. TM1040 (p)*
		Sulfitobacter	*Sulfitobacter sp. EE36 and NAS-14.1 (c)*
	Alpha subdivision-Rhizobiales	*Bartonella*	*Bartonella henselae str. Houston-1 (c)* *Bartonella quintana str. Toulouse (c)*
		Brucella	*Brucella abortus biovar 1 str. 9-941 (c)* *Brucella melitensis 16M (c)* *Brucella melitensis biovar Abortus 2308 (c)* *Brucella suis 1330 (c)*
		Agrobacterium	*Agrobacterium tumefaciens str. C58 (c)*
		Bradyrhizobium	*Bradyrhizobium japonicum USDA 110 (c)*

	Alpha subdivision-Rhizobiales (continued)	*Mesorhizobium*	*Mesorhizobium loti MAFF303099 (c)* *Mesorhizobium sp. BNC1 (p)*
		Rhizobium	*Rhizobium etli (c)* *Rhizobium leguminosarum bv. Phaseoli (p)*
		Rhodopseudomonas	*Rhodopseudomonas palustris CGA009 (c)*
		Sinorhizobium	*Sinorhizobium meliloti 1021 (c)*
	Alpha subdivision-Sphingomonadales	*Zymomonas*	*Zymomonas mobilis subsp. mobilis ZM4 (c)*
		Erythrobacter	*Erythrobacter sp. NAP1 (c)* *Erythrobacter litoralis HTCC2594 (p)*
		Novosphingobium	*Novosphingobium aromaticivorans DSM 12444 (p)*
	Alpha subdivision-Caulobacterales	*Caulobacter*	*Caulobacter crescentus CB15 (c)*
	Alpha subdivision-Parvularculales	*Parvularcula*	*Parvularcula bermudensis HTCC2503 (c)*
	Beta subdivision-Hydrogenophylales	*Thiobacillus*	*Thiobacillus denitrificans ATCC 25259 (c)*
	Beta subdivision-Methylophilales	*Methylobacillus*	*Methylobacillus flagellatus KT (p)*
	Beta subdivision-Rhodocyclales	*Azoarcus*	*Azoarcus sp. EbN1 (c)*
		Dechloromonas	*Dechloromonas aromatica RCB (c)*
	Beta subdivision-Burkholderiales	*Bordetella*	*Bordetella bronchiseptica RB50 (c)* *Bordetella parapertussis 12822 (c)* *Bordetella pertussis Tohama (c)*
		Chromobacterium	*Chromobacterium violaceum ATCC 12472 (c)*
		Neisseria	*Neisseria gonorrhoeae FA 1090 (c)* *Neisseria meningitidis MC58 (c)* *Neisseria meningitidis Z2491 (c)*
		Rhodoferax	*Rhodoferax ferrireducens DSM 15236 (p)*
		Burkholderia (all sequences are almost identical, except B.fungorum and B.cepacia)	*Burkholderia pseudomallei 1655, 1710a, 668, Pasteur, S13 (p); 1710b, K96243 (c)* *Burkholderia mallei ATCC23344 (c); GB8 horse 4, JHU, NCTC10247, 10229, 10399 (p)* *Burkholderia ambifaria AMMD (p)* *Burkholderia cenocepacia AU 1054 and HI2424 (p)* *Burkholderia cepacia (p)* *Burkholderia fungorum LB400 (p)* *Burkholderia sp. 383 (c)* *Burkholderia thailandensis E264 (c)* *Burkholderia vietnamiensis G4 (p)*

	Beta subdivision-Burkholderiales (continued)	*Polaromonas*	*Polaromonas sp. JS666 (p)*
		Ralstonia	*Ralstonia eutropha JMP134 (c)* *Ralstonia metallidurans CH34 (p)* *Ralstonia solanacearum GMI1000 (c)*
		Rubrivivax	*Rubrivivax gelatinosus PM1 (p)*
	Beta subdivision-Nitrosomonadales	*Nitrospira*	*Nitrosospira multiformis ATCC 25196 (c)*
		Nitrosomonas	*Nitrosomonas europaea ATCC 19718 (c)*
	Delta subdivision-Bdellovibrionales	*Bdellovibrio*	*Bdellovibrio bacteriovorus HD100 (c)*
	Delta subdivision-Desulfobacterales	*Desulfotalea*	*Desulfotalea psychrophila LSv54 (c)*
	Delta subdivision-Desulfovibrionales	*Desulfovibrio*	*Desulfovibrio desulfuricans G20 (c)* *Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough (c)*
	Delta subdivision-Myxococcales	*Anaeromyxobacter*	*Anaeromyxobacter dehalogenans 2CP-C (p)*
	Delta subdivision-Syntrophobacterales	*Syntrophobacter*	*Syntrophus aciditrophicus SB (c)* *Syntrophobacter fumaroxidans MPOB (p)*
	Delta subdivision-Desulfuromonadales	*Geobacter*	*Geobacter metallireducens GS-15 (c)* *Geobacter sulfurreducens PCA (c)*
		Desulfuromonas	*Desulfuromonas acetoxidans DSM 684 (p)*
		Pelobacter	*Pelobacter carbinolicus DSM 2380 (c)* *Pelobacter propionicus DSM 2379 (p)*
	Epsilon subdivision		*Campylobacter coli RM2228 (p)* *Campylobacter jejuni RM1221 (c); subsp. jejuni NCTC 11168 (c)* *Campylobacter lari RM2100 (p)* *Helicobacter hepaticus ATCC 51449 (c)* *Helicobacter pylori 26695, J99 (c)* *Thiomicrospira denitrificans ATCC 33889 (c)* *Wolinella succinogenes DSM 1740 (c)*
	Gamma subdivision-Chromatiales	*Nitrosococcus*	*Nitrosococcus oceani ATCC 19707 (c)*
	Gamma subdivision-Methylococcales	*Methylococcus*	*Methylococcus capsulatus str. Bath (c)*
	Gamma subdivision-Pasteurellales		*Haemophilus ducreyi 35000HP (c)* *Haemophilus influenzae 86-028NP, Rd KW20 (c); R2846, R2866 (p)* *Mannheimia succiniciproducens MBEL55E (c)* *Pasteurella multocida subsp. multocida str. Pm70 (c)*

Gamma subdivision-Enterobacteriales	*Buchnera*	*Buchnera aphidicola str. APS (Acyrthosiphon pisum) (c)* *Buchnera aphidicola str. Bp (Baizongia pistaciae) (c)* *Buchnera aphidicola str. Sg (Schizaphis graminum) (c)*
	Blochmannia	*Candidatus Blochmannia floridanus (c)* *Candidatus Blochmannia pennsylvanicus str. BPEN (c)*
	Photorhabdus	*Photorhabdus luminescens subsp. laumondii TTO1 (c)*
	Wigglesworthia	*Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis (c)*
	Erwinia	*Erwinia carotovora subsp. atroseptica SCRI1043 (c)*
	Escherichia	*Escherichia coli CFT073, K12, O157:H7, O157:H7 EDL933 (c); HS (p)*
	Shigella	*Shigella boydii Sb227 (c) and BS512 (p)* *Shigella dysenteriae Sd197 (c)* *Shigella flexneri 2a 2457T and 301 (c)* *Shigella sonnei Ss046 (c)*
	Salmonella	*Salmonella enterica subsp. enterica serovar Choleraesuis SC-B67 (c),* *serovar Paratyphi A ATCC 9150 (c) ;* *serovar Typhi str. CT18 and Ty2 (c)* *Salmonella typhimurium LT2 (c)*
	Yersinia	*Yersinia pestis CO92 and KIM (c)* *Yersinia pestis biovar Medievalis str. 91001(c)* *Yersinia pseudotuberculosis IP 32953 (c)*
Gamma subdivision-Legionellales	*Coxiella*	*Coxiella burnetii RSA 493 (c)*
	Legionella	*Legionella pneumophila Lens and Paris (c)* *Legionella pneumophila subsp. pneumophila Philadelphia 1 (c)*
Gamma subdivision-Thiotrichales	*Thiomicrospira*	*Thiomicrospira crunogena XCL-2 (c)*
	Francisella	*Francisella tularensis subsp. tularensis SCHU S4 (c)*
Gamma subdivision-Pseudomonadales	*Acinetobacter*	*Acinetobacter sp. ADP1 (c)*
	Psychrobacter	*Psychrobacter arcticus 273-4 (c)*
	Pseudomonas	*Pseudomonas aeruginosa PAO1(c) and UCVPP (p)* *Pseudomonas fluorescens PfO-1 (c)* *Pseudomonas putida KT2440 (c)* *Pseudomonas syringae pv. phaseolicola 1448A (c)*
	Azotobacter	*Azotobacter vinelandii AvOP (p)*

	Gamma subdivision-Vibrionales	*Photobacterium*	*Photobacterium profundum SS9 (c)*
		Vibrio	*Vibrio fischeri ES114 (c)* *Vibrio sp. MED222 (c)* *Vibrio splendidus 12B01 (c)* *Vibrio cholerae O1 biovar eltor str. N16961 (c)* *Vibrio parahaemolyticus RIMD 2210633 (c)* *Vibrio vulnificus CMCP6 and YJ016 (c)*
	Gamma subdivision-Xanthomonadales	*Xanthomonas*	*Xanthomonas oryzae pv. oryzae KACC10331 (c)* *Xanthomonas campestris pv. campestris 8804 and ATCC 33913 (c)* *Xanthomonas campestris pv. vesicatoria 85-10 (c)* *Xanthomonas axonopodis pv. citri str. 306 (c)*
		Xylella	*Xylella fastidiosa 9a5c (c)* *Xylella fastidiosa Temecula1 (c)* *Xylella fastidiosa Ann-1 (p)*
	Gamma subdivision-Alteromonadales	*Pseudoalteromonas*	*Pseudoalteromonas haloplanktis TAC125 (c)* *Pseudoalteromonas atlantica T6c (p)*
		Colwellia	*Colwellia psychrerythraea str. 34H (c)*
		Idiomarina	*Idiomarina loihiensis L2TR (c)* *Idiomarina baltica OS145 (p)*
		Saccharophagus	*Saccharophagus degradans 2-40 (p)*
		Shewanella	*Shewanella oneidensis MR-1 (c)* *Shewanella baltica OS155 (p)* *Shewanella denitrificans OS217 (p)* *Shewanella frigidimarina NCIMB 400 (p)*
	Gamma subdivision-Oceanospirillales	*Oceanospirillum*	*Oceanospirillum sp. MED92 (c)*
		Marinomonas	*Marinomonas sp. MED121 (c)*
		Chromohalobacter	*Chromohalobacter salexigens DSM 3043 (p)*
		Hahella	*Hahella chejuensis KCTC 2396 (c)*
Cyanobacteria	Nostocales		*Anabaena variabilis ATCC 29413 (c)* *Nostoc punctiforme PCC 73102 (p)* *Nostoc sp. PCC 7120 (c)*
	Prochlorales		*Prochlorococcus marinus MIT9211, MIT9312, MIT9313, NATL2A (c)* *Prochlorococcus marinus subsp. marinus str. CCMP1375 (c)* *Prochlorococcus marinus subsp. pastoris str CCMP1986 (c)*

Cyanobacteria (continued)	Chroococcales	*Crocosphaera*	*Crocosphaera watsonii WH 8501 (p)*
		Thermosynechococcus	*Thermosynechococcus elongatus BP-1 (c)*
		Synechococcus	*Synechococcus sp. WH8102 and CC9902 (c)* *Synechococcus elongatus PCC6301 (c) and PCC7942 (p)* *Synechococcus sp. str.CC9605 (c)*
		Synechocystis	*Synechocystis sp. str. PCC 6803 (c)*
	Gloeobacterales	*Gloeobacter*	*Gloeobacter violaceus PCC 7421 (c)*
	Oscillatoriales	*Trichodesmium*	*Trichodesmium erythraeum IMS101 (p)*
Bacteroidetes/Chlorobi	Bacteroidetes	*Bacteroides*	*Bacteroides fragilis NCTC 9343 (c)* *Bacteroides fragilis YCH46 (c)* *Bacteroides thetaiotaomicron VPI-5482 (c)*
		Cytophaga	*Cytophaga hutchinsonii (p)*
		Porphyromonas	*Porphyromonas gingivalis W83 (c)*
		Cellulophaga	*Cellulophaga sp. MED134 (c)*
		Flavobacterium	*Flavobacterium sp. MED217 (c)*
		Salinibacter	*Salinibacter ruber DSM13855 (c)*
		Tenacibaculum	*Tenacibaculum sp. MED152 (c)*
	Chlorobia	*Chlorobium*	*Chlorobium tepidum TLS (c)* *Chlorobium limicola DSM 245 (p)* *Chlorobium chlorochromatii CaD3 (c)* *Chlorobium phaeobacteroides DSM 266 (p)*
		Pelodictyon	*Pelodictyon luteolum DSM 273 (c)* *Pelodictyon phaeoclathratiforme BU-1 (p)*
		Synechococcus	*Synechococcus sp. CC9902 (c)* *Synechococcus sp. WH 8102 (c)*
		Prosthecochloris	*Prosthecochloris vibrioformis DSM 265 (p)* *Prosthecochloris aestuarii DSM 271 (p)*
Fibrobacteres/Acidobacteria			*Solibacter usitatus Ellin6076 (p)*
Planctomycetes			*Rhodopirellula baltica SH 1 (c)*

Supplemental Table 1

Table representing all screened prokaryotic organisms. Grey background represents organisms without CP. White background: 1 CP per organism. Black background: 2 CP per organism. (c): genome sequencing complete. (p): genome sequencing partial or not started.

Kingdom-Phylum	Class-Order	Genus	Species
Fungi-Ascomycetes	Eurotiomycetes	Aspergillus	*Aspergillus fumigatus* *Aspergillus nidulans*
		Penicillium	*Penicillium marneffei*
	Leotiomycetes	Blumeria	*Blumeria graminis*
	Sordariomycetes	Gibberella	*Gibberella zeae PH-1*
		Magnaporthe	*Magnaporthe grisea 70-15*
		Neurospora	*Neurospora crassa OR74A*
		Hypocrea (Trichoderma)	*Trichoderma reesei QM9414*
	Unclassified	Mycosphaerella	*Mycosphaerella graminicola IPO323*
Fungi-Basidiomycota	Ustilaginomycetes	Ustilago	*Ustilago maydis 521*

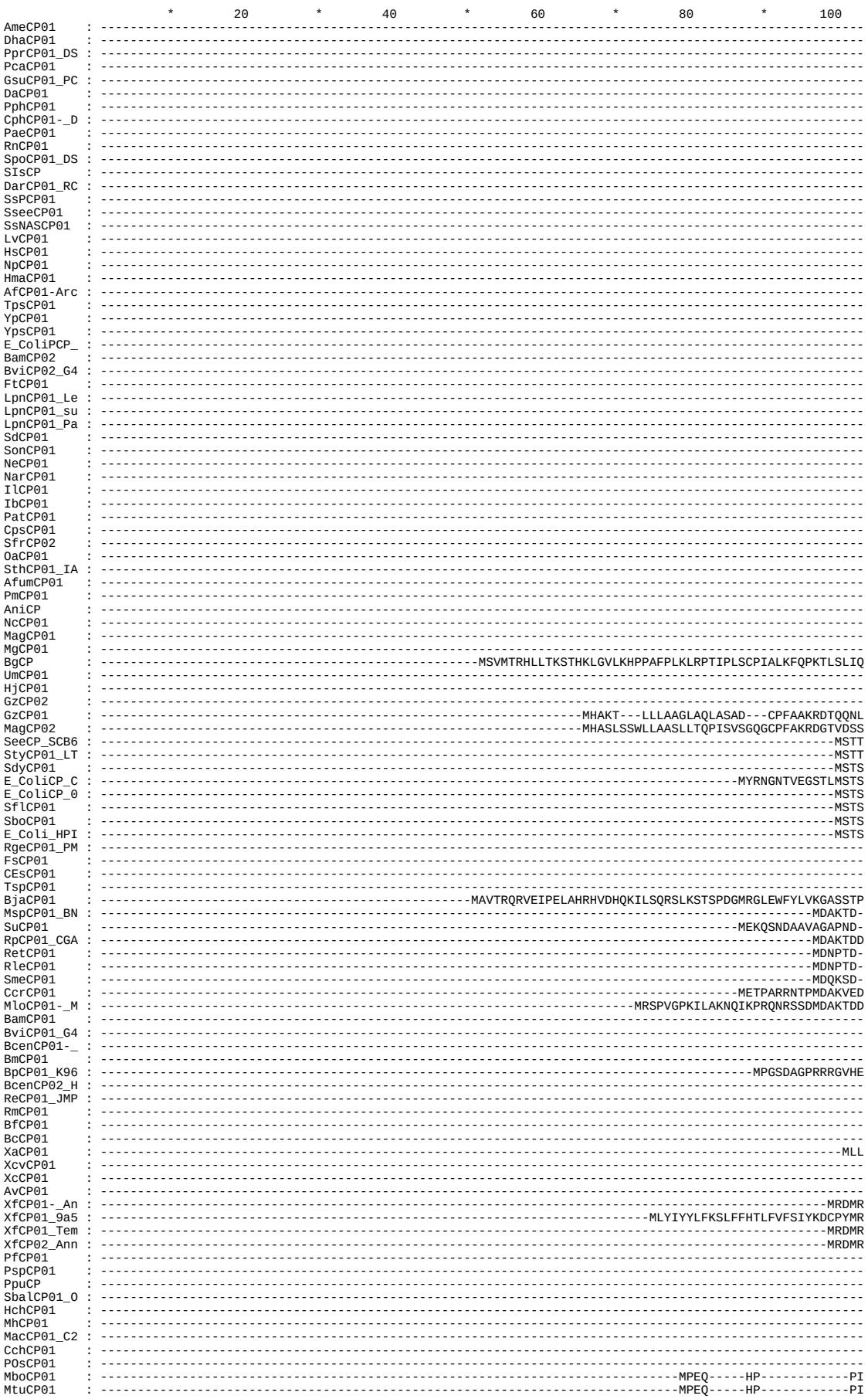
Supplemental Table 2

List of fungal organisms possessing a catalase-peroxidase. Black background represents fungi encoding two CP.

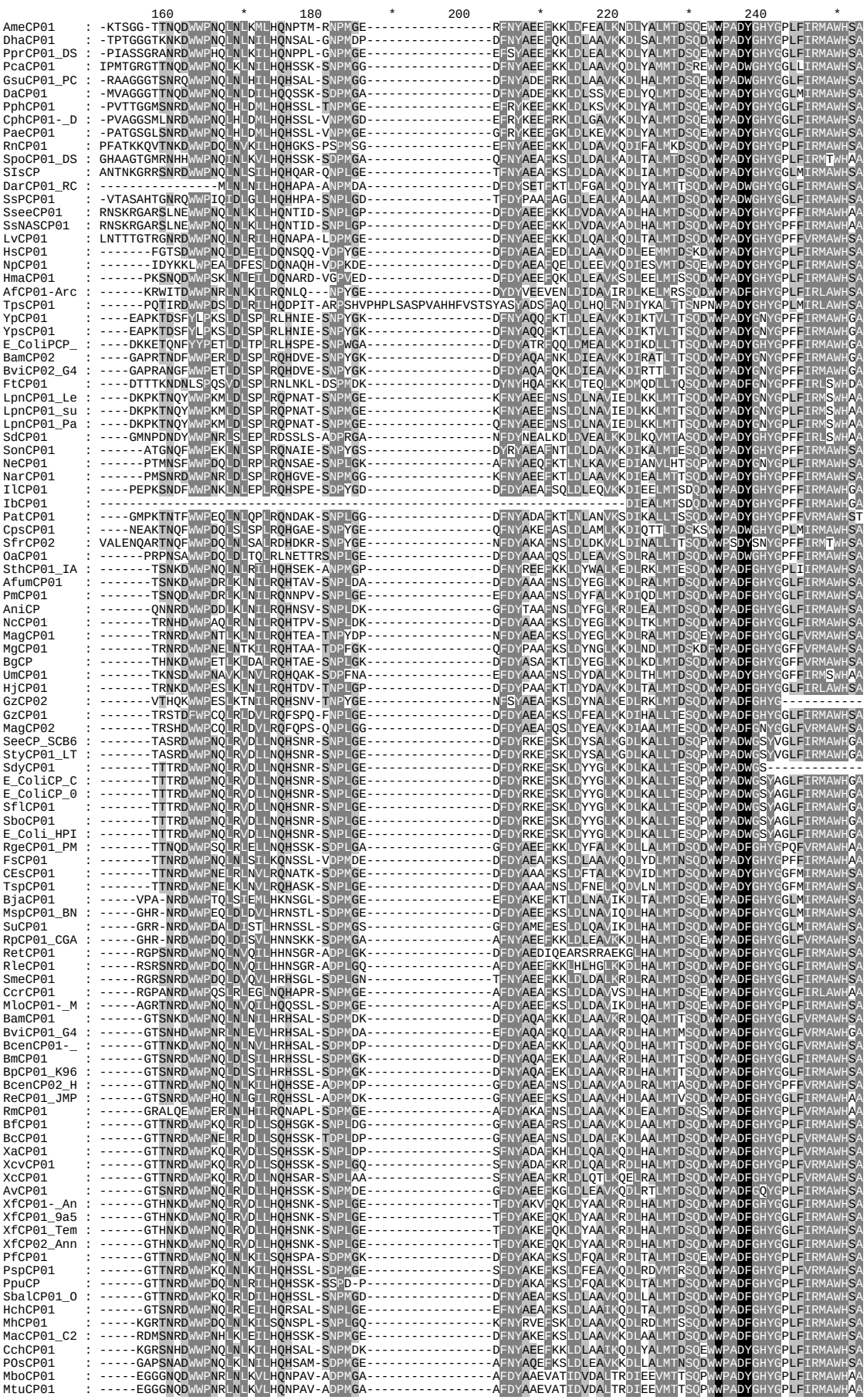
Kingdom-Phylum	Class-Order	Genus	Species
Protists-Alveolates	Dinophyceae	Heterocapsa	*Heterocapsa triquetra*
		Karenia	*Karenia brevis*
Protists-Stramenopiles	Bacillariophyceae	Phaeodactylum	*Phaeodactylum tricornutum*
	Coscinodiscophyceae	Thalassiosira	*Thalassiosira pseudonana CCMP1335*
	Oomycetes	Phytophthora	*Phytophthora infestans* *Phytophthora ramorum* *Phytophthora sojae* *Phytophthora parasitica*
Viridiplantae	Chlorophyceae	Chlamydomonas	*Chlamydomonas reinhardtii CC-503 cw92*

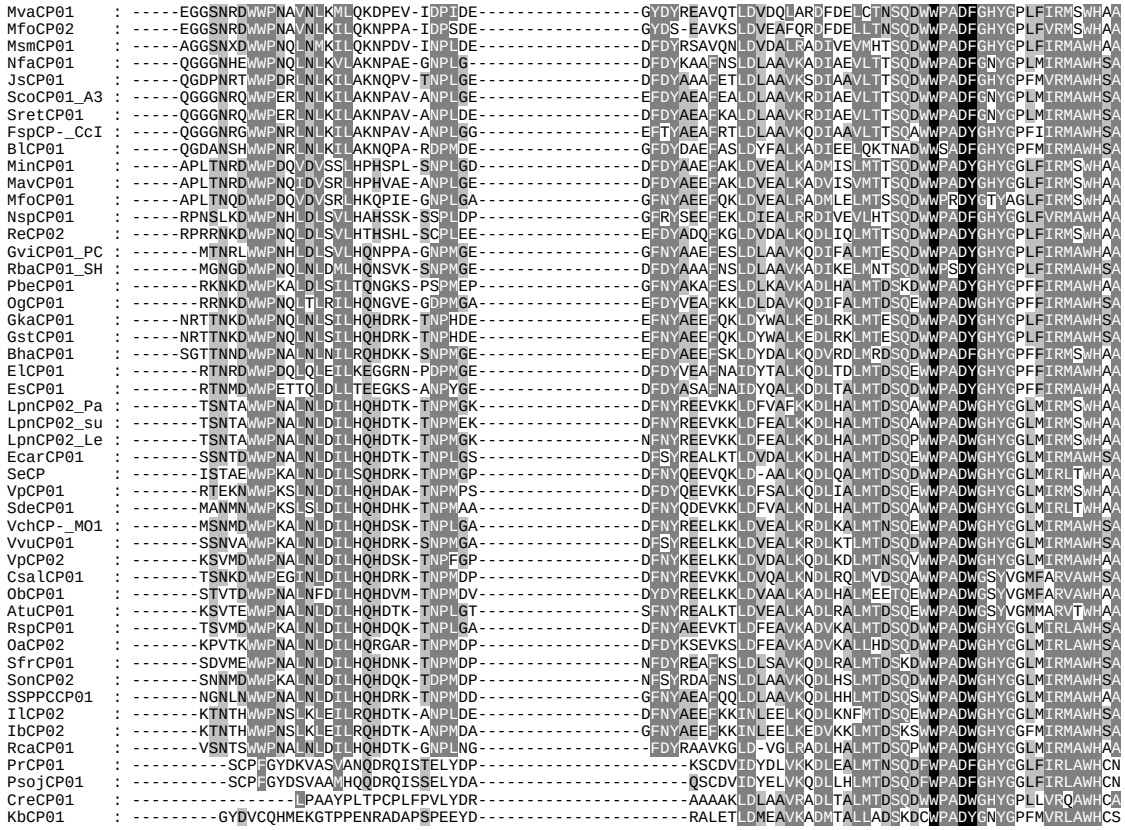
Supplemental Table 3

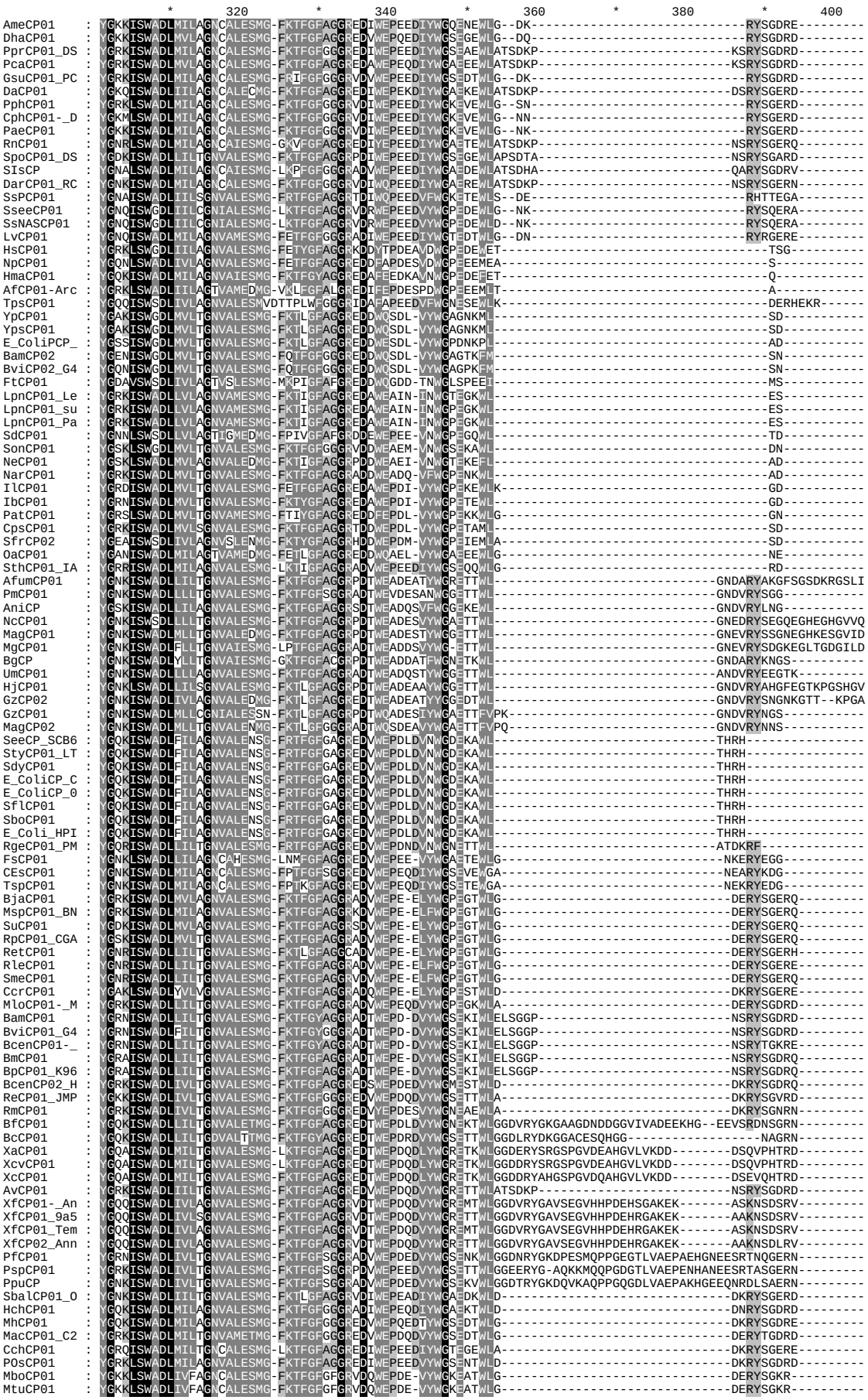
List of all prokaryotic organisms (except fungi) encoding a catalase-peroxidase.

