



Article scientifique

Article

1995

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Les différenciations génétiques des populations humaines révélées par le
chromosome Y correspondent étroitement à leurs apparentements
linguistiques

Poloni, Estella S.

How to cite

POLONI, Estella S. Les différenciations génétiques des populations humaines révélées par le chromosome Y correspondent étroitement à leurs apparentements linguistiques. In: Revue trimestrielle des droits de l'homme, 1995, vol. 5, p. 82–85.

This publication URL: <https://archive-ouverte.unige.ch/unige:14976>

- CAVALLI-SFORZA (L.), MINCH (E.), MOUNTAIN (J.). 1992. Coevolution of genes and languages revisited. *Proceedings of the National Academy of Science*, vol. 89, 5620-5624.
- CAVALLI-SFORZA (L.), PIAZZA (A.), MENOZZI (P.), MOUNTAIN (J.). 1988. Reconstruction of human evolution: bringing together genetic, archaeological and linguistic data. *Proceedings of the National Academy of Science*, 85, 6002-6006.
- EXCOFFIER (L.), HARDING (R.), SOKAL (R.R.), PELLEGRINI (B.), SANCHEZ-MAZAS (A.). 1991. Spatial differentiation of RH and GM haplotype frequencies in Sub-Saharan Africa and its relation to linguistic affinities. *Human biology*, 63, 3, 273-297.
- EXCOFFIER (L.), PELLEGRINI (B.), SANCHEZ-MAZAS (A.), SIMON (C.), LANGANEY (A.). 1987. Genetics and history of Sub-Saharan Africa. *Yearbook of physical anthropology*, 30, 151-194.
- FLEMING (H.C.). 1987. Toward a definitive classification of the world's languages. *Diachronica*, 4, 159-223.
- GREENBERG (J.H.), RUHLEN (M.). 1992. Linguistic Origins of Native Americans. *Scientific American*, nov., 60-65.
- GREENBERG (J.H.), TURNER (C.G.), ZEGURA (S.L.). 1986. The Settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Current anthropology*, 27, 5, 477-497.
- KAISER (M.), SHEVOROSHIKIN (V.). 1988. Nostratic. *Annual review of anthropology*, 17, 309-329.
- PEIROS (I.). 1989. The linguistic situation in Southeast Asia. In: SHEVOROSHIKIN (V.), ed. *Reconstructing languages and cultures. Materials from the first int. interdisciplinary symposium on language and prehistory (Ann Arbor, 1989)*. Bochum: Universitätsverlag N. Brockmeyer, 66-69.
- REID (L. A.). 1984-85. Benedict's Austro-Tai Hypothesis: an evaluation. *Asian perspective*, 26, 1, 19-34.
- RUHLEN (M.). 1987. *A guide to the world's languages*. London: E. Arnold.
- RUHLEN (M.). 1991. Postscript 1991. In: *Postscript to the paperback edition of A guide to the world's languages*. Stanford: Stanford Univ. Press, 397-399.
- RUHLEN (M.). 1994a. *The origin of language: tracing the evolution of the mother tongue*. New York: J. Wiley.
- RUHLEN (M.). 1994b. *On the origin of languages: studies in linguistic taxonomy*. Stanford: Stanford Univ. Press.
- SOKAL (R.), ODEN (N.), LEGENDRE (P.), FORTIN (M.J.), KIM (J.), VAUDOR (A.). 1989. Genetic differences among language families in Europe. *Am. J. of physical anthropology*, 79, 489-502.
- SOKAL (R.), ODEN (N.), THOMSEN (B.). 1988. Genetic changes across language boundaries in Europe. *Am. J. of physical anthropology*, 76, 337-361.
- SOKAL (R.), ODEN (N.), THOMSEN (B.). 1992. Origins of the Indo-Europeans: genetic evidence. *Proceedings of the National Academy of Science*, 89, 7669-7673.