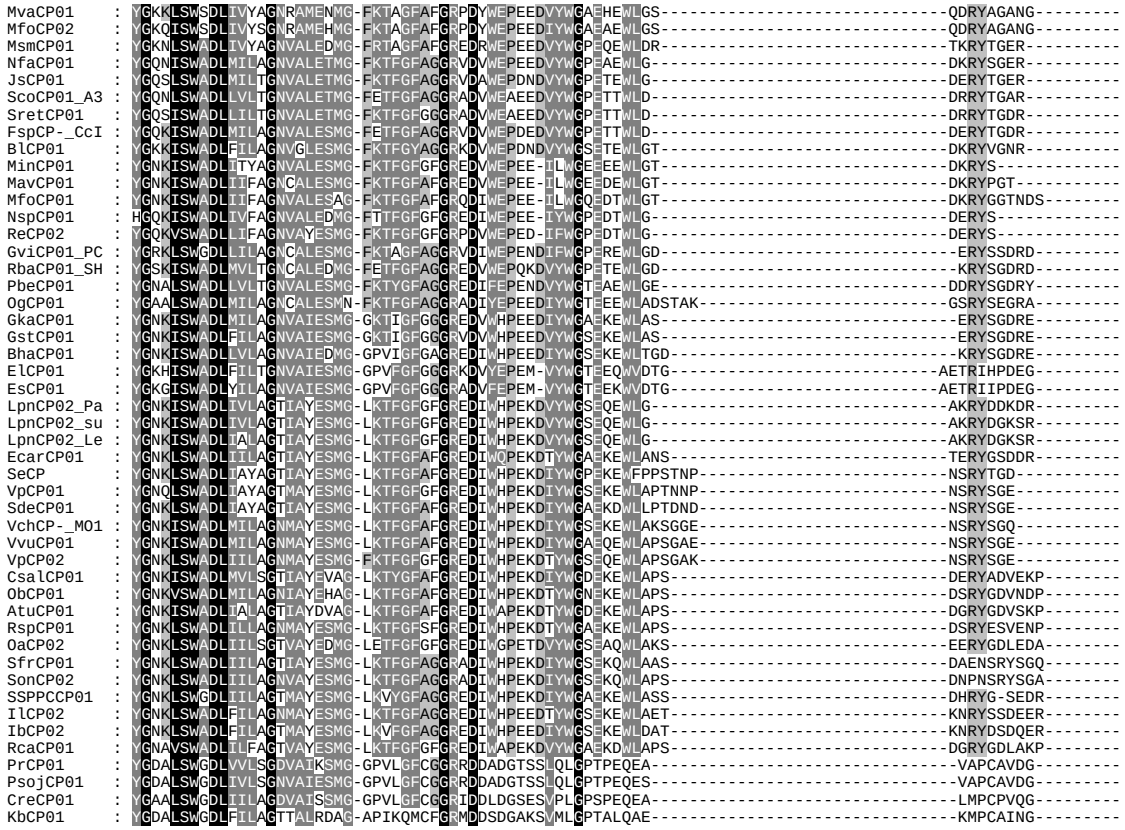

* 20 * 40 * 60 * 80 * 100
AmeCP01 :
DhaCP01 :
PprCP01_DS :
PcaCP01 :
GsuCP01_PC :
DaCP01 :
PphCP01 :
CphCP01-_D :
PaeCP01 :
RnCP01 :
SpoCP01_DS :
SIsCP :
DarCP01_RC :
SsPCP01 :
SseeCP01 :
SsNASCP01 :
LvCP01 :
HsCP01 :
NpCP01 :
HmaCP01 :
AfCP01-Arc :
TpsCP01 :
YpCP01 :
YpsCP01 :
E_ColiPCP_ :
BamCP02 :
BviCP02_G4 :
FtCP01 :
LpnCP01_Le :
LpnCP01_su :
LpnCP01_Pa :
SdCP01 :
SonCP01 :
NeCP01 :
NarCP01 :
IlCP01 :
IbCP01 :
PatCP01 :
CpsCP01 :
SfrCP02 :
OaCP01 :
SthCP01_IA :
AfumCP01 :
PmCP01 :
AniCP :
NcCP01 :
MagCP01 :
MgCP01 :
BgCP : MSVMTRHLLTKSTHKLGVLKHPPAFPLKLRPTIPLSCPIALKFQPKTLSLIQ
UmCP01 :
HjCP01 :
GzCP02 :
GzCP01 : MHAKT---LLLAAGLAQLASAD---CPFAAKRDTQQNL
MagCP02 : MHASLSSWLLAASLLTQPISVSGQGCPFAKRDGTVDSS
SeeCP_SCB6 : MSTT
StyCP01_LT : MSTT
SdyCP01 : MSTS
E_ColiCP_C : MYRNGNTVEGSTLMSTS
E_ColiCP_0 : MSTS
SflCP01 : MSTS
SboCP01 : MSTS
E_Coli_HPI : MSTS
RgeCP01_PM :
FsCP01 :
CEsCP01 :
TspCP01 :
BjaCP01 : MAVTRQRVEIPELAHRHVDHQKILSQRSLKSTSPDGMRGLEWFYLVKGASSTP
MspCP01_BN : MDAKTD-
SuCP01 : MEKQSNDAAVAGAPND-
RpCP01_CGA : MDAKTDD
RetCP01 : MDNPTD-
RleCP01 : MDNPTD-
SmeCP01 : MDQKSD-
CcrCP01 : METPARRNTPMDAKVED
MloCP01-_M : MRSPVGPKILAKNQIKPRQNRSSDMDAKTDD
BamCP01 :
BviCP01_G4 :
BcenCP01-_ :
BmCP01 :
BpCP01_K96 : MPGSDAGPRRRGVHE
BcenCP02_H :
ReCP01_JMP :
RmCP01 :
BfCP01 :
BcCP01 :
XaCP01 : MLL
XcvCP01 :
XcCP01 :
AvCP01 :
XfCP01-_An : MRDMR
XfCP01_9a5 : MLYIYYLFKSLFFHTLFVFSIYKDCPYMR
XfCP01_Tem : MRDMR
XfCP02_Ann : MRDMR
PfCP01 :
PspCP01 :
PpuCP :
SbalCP01_O :
HchCP01 :
MhCP01 :
MacCP01_C2 :
CchCP01 :
POsCP01 :
MboCP01 : MPEQ-----HP----------PI
MtuCP01 : MPEQ-----HP----------PI

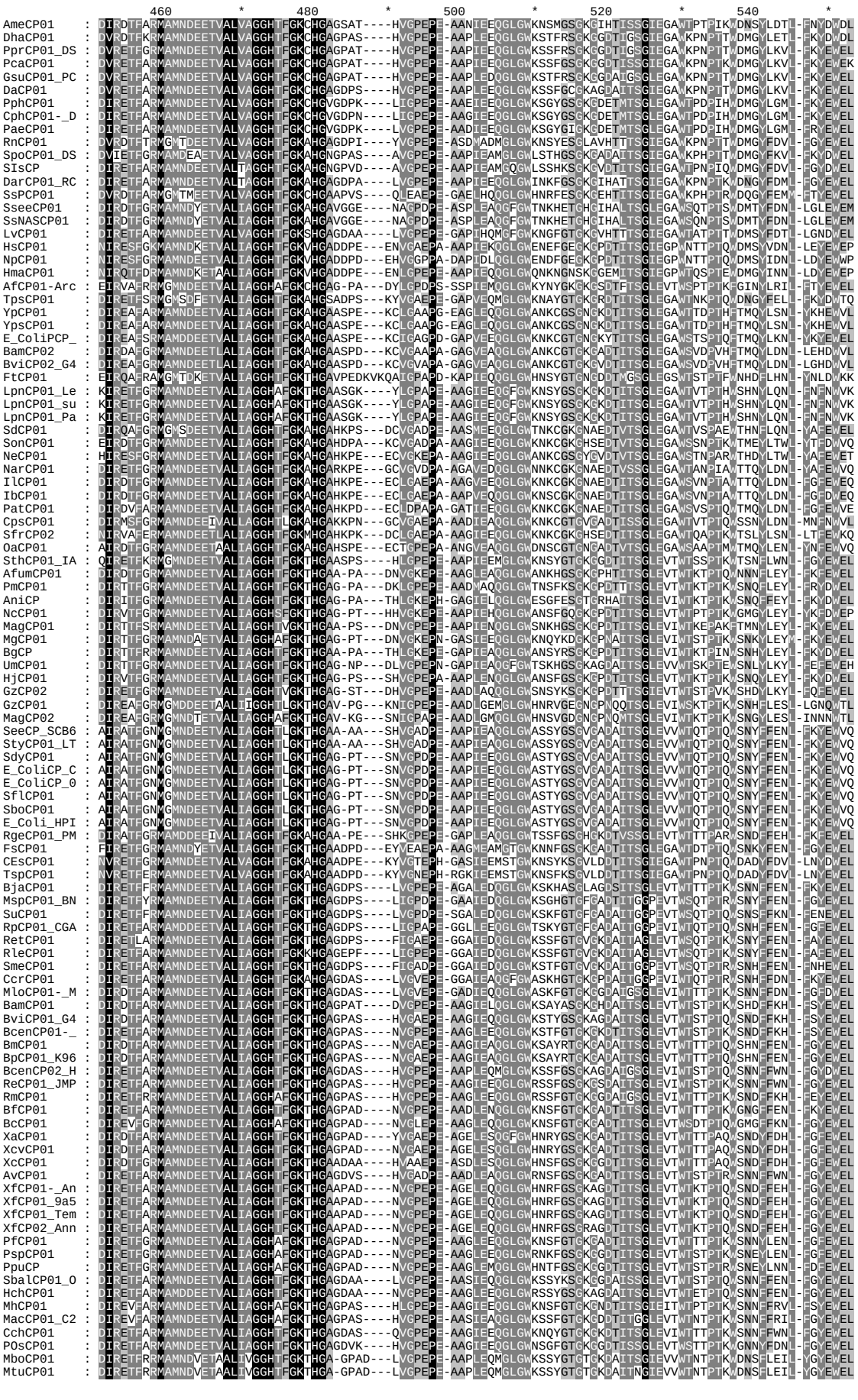
```
MvaCP01    : ----------------------------------------------------------------------------MPEA---TEHP------------PI
MfoCP02    : ----------------------------------------------------------------------------MAEA---ETHP------------PI
MsmCP01    : ----------------------------------------------------------------------------MPEDRPIEDSP------------PI
NfaCP01    : ----------------------------------------------------------------------------MSVE-----HP------------PI
JsCP01     : ------------------------------------------------------------------------------------------------------N
ScoCP01_A3 : ----------------------------------------------------------------------------MSEN-----HD------------AI
SretCP01   : ----------------------------------------------------------------------------MTEN-----HD------------AI
FspCP-_CcI : ----------------------------------------------------------------------------MSEN-----HD------------AV
BlCP01     : ----------------------------------------------------------------------------MTES-------------------ST
MinCP01    : --------------------------------------------------------------------------MSSDTSSSRPPQ-------------P
MavCP01    : --------------------------------------------------------------------------MSSDTSASRPPQ-------------P
MfoCP01    : ------------------------------------------------------------------------MPPNTPDASDARPPQ-------------A
NspCP01    : ------------------------------------------------------------------------MTDSQDNRTPESPQGVDRKAEGGCPVLH
ReCP02     : ----------------------------------------------------------------------------MTDKQH-------------------
GviCP01_PC : ------------------------------------------------------------------------------------------------------
RbaCP01_SH : MRLVHGKDAVNALDVSWLSPILVRDPTSPSESAWDLFATFMRTRTMKSTATNMNQHPTTAQICRRFATRLFRGTVSIRPLTLTLGCLAASASASLVAQETATN
PbeCP01    : ------------------------------------------------------------------------------------------------------
OgCP01     : ------------------------------------------------------------------------------------------------------
GkaCP01    : ------------------------------------------------------------------------------------------------------
GstCP01    : ------------------------------------------------------------------------------------------------------
BhaCP01    : ------------------------------------------------------------------------------------------------------
ElCP01     : ------------------------------------------------------------------------------------------------------MQ
EsCP01     : ------------------------------------------------------------------------------------------------------
LpnCP02_Pa : ------------------------------------------------------------------------------------------------------
LpnCP02_su : ----------------------------------------------------------------------------------------------MNYDLRPGVS
LpnCP02_Le : ------------------------------------------------------------------------------------------------------
EcarCP01   : ------------------------------------------------------------------------------------------------------
SeCP       : ------------------------------------------------------------------------------------------------------
VpCP01     : ------------------------------------------------------------------------------------------------------
SdeCP01    : ------------------------------------------------------------------------------------------------------
VchCP-_MO1 : ------------------------------------------------------------------------------------------------------
VvuCP01    : ------------------------------------------------------------------------------------------------------
VpCP02     : ------------------------------------------------------------------------------------------------------
CsalCP01   : ------------------------------------------------------------------------------------------------------
ObCP01     : ------------------------------------------------------------------------------------------------------
AtuCP01    : ------------------------------------------------------------------------------------------------------
RspCP01    : ------------------------------------------------------------------------------------------------------
OaCP02     : ------------------------------------------------------------------------------------------------------
SfrCP01    : ------------------------------------------------------------------------------------------------------
SonCP02    : ------------------------------------------------------------------------------------------------------
SSPPCCP01  : ----------------------------------------------------------------------------MGTQPARKLRNRVFPHPHNHRKEK
IlCP02     : ------------------------------------------------------------------------------------------------------
IbCP02     : ------------------------------------------------------------------------------------------------------
RcaCP01    : ------------------------------------------------------------------------------------------------------
PrCP01     : ------------------------------------------------------------------------------------------------------
PsojCP01   : ------------------------------------------------------------------------------------------------------
CreCP01    : ------------------------------------------------------------------------------------------------------
KbCP01     : ------------------------------------------------------------------------------------------------------
```

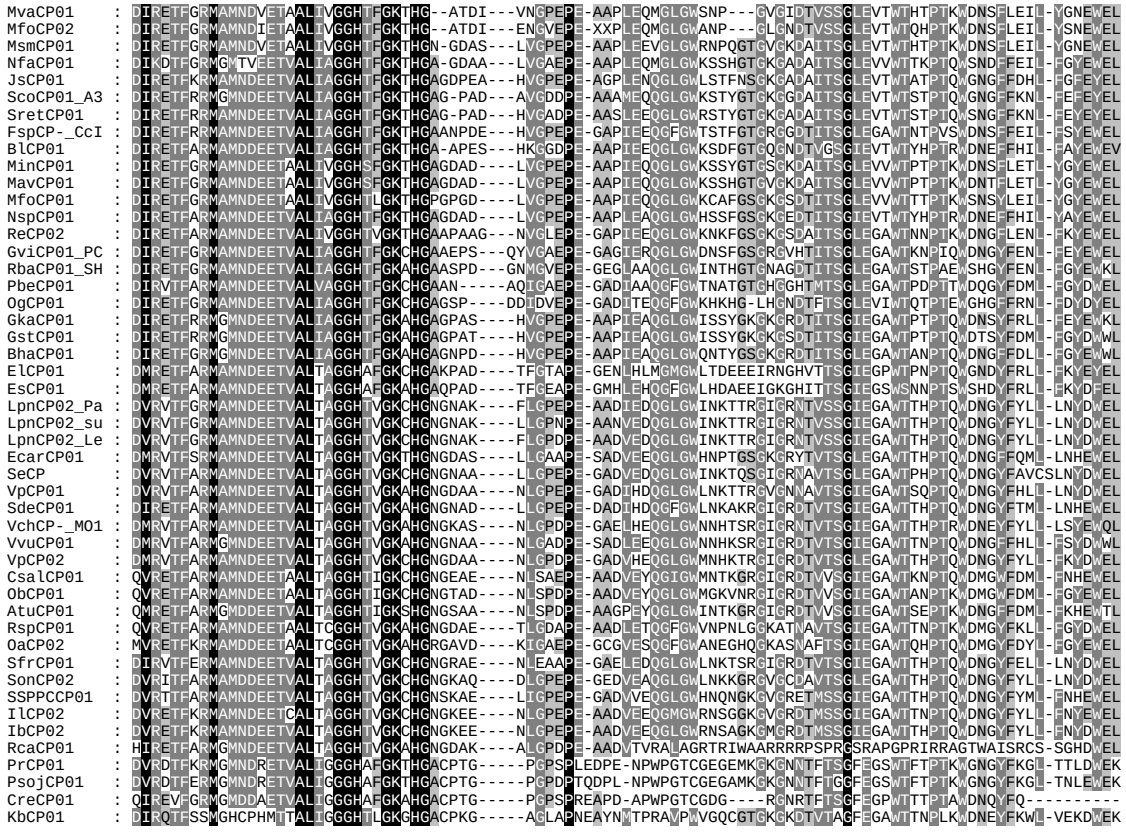


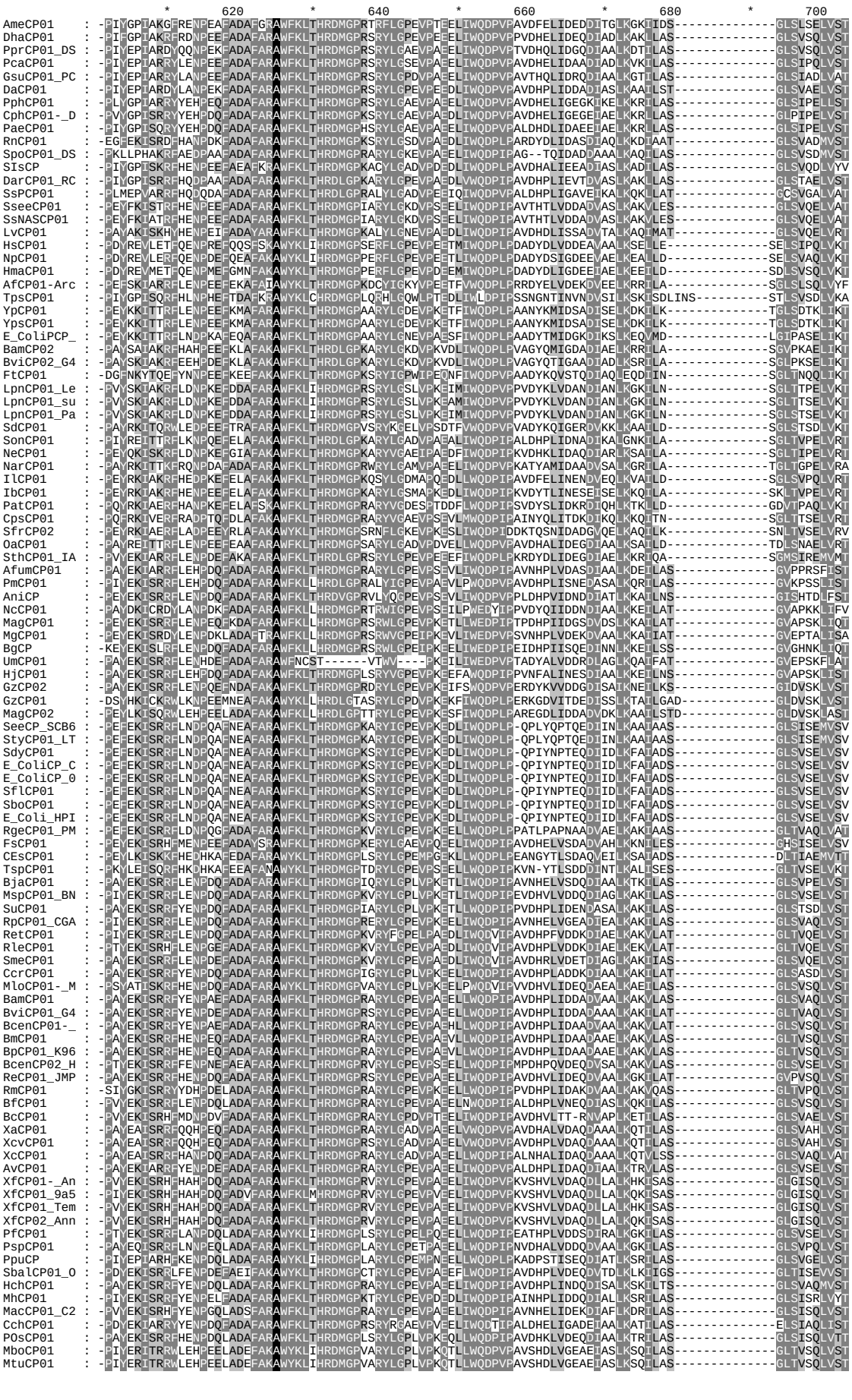


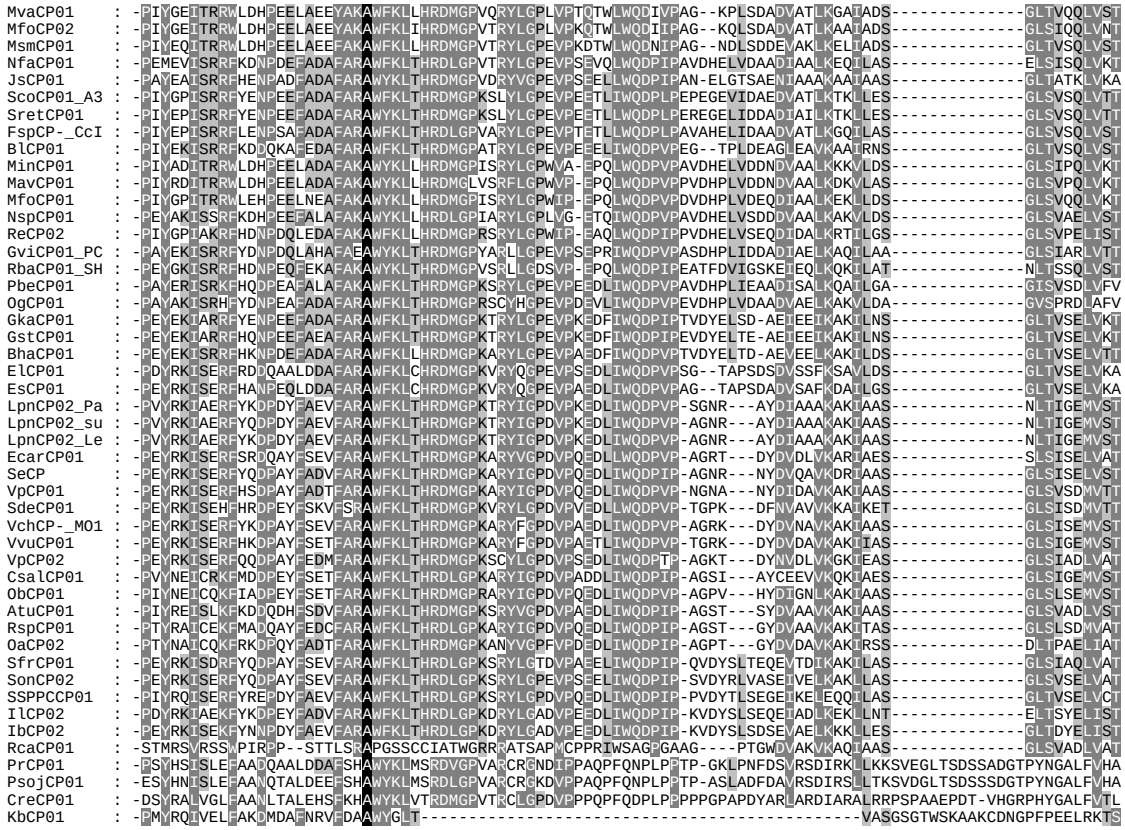


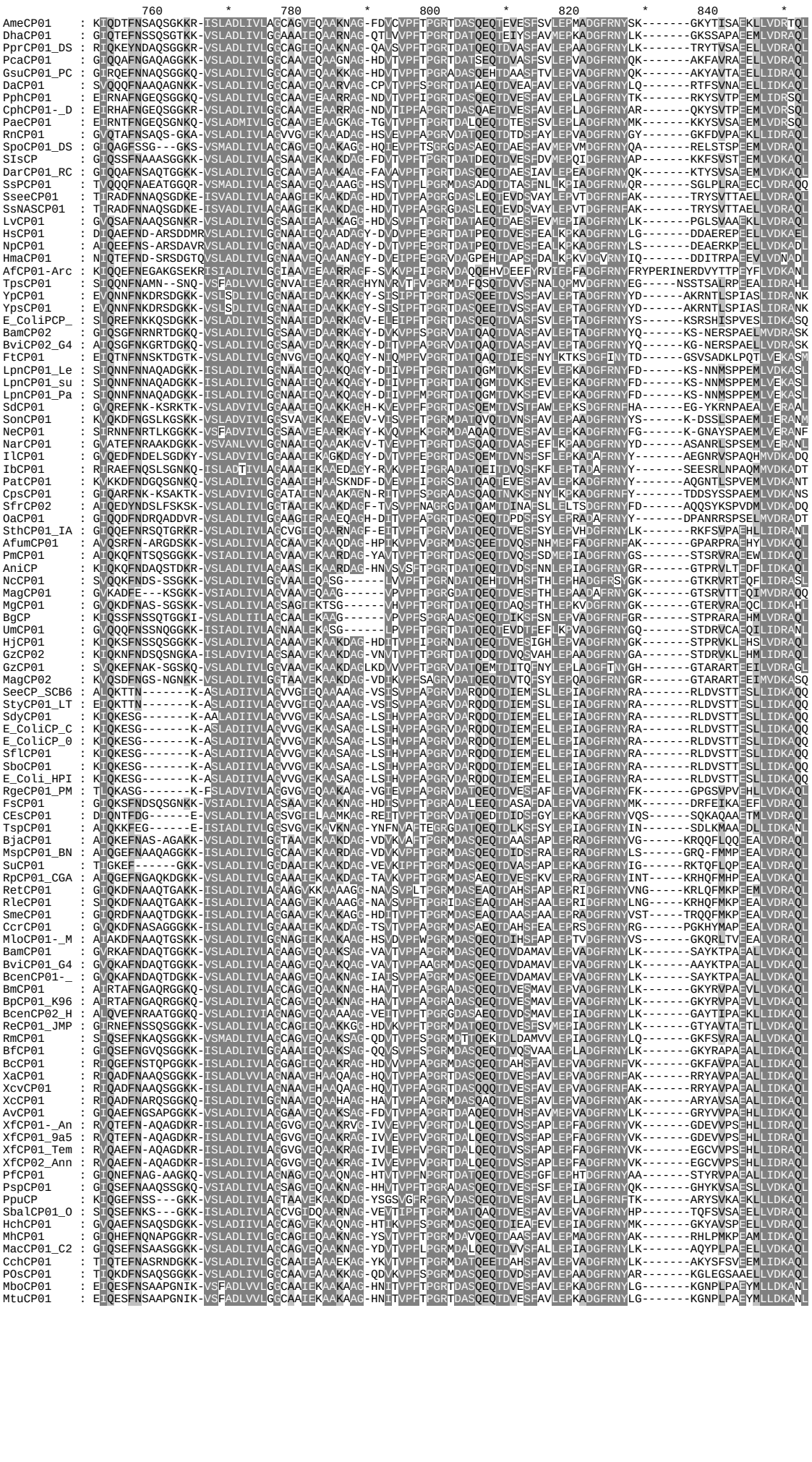


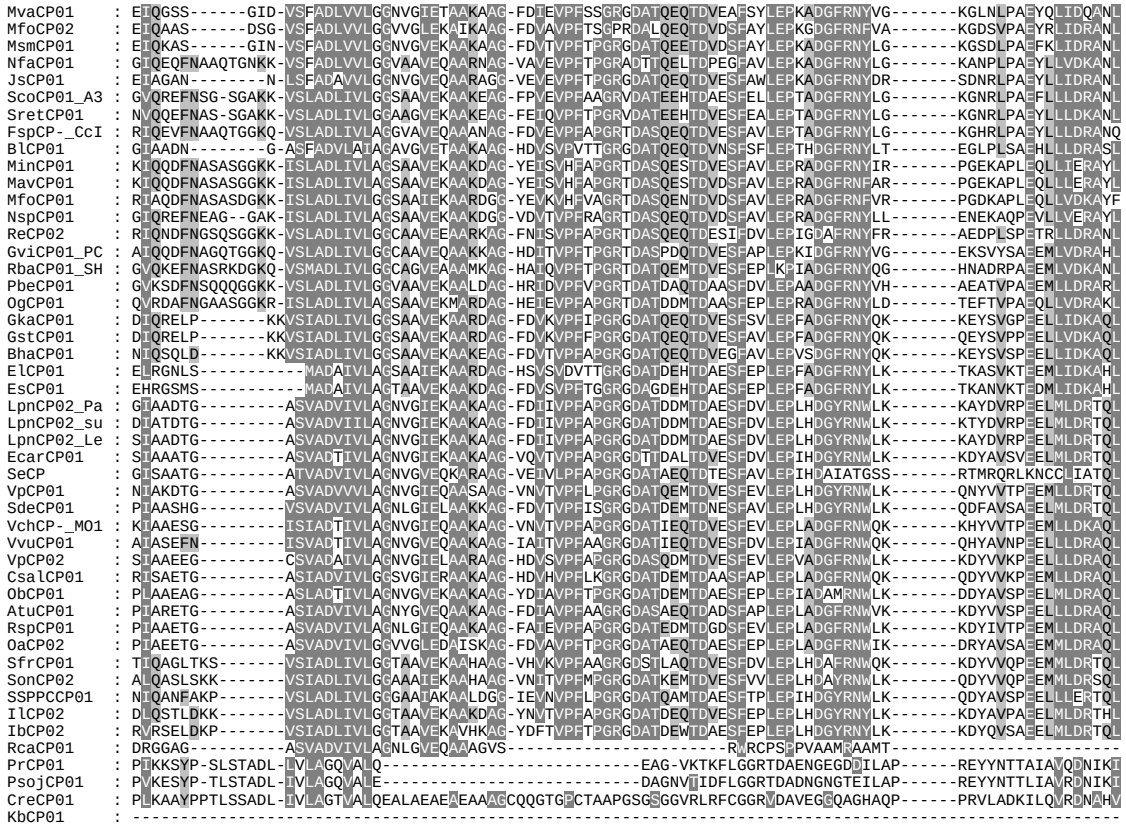


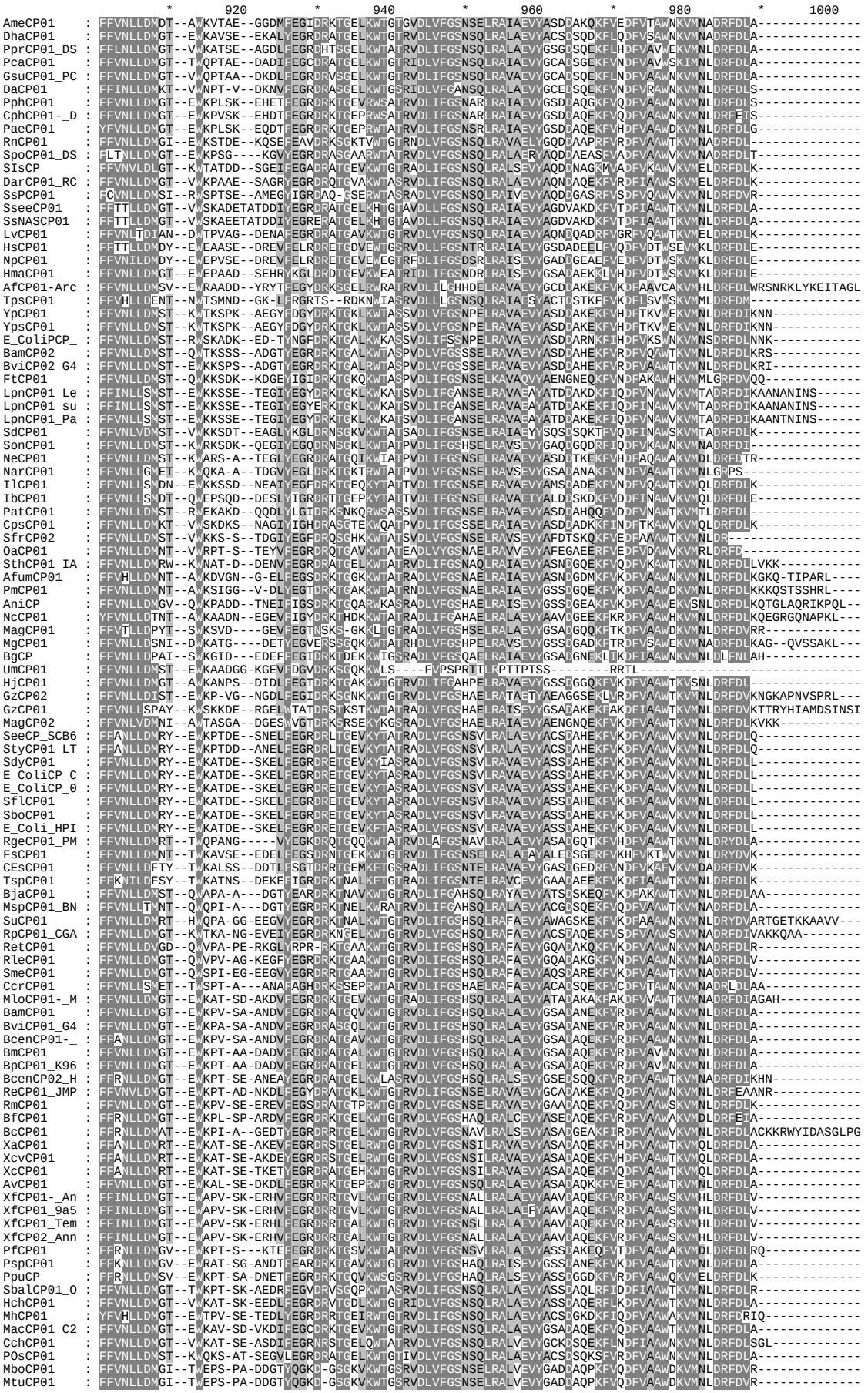


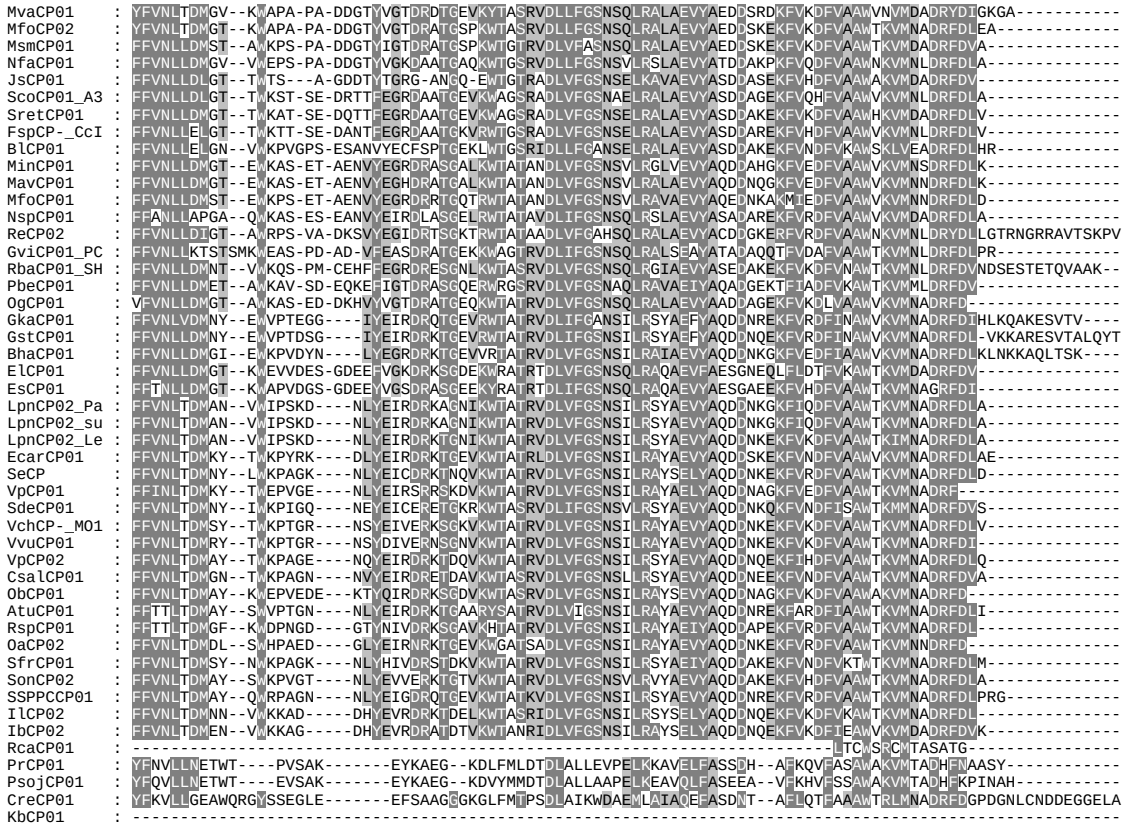












Supplemental figure 1
Alignment of the 152 CP sequences used for figure 2. The names of CP are limited to 10 characters. Black: 100% conservation; dark grey: 80% conservation; light grey: 60% conservation (similarity groups enabled).

C. Article project: “Prokaryotic origins of the non-animal peroxidase superfamily and organellar-mediated transmission to eukaryotes”

The hypothesis of the appearance of catalase-peroxidases is also discussed in a second article project: “**Prokaryotic origins of the non-animal peroxidase superfamily and organellar-mediated transmission to eukaryotes**”. The “superfamily” study especially focuses on the major role that organelles have played, by various symbiotic combinations, to the spreading of peroxidases.

Prokaryotic origins of the non-animal peroxidase superfamily and organellar-mediated transmission to eukaryotes

Research Article

Filippo Passardi[1], Nenad Bakalovic[1], Felipe Karam Teixeira[2], Marcia Margis-Pinheiro[2,3], Claude Penel[1] and Christophe Dunand[1]*

[1]Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland. [2]Federal University of Rio de Janeiro, Department of Genetics, Institute of Biology, Brazil. [3]Federal University of Rio Grande do Sul, Department of Genetics, Brazil.

Christophe Dunand, Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland, Tel.: +41223793012; Fax +41223793017; e-mail: christophe.dunand@bota.unige.ch

Key words: peroxidase superfamily, gene transfer, evolution, organellar-mediated transmission

Abbreviations: APX: ascorbate peroxidase, CCP: cytochrome C peroxidase, CP: catalase-peroxidase, LiP: lignin peroxidase, MnP: manganese peroxidase, VP: versatile peroxidase.

Abstract

The superfamily of plant, fungal and bacterial peroxidases is known to be present in a wide variety of living organisms. Extensive searches across sequencing projects allowed better defining the organisms containing sequences of this superfamily. The class I (cytochrome C peroxidase, ascorbate peroxidase and catalase peroxidase) initially found in fungi, bacteria and plants has been shown to be also present in various protists. CCP sequences were detected in most of mitochondria-containing organisms excepted in the green plants, which only

possess ascorbate peroxidases. The APX encoding sequences previously restricted to the green plants were found in chloroplastic protists, which acquired chloroplasts by secondary endosymbiosis. The CP sequences known to be present in Prokaryotes and in Ascomycetes have also been detected in some basidiomycetes and occasionally in some protists. Class II peroxidases, which are related to the lignin biodegradation, have been shown to be present in the homobasidiomycetes. In the present study, class II peroxidases have been only detected in three orders and degenerated forms have been found in different Pezizomycota orders. Class III peroxidases, specific for higher plants, were so far related to the emergence of the land plants, but we found that they appeared earlier before in some green algae. Presence of peroxidases in all major phyla (except animals) makes them powerful markers for a better understanding of the early events of evolution, which led to the appearance of the ancestors of each eukaryotic group.

Introduction

Oxygen is necessary for most of living organisms but, when not properly controlled, can become lethal. Aerobes need to deal with dioxygen and derivative molecules such as the reactive oxygen species (ROS). For this purpose, they possess a large panel of proteins, including the haem-containing enzymes peroxidases. The term “peroxidases” is used to describe a very large group of enzymes which reduce peroxide and oxidise a wide variety of substrates, such as lignin subunits, lipid membranes or some amino acid side chains. According to Karen Welinder (Welinder, 1992; Dunford, 1999), the so-called "superfamily of plant, fungal and bacterial peroxidases" includes three classes of structurally related peroxidases. The main common feature of the peroxidases belonging to this superfamily is that they all possess a heme moiety (ferriprotoporphyrin IX), a conserved histidine as proximal ligand and conserved arginine and histidine residues as distal ligands. Through the common peroxidative cycle, heme-containing peroxidases use H_2O_2 as an acceptor and various substrates as donors, which will be converted into radicals. In this cycle, the iron of peroxidases is converted from Fe(III) to Fe(IV).

Class I peroxidases have been found in plants, fungi and prokaryotes (Zamocky, 2004). They are not glycosylated, do not have a signal peptide, nor calcium ions, nor disulphide bridges. Three EC numbers belong to this class: EC 1.11.1.5 (Ferrocytochrome-c:hydrogen-peroxide oxidoreductase), EC 1.11.1.6 (Hydrogen peroxide:hydrogen-peroxide oxidoreductase, or more commonly catalase peroxidase) and EC 1.11.1.11 (L-Ascorbate:hydrogen-peroxide oxidoreductase).

Cytochrome c peroxidases (CCP) play a major role in the scavenging of the H_2O_2 generated during the aerobic respiratory process through oxidation of cytochrome c in the mitochondrial intermembrane space. Despite this crucial function, absence of CCP in yeast has not prevented cell viability nor respiration (Kwon et al., 2003). CCP have been shown both in Viridiplantae and fungi, but not in bacteria, which possess di haem CCP unrelated to the peroxidase superfamily.

Catalase-peroxidases (CP) have been shown to be present in Prokaryotes and fungi. They typically have a dual catalytic activity, acting both as a catalase and as a peroxidase. Evolutionarily, CP have the unique feature of containing two peroxidase-like domains, an N- and a C-terminal domain. They are probably issued from the fusion of two copies of a primordial gene, one copy of which also gave rise to ascorbate and cytochrome c peroxidases

(Zamocky, 2004). Only the catalase-peroxidase N-terminal domain contains a heme and is fully functional, whereas the C-terminal domain has no known catalytic activity (Yamada et al., 2002). CP are mainly found in prokaryotes: they have also been shown to be present in some members of the Ascomycetes, probably following a lateral gene transfer (Zamocky, 2004). Despite the same EC number, catalase-peroxidases and monofunctional catalases exhibit no sequence similarity which could suggest different origin (Klotz and Loewen, 2003).

Ascorbate peroxidases (APX) have been detected in chloroplastic organisms and putatively in cyanobacteria (Miyake et al., 1991). They show a particularly strong specificity for the electron donor ascorbate. They play a key role in the elimination of toxic amounts of intracellular H_2O_2 (Asada, 1999). In higher plants, APX are separated, according to their cellular localization, into cytosolic, peroxisomal and chloroplastic (Teixeira et al., 2004). Algae also possess a subgroup of APX localized in chloroplasts and whose sequences are closer to the chloroplastic subgroup than to the cytosolic one (Kitajima et al., 2002; Zamocky, 2004). Finally, putative APX sequences are also present in Euglenozoa, described either as cytoplasmic or as endoplasmic reticulum-bound form (Ishikawa et al., 1996; Wilkinson et al., 2002).

The Class II peroxidases are secreted for plant lignin degradation by a particular type of fungi, the homobasidiomycetes. These proteins mainly consist in **lignin peroxidases** (LiP, EC 1.11.1.14) and **manganese peroxidases** (MnP, EC 1.11.1.13). In contrast with class I peroxidases, but similarly to class III, they are glycosylated, contain calcium ions and disulphide bridges as well as a peptide signal directing the protein to the endoplasmic reticulum for excretion. LiP and MnP are distinct from the class I and III peroxidases in their ability to oxidize molecules with high redox potentials such as lignin. LiP and MnP are essential in order to degrade the dead plant biomass and recycle the soil. Lignin peroxidases have a wide variety of substrates and participate, for instance, in aromatic ring-opening reactions and side-chain cleavages that lead to degradation of lignin (Gold and Alic, 1993; Conesa et al., 2002). Both LiP and MnP contain a ferric heme and work through the same mechanism during the peroxidative cycle. However, they radically differ in their mode of action: whereas LiP directly attacks lignin bonds, MnP takes an electron from Mn(II), resulting into Mn(III) that will diffuse from the enzyme surface and attack lignin (Gold and Alic, 1993). A third class, named **versatile peroxidase** (no defined EC number exists so far)

(Martinez et al., 1996; Moreira et al., 1998; Ruiz-Dueñas et al., 2001) encompass multifunctional peroxidases sharing catalytic properties of LiP and MnP.

The class III peroxidases or secreted plant peroxidases (EC 1.11.1.7) are only found in plants, where they have been shown to form large multigenic families (Passardi et al., 2004; Bakalovic et al., 2006). Although their primary sequence differs in some points from the two other classes of peroxidases, their three-dimensional structures are very similar and, like class II, they also possess calcium ions, disulphide bonds and a peptide signal for endoplasmic reticulum routing. Some of them have an additional C-terminal extension that is suspected to direct the protein to the vacuole (Neuhaus, 1996; Matsui et al., 2003). Class III peroxidases probably appeared in algae just before land colonisation but at this time, no sequence allows confirming this hypothesis. They have possibly played a critical role during this major event of evolution, either by allowing formation of rigid plant structures or by adapting the organism to a more oxygenated environment (Duroux and Welinder, 2003; Passardi et al., 2004). Class III peroxidases are additionally able to perform a second cycle, called "hydroxylic", which is distinct from the peroxidative one (Liszkay et al., 2003; Passardi et al., 2004). During the hydroxylic cycle, peroxidases pass through a Fe (II) state and use mainly the superoxide anion ($O_2^{\bullet-}$) in order to generate hydroxyl radicals ($^{\bullet}OH$). Class III peroxidases, with the help of both cycles, are known to participate to many different processes in plants, from germination to senescence, such as auxin metabolism, cell wall elongation, stiffening or protection against pathogens (Hiraga et al., 2001; Passardi et al., 2004). The precise *in vivo* role of a single peroxidase has however not been found yet, mainly because of the wide range of peroxidase substrates and of the probable functional redundancy of some of these proteins.

Since 1992, numerous publications report independent studies of the different heme peroxidase classes or the situation in a single species, but the exhaustive phylogeny of the complete family has never been studied. The availability of an increasing number of genomic and EST sequencing projects gave us the opportunity to search for peroxidase encoding sequences in all the major phyla of living organisms. For each class, using phylogenetic tree analysis, we have gone back into evolution and searched for the emergence of the diverse peroxidase families from a putative ancestral sequence. We have also tried to correlate the presence or the absence of particular peroxidases with specific roles and with particular species.

Our global phylogenetic analysis of the whole family revealed that APX sequences are absent from cyanobacteria and specific to chloroplastic organisms (green plants and protists) and that CCP sequences are detected in all mitochondria-containing organisms excepted animals and parasites. The CP distribution within the prokaryotes is very heterogeneous and can follow the normal phylogenetic evolution or rather be dependent on horizontal gene transfers. The presence of CP sequences in particular fungal groups and occasionally in various unicellular organisms is related to isolated gene transfers. The Class II peroxidases, initially restricted to Homobasidiomycetes, have been also found in Ascomycetes. With new sequences from green algae, the origin of class III peroxidase has been restricted and localised just before the emergence of the Embryophytes.

Results and discussion

Class I

Class I peroxidases form the largest class of the superfamily and can be found in all living organisms excepted in the animals. Ascorbate and cytochrome C peroxidases are closely related (at least 50% of similarity between distant organisms such as *Rhodophyta* (CmeAPX01) and Ascomycota species (YlCCP03)) and the 3D-structures of APX and CCP are nearly the same (Patterson and Poulos, 1995). APX were found in chloroplast-containing organisms and CCP in most of mitochondria-containing organisms. Class I sequences were absent from Diplomonads, Parabasalids, Apicomplexa, animals and Amoebozoa (Fig. 1). Except for the land plants, which only contain APX, other chloroplastic organisms contain both APX and CCP sequences (Table 1 and Fig. 1). Punctual exceptions to this rule (symbolized with question marks in Fig 1 and Table 1) can be explained by the small size of libraries. Due to the high degree of similarity between APX and CCP sequences, both sequences have evolved from the same ancestral sequence. In many cases, hybrid sequences containing APX and CCP specific motifs (APX/CCP column in Table 1) were hardly classified by BLAST comparisons and phylogenetic alignments. As shown in figure 2 (NJ tree), three representative sequences of the APX/CCP column (complete sequences: EgrAPX01, TpsAPX01 and PparAPX01) form a well supported group which appear at the origin of the APX branch. ML phylogenetic tree supports the position of this group (data not shown). The relatively low number of CCP encoding sequences in Basidiomycetes (*Coprinus*, *Phanerochaete, Ustilago* and *Cryptococcus*) was related to the very low number of sequences

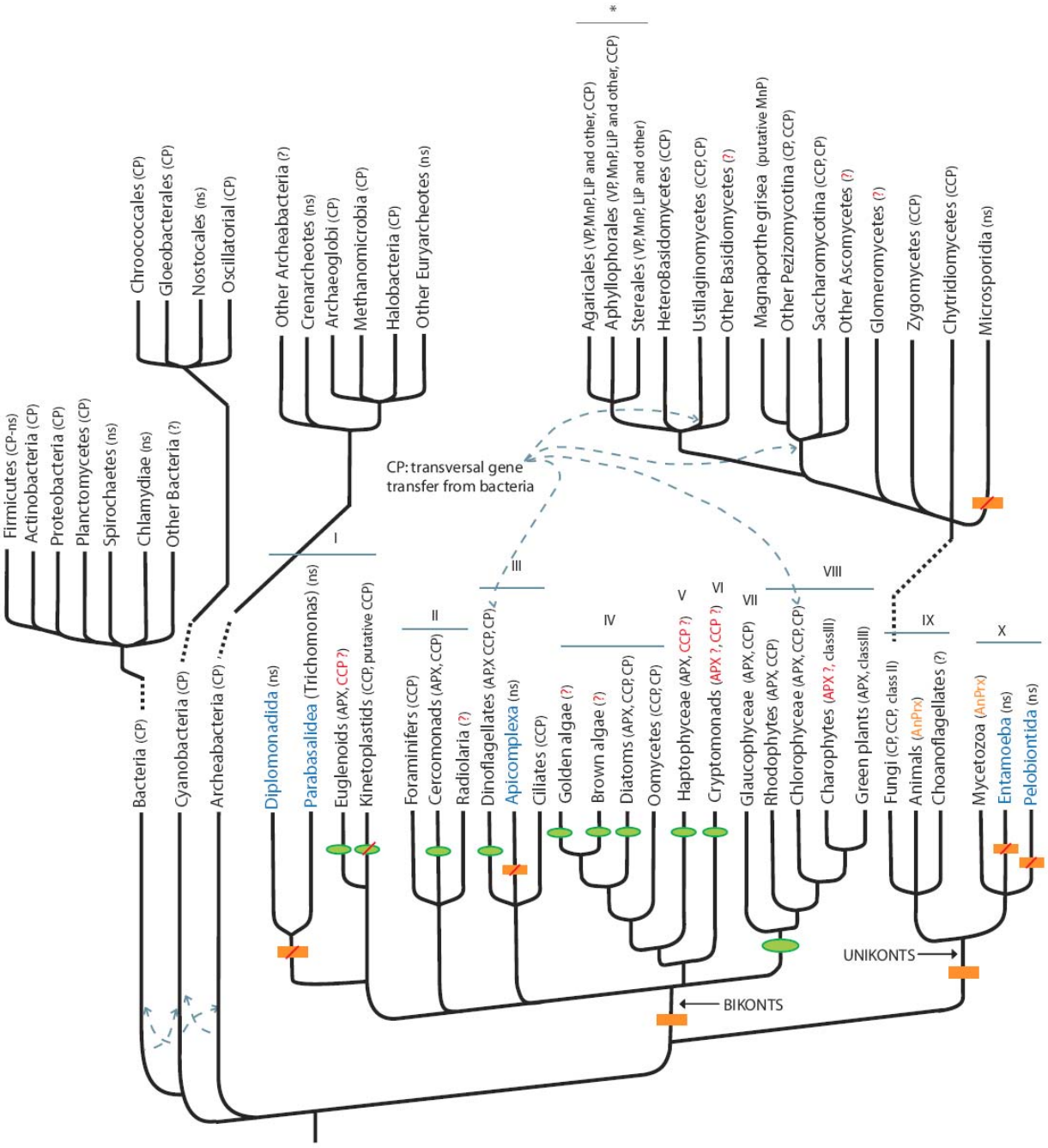


Figure 1

Hypothetical phylogeny of the major phyla of living organisms. Endosymbiotic events which lead to the appearance of mitochondria and chloroplasts are represented with bolded lines. Acquisitions of chloroplast from a secondary endosymbiosis are represented with smaller icons. Losses of plasts are represented with red-slash icon. Lateral gene transfers are represented with dashed lines. (continued next page)

Figure 1 (continued)
When at least one peroxidase encoding gene was detected, the membership of the sequence is specified in brackets for each phylum. APX: ascorbate peroxidase, AnPrx: animal peroxidase, CCP: cytochrome C peroxidase, CP: catalase-peroxidase, LiP: lignin peroxidase, MnP: manganese peroxidase, VP: versatile peroxidase, "?": absence of sequence probably due to the low number of sequences available in the database and ns: absence of sequence. When the presence of a particular class is expected but not detected, the class appears in red and is followed by "?". I: Excavata, II: Rhizaria, III: Alveolata, IV: Stramenopiles, V: Haptophyceae, VI: Cryptomonads, VII: Glaucophyceae, VIII: Viridiplantae, IX: Opisthokonts, X: Amoebozoa, *: Homobasidiomycetes. The phylum marked in blue stands for parasitic organisms.

currently available (Fig. 1). Curiously, CCP were absent from the completed genome of *Encephalitozoon cuniculi* (Katinka et al., 2001). This intracellular parasite belongs to the group of Microsporidia, which are considered as the most basal group of fungi, or the closest sister group to fungi (Bruns et al., 1992; Tanabe et al., 2005). Absence of CCP in Microsporidia adds some evidence for the second hypothesis.

Only one or two CCP encoding sequences can be found per organism whereas several isoforms of APX can be detected in most of the *Viridiplantae* (6 and 8 isoforms in *Arabidopsis* and *Oryza* respectively). The localization of APX can be predicted as chloroplastic, peroxisomal or cytoplasmic, as it can clearly be seen in the class I phylogenetic tree (Fig. 2). APX encoding sequences from unicellular algae, *Rhodophyta* (CCriAPX01, CmeAPX01, GpAPX, PyAPX) and *Chlamydomonas* (CreAPX01 and CspAPX01), were at the base of the chloroplastic APX subbranch (Fig. 2 and 4). *Trypanosoma* and *Leishmania* sequences (LmAPX and TcrAPX), described as APX proteins (Wilkinson et al., 2002; Ivens et al., 2005), appeared at the origin of the APX/CCP branch (Fig. 2 and 4). Despite their names, these two APX sequences share similar motifs with CCP sequences from various organisms, so it would be more correct to consider them as hybrid APX/CCP or as ancestral class I sequences. New recognition profiles have been used to clearly distinguish *in silico* APX from CCP sequences (unpublished data).

Concerning catalase-peroxidases (CP), genomic sequences are present in archaea, bacteria and ascomycota and marginally in a few members of the protist and plant kingdoms (Fig. 1

Families	Species	APX chloro.	APX non chloro.	APX / CCP	CCP
Rhodophyta (ii)	*Chondrus crispus*	?	?	CcriAPX01	?
				CcriAPX02	
	Cyanidioschyzon merolae	CmeAPX01		CmeAPx02	?
	Galdieria partita	?	?	GpAPX	?
	Galdieria sulphuraria		GsuAPX01	GsuAPX02	GsuCCP01
				GsuAPX03	
				GsuAPX04	
	Porphyra yezoensis	?	?	PyAPX	?
Diatoms (iv)	*Thalassiosira pseudonana*	TpsAPX02		TpsAPX01	TpsCCP01
				TpsAPX03	TpsCCP02
					TpsCCP03
					TpsCCP04
	Phaeodactylum tricornutum	PtrAPX02			PtrCCP01
		PtrAPX01			PtrCCP02
					PtrCCP03
					PtrCCP04
Cercomonads (iv)	*Bigelowiella natans*	BnaAPx01		BnaCCP01	BnaCCP02
Dinoflagellates (iii)	*Heterocapsa triquetra*	HtrAPx01		?	?
		HtrAPx02			
	Karenia brevis	?		?	CCP ? (C0064570)
	Amphidinium carterae	?		AcaAPX01	?
	Alexandrium tamarense	AtamAPX		?	AtamCCP
Euglenoids (iii)	*Euglena gracilis*	?		EgrAPx01	?
Haptophyceae (ii)	*Prymnesium parvum*	?		PparAPx01	?
	Isochrysis galbana	?		?	IgCCP01
Glaucocystophyceae (ii)	*Glaucocystis nostochinearum*	?		GnoAPx01	GnoCCP01
					GnoCCP02
	Cyanophora paradoxa	?		?	?
Chlorophyceae (ii)	*Acetabularia*	AacAPx01		AacAPx02	?
		AacAPx03			
	Chlamydomonas reinhardtii	CreAPx02		CreAPx01	CreCCP01
	Chlorella	ClAPX01/ClAPX01-bis		?	?
	Dunaliella salina	DsaAPX01		?	?
	Scenedesmus obliquus	SobAPX01		?	SobCCP01
Cryptomonadaceae (iv)		?		?	?

Table 1. CCP and APX sequences composition of chloroplastic organisms (other than green plants). ?: absence of sequences probably due to the small size of the database. The number in brackets after the family name corresponds to the number of membranes surrounding the chloroplast. APX/CCP column: sequences that possess both APX and CCP motifs. Names of sequences follow the nomenclature used in the PeroxiBase (Bakalovic et al. 2006).

and 4). In bacteria and archaebacteria, their non homogenous distribution does not always follow a phylogenetic transmission but is also due to lateral gene transfer. CP are intron-free also in eukaryotic organisms, and the similarities observed between sequences are over 70%. This information confirms the idea already proposed by Zamocky (Zamocky, 2004) that presence of CP sequences in fungi is related to a bacterial gene transfer. When aligning separately the N-terminal part and the C-terminal part of CP with CCP or APX, it can be seen that the N-terminal part has preserved the critical hemoperoxidase residues responsible for

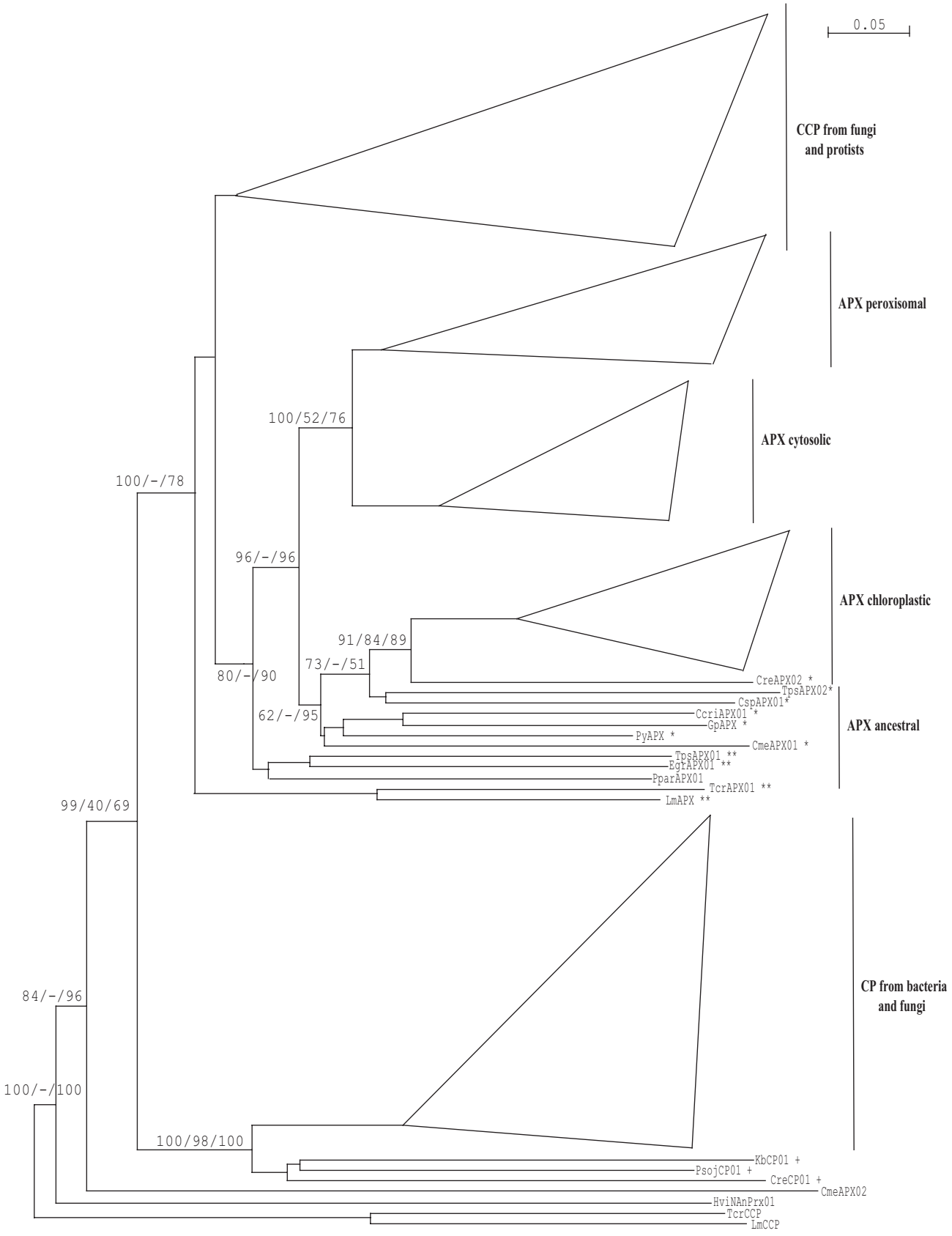

0.05
CCP from fungi and protists
APX peroxisomal
APX cytosolic
APX chloroplastic
APX ancestral
CP from bacteria and fungi
100/52/76
100/-/78
96/-/96
91/84/89
73/-/51
80/-/90
62/-/95
99/40/69
84/-/96
100/-/100
100/98/100
CreAPX02 *
TpsAPX02*
CspAPX01*
CcriAPX01 *
GpAPX *
PyAPX *
CmeAPX01 *
TpsAPX01 **
EgrAPX01 **
PparAPX01
TcrAPX01 **
LmAPX **
KbCP01 +
PsojCP01 +
CreCP01 +
CmeAPX02
HviNAnPrx01
TcrCCP
LmCCP

Figure 2 (previous page)
Neighbor-Joining tree of class I peroxidases based on predicted sequences. Only complete sequences have been used to build the tree: 54 for APX, 22 for CCP and 35 for CP. Numbers represent bootstrap values obtained from distance, MP and ML methods. APX: ascorbate peroxidase, CCP: cytochrome C peroxidase, CP: catalase-peroxidase, *: unicellular Viridiplantae, **: chloroplastic protists, +: CP sequences from other organisms than bacteria/fungi.

catalytic activity (Yamada et al., 2002), whereas they are absent from the C-terminal part. Nevertheless, the C-terminal part has preserved quite some residues, whose function is so far not known. The N-terminal part has also added almost 100 new residues that are forming large gaps when aligned to CCP and APX.

Class II

From our exhaustive data mining, class II encoding sequences were only found in three saprophytic Homobasidiomycetes: Agaricales, Aphyllophorales and Sterales. Four independent families of class II sequences can be detected: in these orders, the manganese peroxidase (MnP), the lignin peroxidase (LiP), the versatile peroxidase (VP) and a fourth unnamed group (Fig. 3). The four families are well defined in a phylogenetic manner even if the sequences are closely related. The node separating MnP from the other main cluster (bootstrap “70/-/-“) is not strongly supported by MP and ML methods. However, on MP and ML trees, both MnP and “new ligninase” groups are well defined. Key residues related to the catalytic specificity (Martinez, 2002) are well conserved within each family and new recognition profiles help us drawing up hypotheses concerning the chronological evolution of the different families (Bakalovic et al., 2006). Several sequences from various fungi (TvMRP, GaMnP01, TceLiP01, ArMnP and CcCIP), which contain some of the major ligninase motifs, are not included in any of the four groups and all come at the base of the class II tree without forming a well-defined group which appeared at the origin of the ligninase tree (Fig. 3 and 4). The position of this group and the others are well supported with by the ML and MP analysis. Three putative lignin peroxidase encoding sequences could also be found in *Magnaporthe grisea*, (MapP01, 02, 03) and in Gibberella zeae (GzP01), two Pezizomycota. These sequences contain motifs conserved between ligninases but show low level of identity on the whole sequence. Probably following the co-evolution between the ligninolytic activity and the

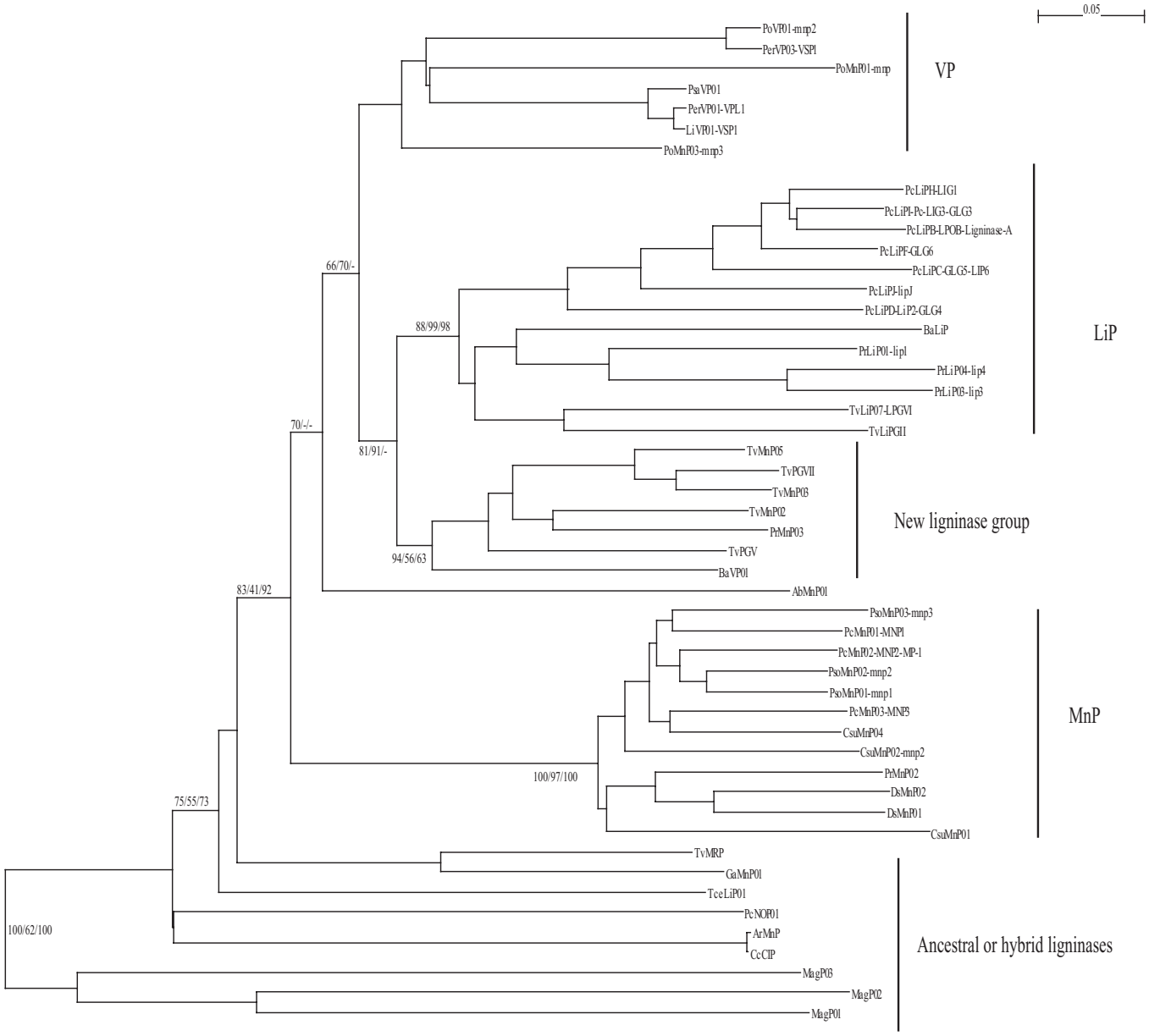

0.05
PoVP01-mnp2
PerVP03-VSP1
PoMnP01-mnp
PsaVP01
PerVP01-VPL1
LiVP01-VSP1
PoMnP03-mnp3
VP
PcLiPH-LIG1
PcLiPI-Pc-LIG3-GLG3
PcLiPB-LPOB-Ligninase-A
PcLiPF-GLG6
PcLiPC-GLG5-LIP6
PcLiPJ-lipJ
PcLiPD-LiP2-GLG4
BaLiP
PrLiP01-lip1
PrLiP04-lip4
PrLiP03-lip3
TvLiP07-LPGVI
TvLiPGII
LiP
66/70/-
88/99/98
70/-/-
81/91/-
TvMnP05
TvPGVII
TvMnP03
TvMnP02
PrMnP03
TvPGV
BaVP01
94/56/63
New ligninase group
AbMnP01
83/41/92
PsoMnP03-mnp3
PcMnP01-MNP1
PcMnP02-MNP2-MP-1
PsoMnP02-mnp2
PsoMnP01-mnp1
PcMnP03-MNP3
CsuMnP04
CsuMnP02-mnp2
PrMnP02
DsMnP02
DsMnP01
CsuMnP01
MnP
100/97/100
75/55/73
TvMRP
GaMnP01
TceLiP01
PcNOP01
ArMnP
CcCIP
100/62/100
MagP03
MagP02
MagP01
Ancestral or hybrid ligninases

Figure 3 (previous page)
Neighbor-Joining tree of class II peroxidase based on predicted sequences. Only complete sequences have been used to build the tree (50 sequences). Numbers represent the bootstrap values obtained from distance, MP and ML methods. LiP: lignin peroxidase, MnP: manganese peroxidase, nd: not determined family, VP: versatile peroxidase

diversity of the plant cell wall composition, ligninolytic fungi contain various copies of the same family and/or of different families (Table 2). Within the same species, LiP and MnP genes are organised in clusters (e.g. PcLiP from *Phanerochaete chrysosporium* and TvMnP from *Trametes versicolor*), suggesting recent gene duplication events (Stewart and Cullen, 1999). The repartition of the different families and number of copies within each organism may reflect the ligninolytic capacity and specifity of fungi (Table 2). The phylogenetic evolution is directly related to the particular function of the ligninase and to the implication of the Basidiomycetes in the biodegradation of the lignin from plants. From the tree constructed here it cannot be concluded which one of the four different class II peroxidases came first. This question could be solved by obtaining more data on genomes of early homobasidiomycetes (cantharellales, phallales, hymenochaetales) (Binder and Hibbett, 2002; Moncalvo et al., 2002).