Les différenciations génétiques des populations humaines révélées par le chromosome Y correspondent étroitement à leurs apparentements linguistiques

Dans le génome humain, il y a deux régions qui se distinguent du reste parce qu'elles ne subissent pas de recombinaison, et sont héritées uniquement par un sexe: l'ADN mitochondrial (ADNmt), qui ne recombine pas et est transmis maternellement, et le chromosome Y, transmis paternellement, et qui, dans sa plus grande partie, ne recombine pas non plus. Dans de telles portions du génome, il est possible de retracer directement, de fille en mère pour l'ADNmt, ou de fils en père, pour le chromosome Y, l'origine et la transmission d'une mutation. L'absence de recombinaison qui caractérise ces régions génomiques les a rendues très attractives pour l'étude de la variabilité génétique humaine. En effet, les modèles de différenciation des populations, qui dérivent des modèles d'évolution de leurs gènes, sont fortement simplifiés en l'absence de recombinaison.

L'étude intensive du polymorphisme de l'ADNmt durant cette dernière décennie a donné lieu à diverses hypothèses sur l'histoire des populations humaines, au nombre desquelles figure la très médiatisée, mais aussi très controversée, théorie de l'Eve Africaine (Cann *et al.* 1987; Excoffier et Langaney 1989; Templeton 1993). L'identification de systèmes génétiques polymorphes sur le chromosome Y est, par contre, beaucoup plus récente. L'analyse de ces systèmes permettra sans doute, dans les années à venir, d'infirmer, de nuancer, ou de confirmer certaines hypothèses avancées lors de l'étude du patrimoine génétique mitochondrial.

Le polymorphisme du chromosome Y

A l'heure actuelle, les techniques de la biologie moléculaire (en particulier l'amplification par PCR et le séquençage de l'ADN) permettent d'identifier un nombre toujours grandissant de systèmes polymorphes sur le chromosome Y (Roewer *et al.* 1992; Mathias *et al.* 1994; Seielstad *et al.* 1994; Hammer 1995; Whitfield *et al.* 1995; Underhill *et al.* 1996). La variabilité de ces systèmes est cependant encore peu documentée à l'échelle de l'espèce humaine (voir cependant Hammer *et al.* 1997). Echappe à ce constat un polymorphisme, nommé p49a,f/TaqI (Ngo *et al.* 1986; Torroni *et al.* 1990; Spurdle et Jenkins 1992), qui a déjà été testé dans 80 échantillons de populations humaines d'origines diverses. Nous avons donc choisi d'entreprendre une revue et une analyse globales des données existant sur ce polymorphisme (Poloni *et al.* 1997).

Le système p49a,f/TaqI est un polymorphisme de longueur de fragments de restriction (PLFR), spécifique au locus DYS1 du chromosome Y. Ce type de polymorphisme

est dû à la présence, ou non, sur le locus, de sites de reconnaissance d'un enzyme de restriction, ici en l'occurrence, l'enzyme *TaqI*.

Un protocole de laboratoire en trois étapes successives met en évidence le polymorphisme: (1) digestion de l'ADN par l'enzyme de restriction *TaqI*; (2) marquage du locus *DYS1* sur les fragments d'ADN digéré par la sonde spécifique p49a,f; (3) visualisation des fragments polymorphes sur un gel d'électrophorèse.

Environ une trentaine de fragments de restriction distincts ont été décrits. Ces fragments peuvent être soit présents, soit absents, et leur combinaison, chez chaque individu, fournit un haplotype. A ce jour, 144 haplotypes différents ont été observés. Ils sont présentés sur le Web, à <http://anthropologie.unige.ch/y-chromosome/>.

Fréquences mondiales et analyses

Nous nous sommes intéressés à la distribution des fréquences de ces haplotypes dans les populations humaines. La première partie de notre étude a donc consisté à collecter, dans la littérature, tous les échantillons de populations humaines testées pour le polymorphisme p49a,f/*TaqI*. Un tri basé sur la taille et l'homogénéité ethnique de ces échantillons nous a permis d'en sélectionner 58 qui satisfaisaient nos critères de qualité d'identification, de typage et d'effectifs. Ces 58 échantillons totalisent 3767 chromosomes Y, et sont distribués sur les cinq continents: principalement en Europe (23 échantillons) et en Afrique (27), mais aussi en Asie (5), Océanie (2), et Amérique (1).