Class III

The large number of sequences entered in the PeroxiBase (Bakalovic et al., 2006) allows confirming the hypothesis, already suggested in previous publications, that the class III appeared with the emergence of the land plants and the number of isoforms increased together with the complexity of the organisms (Passardi et al., 2004). A partial sequence recently found in a green alga (the Zygnemophycea *Closterium peracerosum-strigosum-littorale* complex, CpslPrx01) close to the origin of embryophytes (McCourt et al., 2004) could restrict the area of the class III appearance. In this respect, sequences from *Chlamydomona*s (CrePrx01), *Hydra viridis* (HvinAnPrx01) and Rhodophyta (CmeAPX02), which contain some class III conserved motifs but show low level of identity on the whole sequence, could be hybrid sequences between the actual class III and its ancestral sequence. These sequences appear at the origin of the class III branch on the NJ tree (Fig. 4). The basal position of these sequences is also supported by the bootstrap values obtained with the MP method (Fig. 4).

Some key organisms to track the evolution of the class II and III

In *Hydra viridis* a peroxidase encoding sequence has been detected (HvinAnPrx01). The presence of this sequence in an animal is probably due to an ancestral symbiosis rather than the current one with *Chlorella*, which also possesses an APX, but not particularly similar to HvinAnPrx01 (Habetha and Bosch, 2005). This sequence is indeed also present in non symbiotic species such as *H. magnipapillata* as well as in other symbiotic ones such as *H. vulgaris*.

	Aphyllophorales												Agaricales				
	Antrodia cinnamomea	Bjerkandera adusta	Bjerkandera sp. B33/3	Ceriporiopsis subvermispora	Coriolopsis gallica	Dichomitus squalens	Phanerochaete chrysosporium	Phanerochaete sordida	Phlebia radiata	Pycnoporus coccineus	Trametes cervina	Trametes versicolor	Laccaria	Lepista irina	Pleurotus eryngii	Pleurotus ostreatus	Pleurotus sapidus
MnP				**5**	**1**	**2**	**4**	**3**	**1**							**1**	
VP			**3**										**1**	**1**	**3**	**3**	**1**
LiP		**1**		**2**			**10**	**2**	**3**	**1**		**3**					
New ligninase	**1**	**1**							**1**			**7**					
Ancestral							**1**				**1**	**1**					

Table 2

Distribution of the different class II peroxidase isoforms in relation to the different lignolytic fungus families. Numbers correspond to the number of partial (in brackets) or complete sequences. LiP: lignin peroxidase, MnP: manganese peroxidase, nd: not determined family, VP: versatile peroxidase.

HvinAnPrx01, HvuAPX and HmAPX have probably originated from a class I peroxidase (APX or CCP) horizontally transferred from an ancestral symbiont. The sequences showed the same percentage of identity with the APX and the class III peroxidase sequences, but two cysteine residues characteristic of the class III and implicated in a short disulfide bridge near the distal histidine residue are present in the *Hydra* sequences. Therefore, we think that the current sequence HvinAnPrx01 has indeed evolved towards a class III peroxidase. Additionally, we have found in *Chlamydomonas reinhardtii* two hybrid sequences named CrePrx01 and 02. These proteins were predicted as fusion proteins with a middle region containing similar motifs to class III peroxidases, and N- and C-term regions similar to a

multicopper oxidase. The "peroxidase" region cannot however be considered as a class III peroxidase: it looks more like a hybrid between a class I and a class III peroxidase (Fig. 4). CrePrx01 has been shown to be expressed at least for the multicopper oxidase region (accession BI527303).

Trypanosoma and *Leishmania*, two chloroplast-free members of the *Euglenozoa* branch, contain one APX/CCP encoding sequence and one putative CCP-like sequence. All the motifs suitable to determine if the sequences are APX/CCP are not detectable. We have shown that presence of APX is strongly correlated with presence of chloroplasts: this may suggest that there was a loss of chloroplast from an ancestor of this organism (Hannaert et al., 2003). An alternative hypothesis claims that these organisms are not issued from a chloroplastic ancestor (Leander, 2004), hence acquisition of APX has probably occurred by gene transfer that occurred during parasitism of plants (Waller et al., 2004).

In the fungal branch, *Magnaporthe grisea* is a good candidate to study the evolution of the peroxidase as it contains one putative APX, two CCP, 2 CP, and 3 hybrid sequences MagP01, 02 and 03. Those three sequences showed similar percentage of identity with other ligninases and with APX and indeed appeared well separated from the class II branch (Fig. 3 and 4).

The "hybrid" peroxidases described here above are very difficult to place in one of the three classes of the superfamily. They all share class I motifs, but also class II or III motifs: they may hence be considered as remnants of intermediate forms between class I and class II/III peroxidases. Collecting more such sequences from future databases will certainly help understanding what critical steps allowed conversion of class I peroxidases to the two other classes.

Fusion proteins

During our extensive search for peroxidases, we came in two occasions across sequences that are fusion proteins: GzCP01 (*Gibberella zeae*) and CrePrx01/02 (*Chlamydomonas reinhardtii*). GzCP01 is a CP linked to a cytochrome P450 protein, and CrePrx01/02 are associated to a multicopper oxidase motif. Only CrePrx01 was found to be expressed so far. *Gibberella zeae* also has one separate CP gene (GzCP02), and several separate P450 genes (107 are listed in the P450 database: http://drnelson.utmem.edu/CytochromeP450.html). GzCP02 is located on the same chromosome as GzCP01, whereas the "separate" P450s are scattered on various

chromosomes. Both cytochrome P450 and multicopper oxidases are involved in redox reactions (Nebert and Gonzalez, 1987; Nakamura and Go, 2005), hence proximity of peroxidases suggests a possible cooperation in specific redox mechanisms. To our knowledge, this is the first time that such a fusion event has been reported in the peroxidase superfamily.

Organisms without peroxidases

Peroxidases are absent from a large variety of bacteria. They do not seem to be essential to their survival, since in many bacterial genera, some species possess CP and others not. It even happened that within one species, some strains contain CP sequences and others not (Passardi et al., 2006). When peroxidases are absent, other proteins are clearly able to cope with ROS production, such as superoxide dismutase, catalase or chloroperoxidase, which were frequently found in bacteria lacking CP sequence.

None of the currently known amitochondriate eukaryotic organisms (Apicomplexa, Diplomonadida, Parabasalidea, and Microsporidia, as well as some Amoebozoa), showed any peroxidase sequence after extensive BLAST searches, despite they possess “reduced” form of mitochondria, either mitosomes or hydrogenosomes (van der Giezen et al., 2005), and that it is becoming accepted that they once possessed true mitochondria (Embley et al., 2003; Hrdy et al., 2004; Gray, 2005).

Finally, among non-parasitic and mitochondriate organisms, Mycetozoa, animals and choanoflagellates are the only organisms without peroxidases belonging to the superfamily (Fig. 1). Both animals and Mycetozoa have another set of peroxidases belonging to the animal peroxidase superfamily (myeloperoxidase, lactoperoxidase, eosinophil peroxidase and thyroid peroxidase: prosite pattern PS50292), which are likely to replace the non-animal superfamily peroxidases. Concerning choanoflagellates, we could not find even animal peroxidases, probably because of the little amount of data known so far. Curiously, there is no similarity between the two kinds of peroxidases, hence we cannot predict if animal peroxidases are phylogenetically related to non-animal peroxidases. The only exceptions were the *Hydra* sequences, similar to class I peroxidases: they are clearly coming from an ancestral symbiosis with green algae, and are not remnants of a hypothetical transition form between non-animal and animal peroxidases.

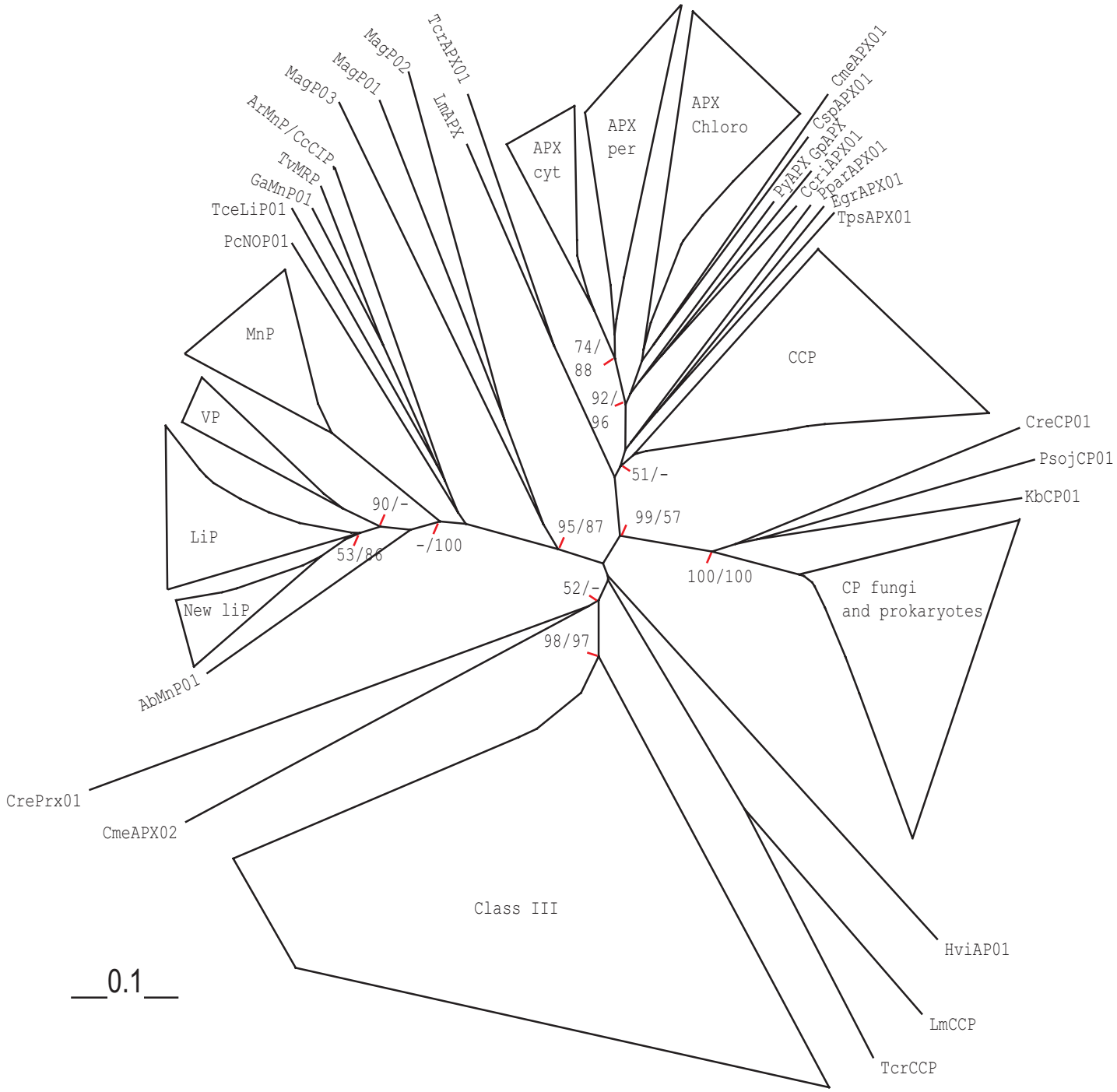

TcrAPX01
MagP02
MagP01
MagP03
ArMnP/CccIP
TvMRP
GaMnP01
TceLiP01
PcNOP01
LmAPX
APX cyt
APX per
APX Chloro
CmeAPX01
CspAPX01
GpAPX
PyAPX
CcriAPX01
PparAPX01
EgrAPX01
TpsAPX01
MnP
VP
LiP
New liP
AbMnP01
74/ 88
92/ 96
CCP
51/-
CreCP01
PsojCP01
KbCP01
90/-
53/86
-/100
95/87
99/57
100/100
CP fungi and prokaryotes
52/-
98/97
CrePrx01
CmeAPX02
Class III
HviAP01
LmCCP
TcrCCP
__0.1__

Figure 4 (previous page)
Cluster-tree of the non animal peroxidase superfamily. Complete sequences from each class (110 sequences for the class I and 50 sequences for the class II and III) have been used to build phylogenetic tree. Schematic representation has been used for the cluster of sequences. Triangles were formed by linking the longer branch of each group to the shortest branch. Isolated sequences have been excluded from the cluster. Numbers represent the bootstrap values obtained from distance and ML methods.

Conclusion

Origins of the superfamily of non-animal peroxidases

APX and CCP show a high level of identity and except for green plants, which contain only APX and for fungi, which contain only CCP, they can be detected together in the same organism (Table 1 and Fig. 1). The class I sequences, present in virtually all organisms but not in animals, have probably evolved from a common sequence belonging to a prokaryotic ancestor (Fig. 5).

It seems therefore that, hundreds of millions years ago, ancestral bacteria (and less likely archaebacteria, which are more recent) possessed a non-duplicated peroxidase form but no present-day prokaryotic organism can confirm this hypothesis, as in current bacteria only catalase-peroxidase fusion forms could be found. An alternative hypothesis (Klotz and Loewen, 2003) proposes that catalase-peroxidases are the ancestral form of class I, and that APX and CCP were issued from the separation of catalase-peroxidase N and C terminus by intronic insertion in an ancestral organism. In order to distinguish which hypothesis is true, it would be necessary to date the fusion event, and see if it occurred before of after organellar acquisition by eukaryotes. Without prokaryotic organisms containing the non fused form and with the multiple gene transfers frequently occurring among prokaryotes, it will be difficult to date this event. We think however that our hypothesis is more probable, as it appears in our study that eukaryotes gained CCP and APX sequences in two separate steps, respectively mitochondria acquisition from α-proteobacteria and, later, chloroplast acquisition from cyanobacteria. (Gray, 1999; McFadden, 2001; Cavalier-Smith, 2004). The two organelles were encoding ancestral CCP and APX sequences separately. In consequence, if an intron

was inserted in the CP sequence in one organelle, each eukaryotic organism should contain both APX and CCP sequences and it would be quite improbable that another intron would have been inserted at the same position in the second organelle.

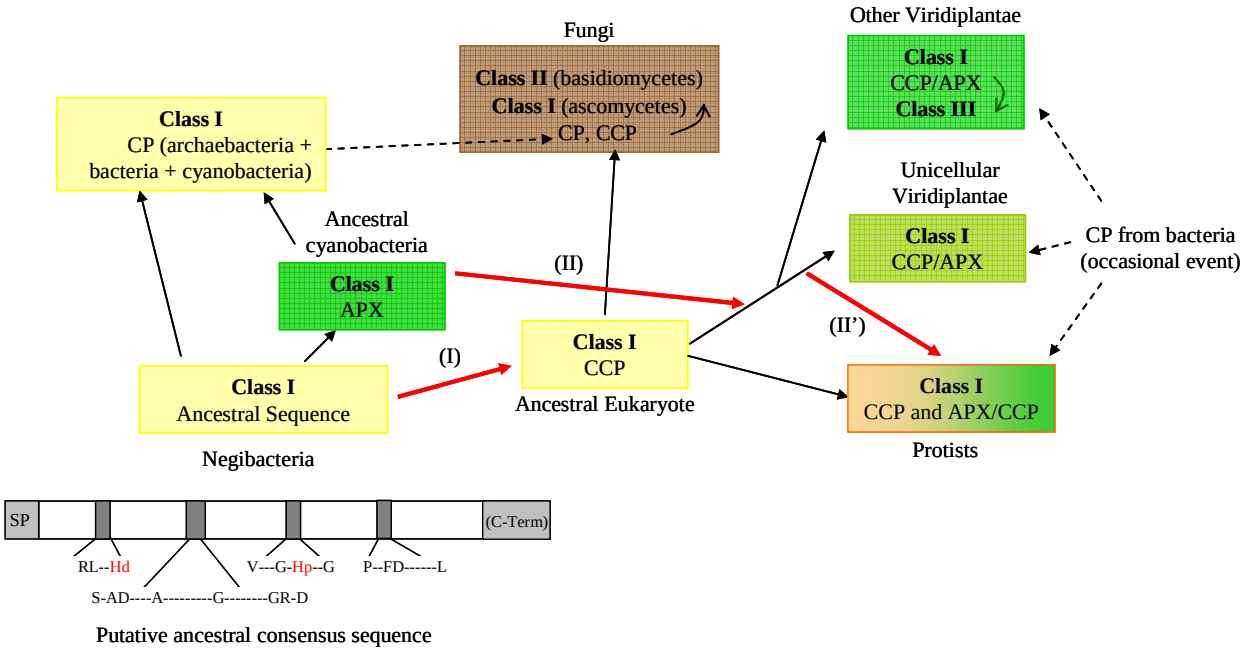


Figure 5

Model for the emergence and the evolution of the three peroxidase classes from one ancestral sequence. Arrows represent the putative evolution, red arrows represent the major endosymbiosis events (I and II for primary and secondary endosymbiosis respectively) and dashed arrows the putative CP sequence transfer.

SP: signal peptide; Hd and Hp: distal and proximal histidines.

Current theories state that eukaryotes are split into two groups, termed Unikonts and Bikonts (Cavalier-Smith, 2002). The first group contains animals, fungi and Amoeba, while the second group consists in plants, algae and other protists. The observation that CCP are absent from Unikonts (but present in fungi) suggests the following hypothesis: an earlier form of mitochondrion lacking CCP was acquired in a first step by a Unikont ancestor. A more evolved form of mitochondrion encoding CCP was then transferred to a Bikont ancestor. The timeline of this transfer is supported by the fact that Unikonts contain separate dihydrofolate

reductases (DHFR) and thymidylate synthase (TS) genes, whereas Bikonts exclusively possess a fused DHFR-TS gene (Stechmann and Cavalier-Smith, 2003). Fungi would have then acquired CCP by a subsequent lateral gene transfer after the emergence of the Microsporidia branch, which does not contain CCP sequences. Although there is no support yet for two different mitochondrial acquisition steps, it seems to us that this hypothesis is the most probable, in order to be consistent with the current view that mitochondria were acquired by all existing eukaryotes (Embley and Martin, 2006). In land plants, CCP probably evolved to class III peroxidases. We indeed observed that all land plants, which encode class III peroxidases, do not have any CCP gene. Similarly, the charophycean alga *Closterium peracerosum* encoded a class III peroxidase, but no CCP gene was found. Both peroxidases are however rather divergent: besides differences in their primary sequence, CCP lacks the four disulfide bridges and the two calcium ions of class III peroxidases. In this respect, the "hybrid" peroxidase found in *Chlamydomonas* (CrePrx01) may represent an intermediate form between CCP and class III peroxidases. We also suggest that animal peroxidases were acquired by a Unikont ancestor, but rather through gene transfer than through organellar transfer, which explains their absence in fungi. Secondary endosymbiosis, which led to unicellular chloroplastic organisms, increased the number of class I isoforms contained in these organisms. Peroxidase genes were further transferred to the nuclear genome, as it is well known for many other genes (Martin and Herrmann, 1998).

Present CCP and APX sequences originated from ancestral peroxidase genes related to two independent endosymbiosis events that led to the appearance of mitochondria and chloroplasts respectively. CCP duplications within an organism are the exception rather than the rule, and their significance is not clearly known. Regarding APX however, the two well defined plant cytoplasmic and chloroplastic APX have the same origin and result from duplication events in the basal Viridiplantae. In the case of Euglenozoa, chloroplasts may have been gained not by an early ancestor of the whole group, but only by some members (Euglenids) in later endosymbiotic events (McFadden, 2001; Leander, 2004). The presence of peroxisomal APX isoforms only in land plants reinforces the idea of more recent duplications in the higher plants. Probably, these isoforms originated from cytosolic isoforms that have gained a last exon encoding the membrane-bound domain and the targeting peptide (Teixeira et al., 2004). The appearance of the chloroplastic alternative splicing genes in certain lineages of eudicots (Cucurbitaceae, Fabaceae, Caryophyllales and some Solanaceae), although some

eudicots (Brassicaceae and Solanaceae) have distinct genes encoding stromal and thylakoid-bound isoforms, confirms the active evolution of the class I family in the green plants. Repartition of APX in cytoplasmic, peroxisomal and chloroplastic is certainly due to subfunctionalisation, and less likely to neofunctionalisation, as their unique known function is detoxification of the cell. We cannot discard, though, the gain of new functions (neofunctionalisation) as a consequence of the repartition of APX in different groups. Subfunctionalisation can lead to very subtle variations of APX localisation, such as the separation observed between stromal and thylakoid-bound chloroplastic APX in *Arabidopsis thaliana* (respectively AtAPX05 and AtAPX06) or in *Lycopersicon esculentum* (respectively LeAPX 07 and LeAPX08), which at last gives rise to organisms bearing up to 8 APX in their genome (*Glycine max*).

Following the extensive search of peroxidase sequences in all living organisms, we could not find any peroxidase sequence in several distinct groups, from animals to amitochondriate parasites and some bacteria. Organellar origin of peroxidase genes confers a great advantage to an organism acquiring these organelles, as both mitochondria and chloroplasts are responsible for ROS formation, and both APX and CCP are able to detoxify cells from these reactive molecules. If peroxidases played such an important role in detoxification and, later in evolution, in many other functions, why did so many organisms lose (or not acquire) this gene? What other gene did they find in order to cope with the loss of peroxidases? Did the parasitic organisms use the host machinery to deal with the ROS production? It is certain that peroxidases are not the only proteins to be able to protect organisms against oxidative damage. Other peroxidases such as superoxide dismutase (EC 1.15.1.1), catalase (EC 1.11.1.6), glutathione peroxidase (EC 1.11.1.9), chloroperoxidase (EC 1.11.1.10) or peroxiredoxin (EC 1.11.1.15) were found in bacteria lacking CP (Passardi, 2006, submitted). Some of these enzymes are also present in amitochondriate and/or parasitic eukaryotes lacking non-animal superfamily peroxidases. Parasitic organisms probably also deal with elimination of ROS by using their host enzymes. Finally, animals and Mycetozoa possess peroxidases belonging to the “animal superfamily”.

Interestingly, a study on yeast (*Saccharomyces cerevisiae*) showed that disruption of its unique CCP gene was not essential for cell viability or respiration, though it rendered yeast more sensitive to H_2O_2 (Kwon et al., 2003). Although CP and CCP may not be essential to viability of, respectively, bacterial organisms or yeast, peroxidases certainly play a major role

in plants. The ancestral class I sequence probably played a crucial role in the bacteria that gave rise to chloroplasts and mitochondria. Class I peroxidases were preserved by evolution, suggesting that they are still bringing some advantage to the organisms that retained them.

Evolution of class I peroxidases to classes II and III

Specific duplication events have occurred for APX and class III in *Viridiplantae* and for class II in fungi. The conservation of duplicated genes could be related to the neofunctionalisation or subfunctionalisation but also specific to a particular phylum or branch. In fungi, class II peroxidases are only present in one saprophytic fungal family and are organized in clusters of highly conserved and duplicated sequences. The class II could have evolved from the existing class I together with the appearance and the diversification of lignin polymers in the plant cell wall. The ancestor of class II peroxidases is most probably a cytochrome C peroxidase that, in response to the very high redox potentials of lignin polymers, had to modify key amino acid residues able to oxidise their new targets. Appearance of class III peroxidases from class I is less clear, but may be related to the emergence of land plants, and hence could be due to the dramatic modifications of growth conditions (new organs, structures, atmospheric conditions) (Passardi et al., 2004). Interestingly, whenever an organism possesses class III peroxidases, it does not have any CCP encoding sequences. Therefore, it is possible that class I, which do not have disulfide bridges, are at the origin of class II and III peroxidases, which in contrast both form four disulfide bridges. In this respect, sequences from the ascomycetous (MagP01, 02 and 03), Chlamydomonas (CrePrx01 and 02) and HvinAnPrx01 may represent intermediary forms between CCP and class II/III peroxidases (Fig. 4). When searching for cysteine residues, it appears that these sequences contain less than three disulfide bridges. As class III peroxidases possess four disulfide bridges, while class I have none, this observation would be in favour of our theory for the origin of class III peroxidases. The discovery of two hybrid class I-class III peroxidases in *Chlamydomonas reinhardtii* and their phylogenetic position at the origin of the class III peroxidase branch suggests that other reasons than land colonisation led to the appearance of this class III. The number of class III peroxidase encoding genes has dramatically increased in land plants along with evolution: the most recent and evolved plants bear often more than 100 peroxidases genes (Bakalovic et al., 2006). Several reasons account for the “explosion” of class III peroxidases: organ diversification, climatic changes,

colonisation of new biotopes, constant appearance of new pathogens and the human impact on cultivated plants (Passardi et al., 2004).

The need for a revision of the peroxidase superfamily

In 1992, Karen Welinder made a description of the "plant, fungal and bacterial peroxidase superfamily", with the available information. Since then, extensive genome sequencing of very diverse organisms has allowed discovering that peroxidases from Welinder's superfamily are also well-represented in new classes of organisms. We therefore chose to update the peroxidase classification and propose the new name "superfamily of non-animal peroxidases". This denomination does not predict which organisms will have peroxidase encoding genes or not, but it reflects the fact that peroxidases can be present in almost any living organism.

The peroxidase sequences found in protistean organisms have evolved on their own, and their current form does not allow at first sight to precisely classify them, even after enzymatic activity experiments. In the next years, the amount of data on protistean organisms will increase, and many more sequences of this kind will appear that will help us to draw up a more precise evolution of the non animal peroxidase superfamily. It is probable that with the constant discovering of these "hybrid" forms of peroxidases, new classes will need to be defined. However its evolution is better understood now, the superfamily of non-animal peroxidases remains an exciting and challenging domain of gene evolution.

Experimental procedures

Retrieval of the peroxidase sequences

A part of the protein sequences used for the phylogenetic study was obtained from the UniProt database (http://www.expasy.uniprot.org/). In a second time, in order to find peroxidase sequences in a large variety of organisms, class I and III sequences from *Arabidopsis* and class I and II sequences from *Phanerochaete* were used as input sequences in tBLASTn searches within different databases. Peroxidases were sought for in the NCBI website (www.ncbi.nih.gov/BLAST) and in several specialized databases such as the PEP EST database (www.moss.leeds.ac.uk), the PlantGDB database (www.plantgdb.org/), the DOE Joint Genome Institute (JGI) website (genome.jgi-psf.org/). Some catalase-peroxidases

were also found trough the J. Craig Venter Institute BLAST server (https://research.venterinstitute.org/blast/). The non-annotated sequences were analyzed for the presence of the gene with different programs such as FGenesh (http://www.softberry.com/berry.phtml) and GenScan (http://genes.mit.edu/GENSCAN.html). The corresponding coding sequence (CDS) was translated with the "translate" tool on Expasy (http://us.expasy.org/tools/dna.html) and controlled for presence of specific motifs. Sequences of protein used in this study can be found at this following link: http://peroxidase.isb-sib.ch. (Bakalovic et al., 2006).

Comprehensive phylogenetic analysis of peroxidase sequences

Among the high number of available peroxidase sequences in the databases, for each class complete sequences representative of each phylum were used to build the global tree containing all three classes. For each independent tree, a large number of sequences have been aligned. Peroxidase protein sequences were aligned using Clustal W (Thompson et al., 1994). Only a part of the sequence containing highly conserved residues was used for the alignment. For this reason, signal peptide and C-term extension were removed. Regarding the CP sequences, only the N-term was used for the alignment. The alignment was further inspected and visually adjusted, and realigned with Clustal X. The distance tree was constructed with the NEIGHBOR option of the PHYLIP 3.6a3 package (Felsenstein, 2004) under the JTT substitution frequency matrix, and 1000 bootstrap replicates were carried out. Maximum parsimony trees were built with the PHYLO_WIN software, using 100 bootstrap replicates (Galtier et al., 1996). Maximum likelihood trees were inferred with the PHYML algorithm, under the JTT substitution frequency matrix, by using BIONJ starting tree (Guindon and Gascuel, 2003). Njplot software was used to visualize phylogenetic trees and BioEdit software to obtain the different consensus sequences.

Acknowledgments

We thank Jan Pawlowski for his critical reading and constructive suggestions and Laurent Falquet for the use of his unpublished profiles. The financial support of the Swiss National Science Foundation (grant 31-068003.02) to C.P. and C.D. is gratefully acknowledged; N.B. is paid by the Office Cantonal de l'Emploi.

D. Discussion and prospects

Following the study on the repartition of the peroxidase superfamily among living organisms and the exhaustive search of catalase-peroxidases in bacteria, it clearly appears that class I, II and III peroxidases originate from a same ancestor. The primitive peroxidase sequence remains an enigma. Nevertheless, the extensive search of peroxidases and the centralisation of this information in the PeroxiBase allowed going forward in this investigation. For instance, it seems that the ancient fusion event that gave rise to actual CP occurred before organellar transfer of CP. Furthermore, spreading of peroxidases also affected the protistean reign, and issued into novel forms that are difficult to categorize and may necessitate the creation of additional classes of peroxidases. The almost ubiquitous presence of peroxidases gives rise to paradoxical conclusions regarding their function. For instance, in bacteria, although most of aerobic bacteria were found with a catalase-peroxidase, some did not encode it. One explanation proposed was that other types of peroxidases could replace them: in this case, why keeping catalase-peroxidases? In green plants, an opposite paradox is reached by the very large multigenic families encountered with class III peroxidases. In our laboratory, despite having several lines of homozygous peroxidase T-DNA mutants in *Arabidopsis*, we observed that mutating one class III peroxidase usually does not dramatically affect the plant growth and development. Then why encoding so many? These questions may find some answers thanks to phylogenetic studies such as the ones described in this chapter, although such studies can bring even more questions: When did the first peroxidase appear? Why was the ancestral peroxidase duplicated? Why are there three classes of peroxidases? Why not only one? Why are there no peroxidases in animals? The peroxidase superfamily is still a widely open subject for fascinating studies and shows that searching for very ancient events can help understanding today’s observations.

6

Conclusions

A. What can a root tip tell to a phylogenetic branch?

Most of my PhD bench work consisted in applying different strategies in order to identify a function for a single peroxidase in *Arabidopsis*, namely AtPrx33. As it could have been expected from the redundant role of peroxidases, the phenotype of *atprx33* mutant was not particularly striking, although it was identified as a "short root" phenotype. Following studies with calli, it appeared that a single peroxidase can have various functions. This observation complicates the already tedious problem of functional redundancy. If one peroxidase can have different roles, then, again, why are there so many peroxidases in a plant? A part of the answer may come from the total lack of conservation of peroxidase promoters, which suggests that a same role can be fulfilled by many peroxidases, but not at the same time or at the same place. Considering the important amount of work needed to characterise the function of one peroxidase in one plant, it is quite unimaginable to consider studying plant peroxidases one by one. Even if it could prove necessary, especially for practical applications (defence of crops against pathogens, optimisation of growth conditions), this goal will be facilitated by phylogenetic studies and the constant update of the PeroxiBase. The work performed with AtPrx33, for instance, can give some guidelines to researchers interested in root development of other plants: it would not be surprising that the closest orthologs of AtPrx33 in other Brassicaceae (HRPC1B, BnPrx34) could also be involved in root elongation and cell differentiation. In this respect, the promoter sequence will be an essential point that should be investigated before assuming such hypotheses. In the article project "***Arabidopsis* peroxidase AtPrx33: a new actor on the scene of cell differentiation**" (figure 6), the mutant *atprx32* (second closest paralog to AtPrx34, with 88% identity) showed a delayed differentiation, confirming the assumption that close paralogs or orthologs can have the same functions. However, the two other peroxidases knock-out mutants (*atprx20* and *atprx47*) found to have a delayed differentiation were unrelated to AtPrx34. A particular common motif might be present in the protein sequence of AtPrx33, AtPrx32, AtPrx20 and AtPrx47. Alternatively, a specific pattern in the promoter sequence may favour expression of these peroxidases in calli. The PeroxiBase could hence be improved by addition of promoter sequences.

More generally, what root tips can tell to phylogenetic branches can be applied to any organ: one of the goals and assets of the PeroxiBase is precisely the presence of the fields "inducers", "repressors" and "tissue type". Together with the creators and developers of the

PeroxiBase, we hope that these indications will bring valuable tools for researchers to facilitate orientation of works destined to understand functions of single peroxidases.

Linking a peroxidase sequence (and its promoter) to a specific function and organ expression will be of particular interest in early organisms, as soon as enough data becomes available. Appearance of roots, vascular tissues, seeds or other major breakthroughs of evolution was probably accompanied by the emergence of new peroxidases. By comparing the peroxidases present in two (or more) organisms, provided that they possess sufficiently large EST libraries, it will be possible to identify specific groups of peroxidases and try and relate them with appearance of organs or other particular traits. For instance, when comparing peroxidase genes from *Arabidopsis thaliana* and *Oryza sativa*, we identified four OsPrx clusters specific to rice. In the cluster named "V.1" (see figure 4 in: "**The class III peroxidase multigenic family in rice and its evolution in land plants**"), one putative rice-specific (or possibly monocotyledon-specific) motif was even detected. Similarly, Christophe and Nenad compared all peroxidase sequences available from rice, wheat, soybean, *Medicago*, *Arabidopsis* and tomato, in an unpublished study. The tree obtained is rather cumbersome (it measured almost two metres!), but it confirmed presence of legume-specific (*Medicago* and soybean), cereal-specific (rice and wheat) and dicotyledon-specific (*Arabidopsis* and tomato) groups. Mixed groups containing peroxidases from all six organisms are also interesting: they probably contain more ancient peroxidases. Within these mixed groups, AtPrx20 and AtPrx47, the two peroxidases previously mentioned and shown to delay callus differentiation, were found. This observation is in line with a more ancient origin linked to organ formation. Concerning AtPrx33, however, it belonged to an *Arabidopsis*-specific cluster: roots have not yet told everything they know to phylogenetic branches…

B. Futuristic thoughts…

The functional study of AtPrx33 and AtPrx34 showed that both peroxidases play a role in root growth, and that AtPrx33 has some interaction with auxin. If the project on these two proteins could be continued, it would be interesting to obtain AtPrx34::*GFP* plants. Both peroxidases are extremely similar, and the rationale for keeping both genes in *Arabidopsis* should be found in their expression profiles. A possible variant of interest could be AtPrx34::*DsRed*, for co-localisation studies. DsRed is a red fluorescent protein from the coral-like anemone *Discosoma sp.* (Matz et al., 1999) and has absorption and emission wavelengths well distinct from GFP. Alternatively, the close paralog AtPrx32 could also be studied. While writing my thesis, I learned that a former PhD student, Michael Tognolli, had obtained AtPrx32::*GFP* transgenic plants. However, the AtPrx32 promoter fragment used was only covering the first 290 bp of the wild-type promoter (compared to 1200 bp for AtPrx33::*GFP*). In his thesis (pp. 90-91), he mentioned that roots of transgenic seedlings did not show any fluorescence up to two weeks after germination. He did not report fluorescence at the root tips of the central root, but only at the tip of adventitious roots. My transgenic AtPrx33::*GFP* plants also fluoresced at the root tip of adventitious roots (data not shown). Michael's experiments were indeed performed with low specificity filters on another microscope: he excited roots with a wavelength range of 450-490 nm (with the Axioplan, I used 475±2nm) and filtered emission with a longpass filter >515nm (I used a narrow EGFP filter of 510±10nm). I hence verified Michael's results by observing one week-old Atprx32::*GFP* seedlings: the results obtained confirm absence of the GFP signal in root tips, but also in the whole seedlings, with an exception in the region at the hypocotyl-root junction. As both peroxidases are close paralogs, this observation supports the idea that close peroxidases were retained by the plant because they are expressed in different regions and/or at different times.

By comparing AtPrx33::*GFP* observation to X-Gluc staining of GH3::*GUS* transgenic plants, I tried to establish a relationship between peroxidase expression and auxin distribution. Whereas elongation zone seems to be the place of choice for an interaction between AtPrx33 and auxin, the correlation was much less clear when different chemical treatments were applied. One explanation could be that the chemicals used have a wide range of consequences on many other mechanisms in the root. However, the use of GH3::*GUS* may also be a source of erroneous conclusions, as stated in the discussion of the article project "**The intricate**

relationship of auxin and AtPrx33 peroxidase: a dance or a glance?". In the future, it would be interesting to compare GUS expression with an antibody against auxin, which is generally more reliable than a promoter (discussed in chapter 2).

The search for ancestral sequences allowed finding some interesting candidates, such as AtPrx29, or the algal CpslPrx01. Despite these peroxidases may have a crucial importance in the development of specific structures or organs, my phylogenetic analyses were not able to draw any conclusions about their role. In the future, an interesting project could be transformation of *Chlamydomonas reinhardtii* with an "ancestral" peroxidase sequence. This alga is indeed well characterised, readily available, and molecular biology techniques can be easily applied. More importantly, it is a representative of the phylum Chlorophyta, which lacks class III peroxidase sequences. Once transformed, parameters such as growth rate in water and out of water, cell wall composition or ROS production and consumption could be monitored. Such analysis could provide unique information on the role of ancestral peroxidases in algae, and maybe bring some clues to explain why Chlorophyta, while able to grow on land (Lange et al., 1992; Lewis and Flechtner, 2002), did not evolve to more complex organisms. A similar approach could be envisaged with the peroxidase AtPrx42 transformed in cereals or cryptogams. Regarding cereals, transgenic AtPrx42-expressing plants may bring interesting clues for explaining the sudden disappearance of AtPrx42 in this group. Attention should be focused especially on seeds, since *atprx42* knock-out mutations in *Arabidopsis* seem to be lethal (chapter 4). Regarding later developmental stages, it will be necessary to use an inducible promoter coupled to an RNAi construct for *AtPrx42*, a project that is currently developed by Marion.

Phylogenetic studies demonstrated to be powerful tools that can even replace some bench work experiments. The most striking example was the discovery of the algal class III peroxidase, suddenly invalidating three years of speculations based on experimental evidence. The lack of guaiacol oxidase activity in *Nitella hyalina* and the three Zygnemophyceae analysed in the laboratory, as well as the tenuous activity observed in *Coleochaete scutata*, made us predict appearance of the first class III peroxidases in charophycean algae. Since this event should have occurred in an ancestor of *Chara*, I was not particularly surprised to detect several peroxidase isoforms during my experiments with this alga: it was easily conceivable that a few gene duplications would have happened in the 500 million years (McCourt et al., 2004) following appearance of the first Charales.

Limitations of phylogenetic studies are encountered when the gene of interest is not encoded by an organism: in that case, they can be combined with bench work and give rise to very interesting projects. Following the study of catalase-peroxidases (CP) in bacteria, an idea was to transform anaerobic bacteria lacking CP with a catalase-peroxidase gene, in order to assess the importance of this peroxidase in resistance to aerobic conditions. The candidate of choice was *Clostridium*: various species of this bacterium are commercially available and efficient transformation protocols have been developed (Allen and Blaschek, 1988; Tyurin et al., 2004). Moreover, it is known that *Clostridium* is not able to grow in aerobic conditions. However, in microoxic environments (around 1% oxygen), several studies have shown that this bacterium was able to consume oxygen (Kawasaki et al., 1998; Karnholz et al., 2002). If CP is truly able to help anaerobic bacteria coping with aerobic conditions, then we could study growth rate and resistance of this novel artificial *Clostridium* line in various oxygenic environments, and "mimic" a possible evolutionary event.

Many questions remain unanswered concerning peroxidases. The advent of molecular biology and bioinformatics were not sufficient to solve crucial points, such as the exact role and mechanism of action of a single class III peroxidase in a plant, the event that led catalase-peroxidases to appear in bacteria, where so many other peroxidases can fulfil similar detoxifying functions, or the origin of the superfamily. Sequencing of more organisms is definitely needed, as well as an extension of peroxidase promoter studies and the development of reliable and sharp techniques to identify peroxidase *in vivo* substrates such as auxin or the various types of reactive oxygen species. Better defining the function of single class III peroxidases is an essential task that will bring precious information about the raison d'être of these enzymes and, more generally, of plant multigenic families.

7

Literature

Abel S, Theologis A (1996) Early genes and auxin action. Plant Physiol **111:** 9-17

Allen SP, Blaschek HP (1988) Electroporation-induced transformation of intact cells of *Clostridium perfringens*. Appl Environ Microbiol **54:** 2322-2324

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science **301:** 653-657

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends Plant Sci **5:** 22-29

Al-Shehbaz IA, O'Kane SLJ (2002) Taxonomy and phylogeny of Arabidopsis (Brassicaceae). American Society of Plant Biologists, *In* The Arabidopsis book, pp 1-22

Asada K (1999) The water-water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. Annu Rev Plant Physiol Plant Mol Biol **50:** 601-639

Auderset Joye D, Castella E, Lachavanne JB (2002) Occurrence of Characeae in Switzerland over the last two centuries (1800-2000). Aquatic Botany **72:** 369-385

Avsian-Kretchmer O, Cheng JC, Chen L, Moctezuma E, Sung ZR (2002) Indole acetic acid distribution coincides with vascular differentiation pattern during Arabidopsis leaf ontogeny. Plant Physiol **130:** 199-209

Aylward N, Bofinger N (2005) Possible origin for porphin derivatives in prebiotic chemistry--a computational study. Orig Life Evol Biosph **35:** 345-368

Aylward N, Bofinger N (2006) A plausible prebiotic synthesis of pyridoxal phosphate: vitamin B6 - a computational study. Biophys Chem **123:** 113-121

Badiani M, Felici M, Luna M, Artemi F (1983) Laccase assays by means of high-performance liquid chromatography. Anal Biochem **133:** 275-276

Bakalovic N, Passardi F, Ioannidis V, Cosio C, Penel C, et al. (2006) PeroxiBase: a class III plant peroxidase database. Phytochemistry **67:** 534-539

Baluska F, Hlavacka A, Samaj J, Palme K, Robinson DG, et al. (2002) F-actin-dependent endocytosis of cell wall pectins in meristematic root cells. Insights from brefeldin A-induced compartments. Plant Physiol **130:** 422-431

Bandyopadhyay P, Steinman HM (2000) Catalase-peroxidases of *Legionella pneumophila*: cloning of the katA gene and studies of KatA function. J Bacteriol **182:** 6679-6686

Becker D, Hedrich R (2002) Channelling auxin action: modulation of ion transport by indole-3-acetic acid. Plant Mol Biol **49:** 349-356

Beemster GT, De Vusser K, De Tavernier E, De Bock K, Inze D (2002) Variation in growth rate between Arabidopsis ecotypes is correlated with cell division and A-type cyclin-dependent kinase activity. Plant Physiol **129:** 854-864

Beloqui A, Pita M, Polaina J, Martinez-Arias A, Golyshina OV, et al. (2006) Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. J Biol Chem **281:** 22933-22942

Benkova E, Michniewicz M, Sauer M, Teichmann T, Seifertova D, et al. (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell **115:** 591-602

Bent AF (2000) Arabidopsis in planta transformation. Uses, mechanisms, and prospects for transformation of other species. Plant Physiol **124:** 1540-1547

Berglund GI, Carlsson GH, Smith AT, Szoke H, Henriksen A, et al. (2002) The catalytic pathway of horseradish peroxidase at high resolution. Nature **417:** 463-468

Bertaccini A, Franova J, Botti S, Tabanelli D (2005) Molecular characterization of phytoplasmas in lilies with fasciation in the Czech Republic. FEMS Microbiol Lett **249:** 79-85

Bertani G (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol **62:** 293-300

Bertrand T, Eady NA, Jones JN, Jesmin NA, Nagy JM et al. Crystal structure of *Mycobacterium tuberculosis* catalase-peroxidase. J Biol Chem **279:** 38991-38999

Bierfreund NM, Reski R, Decker EL (2003) Use of an inducible reporter gene system for the analysis of auxin distribution in the moss *Physcomitrella patens*. Plant Cell Rep **21:** 1143-1152

Binder M, Hibbett DS (2002) Higher-level phylogenetic relationships of Homobasidiomycetes (mushroom-forming fungi) inferred from four rDNA regions. Mol Phylogenet Evol **22:** 76-90

Blazquez MA, Soowal LN, Lee I, Weigel D (1997) LEAFY expression and flower initiation in Arabidopsis. Development **124:** 3835-3844

Boerjan W, Cervera MT, Delarue M, Beeckman T, Dewitte W, et al. (1995) Superroot, a recessive mutation in Arabidopsis, confers auxin overproduction. Plant Cell **7:** 1405-1419

Boutte Y, Crosnier MT, Carraro N, Traas J, Satiat-Jeunemaitre B (2006) The plasma membrane recycling pathway and cell polarity in plants: studies on PIN proteins. J Cell Sci **119:** 1255-1265

Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, et al. (2001) Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. Plant Cell **13:** 1499-1510

Bruns TD, Vilgalys R, Barns SM, Gonzalez D, Hibbett DS, et al. (1992) Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Mol Phylogenet Evol **1:** 231-241

Cameron RE (1960) Communities of soil algae accurring in the Sonoran Desert in Arizona. J Arizona Acad Sciences **1:** 85-88

Carpena X, Loprasert S, Mongkolsuk S, Switala J, Loewen PC et al. (2003) Catalase-peroxidase KatG of *Burkholderia pseudomallei* at 1.7Å resolution. J Mol Biol **327:** 475-489

Carpin S, Crevecoeur M, de Meyer M, Simon P, Greppin H, et al. (2001) Identification of a Ca^{2+}-pectate binding site on an apoplastic peroxidase. Plant Cell **13:** 511-520

Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J **3:** 1-30

Carrington JC, Freed DD (1990) Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. J Virol **64:** 1590-1597

Carter C, Pan S, Zouhar J, Avila EL, Girke T, et al. (2004) The vegetative vacuole proteome of *Arabidopsis thaliana* reveals predicted and unexpected proteins. Plant Cell **16:** 3285-3303

Cary A, Uttamchandani SJ, Smets R, Van Onckelen HA, Howell SH (2001) Arabidopsis mutants with increased organ regeneration in tissue culture are more competent to respond to hormonal signals. Planta **213:** 700-707

Casal JJ, Whitelam GC, Smith H (1990) Phytochrome control of extracellular peroxidase activity in mustard internodes: correlation with growth, and comparison with the effect of wounding. Photochemistry and Photobiology **52:** 165-172

Casal JJ (2000) Phytochromes, cryptochromes, phototropin: photoreceptor interactions in plants. Photochem Photobiol **71:** 1-11

Casal JJ, Luccioni LG, Oliverio KA, Boccalandro HE (2003) Light, phytochrome signalling and photomorphogenesis in Arabidopsis. Photochem Photobiol Sci **2:** 625-636

Castillo FJ, Celardin F, H G (1984) Peroxidase assay in plants - interference by ascorbic-acid and endogenous inhibitors in *Sedum* and *Pelargonium* enzyme extracts. Plant Growth Regulation **2:** 69-75

Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol **52:** 597-354

Cavalier-Smith T (2004) Only six kingdoms of life. Proc Biol Sci **271:** 1251-1262

Chance B, Maehly AC (1955) Assay of catalases and peroxidases. Methods in Enzymology **2:** 764-775

Che P, Lall S, Nettleton D, Howell SH (2006) Gene expression programs during shoot, root, and callus development in Arabidopsis tissue culture. Plant Physiol **141:** 620-637

Chen SX, Schopfer, P (1999) Hydroxyl-radical production in physiological reactions. A novel function of peroxidase. Eur J Biochem **260:** 726-735

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, et al. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res **31:** 3497-3500

Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, et al. (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. Genes Dev **17:** 1043-1054

Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, et al. (2001) Massive gene decay in the leprosy bacillus. Nature **409:** 1007-1011

Conesa A, Punt PJ, van den Hondel CA (2002) Fungal peroxidases: molecular aspects and applications. J Biotechnol **93:** 143-158

Cordewener J, Booij H, van der Zandt H, van Engelen F, van Kammem A, et al. (1991) Tunicamycin-inhibited carrot somatic embryogenesis can be restored by secreted cationic peroxidase isoenzymes. Planta **184:** 478-486

Cosgrove DJ (2000) Loosening of plant cell walls by expansins. Nature **407:** 321-326

Dagan T, Blekhman R, Graur D (2006) The "domino theory" of gene death: gradual and mass gene extinction events in three lineages of obligate symbiotic bacterial pathogens. Mol Biol Evol **23:** 310-316

Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell **101:** 543-553

de Candolle AP (1821) Regni Vegetabilis Systema Naturale. Treuttel and Würtz, Paris

Debeaujon I, Leon-Kloosterziel KM, Koornneef M (2000) Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiol **122:** 403-414

DeLano WL (2002) The PyMol molecular graphics system. DeLano Scientific, San Carlos, CA, USA

Dharmasiri N, Dharmasiri S, Estelle M (2005) The F-box protein TIR1 is an auxin receptor. Nature **435:** 441-445

DiTomaso JM, Hart JJ, Kochian LV (1993) Compartmentation analysis of Paraquat fluxes in maize roots as a means of estimating the rate of vacuolar accumulation and translocation to shoots. Plant Physiol **102:** 467-472

Dobrindt U, Hochhut B, Hentschel U, Hacker J Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol **2:** 414-424

Dowd PF, Lagrimini LM, Herms DA (1998) Differential leaf resistance to insects of transgenic sweetgum (*Liquidambar styraciflua*) expressing tobacco anionic peroxidase. Cell Mol Life Sci **54:** 712-720

Duarte-Vazquez MA, Garcia-Almendarez BE, Regalado C, Whitaker JR (2001) Purification and properties of a neutral peroxidase isozyme from turnip (*Brassica napus* L. var. Purple Top White Globe) roots. J Agric Food Chem **49:** 4450-4456

Dunand C, Tognolli M, Overney S, von Tobel L, De Meyer M, et al. (2002) Identification and characterisation of Ca^{2+}-pectate binding peroxidases in *Arabidopsis thaliana*. J Plant Physiol **159:** 1165-1171

Dunford HB (1999) Heme peroxidase nomenclature. Plant Peroxidase Newsletter **13:** 65-71

Dunn MA, White AJ, Vural S, Hughes MA (1998) Identification of promoter elements in a low-temperature-responsive gene (blt4.9) from barley (*Hordeum vulgare* L.). Plant Mol Biol **38:** 551-564

Duroux L, Welinder KG (2003) The peroxidase gene family in plants: a phylogenetic overview. J Mol Evol **57:** 397-407

Edwards SL, Raag R, Wariishi H, Gold MH, Poulos TL (1993) Crystal structure of lignin peroxidase. Proc Natl Acad Sci U S A **90:** 750-754

El-Lithy ME, Clerkx EJ, Ruys GJ, Koornneef M, Vreugdenhil D (2004) Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. Plant Physiol **135:** 444-458

Embley TM, van der Giezen M, Horner DS, Dyal PL, Bell S, et al. (2003) Hydrogenosomes, mitochondria and early eukaryotic evolution. IUBMB Life **55:** 387-395

Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. Nature **440:** 623-630

Erman JE, Vitello LB (2002) Yeast cytochrome c peroxidase: mechanistic studies via protein engineering. Biochim Biophys Acta **1597:** 193-220

Faguy DM, Doolittle WF (2000) Horizontal transfer of catalase-peroxidase genes between archaea and pathogenic bacteria. Trends Genet **16:** 196-197

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6. University of Washington, Seattle

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, et al. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391:** 806-811

Fokin SI (2004) Bacterial endocytobionts of ciliophora and their interactions with the host cell. Int Rev Cytol **236:** 181-249

Fournier M, Zhang Y, Wildschut JD, Dolla A, Voordouw JK, et al. (2003) Function of oxygen resistance proteins in the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* hildenborough. J Bacteriol **185:** 71-79

Frank M, Rupp HM, Prinsen E, Motyka V, Van Onckelen H, et al. (2000) Hormone autotrophic growth and differentiation identifies mutant lines of Arabidopsis with altered cytokinin and auxin content or signaling. Plant Physiol **122:** 721-729

Friml J (2003) Auxin transport - shaping the plant. Curr Opin Plant Biol **6:** 7-12

Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, et al. (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature **438:** 1105-1115

Gallagher SR, ed (1992) GUS Protocols. Using the GUS gene as a reporter of gene expression. Academic press

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci **12:** 543-548

Garcia-Vallve S, Romeu A, Palau J (2000) Horizontal gene transfer in bacterial and archaeal complete genomes. Genome Res **10:** 1719-1725

Garrity GM, ed (2001-2005) Bergey's Manual of Systematic Bacteriology, 2nd Ed Vol 1 and 2. Springer

Gaspar T, Penel C, Thorpe T, Greppin H (1982) Physiological processes mediated by peroxidases. *In* Peroxidases 1970-1980. A survey of their biochemical and physiological roles in higher plants. University of Geneva, Geneva, pp 90-97

Gazaryan IG, Lagrimini LM (1996) Purification and unusual kinetic properties of a tobacco anionic peroxidase. Phytochemistry **41:** 1029-1034

Gazaryan IG, Lagrimini LM, Ashby GA, Thorneley RN (1996) Mechanism of indole-3-acetic acid oxidation by plant peroxidases: anaerobic stopped-flow spectrophotometric studies on horseradish and tobacco peroxidases. Biochem J **313 (Pt 3):** 841-847

Geldner N, Friml J, Stierhof YD, Jurgens G, Palme K (2001) Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. Nature **413:** 425-428

Goda H, Sawa S, Asami T, Fujioka S, Shimada Y, et al. (2004) Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in Arabidopsis. Plant Physiol **134:** 1555-1573

Gogarten JP, Townsend JP (2005) Horizontal gene transfer, genome innovation and evolution. Nat Rev Microbiol **3:** 679-687

Gold MH, Alic M (1993) Molecular biology of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Microbiol Rev **57:** 605-622

Grace ML, Chandrasekharan MB, Hall TC, Crowe AJ (2004) Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter. J Biol Chem **279:** 8102-8110

Gray MW (1999) Evolution of organellar genomes. Curr Opin Genet Dev **9:** 678-687

Gray MW (2005) Evolutionary biology: the hydrogenosome's murky past. Nature **434:** 29-31

Gray WM, Kepinski S, Rouse D, Leyser O, Estelle M (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. Nature **414:** 271-276

Grossart HP, Levold F, Allgaier M, Simon M, Brinkhoff T (2005) Marine diatom species harbour distinct bacterial communities. Environ Microbiol **7:** 860-873

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52:** 696-704

Gyula P, Schafer E, Nagy F (2003) Light perception and signalling in higher plants. Curr Opin Plant Biol **6:** 446-452

Habetha M, Bosch TC (2005) Symbiotic Hydra express a plant-like peroxidase gene during oogenesis. J Exp Biol **208:** 2157-2165

Hagen G, Kleinschmidt AJ, Guilfoyle T (1984) Auxin-regulated gene expression in intact soybean hypocotyl and excised hypocotyl sections. Planta **162:** 147-153

Hagen G, Martin G, Li Y, Guilfoyle TJ (1991) Auxin-induced expression of the soybean GH3 promoter in transgenic tobacco plants. Plant Mol Biol **17:** 567-579

Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol Biol **49:** 373-385

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Mol Biol **25:** 989-994

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser **41:** 95-98

Hannaert V, Saavedra E, Duffieux F, Szikora JP, Rigden DJ, et al. (2003) Plant-like traits associated with metabolism of *Trypanosoma* parasites. Proc Natl Acad Sci U S A **100:** 1067-1071

Hassan HM, Fridovich I (1979) Paraquat and *Escherichia coli*. Mechanism of production of extracellular superoxide radical. J Biol Chem **254:** 10846-10852

Heinemann JA, Sprague GF, Jr. (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature **340:** 205-209

Heloir MC, Kevers C, Hausman JF, Gaspar T (1996) Changes in the concentrations of auxins and polyamines during rooting of *in vitro*-propagated walnut shoots. Tree Physiol **16:** 515-519

Hiraga S, Yamamoto K, Ito H, Sasaki K, Matsui H, et al. (2000) Diverse expression profiles of 21 rice peroxidase genes. FEBS Lett **471:** 245-250

Hiraga S, Sasaki K, Ito H, Ohashi Y, Matsui H (2001) A large family of class III plant peroxidases. Plant Cell Physiol **42:** 462-468

Hoffmann MH, Glass AS, Tomiuk J, Schmuths H, Fritsch RM, et al. (2003) Analysis of molecular data of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae) with Geographical Information Systems (GIS). Mol Ecol **12:** 1007-1019

Hou ZX, Huang WD (2005) Immunohistochemical localization of IAA and ABP1 in strawberry shoot apexes during floral induction. Planta **222:** 678-687

Hrdy I, Hirt RP, Dolezal P, Bardonova L, Foster PG, et al. (2004) *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature **432:** 618-622

Intapruk C, Higashimura N, Yamamoto K, Okada N, Shinmyo A, et al. (1991) Nucleotide sequences of two genomic DNAs encoding peroxidase of *Arabidopsis thaliana*. Gene **98:** 237-241

Ishikawa T, Takeda T, Kohno H, Shigeoka S (1996) Molecular characterization of *Euglena* ascorbate peroxidase using monoclonal antibody. Biochim Biophys Acta **1290:** 69-75

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. Science **309:** 436-442

Jain R, Rivera MC, Moore JE, Lake JA (2002) Horizontal gene transfer in microbial genome evolution. Theor Popul Biol **61:** 489-495

Jain R, Rivera MC, Moore JE, Lake JA (2003) Horizontal gene transfer accelerates genome innovation and evolution. Mol Biol Evol **20:** 1598-1602

Jones MA, Raymond MJ, Smirnoff N (2006) Analysis of the root-hair morphogenesis transcriptome reveals the molecular identity of six genes with roles in root-hair development in Arabidopsis. Plant J **45:** 83-100

Karnholz A, Kusel K, Gossner A, Schramm A, Drake HL (2002) Tolerance and metabolic response of acetogenic bacteria toward oxygen. Appl Environ Microbiol **68:** 1005-1009

Katagiri N, Tsutsumi Y, Nishida T (1995) Correlation of brightening with cumulative enzyme activity related to lignin biodegradation during biobleaching of kraft pulp by white rot fungi in the solid-state fermentation system. Appl Environ Microbiol **61:** 617-622

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, et al. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature **414:** 450-453

Kawano T (2003) Roles of the reactive oxygen species-generating peroxidase reactions in plant defense and growth induction. Plant Cell Rep **21:** 829-837

Kawaoka A, Matsunaga E, Endo S, Kondo S, Yoshida K, et al. (2003) Ectopic expression of a horseradish peroxidase enhances growth rate and increases oxidative stress resistance in hybrid aspen. Plant Physiol **132:** 1177-1185

Kawasaki S, Nakagawa T, Nishiyama Y, Benno Y, Uchimura T, et al. (1998) Effect of oxygen on the growth of *Clostridium butyricum* (type species of the genus *Clostridium*), and the distribution of enzymes for oxygen and for active oxygen species in Clostridia. J Ferm Bioeng **86:** 368-372

Kawasaki S, Watamura Y, Ono M, Watanabe T, Takeda K, et al. (2005) Adaptive responses to oxygen stress in obligatory anaerobes *Clostridium acetobutylicum* and *Clostridium aminovalericum*. Appl Environ Microbiol **71:** 8442-8450

Kay LE, Basile DV (1987) Specific peroxidase isoenzymes are correlated with organogenesis. Plant Physiol **84:** 99-105

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature **389:** 33-39

Kim SS, Choi SY, Park JH, Lee DJ (2004) Regulation of the activity of Korean radish cationic peroxidase promoter during dedifferentiation and differentiation. Plant Physiol Biochem **42:** 763-772

Kis M, Burbridge E, Brock IW, Heggie L, Dix PJ, et al. (2004) An N-terminal peptide extension results in efficient expression, but not secretion, of a synthetic horseradish peroxidase gene in transgenic tobacco. Ann Bot (Lond) **93:** 303-310

Kitajima S, Ueda M, Sano S, Miyake C, Kohchi T, et al. (2002) Stable form of ascorbate peroxidase from the red alga *Galdieria partita* similar to both chloroplastic and cytosolic isoforms of higher plants. Biosci Biotechnol Biochem **66:** 2367-2375

Kjaersgard IV, Jespersen HM, Rasmussen SK, Welinder KG (1997) Sequence and RT-PCR expression analysis of two peroxidases from *Arabidopsis thaliana* belonging to a novel evolutionary branch of plant peroxidases. Plant Mol Biol **33:** 699-708

Klockgether J, Reva O, Larbig K, Tummler B (2004) Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. J Bacteriol **186:** 518-534

Klotz MG, Loewen PC (2003) The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota. Mol Biol Evol **20:** 1098-1112

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). Mol Biol Evol **17:** 1483-1498

Kowalczyk M, Sandberg G (2001) Quantitative analysis of indole-3-acetic acid metabolites in Arabidopsis. Plant Physiol **127:** 1845-1853

Kwon M, Chong S, Han S, Kim K (2003) Oxidative stresses elevate the expression of cytochrome c peroxidase in *Saccharomyces cerevisiae*. Biochim Biophys Acta **1623:** 1-5

Landolt E (1977) Oekologische Zeigerwerte zur Schweizer Flora. Veröffentlichungen des Geobotanischen Institutes der ETH, Stiftung Rübel, Zürich, Switzerland

Lange OL, Kidron GJ, Budel B, Meyer A, Kilian E, et al. (1992) Taxonomic composition and photosynthetic characteristics of the biological soil crusts covering sand dunes in the western Negev desert. Funct Ecol **6:** 519-527

Lariguet P, Boccalandro HE, Alonso JM, Ecker JR, Chory J, et al. (2003) A growth regulatory loop that provides homeostasis to phytochrome A signaling. Plant Cell **15:** 2966-2978