Nous avons estimé des indices de structure génétique des populations (F_{ST} , F_{CT}) par l'approche AMOVA (Excoffier *et al.* 1992). Ces indices, qui mettent en évidence les variations du polymorphisme entre populations, se basent à la fois sur les distributions des fréquences haplotypiques et sur les différences moléculaires entre haplotypes. Nous avons aussi estimé des distances génétiques entre paires de populations, distances qui nous ont servi pour une analyse multivariée (multidimensional scaling analysis).

Il est intéressant de noter à ce point que seuls quelques haplotypes (4, 5, 7, 8, 11, 12, 15) parmi les 144 haplotypes décrits, sont observés à des fréquences substantielles (5% ou plus) dans plusieurs échantillons, ou dans plusieurs régions continentales. Dans leur vaste majorité, les autres haplotypes du polymorphisme p49a,f/*TaqI* ne sont observés que dans un échantillon, et souvent chez un seul individu. Les haplotypes 5 et 11 sont présents dans presque tous les échantillons analysés: l'un ou l'autre pourrait représenter la forme haplotypique ancestrale du locus *DYS1* (ces deux haplotypes ne diffèrent que par un fragment). L'haplotype 4 est prépondérant dans les populations Niger-Congo, les haplotypes 7 et 8

sont particulièrement communs dans les populations du pourtour Méditerranéen, et l'haplotype 15 est fréquemment observé dans les échantillons Européens.

Relation entre affinités génétiques et linguistiques des populations

Le polymorphisme du locus *DYS1* révèle une forte structure des populations ($F_{ST} = 0,23$, $P < 0,001$). Cette structure génétique correspond à la différenciation des populations humaines en familles linguistiques. Comme c'est le cas général du polymorphisme humain, la plus grande variabilité génétique du système s'explique par les différences entre individus d'une même population (70%). Cependant, les différences génétiques entre familles de langues rendent compte de 25% de la diversité totale, alors que seul 5% de cette variabilité est due aux différenciations génétiques des populations dans les familles linguistiques.

Sur la figure 1 sont montrées les affinités génétiques entre populations mises en évidence par l'analyse multivariée. Le lien entre affinités génétiques et apparentements linguistiques est directement évident sur cette figure. Trois groupes principaux de populations génétiquement proches sont observés, qui correspondent à trois familles de langues bien définies: un groupe Niger-Congo, un groupe Afro-Asiatique, et un groupe Indo-Européen. Une comparaison des niveaux de différenciation observés dans ces trois groupes montre que les populations Niger-Congo sont les moins différenciées ($F_{ST} = 0,04$, $P < 0,001$). Un niveau plus élevé et comparable est observé dans les groupes Afro-Asiatique ($F_{ST} = 0,09$, $P < 0,001$) et Indo-Européen ($F_{ST} = 0,07$, $P < 0,005$). Les deux échantillons de populations khoisanes ne forment pas un groupe bien défini dans l'espace multivarié, et sont en effet fortement différenciés l'un de l'autre ($F_{ST} = 0,13$, $P < 0,005$). La correspondance entre différenciations génétiques et linguistiques mise en évidence à l'échelle mondiale est visible aussi à l'échelle continentale. En Afrique, par exemple, la différenciation des populations selon leur affiliation aux familles Niger-Congo, Afro-Asiatique ou Khoisane rend compte de 33% ($F_{CT} = 0,33$, $P < 0,001$) de la variabilité totale du système p49a,f/*TaqI*. Il est difficile d'interpréter la position des échantillons amerindien, chinois, aborigène australien et polynésien dans l'analyse multivariée (ronds blancs sur la fig. 1): un échantillonnage supplémentaire en Asie, Océanie et Amérique est en effet nécessaire. De fait, étant donné le faible nombre d'échantillons représentant ces continents, aucune relation entre génétique et linguistique n'a pu être mise en évidence.