Larkin JC, Young N, Prigge M, Marks MD (1996) The control of trichome spacing and number in Arabidopsis. Development **122:** 997-1005

Larkin PJ, Gibson JM, Mathesius U, Weinman JJ, Gartner E, et al. (1996) Transgenic white clover. Studies with the auxin-responsive promoter, GH3, in root gravitropism and lateral root development. Transgenic Res **5:** 325-335

Leander BS (2004) Did trypanosomatid parasites have photosynthetic ancestors? Trends Microbiol **12:** 251-258

Leverone LA, Kossenjans W, Jayasimihulu K, Caruso JL (1991) Evidence of zein-bound indoleacetic acid using gas chromatography-selected ion monitoring-mass spectrometry analysis and immunogold labeling. Plant Physiol **96:** 1070-1075

Lewis LA, Flechtner VR (2002) Green algae (Chlorophyta) of desert microbiotic crusts: diversity of North American taxa. Taxon **51:** 443-451

Lewis LA, McCourt RM (2004) Green algae and the origin of land plants. Am J Bot **91:** 1535-1556

Lewis LA, Lewis PO (2005) Unearthing the molecular phylodiversity of desert soil green algae (Chlorophyta). Syst Biol **54:** 936-947

Li L, Xu J, Xu ZH, Xue HW (2005) Brassinosteroids stimulate plant tropisms through modulation of polar auxin transport in *Brassica* and *Arabidopsis*. Plant Cell **17:** 2738-2753

Liszkay A, Kenk B, Schopfer P (2003) Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth. Planta **217:** 658-667

Liu ZB, Ulmasov T, Shi X, Hagen G, Guilfoyle TJ (1994) Soybean GH3 promoter contains multiple auxin-inducible elements. Plant Cell **6:** 645-657

Ljung K, Hull AK, Celenza J, Yamada M, Estelle M, et al. (2005) Sites and regulation of auxin biosynthesis in Arabidopsis roots. Plant Cell **17:** 1090-1104

Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and Why? Plant Physiol **118:** 9-17

Martinez AT (2002) Molecular biology and structure function of lignin-degrading heme peroxidase. Enz Microb Technol **30:** 425-444

Martinez AT, Speranza M, Ruiz-Duenas FJ, Ferreira P, Camarero S, et al. (2005) Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int Microbiol **8:** 195-204

Martinez MJ, Ruiz-Duenas FJ, Guillen F, Martinez AT (1996) Purification and catalytic properties of two manganese peroxidase isoenzymes from *Pleurotus eryngii*. Eur J Biochem **237:** 424-432

Mathesius U, Schlaman HRM, Spaink HP, Sautter C, Rolfe BG, et al. (1998) Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. Plant J **14:** 23-34

Mathesius U (2001) Flavonoids induced in cells undergoing nodule organogenesis in white clover are regulators of auxin breakdown by peroxidase. J Exp Bot **52:** 419-426

Matsui T, Nakayama H, Yoshida K, Shinmyo A (2003) Vesicular transport route of horseradish C1a peroxidase is regulated by N- and C-terminal propeptides in tobacco cells. Appl Microbiol Biotechnol **62:** 517-522

Matsui T, Hori M, Shizawa N, Nakayama H, Shinmyo A, et al. (2006) High-efficiency secretory production of peroxidase C1a using vesicular transport engineering in transgenic tobacco. J Biosci Bioeng **102:** 102-109

Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, et al. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. Nat Biotechnol **17:** 969-973

Mauro ML, De Lorenzo G, Costantino P, Bellincampi D (2002) Oligogalacturonides inhibit the induction of late but not of early auxin-responsive genes in tobacco. Planta **215:** 494-501

Mazel A, Levine A (2001) Induction of cell death in arabidopsis by superoxide in combination with salicylic acid or with protein synthesis inhibitors. Free Radic Biol Med **30:** 98-106

McBride KE, Summerfelt KR (1990) Improved binary vectors for Agrobacterium-mediated plant transformation. Plant Mol Biol **14:** 269-276

McCourt RM (1995) Green algal phylogeny. Trends Ecol Evol **10:** 159-163

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends Ecol Evol **19:** 661-666

McFadden GI (2001) Primary and secondary endosymbiosis and the origin of plastids. J Phycol **37:** 951-959

McInnis SM, Costa LM, Gutierrez-Marcos JF, Henderson CA, Hiscock SJ (2005) Isolation and characterization of a polymorphic stigma-specific class III peroxidase gene from *Senecio squalidus* L. (Asteraceae). Plant Mol Biol **57:** 659-677

Mertens R, Eberle J, Arnscheidt A, Ledebur A, Weiler EW (1985) Monoclonal antibodies to plant growth regulators. II. Indole-3-acetic acid. Planta **166:** 389-393

Meyer RC, Torjek O, Becher M, Altmann T (2004) Heterosis of biomass production in Arabidopsis. Establishment during early development. Plant Physiol **134:** 1813-1823

Miller JG, Fry SC (2001) Characteristics of xyloglucan after attack by hydroxyl radicals. Carbohydr Res **332:** 389-403

Minor JL (1991) Location of lignin-bonded pectic polysaccharides. J Wood Chem Technol **11:** 159-169

Mitchell RJ, Garrett HE, Cox GS, Atalay A (1986) Boron and ectomycorrhizal influences on indole-3-acetic acid levels and indole-3-acetic acid oxidase and peroxidase activities of *Pinus echinata* Mill. roots. Tree Physiol **1:** 1-8

Miyake C, Michihata F, Asada K (1991) Scavenging of hydrogen peroxide in prokaryotic and eukaryotic algae: acquisition of ascorbate peroxidase during the evolution of cyanobacteria. Plant Cell Physiol **32:** 33–43

Moncalvo JM, Vilgalys R, Redhead SA, Johnson JE, James TY, et al. (2002) One hundred and seventeen clades of euagarics. Mol Phylogenet Evol **23:** 357-400

Moreira MT, Feijoo G, Mester T, Mayorga P, Sierra-Alvarez R, et al. (1998) Role of organic acids in the manganese-independent biobleaching system of *Bjerkandera sp.* strain BOS55. Appl Environ Microbiol **64:** 2409-2417

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio-assay with tobacco tissue culture. Physiol Plant **15:** 473-497

Nakamura K, Go N (2005) Function and molecular evolution of multicopper blue proteins. Cell Mol Life Sci **62:** 2050-2066

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. Plant Cell **2:** 279-289

Nebert DW, Gonzalez FJ (1987) P450 genes: structure, evolution, and regulation. Annu Rev Biochem **56:** 945-993

Neuhaus JM (1996) Protein targeting to the plant vacuole. Plant Physiol Biochem **34:** 217-221

Nicholas KB, Nicholas Jr. HB (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. v. 2.6.002. distributed by the authors

Nicholls P, Fita I, Loewen, PC (2001) Enzymology and structure of catalases. Adv Inorg Chem **51**: 51-106

Nicolaisen M, Johansen E, Poulsen GB, Borkhardt B (1992) The 5' untranslated region from pea seedborne mosaic potyvirus RNA as a translational enhancer in pea and tobacco protoplasts. FEBS Lett **303:** 169-172

Nilsson O, Moritz T, Sundberg B, Sandberg G, Olsson O (1996) Expression of the *Agrobacterium rhizogenes* rolC gene in a deciduous forest tree alters growth and development and leads to stem fasciation. Plant Physiol **112:** 493-502

Nishihara N, Horiike S, Takahashi T, Kosaka T, Shigenaka Y, et al. (1998) Cloning and characterization of endosymbiotic algae isolated from *Paramecium bursaria*. Protoplasma **203:** 91-99

Nordström AC, Eliasson L (1991) Levels of endogenous indole-3-acetic acid and indole-3-acetylaspartic acid during adventitious root formation in pea cuttings. Physiol Plant **82:** 599-605

Obinger C, Regelsberger G, Engleder M, Dietmann D, Jakopitsch C, et al. (1999) A transient-state kinetic investigation of reactions of wild-type and recombinant cyanobacterial catalase-peroxidases with peroxides, phenols and anilines. Plant Peroxidase Newsletter **13:** 73-81

Ogawa N, Okamura H, Hirai H, Nishida T (2004) Degradation of the antifouling compound Irgarol 1051 by manganese peroxidase from the white rot fungus *Phanerochaete chrysosporium*. Chemosphere **55:** 487-491

Overmyer K, Tuominen H, Kettunen R, Betz C, Langebartels C, et al. (2000) Ozone-sensitive Arabidopsis *rcd1* mutant reveals opposite roles for ethylene and jasmonate signaling pathways in regulating superoxide-dependent cell death. Plant Cell **12:** 1849-1862

Passardi F, Longet D, Penel C, Dunand C (2004) The class III peroxidase multigenic family in rice and its evolution in land plants. Phytochemistry **65:** 1879-1893

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci **9:** 534-540

Passardi F, Cosio C, Penel C, Dunand C (2005) Peroxidases have more functions than a Swiss army knife. Plant Cell Rep **24:** 255-265

Passardi F, Dobias J, Valério L, Guimil S, Penel C, et al. (2006) Morphological and physiological traits of three major *Arabidopsis thaliana* accessions. J Plant Physiol**:** in press

Passardi F, Favet J, Penel C, Dunand C (2006) Phylogenetic distribution of catalase-peroxidases: are there patches of order in chaos? submitted article

Passardi F, Tognolli M, De Meyer M, Penel C, Dunand C (2006) Two cell wall associated peroxidases from Arabidopsis influence root elongation. Planta **223:** 965-974

Passardi F, Bakalovic N, Teixeira FK, Margis-Pinheiro M et al. Prokaryotic origins of the non-animal peroxidase superfamily and organellar-mediated transmission to eukaryotes. Genomics: in press

Patterson WR, Poulos TL (1995) Crystal structure of recombinant pea cytosolic ascorbate peroxidase. Biochemistry **34:** 4331-4341

Penel C (1976) Activité peroxydasique et développement chez *Spinacia oleracea* (thesis work). Université de Genève, Geneva, Switzerland

Penel C, Gaspar T, Greppin H (1992) Plant peroxidases. University of Geneva, Switzerland

Pickering JW, Powell BL, Wender SH, Smith EC (1973) Ferulic acid: a substrate for two isoperoxidases from *Nicotiana tabacum* tissue cultures. Phytochemistry **12:** 2639-2643

Piontek K, Smith AT, Blodig W (2001) Lignin peroxidase structure and function. Biochem Soc Trans **29:** 111-116

Prestridge DS (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. Comput Appl Biosci **7:** 203-206

Price RA, Palmer JD, Al-Shehbaz IA (1994) Systematic relationships of Arabidopsis: a molecular and morphological perspective. EM Meyerowitz, CR Somerville, *In* Arabidopsis. Cold Spring Harbor Press, pp 7-20

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, et al. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature **409:** 618-622

Quiroga M, Guerrero C, Botella MA, Barcelo A, Amaya I, et al. (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. Plant Physiol **122:** 1119-1127

Ratcliffe D (1965) The geographical and ecological distribution of Arabidopsis and comments on physiological variations. *In* Arabidopsis research international symposium, Göttingen, pp 37-45

Rayle DL, Cleland R (1977) Control of plant cell enlargement by hydrogen ions. Curr Top Dev Biol **11:** 187-214

Reisser W (1988) Signals in the *Paramecium bursari-Chlorella* sp. association. S Scannerini, D Smith, P Bonfante-Fasolo, V Gianinazzi-Pearson, *In* Cell to Cell Signals in Plant, Animal and Microbial Symbiosis. Springer Verlag, Berlin, pp 281-296

Ricard J, Nari J (1966) Contribution to the study of mechanisms of indolylacetic acid degradation by horseradish peroxidase. Biochim Biophys Acta **113:** 57-70

Röbbelen G (1965) The Laibach standard collection of natural races. *Arabidopsis* Info Serv **2:** 36-47

Rohde P, Hincha DK, Heyer AG (2004) Heterosis in the freezing tolerance of crosses between two *Arabidopsis thaliana* accessions (Columbia-0 and C24) that show differences in non-acclimated and acclimated freezing tolerance. Plant J **38:** 790-799

Ruiz-Dueñas FJ, Camarero S, Perez-Boada M, Martinez MJ, Martinez AT (2001) A new versatile peroxidase from *Pleurotus*. Biochem Soc Trans **29:** 116-122

Ryals J, Neuenschwander U, Willits M, Molina A, Steiner HY et al. (1996) Systemic acquired resistance. Plant Cell **8:** 1809-18189

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd Ed Vol

Sanderson MJ (2003) Molecular data from 27 proteins do not support a Precambrian origin of land plants. Am J Bot **90:** 954-956

Saran M, Bors W (1994) Signalling by $^{\cdot}O_2^-$ and $NO^{\cdot}$: how far can either radical, or any specific reaction product, transmit a message under in vivo conditions? Chem Biol Interact **90:** 35-45

Sarkar G, Sommer SS (1990) The "megaprimer" method of site-directed mutagenesis. Biotechniques **8:** 404-407

Savard L, Li P, Strauss SH, Chase MW, Michaud M, et al. (1994) Chloroplast and nuclear gene sequences indicate late Pennsylvanian time for the last common ancestor of extant seed plants. Proc Natl Acad Sci U S A **91:** 5163-5167

Savitsky PA, Gazaryan IG, Tishkov VI, Lagrimini LM, Ruzgas T, et al. (1999) Oxidation of indole-3-acetic acid by dioxygen catalysed by plant peroxidases: specificity for the enzyme structure. Biochem J **340:** 579-583

Sawa S, Ohgishi M, Goda H, Higuchi K, Shimada Y, et al. (2002) The HAT2 gene, a member of the HD-Zip gene family, isolated as an auxin inducible gene by DNA microarray screening, affects auxin response in Arabidopsis. Plant J **32:** 1011-1022

Sawasaki Y, Inomata K, Yoshida K (1996) Trans-kingdom conjugation between *Agrobacterium tumefaciens* and *Saccharomyces cerevisiae*, a bacterium and a yeast. Plant Cell Physiol **37:** 103-106

Schmidt M, Schneider-Poetsch HA (2002) The evolution of gymnosperms redrawn by phytochrome genes: the Gnetatae appear at the base of the gymnosperms. J Mol Evol **54:** 715-724

Schopfer P, Liszkay A, Bechtold M, Frahry G, Wagner A (2002) Evidence that hydroxyl radicals mediate auxin-induced extension growth. Planta **214:** 821-828

Schwalm K, Aloni R, Langhans M, Heller W, Stich S, et al. (2003) Flavonoid-related regulation of auxin accumulation in *Agrobacterium tumefaciens*-induced plant tumors. Planta **218:** 163-178

Schweikert C, Liszkay A, Schopfer P (2000) Scission of polysaccharides by peroxidase-generated hydroxyl radicals. Phytochemistry **53:** 565-570

Setsukinai K, Urano Y, Kakinuma K, Majima HJ, Nagano T (2003) Development of novel fluorescence probes that can reliably detect reactive oxygen species and distinguish specific species. J Biol Chem **278:** 3170-3175

Shah K, Penel C, Gagnon J, Dunand C (2004) Purification and identification of a Ca^{2+}-pectate binding peroxidase from *Arabidopsis* leaves. Phytochemistry **65:** 307-312

Shannon LM, Kay E, Lew JY (1966) Peroxidase isozymes from horseradish roots. I. Isolation and physical properties. J Biol Chem **241:** 2166-2172

Sharbel TF, Haubold B, Mitchell-Olds T (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Mol Ecol **9:** 2109-2118

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, et al. (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot **53:** 1305-1319

Shin KS, Lee YJ (2000) Purification and characterization of a new member of the laccase family from the white-rot basidiomycete *Coriolus hirsutus*. Arch Biochem Biophys **384:** 109-115

Shinkle JR, Swoap SJ, Simon P, Jones RL (1992) Cell wall free space of *Cucumis* hypocotyls contains NAD and a blue light-regulated peroxidase activity. Plant Physiol **98:** 1336-1341

Sieberer T, Hauser MT, Seifert GJ, Luschnig C (2003) PROPORZ1, a putative Arabidopsis transcriptional adaptor protein, mediates auxin and cytokinin signals in the control of cell proliferation. Curr Biol **13:** 837-842

Siegel BZ, Siegel SM (1970) Anomalous substrate specificities among the algal peroxidases. Am J Bot **57:** 285-287

Siegel BZ, Siegel SM (1986) Differential substrate specificity among peroxidases: a functional view of phyletic relations. University of Geneva, *In* Molecular and physiological aspects of plant peroxidases, Geneva, Switzerland, pp 131-142

Siemering KR, Golbik R, Sever R, Haseloff J (1996) Mutations that suppress the thermosensitivity of green fluorescent protein. Curr Biol **6:** 1653-1663

Skoog F, Miller CO (1957) Chemical regulation of growth and organ formation in plant tissues cultured *in vitro*. Symp Soc Exp Biol **54:** 118-130

Skulachev VP (1998) Cytochrome c in the apoptotic and antioxidant cascades. FEBS Lett **423:** 275-280

Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. Annu Rev Biochem **72:** 449-479

Smulevich G, Jakopitsch C, Droghetti E, Obinger C (2006) Probing the structure and bifunctionality of catalase-peroxidase (KatG). J Inorg Biochem **100:** 568-585

Soltis DE, Soltis P, Chase MW, Mort M, Albach D, et al. (2000) Angiosperm phylogeny inferred from a combined data set of 18S rDNA, rbcL and atpB sequences. Bot J Linn Soc **133:** 381-461

Srinivasa Rao PS, Yamada Y, Leung KY (2003) A major catalase (KatB) that is required for resistance to H_2O_2 and phagocyte-mediated killing in *Edwardsiella tarda*. Microbiology **149:** 2635-2644

Stechmann A, Cavalier-Smith T (2003) The root of the eukaryote tree pinpointed. Curr Biol **13:** R665-666

Stewart P, Cullen D (1999) Organization and differential regulation of a cluster of lignin peroxidase genes of *Phanerochaete chrysosporium*. J Bacteriol **181:** 3427-3432

Strother PK (2000) Cryptospores: the origin and early evolution of the terrestrial flora. Paleontol Soc Pap **6:** 3-20

Sullivan JA, Deng XW (2003) From seed to seed: the role of photoreceptors in Arabidopsis development. Dev Biol **260:** 289-297

Taiz L, Zeiger E (2002) Auxin: the growth hormone. *In* Plant Physiology. Sinauer Associates, Inc., Sunderland (MA) USA, pp 438-441

Takeda H, Kotake T, Nakagawa N, Sakurai N, Nevins DJ (2003) Expression and function of cell wall-bound cationic peroxidase in asparagus somatic embryogenesis. Plant Physiol **131:** 1765-1774

Tanabe Y, Watanabe MM, Sugiyama J (2005) Evolutionary relationships among basal fungi (Chytridiomycota and Zygomycota): Insights from molecular phylogenetics. J Gen Appl Microbiol **51:** 267-276

Teixeira FK, Menezes-Benavente L, Margis R, Margis-Pinheiro M (2004) Analysis of the molecular evolutionary history of the ascorbate peroxidase gene family: inferences from the rice genome. J Mol Evol **59:** 761-770

Thimann KV, Sachs T (1966) The role of cytokinins in the “fasciation” disease caused by *Corynebacterium fascians*. Am J Bot **53:** 731-739

Thomas CM, Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Microbiol **3:** 711-721

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:** 4673-4680

Tian Q, Uhlir NJ, Reed JW (2002) Arabidopsis SHY2/IAA3 inhibits auxin-regulated gene expression. Plant Cell **14:** 301-319

Titball RW, Johansson A, Forsman M (2003) Will the enigma of *Francisella tularensis* virulence soon be solved? Trends Microbiol **11:** 118-123

Tjaden G, Edwards JW, Coruzzi GM (1995) Cis elements and trans-acting factors affecting regulation of a nonphotosynthetic light-regulated gene for chloroplast glutamine synthetase. Plant Physiol **108:** 1109-1117

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene **288:** 129-138

Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, et al. (1996) The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. Plant Cell **8:** 735-746

Tournaire C, Kushnir S, Bauw G, Inze D, Teyssendier de la Serve B, et al. (1996) A thiol protease and an anionic peroxidase are induced by lowering cytokinins during callus growth in *Petunia*. Plant Physiol **111:** 159-168

Tsutsumi Y, Haneda T, Nishida T (2001) Removal of estrogenic activities of bisphenol A and nonylphenol by oxidative enzymes from lignin-degrading basidiomycetes. Chemosphere **42:** 271-276

Tyurin MV, Desai SG, Lynd LR (2004) Electrotransformation of *Clostridium thermocellum*. Appl Environ Microbiol **70:** 883-890

Uknes S, Mauch-Mani B, Moyer M, Potter S, Williams S, et al. (1992) Acquired resistance in *Arabidopsis*. Plant Cell **4:** 645-656

Ulmasov T, Liu ZB, Hagen G, Guilfoyle TJ (1995) Composite structure of auxin response elements. Plant Cell **7:** 1611-1623

Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. Plant Cell **9:** 1963-1971

Ulmasov T, Hagen G, Guilfoyle TJ (1999) Activation and repression of transcription by auxin-response factors. Proc Natl Acad Sci U S A **96:** 5844-5849

Valério L, De Meyer M, Penel C, Dunand C (2004) Expression analysis of the *Arabidopsis* peroxidase multigenic family. Phytochemistry **65:** 1331-1342

van der Giezen M, Tovar J, Clark CG (2005) Mitochondrion-derived organelles in protists and fungi. Int Rev Cytol **244:** 175-225

van der Zwan C, Brodie SA, Campanella JJ (2000) The intraspecific phylogenetics of *Arabidopsis thaliana* in worldwide populations. Syst Bot **25:** 47-59

Vargas M, Kashefi K, Blunt-Harris EL, Lovley DR (1998) Microbiological evidence for Fe(III) reduction on early Earth. Nature **395:** 65-67

Veitch NC (2004) Horseradish peroxidase: a modern view of a classic enzyme. Phytochemistry **65:** 249-259

Veljovic-Jovanovic S, Kukavica B, Stevanovic B, Navari-Izzo F (2006) Senescence- and drought-related changes in peroxidase and superoxide dismutase isoforms in leaves of *Ramonda serbica*. J Exp Bot **57:** 1759-1768

Veselov SY, Kudoyarova GR, Egutkin NL, Gyuli-Zade VZ, Mustafina AR, et al. (1992) Modified solvent partitioning scheme providing increased specificity and rapidity of immunoassay for indole 3-acetic acid. Physiol Plant **86:** 93-96

Volpert R, Osswald W, Elstner EF (1995) Effects of cinnamic acid-derivatives on indole acetic-acid oxidation by peroxidase. Phytochemistry **38:** 19-22

von Arnim AG, Deng XW, Stacey MG (1998) Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene **221:** 35-43

Wada K, Tada T, Nakamura Y, Kinoshita T, Tamoi M et al. (2002) Crystallization and preliminary X-ray diffraction studies of catalase-peroxidase from *Synechococcus* PCC 7942. Acta Crystallogr D Biol Crystallogr **58:** 157-159

Waller RF, McConville MJ, McFadden GI (2004) More plastids in human parasites? Trends Parasitol **20:** 54-57

Wang H, Deng XW (2003) Dissecting the phytochrome A-dependent signaling network in higher plants. Trends Plant Sci **8:** 172-178

Weiler EW (1981) Radioimmunoassay for pmol-quantities of indole-3-acetic acid for use with highly stable [^{125}I]-and [^{3}H]IAA derivatives as radiotracers. Planta **153:** 319-325

Welinder KG (1991) Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. Biochim Biophys Acta **1080**: 215-220

Welinder KG (1992) Superfamily of plant, fungal and bacterial peroxidases. Curr Opin Struct Biol **2:** 388-393

Welinder KG, Justesen AF, Kjaersgard IV, Jensen RB, Rasmussen SK, et al. (2002) Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. Eur J Biochem **269:** 6063-6081

Wellman CH, Osterloff PL, Mohiuddin U (2003) Fragments of the earliest land plants. Nature **425:** 282-285

Wernegreen JJ (2002) Genome evolution in bacterial endosymbionts of insects. Nat Rev Genet **3:** 850-861

Wesenberg D, Kyriakides I, Agathos SN (2003) White-rot fungi and their enzymes for the treatment of industrial dye effluents. Biotechnol Adv **22:** 161-187

Wi SG, Singh AP, Lee KH, Kim YS (2005) The pattern of distribution of pectin, peroxidase and lignin in the middle lamella of secondary xylem fibres in alfalfa (*Medicago sativa*). Ann Bot (Lond) **95:** 863-868

Wier A, Dolan M, Grimaldi D, Guerrero R, Wagensberg J, et al. (2002) Spirochete and protist symbionts of a termite (*Mastotermes electrodominicus*) in Miocene amber. Proc Natl Acad Sci U S A **99:** 1410-1413

Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc Biol Sci **268:** 2211-2220

Wilkinson SR, Obado SO, Mauricio IL, Kelly JM (2002) *Trypanosoma cruzi* expresses a plant-like ascorbate-dependent hemoperoxidase localized to the endoplasmic reticulum. Proc Natl Acad Sci U S A **99:** 13453-13458

Woodward AW, Bartel B (2005) Auxin: regulation, action, and interaction. Ann Bot (Lond) **95:** 707-735

Worley CK, Zenser N, Ramos J, Rouse D, Leyser O, et al. (2000) Degradation of Aux/IAA proteins is essential for normal auxin signalling. Plant J **21:** 553-562

Yamada Y, Fujiwara T, Sato T, Igarashi N, Tanaka N (2002) The 2.0 A crystal structure of catalase-peroxidase from *Haloarcula marismortui*. Nat Struct Biol **9:** 691-695

Yamazaki H, Yamazaki I (1973) The reaction between indole 3-acetic acid and horseradish peroxidase. Arch Biochem Biophys **154:** 147-159

Yanovsky MJ, Casal JJ, Luppi JP (1997) The VLF loci, polymorphic between ecotypes Landsberg *erecta* and Columbia, dissect two branches of phytochrome A signal transduction that correspond to very-low-fluence and high-irradiance responses. Plant J **12:** 659-667

Yoshida K, Kaothien P, Matsui T, Kawaoka A, Shinmyo A (2003) Molecular biology and application of plant peroxidase genes. Appl Microbiol Biotechnol **60:** 665-670

Zamocky M, Janecek S, Koller F (2000) Common phylogeny of catalase-peroxidases and ascorbate peroxidases. Gene **256:** 169-182

Zamocky M (2004) Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. Eur J Biochem **271:** 3297-3309

Zhang R, Zhang CT (2005) Genomic islands in the *Corynebacterium efficiens* genome. Appl Environ Microbiol **71:** 3126-3130

Zhao Y, Dai X, Blackwell HE, Schreiber SL, Chory J (2003) SIR1, an upstream component in auxin signaling identified by chemical genetics. Science **301:** 1107-1110

Zhu Q, Dabi T, Lamb C (1995) TATA box and initiator functions in the accurate transcription of a plant minimal promoter in vitro. Plant Cell **7:** 1681-1689

Zuo J, Niu QW, Chua NH (2000) Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J **24:** 265-273

Appendix

A. Primers

Name	Sequence (5'-3')	Tm	Application/hybridisation target
5'PROCA Bis KpnI	GGGGTACCGCTTGGTTTGGTTTCCATTG	64.2	AtPrx33 promoter
5'UTR-P33	ATCTTCTCCGTTAAACCC	50.2	AtPrx33 promoter
3' promo CA NcoI	CATGCCATGGTTTTCACAAGGAC	57.8	AtPrx33 promoter
5'CA KpnI	GGGGTACCATGCAATTCTCTTCATCTTC	61.0	AtPrx33
3'CA XhoI	CCGTTCTCGAGACATAGAACTTACAAAGTC	61.0	AtPrx33
5'CA PstI	AACTGCAGATGCAATTCTCTTCATCTTC	58.1	AtPrx33
3'CA BamHI	CGCGGATCCACATAGAACTTACAAAGTC	61.1	AtPrx33
5'P33	GATAAATAAGCGGGATCTCA	51.3	Amplicon for AtPrx33
3'P33	GGATAAAACAGTCCAAAACC	51.3	Amplicon for AtPrx33
5'P34	TTGAAAGGTCGTCGAGGAGT	55.5	Amplicon for AtPrx34
5'P34 new	GGGTAAACAGAATCAAC	47.4	Amplicon for AtPrx34
3'P34	TCCAAGTGGACGATGTTGAAG	55.5	Amplicon for AtPrx34
CARTS	CTATTACTTCTTTCACTTGG	49.3	AtPrx33 and AtPrx34 (not specific)
CARTA	CCATTGCGAGGACATTG	52.2	AtPrx33 and AtPrx34 (not specific)
5'mut	ATCCTTCGTCTTCACTTCCATGAC	57.9	AtPrx33 and AtPrx34 (not specific)
3'mut	GTCATGGAAGTGAAGACGAAGGAT	57.9	reverse complement of 5'mut
5'S Bind SacI	TCCGAGCTCGGTAAAAATCAATG	56.0	AtPrx33 and AtPrx34 (not specific)
5'AS Bind SmaI	TCCCCCGGGGGTAAAAATCAATG	59.5	AtPrx33 and AtPrx34 (not specific)
3'AS Bind NcoI	CATGCCATGGTGTTTGTGTGCCATC	61.2	AtPrx33 and AtPrx34 (not specific)
3'S Bind HindIII	CCCAAGCTTTGTTTGTGTGCCATC	59.6	AtPrx33 and AtPrx34 (not specific)
5'MID$_{mut}$ p42	CATTTCCCTT**g**A**g**TT**g**A**t**GAGGCTGCA	61.1	megaprimer for p42 promoter Far$_{mut}$ construct
NcoI 3'CLOSE p42	CATGCCATGGCATCACACCTTTGCCTC	64.3	p42 promoter constructs
BamHI 5'FAR p42	CGGGATCCATATTATTTGTTATAGCC	56.3	p42 promoter constructs
BamHI 5'MID p42	CGGGATCCCTTTATTTAAGGAGGCTG	61.2	p42 promoter constructs
BamHI 5'SIDE p42	CGGGATCCGAGGCTGCATTGCTTGCT	66.1	p42 promoter constructs
5'GH3	TGCTTGAATGCGTCGGCGGCGCCC	68.1	GH3 promoter
3'GH3 end	GGTGAGCTATCACAAATTAATTTC	52.6	GH3 promoter
3'GH3	GCCGCCGACGCATTCAAGCA	61.5	GH3 promoter
5'PL15	GTTACGTCCTGTAGAAACCC	55.4	GUS gene
PL15 reverse	GGGTTTCTACAGGACGTAAC	55.4	GUS gene (reverse complement of 5'PL15)
3'GUS BamHI	GAGCTGGATCCCCTGGCGCCGAG	61.5	GUS gene
3'GUS-no stop BamHI	CTCTTAGGATCCAGAAACTTTATTGCC	58.2	nopaline synthase terminator
5'GFP (4T3)	CAGTGGAGAGGGTGAAG	54.0	GFP gene
3'GFP (4B6)	GCAGCTGTTACAAACTCAA	50.8	GFP gene
pCHF3*35S ter	ACAGTGGTCCCAAAGATGGA	55.4	35S promoter (pCAMBIA 1281Z, pZP222, pCHF3, pAVA393)
pCHF3 35S new	AGCATCGTGGAAAAAGAAGA	51.3	35S promoter (pCAMBIA 1281Z, pZP222, pCHF3, pAVA393)
CF35S	CCCACTATCCTTCGCAAGAC	57.5	35S promoter (pCAMBIA 1281Z, pZP222, pCHF3, pAVA393)
5'pCAMBIA 35S	GCAACGCTCTGTCATCGTTA	55.4	T-DNA left border (pCAMBIA 1281Z, pZP222, pCHF3)
pZP 3'LB	CCCCGAATTAATTCGGCGTT	55.4	T-DNA left border (pZP222, pCHF3)
CD5 RB	GCTCAGGATCCGATTGTCGTTTCCCGCCTT	66.5	T-DNA right border (pZP222, pCHF3)
M13 forward(-20)	GTAAAACGACGGCCAGT	52.0	pCAMBIA 1281Z, pZP222, pCHF3, pAVA393, pCRII, pGEM-T, pBluescrip
M13 reverse(-48)	GGATAACAATTTCACACAGGA	51.7	pCAMBIA 1281Z, pZP222, pCHF3, pAVA393, pCRII, pGEM-T, pBluescrip
T7	TAATACGACTCACTATAGGG	51.3	pET29a, pAVA393, pCRII, pGEM-T, pBluescript
T3	AATTAACCCTCACTAAAGGG	51.3	pAVA393, pBluescript
SP6	GATTTAGGTGACACTATAG	48.7	pCRII, pGEM-T
Cons	GNCTNCWYTTCCACGAYTGYTTYGT	59.6	amplification of class III peroxidases (not specific)
Sca	ANRATRTCRGCRCAWGANAC	53.4	amplification of class III peroxidases (not specific)
AP1	CCATCCTAATACGACTCACTATAGGGC	58.1	adaptor molecules (chapter 3)
AP2	ACTCACTATAGGGCTCGAGCGGC	63.1	adaptor molecules (chapter 3)
5'PR-1	ATGAATTTTACTGGCTATT	44.3	PR-1 gene (486 bp with 3'PR-1) involved in resistance against pathogens
3'PR-1	AGTATGGCTTCTCGTTCAC	52.9	PR-1 gene

Tm: melting temperature [°C] (G/C method).

Engineered restriction sites are underlined.

All primers were obtained from Microsynth (Switzerland).

N = A/T/C/G

W = A/T

Y = T/C

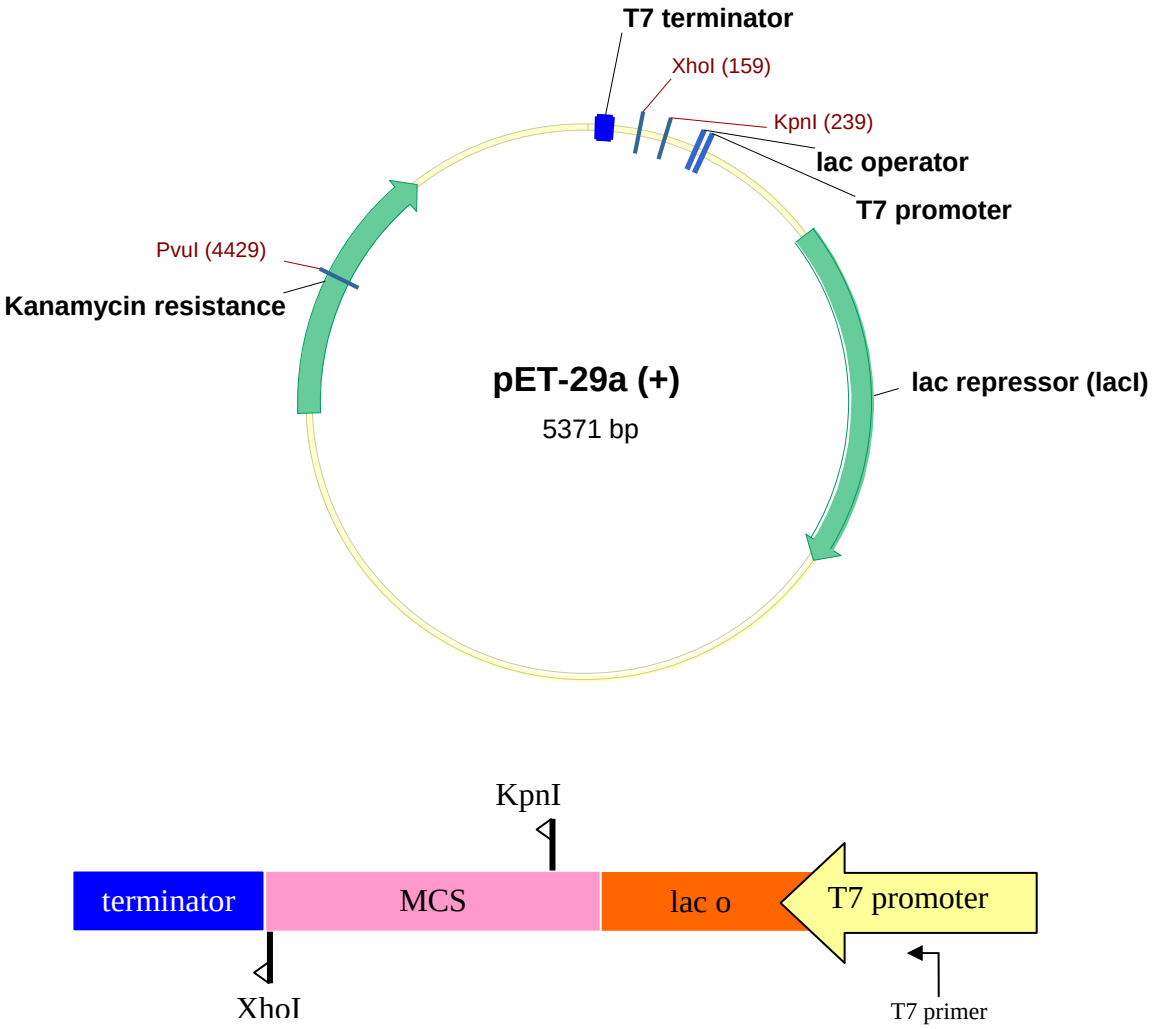
R = G/A

B. Plasmids

Single restriction sites are marked in dark red.

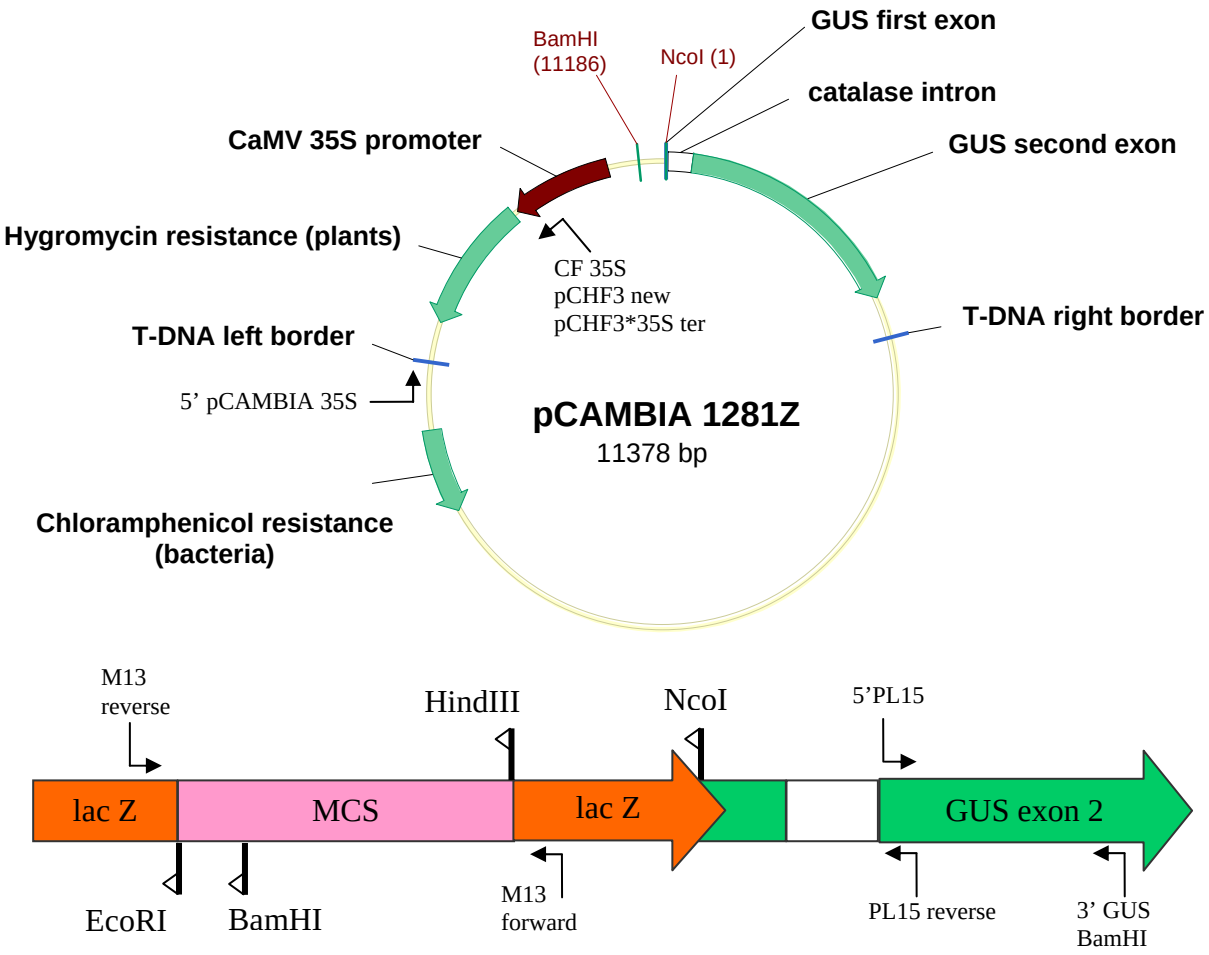
Restriction sites are denoted by flags in the vector fragment close-ups.

Primers positions are indicated by arrows.



pET-29a (+) has been used for expression of *AtPrx33* cDNA in *Escherichia coli* (see further comments in chapter 2). *AtPrx33* cDNA has been cloned with KpnI and XhoI.

Full sequence is available in the Novagen website (http://www.merckbiosciences.co.uk/docs/NDIS/69871-000.html).



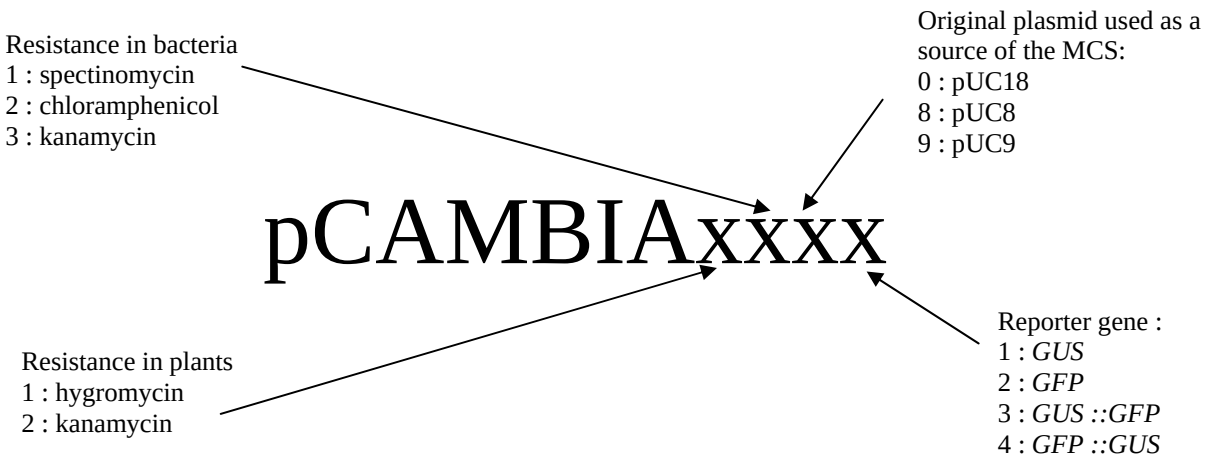
pCAMBIA 1281Z has been used as a binary vector where I introduced *AtPrx42* promoter constructs (see further comments in chapter 4) with BamHI and NcoI.

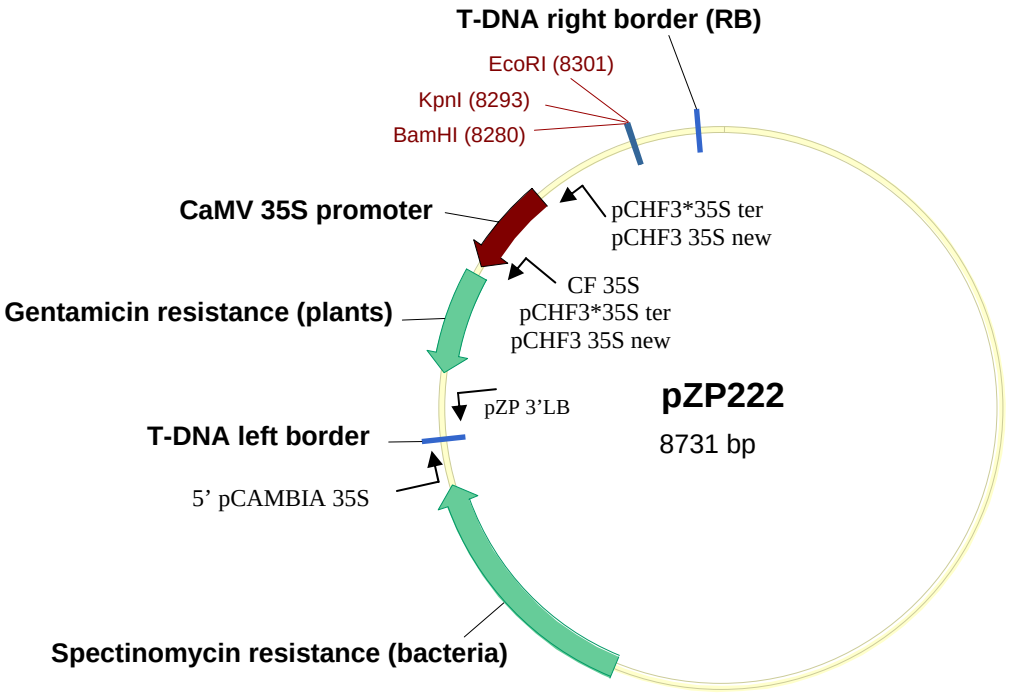
The plasmid is selectable for blue/white screening with the interrupted *lac Z* gene.

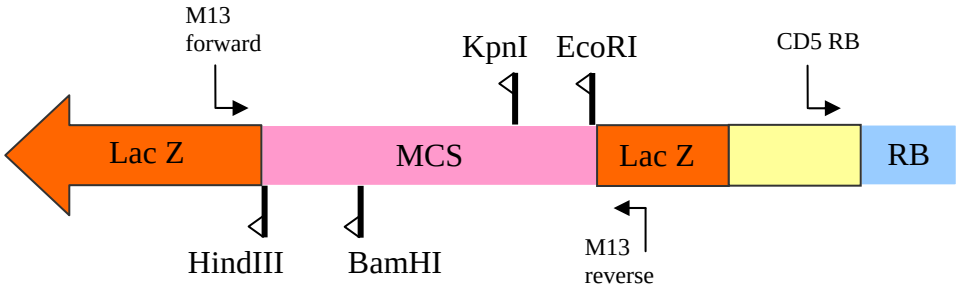
The sequence is available in GenBank, with the accession AF234294.

Full description of pCAMBIA vectors can be found in the pCAMBIA website: http://www.cambia.org/daisy/cambia/materials.html#dsy1105_Vectors

The nomenclature of pCAMBIA vectors has been established as follows (the "Z" in pCAMBIA 1281 Z indicates presence of the *lac Z* gene):





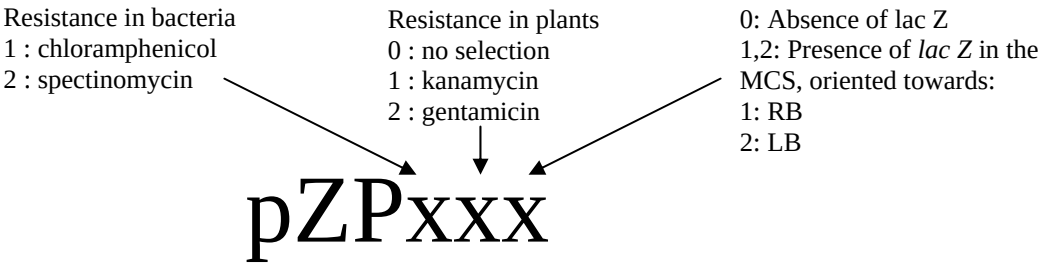


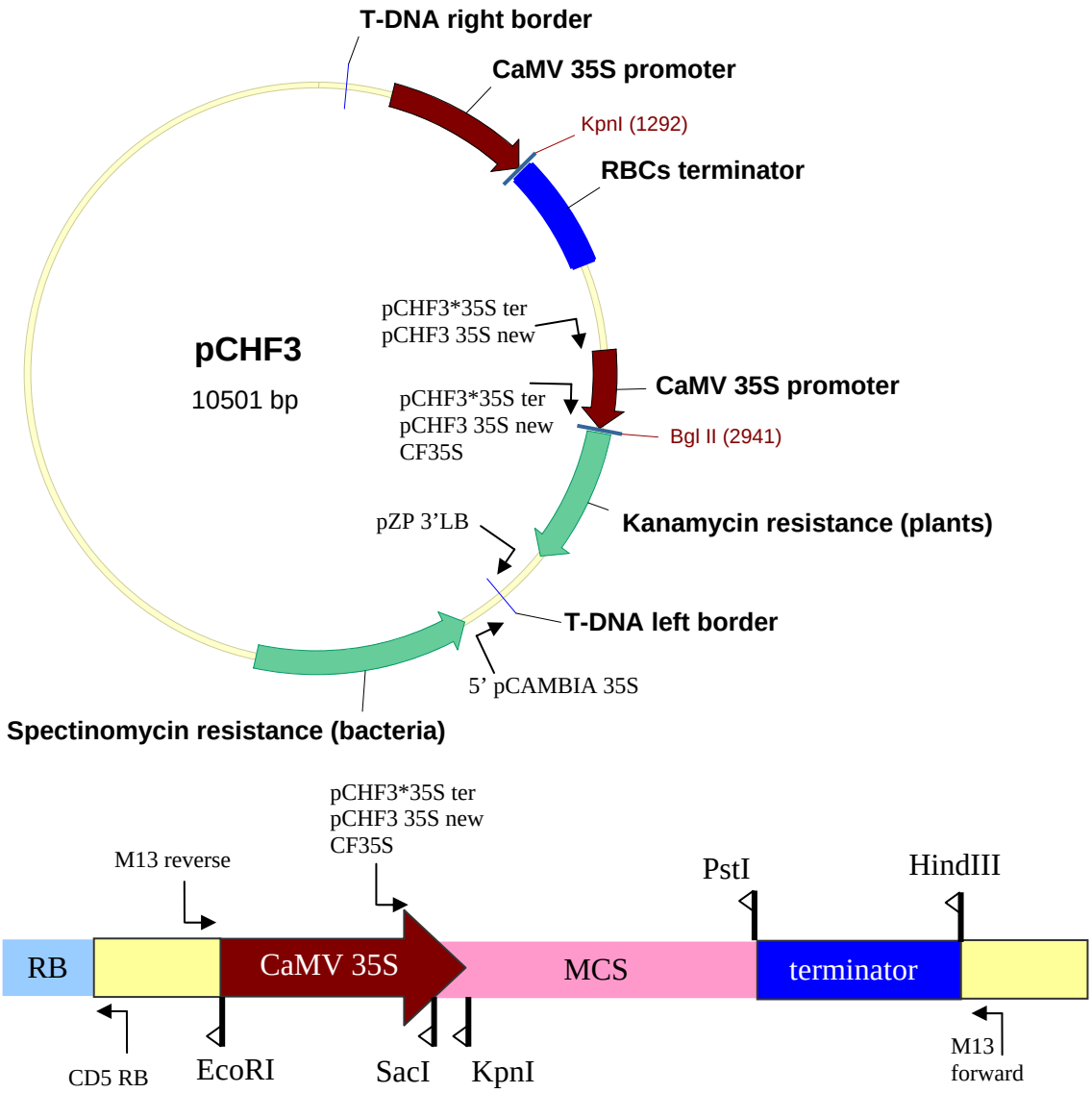
pZP222 has been used as a binary vector: GH3::*GUS* (EcoRI), AtPrx33::*AtPrx33*::*GFP* (KpnI/BamHI) and AtPrx33::*GFP* (KpnI/BamHI) were cloned into this plasmid.

The plasmid is selectable for blue/white screening with the interrupted *lac Z* gene.

Full sequence is available in GenBank, under the accession U10463, and details of the pZP series can be found in the literature (Hajdukiewicz et al., 1994).

The nomenclature of pZP vectors has been established as follows:





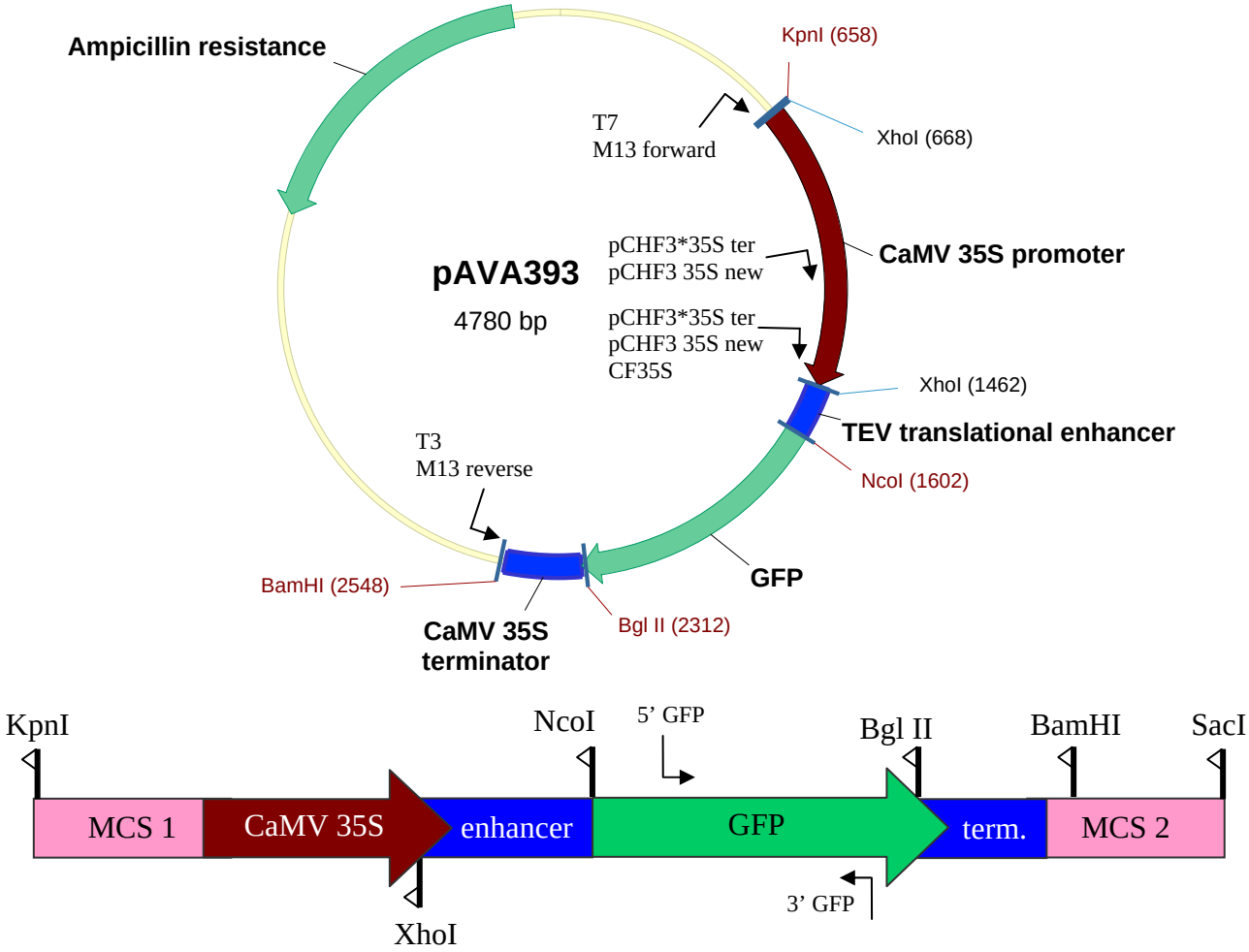
pCHF3 has been used as a binary vector, where I introduced the 35S promoter upstream from the *AtPrx33::GFP* fusion construct, with restriction enzymes KpnI and BglII.

This vector has been made by prof. Christian Fankhauser (University of Lausanne) starting from a pZP plasmid (probably a pZP212, not communicated). The main modifications are the addition of a CaMV 35S promoter and a terminator sequence on both sides of the pZP multiple cloning site. This plasmid allows insertion of the gene of interest downstream from the 35S promoter for constitutive expression.

From the HindIII restriction site until position 10501 (top of the plasmid), the sequence is identical to pZP212.

The full sequence of this plasmid is not available in GenBank and the making of the plasmid is not described in the literature. pCHF3 was a gift from prof. Fankhauser.

RBCs terminator: ribulose-1,5-bisphosphate carboxylase (RuBisCO) small subunit terminator (GenBank M21375).



pAVA 393 has been used to fuse the *AtPrx33* promoter to the *GFP* gene (KpnI/NcoI), as well as the AtPrx33::*AtPrx33* promoter-gene construct (KpnI/XhoI, then Klenow) and the *AtPrx33* gene (KpnI/XhoI, then Klenow).

pAVA393 contains the *mgfp5* gene (GenBank U87973) issued from mutagenesis studies on another GFP mutant gene, *mgfpa* (Siemering et al., 1996). *mgfpa* (Val163 → Ala and Ser175 → Gly) encodes a protein with improved folding characteristics at 37°C resulting in enhanced fluorescence at 37°C. *mgfp5* possesses additional mutations that allow efficient excitation (Ile197 → Thr) with blue light (commonly used in confocal microscopes), as well as enhanced fluorescence in *Arabidopsis thaliana* (elimination of potential intron splice sites by mutating the third position of several codons).

This vector is intended for cloning amino-terminal fusions to *GFP* via NcoI and carboxy-terminal residues via BglII. For promoter studies, the desired sequence should be inserted with XhoI (35S promoter will be excised): the tobacco etch virus (TEV) enhancer was shown to be effective downstream from promoter sequences (Carrington and Freed, 1990; Nicolaisen et al., 1992).

The sequence is not available in GenBank, but only upon request to the author (http://www.bio.utk.edu/vonarnim/vectors.html). However, apart from the *gfp* gene, it is very similar to pAVA319 (GenBank accession AF078810; based on *mgfp4*, GenBank U87624).

Full description of pAVA vectors can be found in the literature (von Arnim et al., 1998).

Other plasmids used

pGEM-T, **pGEM-T easy** (Promega) **and pCRII** (Invitrogen)
These plasmids are linearised and designed for direct ligation (no restriction enzyme needed) of PCR products. A thymidine is added to the 3'end of each extremity of the linearised plasmid, and hybridises to the adenines added by DNA polymerase to the 3'ends of PCR amplified products. The insertion region is located in the middle of a multiple cloning site. Both pGEM plasmids are almost identical, with the exception of two restriction sites in the MCS (EcoRI and NotI added to pGEM-T easy).

Presence of the plasmid containing the insert can be verified by selection on ampicillin and by amplification with T7/SP6 or M13 forward/M13 reverse primers, which flank the MCS. pCRII bears two resistance genes (ampicillin + kanamycin).

All three plasmids were used for intermediate cloning steps (*AtPrx33* cDNA in pET29a, AtPrx33::*GFP* construct). pGEM-T was also used for isolation of *Chara* peroxidase genes (chapter 3) and as a PCR template for synthesis of the M_m megaprimer (chapter 4).

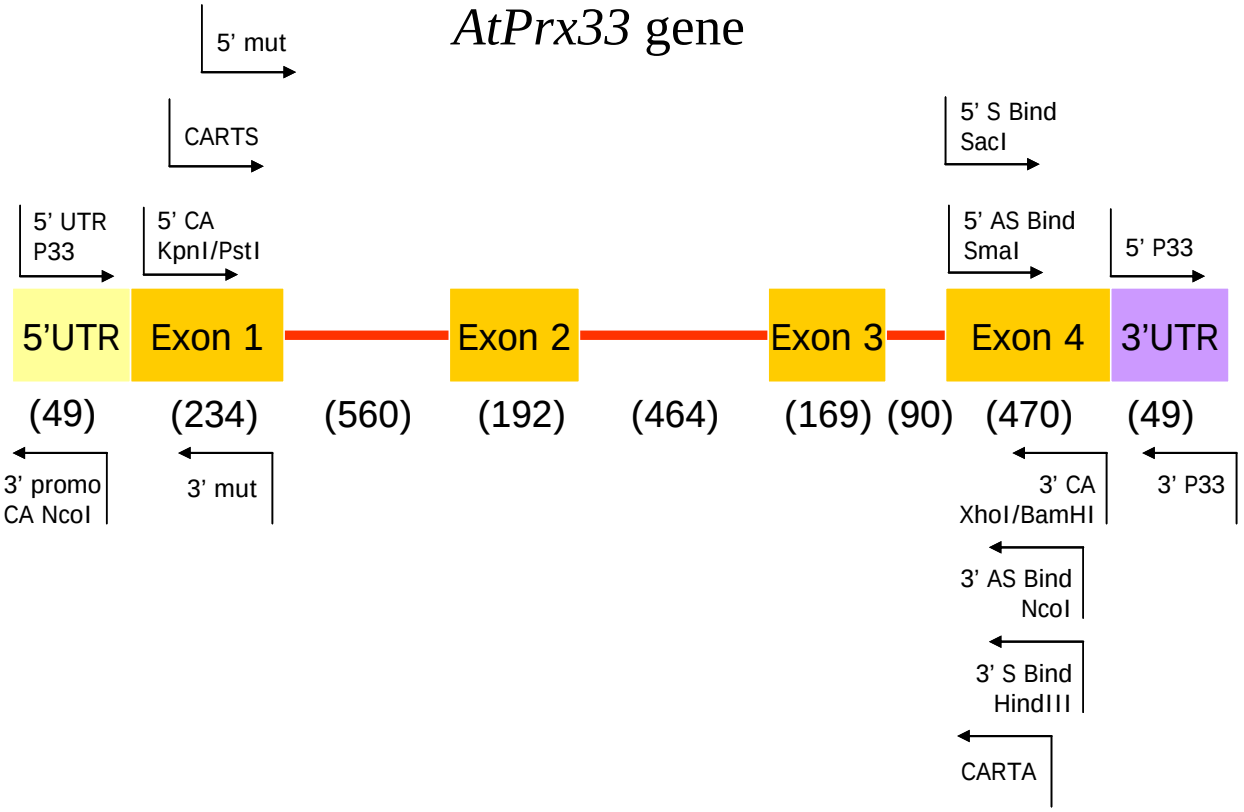
pBluescript II KS (+) (Fermentas Life Sciences)
This plasmid is designed for classical DNA cloning. It possesses a MCS interrupting a *lac Z* gene, and is hence selectable for blue/white colony screening. Its resistance in bacteria is ampicillin. Primers flanking the MCS are M13 forward and T7 on one side; M13 reverse and T3 on the other side.

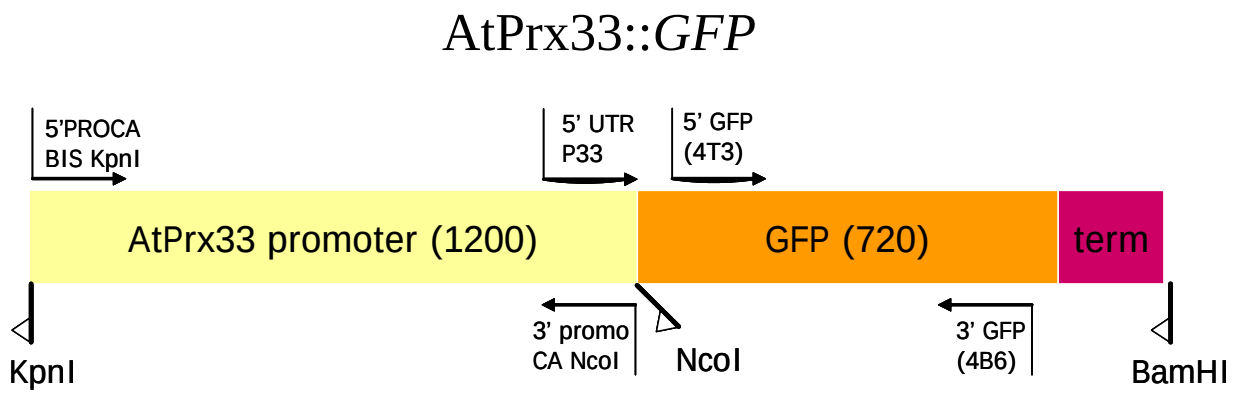
Like the two preceding plasmids, pBluescript has been used for intermediate cloning steps (AtPrx33::*AtPrx33*::*GFP* construct).

pVL1392 (PharMingen)
This plasmid is described in chapter 2. It is designed for protein expression in eukaryotic cells. It bears a resistance gene for ampicillin. Its MCS cannot be amplified with "classical" (SP6, T7, T3, forward, reverse) primers.

C. Detail of the main constructs used

Length of each segment (promoter, gene, other) is indicated into brackets.