Dans certains cas, les affinités génétiques révélées par ce polymorphisme du chromosome Y et les apparentements linguistiques des populations fournissent des

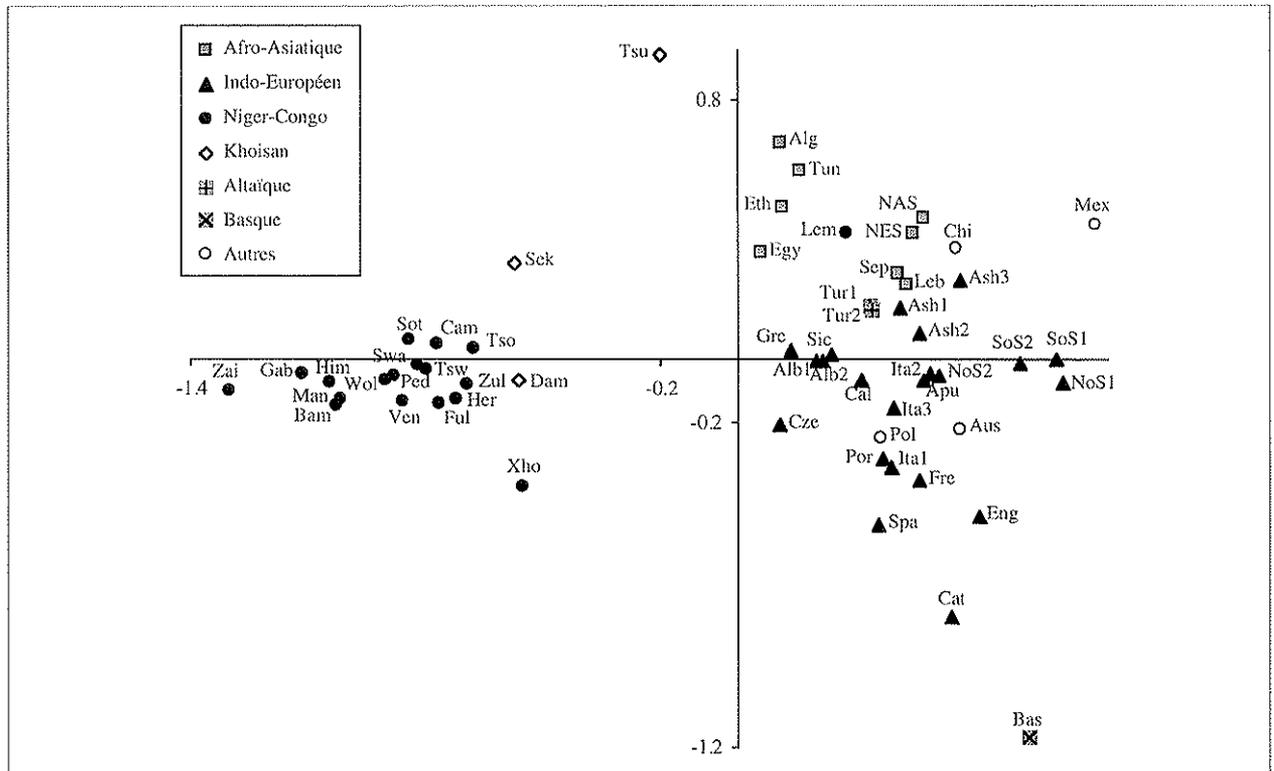


Fig. 1. Analyse multivariée des distances génétiques par paires entre 58 échantillons de populations humaines testés pour le polymorphisme du chromosome Y $p49a,f/TaqI$. Les affiliations linguistiques des échantillons sont indiquées dans l'encadré.

informations contrastantes. C'est le cas, notamment, des Dama et des Lemba d'Afrique du Sud (fig. 1). Les Dama parlent une langue khoisane, mais sont génétiquement très proches des populations Niger-Congo, alors que les Lemba, qui parlent une langue bantoue (Niger-Congo), sont génétiquement proches des populations Afro-Asiatiques. L'étude de la tradition orale Lemba, ainsi que d'autres informations ethnologiques suggère une origine sémitique pour cette population: les Lemba auraient adopté une langue bantoue par suite de leur établissement en Afrique du Sud, mais auraient conservé d'autres traits culturels typiques des populations sémites. Cette hypothèse est étayée par nos résultats génétiques. Dans le cas des Dama, on suppose qu'il s'agit d'une population d'origine Niger-Congo, qui aurait été dominée par une population khoisane dont elle aurait adopté le parler.