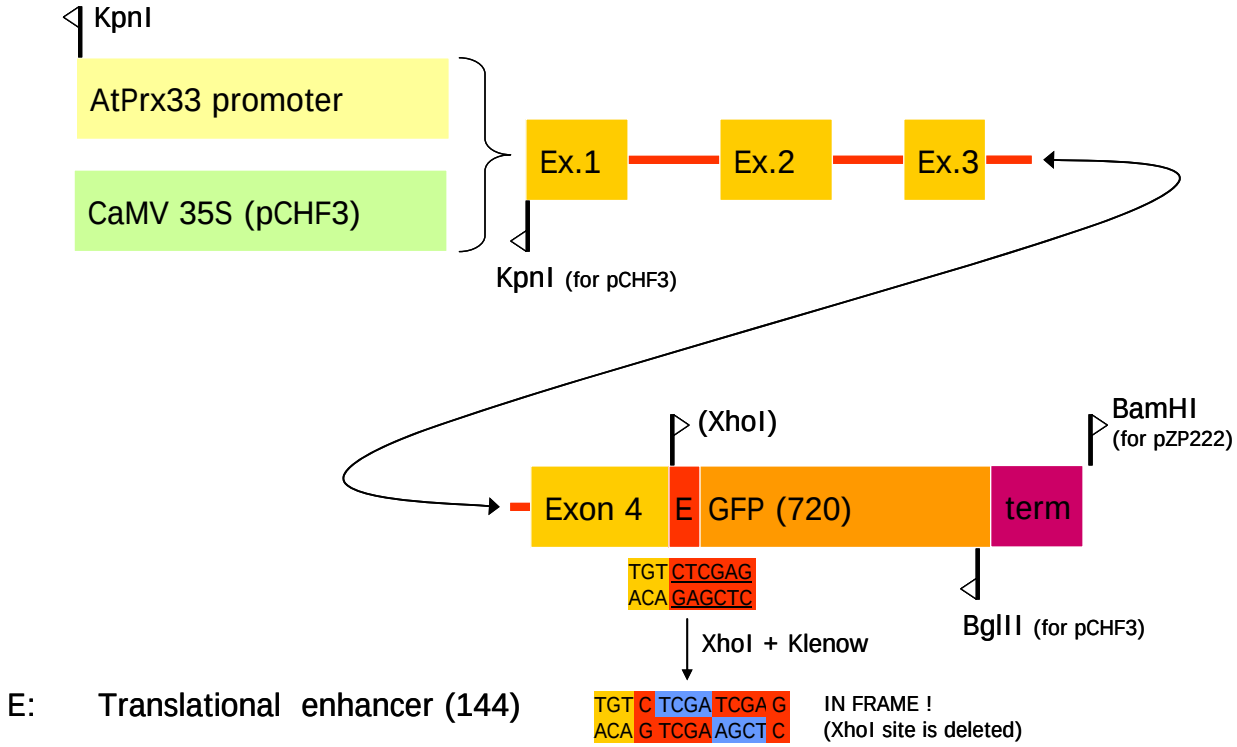




term: CaMV 35S terminator (211)

AtPrx33::*AtPrx33*::*GFP*

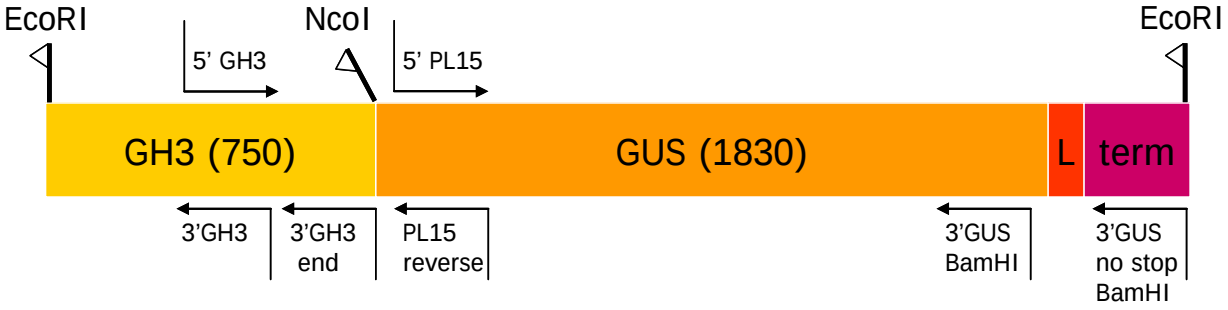
35S::*AtPrx33*::*GFP*



E: Translational enhancer (144)

term: CaMV 35S terminator (211)

GH3::*GUS*



L: linker (70)

term: nos terminator (253)

D. Sequences of the main constructs used

Position of primers (see previous chapter) is underlined.
Double-underline stands for two or more overlapping primers

AtPrx33 gene

5'UTR

atcttctccgttaaacccatttaaatagttttcattgtccttgtgaaaa

AtPrx33 gene

ATGCAATTCTCTTCATCTTCTATTACTTCTTTCACTTGGACAGTTTTAA
TCACAGTGGGATGTCTTATGCTTTGTGCGTCTTTCTCCGATGCTCAACT
TACCCCTACTTTTTACGACACTTCATGTCCTACCGTCACCAACATTGTA
AGAGATACCATTGTCAACGAGCTAAGATCGGACCCTCGTATCGCCGG
GAGCATCCTTCGTCTTCACTTCCATGACTGCTTTGTTAATgtaagatatcacttt
catatatctatagcgttatgaaattactgtacttaatctcttcatcttttaaatattgattaaatcagtgtaaaaaagattatag
ttgttgtcagcattagcaattgtttaggtgtactcttagttgctttcaatttaatgtcattttagcttttcttttatgtttcaaaatt
tggattctgaatgcaaattttaaaggaaaaagtaatttaacatcacgtcaaaattaatcaaacagacaaagaatcaatta
gttttggaatcacgtagtagtataatggaacttattaaatttccaccatattttggtaatataaatttttcttttactgaatttatt
ttagtgaattaaaatgggttagaacttagaagagttaattcctccttacattgttttccttaactcagttacttgcttcagag
gaacaaatggttttaatggatgaaaaaagttcaacgttaaactgttacctttggtaaagatgtaaactatggatgacatta
actaataatttgttaaataatattttggacttttgaaagGGTTGTGATGCTTCGATCTTGTTAGA
CAACACGACATCATTTCGAACAGAGAAAGATGCACTTGGAAATGCAA
ATTCAGCCCGAGGATTTCCAGTGATTGATAGAATGAAAGCTGCGGTG
GAGAGGGCATGCCCAAGAACCGTTTCATGCGCAGATATGCTCACCAT
TGCTGCTCAACAATCTGTCACTTTGgtatgctccattgtcccatattcttttattcattacacacatt
gttcgatataaatatgtcaagtagtcgctttacgtacatgttaaacatgacagtctaagttcgaaaataaaattgttggtat
attttatgtcttttagaatgtcacaaactcccaatatatgtttagaataaaatagacaagtattcaaaattaatcggaaaac
attcaaaacactcccgttgatatttttgataatttaatatcactgagaaaataaggtctcatcactatgatttatgtacctttt
ttagttaacaaaatagaagaaaattccttttcatccaagaaagagaccatggaaattacctttgagttttgagtatataatc
aatagatttataaaacatgtttcagggtaataatattttctaagaaaagaattattattttgaatataatggttaattattatact
atagGCAGGAGGTCCTTCTTGGAAGGTTCCTTTAGGGAGAAGAGACAGC
TTACAAGCATTTCTAGATCTTGCTAACGCAAATCTTCCAGCTCCATTCT
TCACACTTCCACAGCTTAAAGCCAACTTCAAAAATGTTGGCCTCGATC
GTCCTTCTGATCTTGTTGCGCTCTCCGgtaattattaaaaagaaatgaaaccttgtttgatata
gttgtaatttagcaagttaatcgttgatctaattatggtcttccttttgttagGGGCTCACACATTTGGTA
AAAATCAATGTCGATTCATTATGGACAGATTATACAACTTTAGCAACA
CTGGATTACCTGACCCTACACTCAACACTACTTACCTCCAAACTCTTC
GTGGTCAATGTCCTCGCAATGGTAATCAAAGCGTCTTAGTGGATTTCG
ATCTGCGTACGCCTTTGGTTTTCGACAACAAATACTATGTGAATCTTA
AAGAGCAAAAAGGTCTTATCCAGAGCGACCAAGAGTTGTTCTCTAGC
CCCAATGCCACTGACACAATCCCCTTGGTGAGAGCATATGCTGATGGC
ACACAAACAATTCTTCAATGCATTCGTGGAGGCAATGAATAGGATGGG
AAATATTACACCAACTACAGGAACTCAAGGACAAATCAGGTTGAATT
GTAGAGTGGTGAACTCCAACTCTCTACTCCATGATGTGGTGGATATCG
TTGACTTTGTAAGTTCTATGTGA

3' UTR

gaatagttgactcaatatctggcgaccagactatatgttaagataaataagcgggatctcaagttgttacttgagcagg
agcagattgtattggtgtttatagtatccaagtagttctctgtttttttttatattggcctttgatagcgttttgttaatcggtct
aagcttctatggttttggactgttttatcct

AtPrx33::*GFP*

AtPrx33 promoter

ggtaccgcttggtttggtttccattgtttgaatttattgggtatacatattagaccaattagaattccacttgtaaatgcatat
tctggaacaataatgttgactatactaataatttttttcttgggtctaaaacatttagtatacttttcttgaattttatgtaaagtta
ttctgtttttaggacttttatcctcactgaattatgttgtgttaaagtagtagcttatatggtcacatatctcacacgtgtttcc
atttttttttctatctaagttttcggtttttacattgtatacaaaaataaatggcatgccttgggccttggccactgattcacatg
atcacgtagagtttgtgtgtgattcaaatcctagatgagaagatcatccaattgatctaagatgtgtctttcttatatcaca
cgttttattttaataatactcttcagttgggtttgattccatttatgatttatctatgaaaacagaatatgtcctcatgctgatta
attcgtttgtggttgcattttaccatctgtagttttttttgtttactgacaaacaactaagattttatgaattcatttgatgataaa
tagttaatcaaatatattattatgagtgactcgatggcgaaagctggtattatcttatggctgctaaagtctacttttagaa
gaagaaaaaaaaaagggtctgttgcattaccctgttgttgcgactctcctacaataactctcacgtggcccgtttcttggtt
tttgccctattttcacaaaagtaaacagagattttcactagcaaattgggcccaattggagacaacttcatgtggtttgaa
tcgtatatttcagatagattagacatctagtcttctttgatgtgcggtgaggatcttaagtcgctataacgttgtgaaaaaa
gaaaagaaaaagtcactataacatagatcgcaactaaacaaataaaggacactatacatgtacgtaataatattcattt
attttcaacgacaatgttccctagggtggtaacccaagtatttttcatagatattatacttctagcgaaaagccatcaattg
agaaaaaaggcctccacgtttgacccgttttgcctataaaaggattctttaagatatggaccttactc**atcttctccgtt**
aaacccatttaaatagttttcattgtccttgtgaaaa

GFP

ccatgggtaaaggagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaatttt
ctgtcagtggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacc
tgttccctggccaacacttgtcactactttctcttatggtgttcaatgcttttcaagatacccagatcatatgaagcggcac
gacttcttcaagagcgccatgcctgagggatacgtgcaggagaggaccatcttcttcaaggacgacgggaactaca
agacacgtgctgaagtcaagtttgagggagacaccctcgtcaacaggatcgagcttaagggaatcgatttcaagga
ggacggaaacatcctcggccacaagttggaatacaactacaactcccacaacgtatacatcatggcagacaaacaa
aagaatggaatcaaagttaacttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaa
caaaatactccaattggcgatggccctgtccttttaccagacaaccattacctgtccacacaatctgccctttcgaaag
atcccaacgaaaagagagaccacatggtccttcttgagtttgtaacagctgctgggattacacatggcatggatgaac
tatacagatctatctag

CaMV 35S terminator

agtccgcaaaaatcaccagtctctctctacaaatctatctctctctatttttctccagaataatgtgtgagtagttcccagat
aagggaattagggttcttatagggtttcgctcatgtgttgagcatataagaaacccttagtatgtatttgtatttgtaaaata
cttctatcaataaaatttctaattcctaaaaccaaaatccagtga

pAVA 393 vector

cctgcagcccgggggatcc

Restriction sites : BamHI, KpnI, NcoI

Bolded: 5’UTR

AtPrx33::*GFP* (from exon 4)

AtPrx33 exon 4

GGGCTCACACATTTGGTAAAAATCAATGTCGATTCATTATGGACAGAT
TATACAACTTTAGCAACACTGGATTACCTGACCCTACACTCAACACTA
CTTACCTCCAAACTCTTCGTGGTCAATGTCCTCGCAATGGTAATCAAA
GCGTCTTAGTGGATTTCGATCTGCGTACGCCTTTGGTTTTCGACAACA
AATACTATGTGAATCTTAAAGAGCAAAAAGGTCTTATCCAGAGCGAC
CAAGAGTTGTTCTCTAGCCCCAATGCCACTGACACAATCCCCTTGGTG
AGAGCATATGCTGATGGCACACAAACATTCTTCAATGCATTCGTGGAG
GCAATGAATAGGATGGGAAATATTACACCAACTACAGGAACTCAAGG
ACAAATCAGGTTGAATTGTAGAGTGGTGAACTCCAACTCTCTACTCCA
TGATGTGGTGGATATCGTTGACTTTGTAAGTTCTATGT

Enhancer

ctcgatcgagaattctcaacacaacatatacaaaacaaacgaatctcaagcaatcaagcattctacttctattgcagca
atttaaatcatttcttttaaagcaaaagcaattttctgaaaattttcaccatttacgaacgatag

GFP

ccatgggtaaaggagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaatttt
ctgtcagtggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacc
tgttccctggccaacacttgtcactactttctcttatggtgttcaatgcttttcaagatacccagatcatatgaagcggcac
gacttcttcaagagcgccatgcctgagggatacgtgcaggagaggaccatcttcttcaaggacgacgggaactaca
agacacgtgctgaagtcaagtttgagggagacaccctcgtcaacaggatcgagcttaagggaatcgatttcaagga
ggacggaaacatcctcggccacaagttggaatacaactacaactcccacaacgtatacatcatggcagacaaacaa
aagaatggaatcaaagttaacttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaa
caaaatactccaattggcgatggccctgtccttttaccagacaaccattacctgtccacacaatctgccctttcgaaag
atcccaacgaaaagagagaccacatggtccttcttgagtttgtaacagctgctgggattacacatggcatggatgaac
tatacagatctatctag

CaMV 35S terminator

agtccgcaaaaatcaccagtctctctctacaaatctatctctctctatttttctccagaataatgtgtgagtagttcccagat
aagggaattagggttcttatagggtttcgctcatgtgttgagcatataagaaacccttagtatgtatttgtatttgtaaaata
cttctatcaataaaatttctaattcctaaaaccaaaatccagtga

pAVA 393 vector

cctgcagcccgggggatcc

Restriction sites : BamHI, BglII

Shaded in grey: nucleotides added by DNA polymerase I large (Klenow) fragment.

Note: BglII is used for cloning into pCHF3 (CaMV35S::*AtPrx33*::*GFP*)

BamHI is used for cloning into pZP222 (AtPrx33::*AtPrx33*::*GFP*)

GH3::*GUS*

GH3 promoter

gaattcacgaataaagaaaaattaaaagtctcaacaaatgtagtaagagggcaaaaataggctgtaataacttgcaaa
gtgtgcagtgaagttttcttcgtactacgtagaaacttctcagttctttctcacatttctgcccacagggatttggatttcgt
gtattgacgcagttataccatcattaatcttatccttcaatttttataaaattaataaaataaataaaaaattaattaagcttcc
gatcttgactgcctgcttgaatgcgtcggcggcgcccattagtttctcatgccaacacaccctatacgcctaattttgcc
cgagtattactatattggggagaacttttgctgacgtggcgacacatctgggacccacatgtcggccaccatgcacca
tccctggccctcgtgtctcctcaataagctacacaatttgaaacatacacgcaatcctttgtctcaataagttccactcag
gtactgttttctcccgcaaccatgacgtaattctgtaaatcacatgtttcatgctcccaattattttccgcttctataaatacc
tctcccatttcgcaacttttctccatccatactcatcc**acttcttgaaccgtgccttaactaaactagagctagaatta**
gagttagctaccttgcctaattcacaaacgcgtccctctacggctctacctattagctatctttttgtgctgtgat
tgaaattaatttgtgatagctcacc

GUS gene

atggtaagcttagcgggccccgtccgtcctgtagaaaccccaacccgtgaaatcaaaaaactcgacggcctgtggg
cattcagtctggatcgcgaaaactgtggaattgatcagcgttggtgggaaagcgcgttacaagaaagccgggcaatt
gctgtgccaggcagttttaacgatcagttcgccgatgcagatattcgtaattatgcgggcaacgtctggtatcagcgc
gaagtctttataccgaaaggttgggcaggccagcgtatcgtgctgcgtttcgatgcggtcactcattacggcaaagtg
tgggtcaataatcaggaagtgatggagcatcagggcggctatacgccatttgaagccgatgtcacgccgtatgttatt
gccgggaaaagtgtacgtatcaccgtttgtgtgaacaacgaactgaactggcagactatcccgccgggaatggtga
ttaccgacgaaaacggcaagaaaaagcagtcttacttccatgatttctttaactatgccggaatccatcgcagcgtaat
gctctacaccacgccgaacacctgggtggacgatatcaccgtggtgacgcatgtcgcgcaagactgtaaccacgc
gtctgttgactggcaggtggtggccaatggtgatgtcagcgttgaactgcgtgatgcggatcaacaggtggttgcaa
ctggacaaggcactagcgggactttgcaagtggtgaatccgcacctctggcaaccgggtgaaggttatctctatgaa
ctgtgcgtcacagccaaaagccagacagagtgtgatatctacccgcttcgcgtcggcatccggtcagtggcagtga
agggcgaacagttcctgattaaccacaaaccgttctactttactggctttggtcgtcatgaagatgcggacttgcgtgg
caaaggattcgataacgtgctgatggtgcacgaccacgcattaatggactggattggggccaactcctaccgtacct
cgcattacccttacgctgaagagatgctcgactgggcagatgaacatggcatcgtggtgattgatgaaactgctgctg
tcggctttaacctctctttaggcattggtttcgaagcgggcaacaagccgaaagaactgtacagcgaagaggcagtc
aacggggaaactcagcaagcgcacttacaggcgattaaagagctgatagcgcgtgacaaaaaccacccaagcgt
ggtgatgtggagtattgccaacgaaccggatacccgtccgcaaggtgcacgggaatatttcgcgccactggcgga
agcaacgcgtaaactcgacccgacgcgtccgatcacctgcgtcaatgtaatgttctgcgacgctcacaccgatacca
tcagcgatctctttgatgtgctgtgcctgaaccgttattacggatggtatgtccaaagcggcgatttggaaacggcaga
gaaggtactggaaaaagaacttctggcctggcaggagaaactgcatcagccgattatcatcaccgaatacggcgtg
gatacgttagccgggctgcactcaatgtacaccgacatgtggagtgaagagtatcagtgtgcatggctggatatgtat
caccgcgtctttgatcgcgtcagcgccgtcgtcggtgaacaggtatggaatttcgccgattttgcgacctcgcaaggc
atattgcgcgttggcggtaacaagaaagggatcttcactcgcgaccgcaaaccgaagtcggcggcttttctgctgca
aaaacgctggactggcatgaacttcggtgaaaaaccgcagcagggaggcaaacaatga

linker

atcaacaactctcctggcgcaccatcgtcggctacagcctcgggaattgctaccgagctcgaatttcccc

nopaline synthase terminator

gatcgttcaaacatttggcaataaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttg
aattacgttaagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtcccgcaatt
atacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgcgcggtgtcatctatgtta
ctagatcggattc

Restriction sites : EcoRI, NcoI

Bolded: 5’UTR

E. Common protocols and stock solutions

E.1 Polymerase chain reaction (PCR)

E.1.1 Protocol

When not stated otherwise, PCR reactions were performed as follows:

95°C 1’00 (partially purified or pure DNA sample)
5’00 (bacteria)
then

95°C 30”
55°C 30”
72°C 30” per 500bp
} 30-40 cycles (empirical)

final elongation at 72°C (5’00)

E.1.2 Reaction mixture

	stock solution	volume (µl)	final concentration
PCR buffer	10x	2	1x
5'primer	20µM	0.4	0.4µM
3'primer	20µM	0.4	0.4µM
$MgCl_2$	100mM	0.4	2mM
dNTPs	10mM	0.4	0.2mM
DNA sample		0.1-2 or bacteria	
Taq polymerase		0.3	
H_2O		complete to 20µl	

PCR 10x buffer

Tris 200mM
KCl 500mM
$MgCl_2$ 15mM

E.2 Bacteria transformation and culture conditions

E.2.1 Transformation

E.2.1.a Escherichia coli

a. Thaw on ice an aliquot (100-200µl) of competent *E. coli* (usually DH5 α strain, stored at -70°C).
b. Inject 1 to 5µl of plasmid (20-200ng/µl), and mix gently.
c. Incubate on ice during 20-30'.
d. Thermal shock: 42°C during 45"
4°C (ice) during 2'.
e. Add 800µl LB medium, and grow under strong agitation at 37°C during 1 hour.
f. Spin 5' at 3'000 rpm (Heraeus "Pico" Biofuge).
g. Discard supernatant (leave ~100µl in the tube).
h. Resuspend pellet and plate on solid LB agar Petri dish (+antibiotics).
i. Grow at 37°C overnight.

E.2.1.b Agrobacterium tumefaciens

a. Thaw on ice an aliquot (100µl) of competent *A. tumefaciens* (usually ASE strain, stored at -70°C).
b. Inject 1 to 5µl of purified plasmid (50-200ng/µl), and mix gently.
c. Thermal shock: liquid nitrogen during 5'
thaw at 37°C.
d. Add 800µl LB medium, and grow under gentle agitation at 28°C during 2 hours.
e. Spin 5' at 3'000 rpm (Heraeus "Pico" Biofuge).
f. Discard supernatant (leave ~100µl in the tube).
g. Resuspend pellet and plate on solid LB agar Petri dish (+antibiotics).
h. Grow at 28°C during 2 days.

E.2.2 Blue/white colony selection

This selection is possible for plasmids pCAMBIA 1281Z, pZP222, pBluescript KS (+), pGEM-T and pCRII. The principle is the following: the MCS is included in a *lac Z* (β-galactosidase) gene that, upon insertion of a foreign DNA fragment, is rendered inactive. Addition of the inducer IPTG and the colour-changing substrate X-Gal (converted to a blue product by lac Z) allows selecting for positive (white) colonies.

a. When plating transformed *E.coli* cultures, immediately add 35µl X-Gal (20mg/ml in dimethylformamide) and 16µl IPTG (50mg/ml), and streak together with the bacteria.
b. The next day, put plates at 4°C for 1 hour: at low temperatures, blue colour is enhanced.

E.2.3 Plasmid purification (*E.coli*)

a. Screen overnight cultures by PCR.
b. Pick positive colonies and grow them overnight at 37°C in 5ml liquid LB (+antibiotics), under strong shaking.
c. Take 500µl of the overnight culture and keep them at 4°C for glycerol stocks.
d. Centrifuge culture at 3'000rpm during 5 minutes, 4°C (ALC 4239R centrifuge, fixed angle rotor A-18C, i.e. 750g).
e. Discard supernatant and purify plasmid from the bacterial pellet with the Sigma GenElute Plasmid Miniprep kit (PLN70), according to manufacturer's instructions.
f. Usually, the plasmid is eluted in 50µl distilled water and stored at -20°C. Common yields are in the range of 50 to 200ng/µl.
   Note that water is preferable to the Sigma elution solution if further sequencing reactions are needed: the EDTA contained in the latter solution prevents correct activity of the DNA polymerase during sequencing.

E.2.4 Glycerol stocks (*Agrobacterium* and *E.coli*)

Mix 850µl of bacterial liquid culture with 150µl sterile glycerol, and freeze in liquid nitrogen.

Glycerol stocks can be stored at -70°C for several years. For further use, they should never be thawed again (just scratch an aliquot and grow in liquid LB).

E.2.5 Antibiotics

Concentrations used for selection in bacteria:

Ampicillin (pBluescript II KS, pGEM-T, pCRII, pVL1392):	100µg/ml
Chloramphenicol (*E.coli* BL21):	34µg/ml
Chloramphenicol (pCAMBIA 1281Z) :	25µg/ml (*E.coli*)
	10µg/ml (*Agrobacterium*)*
Gentamicin sulfate (*Agrobacterium*)*:	20µg/ml
Kanamycin sulfate (pCRII, pAVA393, *Agrobacterium* ASE):	50µg/ml
Spectinomycin (pZP222, pCHF3):	100µg/ml

* strains LBA4404 and GV3101/2

E.2.6 Lysogeny broth ("Luria-Bertani") medium (LB) (Bertani, 1951)

[liquid]

10g NaCl
10g tryptone
5g yeast extract
complete to 1 l. with water
Autoclave

[solid]

Add 15g of Bacto agar per liter of liquid LB

E.3 Plant growth conditions

E.3.1 On agar

a. Sterilise seeds with bleach 2.4% and Triton X100 0.02% during 8 minutes.
b. Rinse at least three times with sterile water.
c. Sow on ½ MS agar plates.
d. Vernalise for at least two days at 4°C.
e. Transfer plates into the growth room.

Growth conditions:
Photoperiod : 16hrs. day
Temperature : 22±1 °C
Humidity : 60% ± 5

Lamps:

3 lamps of each:

Sylvania Luxline Plus
F58W/840 Cool white De Luxe
and
Sylvania Luxline Plus
F58W/830 Warm white De Luxe

Light intensity: 80 or 180 $\mu E/m^2/sec$

Light intensity was measured with a
radiometer/photometer International light IL1400A and a filter White # 9540

E.3.2 On soil

Seedlings were transferred on soil (Substrate 1, Klasmann-Deilmann GmbH, Germany) when they were at least 1 week-old.
Same growth conditions as for agar (same growth chamber).
Light intensity: >180 $\mu E/m^2/sec$

E.3.3 Murashige-Skoog medium (MS) (Murashige and Skoog, 1962)

For one litre MS (1/2 MS: for two litres)

[Sigma M9274]

Macronutrients (mg)	
Ammonium nitrate (NH_4NO_3)	1’650
Boric acid (H_3BO_3)	6.2
Calcium chloride ($CaCl_2*H_2O$)	332.2
Colbalt chloride ($CoCl_2*6H_2O$)	0.025
Cupric Sulfate ($CuSO_4*5H_2O$)	0.025
Ferrous sulfate ($FeSO_4*7H_2O$)	27.8
Magnesium sulfate ($MgSO_4*7H_2O$)	180.7
Manganese sulfate ($MnSO_4*4H_2O$)	16.9
Potassium iodide (KI)	0.83
Potassium nitrate (KNO_3)	1’900
Potassium phosphate (KH_2PO_4)	170
Sodium molybdate ($Na_2MoO_4*2H_2O$)	0.25
Zinc sulfate ($ZnSO_4*7H_2O$)	8.6
$Na_2EDTA*2H_2O$	37.26

Organic Additives (mg)	
myo-inositol	100
Glycine	2
Sucrose	30’000
Agar	8’000
B vitamins (mg)	
Nicotinic Acid	0.5
Pyridoxine*HCl	0.5
Thiamine*HCl	0.1

E.4 Transformation of Arabidopsis *plants with* Agrobacterium *and selection of transgenic seedlings*

E.4.1 Transformation

a. Scratch an aliquot of glycerol stock (or colony from a LB Petri dish), and grow during 2 days in 25ml LB (+antibiotics) under strong shaking at 28°C.
b. Spin culture (keep an aliquot of culture for further rounds of transformation) in a Falcon 50 tube.
c. Discard supernatant and resuspend pellet in a freshly prepared solution of sucrose 5% and Silwet L-77® 0.1 ‰ (Lehle seeds, USA) to an optical density (600nm) of 0.8.
d. Spray bacteria on adult *Arabidopsis* plants (trim flowers and siliques before transformation). Usually, 40ml of bacterial solution are sufficient for six healthy plants.
e. Lay down plants and cover them for a few hours.
f. Put plant back to the growth chamber.
g. Repeat transformation twice, every other day.

E.4.2 Selection

a. Collect seeds of transformed plants (T0 generation), sterilise them and resuspend them in agar 0.1% (+antibiotics). For a dry seed pellet of less than 150µl, resuspend in 4ml agar and then pour onto a small (9cm diameter) Petri. For dry seed pellet of 150µl to 300µl, resuspend in 10ml agar and then pour onto a big Petri (14cm).
b. Pick up resistant seedlings and grow them on soil.
c. Collect the next generation of seeds (T1) and check their ratio on antibiotics selection: if one insert is present, then the ratio should be 3 resistants for 1 sensitive.
d. Grow resistant T1 seedlings, and collect T2 seeds separately from each plant.
e. Among these plants, one out of three should be homozygous for the transgene: sow each line separately and select for 100% resistant seedlings.

E.4.3 Antibiotics

Concentrations used for selection in plants:

Gentamicin sulfate (pZP222): 75µg/ml
Kanamycin sulfate (pCHF3): 50µg/ml
Hygromycin B (pCAMBIA 1281Z): 25µg/ml

Note that selection on hygromycin works better if performed on dark-grown seedlings: non-transformed seedlings die rapidly (within one week), whereas in light conditions, differences between wild types and transgenic plants appear only after 5 weeks.

F. Abbreviations

2,4 D	2,4-dichlorophenoxyacetic acid
AcNPV	*Autographa californica* nuclear polyhedrosis virus
APRX	anionic peroxidase from zucchini (CpPrx01)
APX	ascorbate peroxidase
ARF	auxin response factor protein
BAC	bacterial artificial chromosome
BAP	6-benzylaminopurine
BFA	brefeldin A
BL	brassinolide
BLAST	Basic Local Alignment Search Tool
CaMV 35S	*Cauliflower mosaic virus*, promoter of the 35S RNA
CCP	cytochrome c peroxidase
CIM	callus inducing medium
Col	Columbia ecotype (*Arabidopsis thaliana*)
CP	catalase peroxidase
CTPP	C-terminal propeptide
DAPI	4',6-diamidino-2-phenylindole
DEAE	diethylaminoethyl
DPI	diphenyleneiodonium
DTT	dithiotreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EST	expressed sequence tag
GFP	green fluorescent protein
GPX	glutathione peroxidase (subgroup of thiol peroxidases)
GUS	β-glucuronidase (*uidA* gene from *E.coli*)
HPF	hydroxyphenyl fluorescein
HRP	horseradish peroxidase
IAA	indole-3-acetic acid
IEF	isoelectric focusing gel
iPA	isopentenyladenine
IPTG	isopropyl-D-thiogalactopyranoside
lac Z	β-galactosidase gene
La*er*	Landsberg *erecta* ecotype (*Arabidopsis thaliana*)
LB	Lysogeny broth (commonly named "Luria-Bertani broth")
	T-DNA left border repeat
LGT	lateral gene transfer
LiP	lignin peroxidase
MCS	multiple cloning site
MITE	miniature inverted-repeat transposable element
ML	maximum likelihood (method of phylogeny estimation)
MnP	manganese peroxidase

MP	maximum parsimony (method of phylogeny estimation)
MUG	4-methylumbelliferyl β-D-glucuronide
MS	Murashige-Skoog medium (plant *in vitro* culture)
MYA	million years ago
NBT	nitrotetrazolium blue chloride
NJ	Neighbor-Joining algorithm
nkat	nanokatal (nanomol formed per second)
nos	nopaline synthase
OD	optical density
PAGE	polyacrylamide gel
PVDF	polyvinylidene difluoride
PX	peroxiredoxin (subgroup of thiol peroxidases)
RB	T-DNA right border repeat
RGP	rice genome project
RIM	root inducing medium
ROS	reactive oxygen species
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium dodecyl sulfate
SIM	shoot inducing medium
SIR	sirtinol
SOD	superoxide dismutase
SOR	superoxide reductase
TIBA	2,3,5-triiodobenzoic acid
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol (Trizma®base, Sigma T6066)
UTR	untranslated region
VP	versatile peroxidase
WAK	wall associated kinase
Ws	Wassilewskija ecotype (*Arabidopsis thaliana*)
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-Gluc	5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid, cyclohexylammonium salt
X-XO	xanthine-xanthine oxidase