Pour obtenir une meilleure idée sur l'histoire évolutive de ces populations, nous avons analysé la structure génétique révélée par le polymorphisme de restriction de l'ADNmt, ce qui nous a aussi permis une comparaison globale des différenciations survenues dans les lignées purement paternelles et maternelles. Cette comparaison n'a pu être réalisée que sur 19 échantillons de populations, les seuls qui avaient aussi été testés pour

l'ADNmt. La structure génétique globale observée pour le chromosome Y est confirmée par l'ADNmt, à l'exception, précisément, de populations comme les Dama. Ces derniers, nous l'avons vu, présentent une affinité génétique étroite avec les populations Niger-Congo pour le polymorphisme du chromosome Y, mais ils s'en différencient nettement au niveau mitochondrial, et semblent plutôt présenter des liens génétiques avec les populations khoisanes. Un tel résultat suggère, d'abord, une origine métissée pour cette population. Il suggère aussi que la composante Niger-Congo dans cette population est principalement d'origine paternelle, alors que la composante khoisane a été transmise plutôt maternellement. Des hypothèses similaires sont suggérées pour d'autres populations, et notamment en Afrique du Sud, pour les Lemba et les Herero.

Si notre étude confirme, encore une fois, la relation étroite entre histoire génétique et histoire linguistique des populations humaines, elle montre par ailleurs la nécessité des comparaisons entre différents systèmes génétiques pour inférer ou tester des hypothèses précises sur l'histoire du peuplement humain.

Estella S. Poloni

Bibliographie

- CANN (R.L.), STONEKING (M.), WILSON (A.C.). 1987. Mitochondrial DNA and human evolution. *Nature*, 325, 31-36.
- EXCOFFIER (L.), LANGANEY (A.). 1989. Origin and differentiation of human mitochondrial DNA. *Am. J. of hum. genetics*, 44, 73-85.
- EXCOFFIER (L.), SMOUSE (P.), QUATTRO (J.M.). 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131, 479-491.
- HAMMER (M.F.). 1995. A recent common ancestry for human Y chromosomes. *Nature*, 378, 376-378.
- HAMMER (M.F.), SPURDLE (A.B.), KARAFET (T.), BONNER (M.R.), WOOD (E.T.), NOVELLETTO (A.), MALASPINA (P.) et al. 1997. The geographic distribution of human Y chromosome variation. *Genetics*, 145, 787-805.
- MATHIAS (N.), BAYES (M.), TYLER-SMITH (C.). 1994. Highly informative compound haplotypes for the human Y chromosome. *Human molecular genetics*, 3, 115-123.
- NGO (K.Y.), VERGNAUD (G.), JOHNSON (C.), LUCOTTE (G.), WEISSENBACH (J.). 1986. A DgA probe detecting multiple haplotypes of the human Y chromosome. *Am. J. of human genetics*, 38, 407-418.
- POLONI (E.S.), SEMINO (O.), PASSARINO (G.), SANTACHIARA BENERECETTI (A.S.), DUPANLOUP (I.), EXCOFFIER (L.). (1997). Human genetic affinities for Y chromosome p49a,ff *TaqI* haplotypes show strong correspondence with linguistics. *Am. J. of human genetics*, 61, 1015-1035.
- ROEWER (L.), AMEMANN (J.), SPURR (N.K.), GRZESCHIK (K.-H.), EPLEN (J.T.). 1992. Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Human Genetics*, 89, 389-394.
- SEHELSTAD (M.T.), HEBERT (J.M.), LIN (A.A.), UNDERHILL (P.A.), IBRAHIM (M.), VOLLRATH (D.), CAVALLI-SFORZA (L.L.). 1994. Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Human molecular genetics*, 3, 2159-2161.
- SPURDLE (A.), JENKINS (T.). 1992. Y chromosome probe p49a detects complex *PvuII* haplotypes and many new *TaqI* haplotypes in Southern Africa. *Am. J. of hum. genetics*, 50, 107-125.
- TEMPLETON (A.R.). 1993. The «Eve» hypothesis: a genetic critique and reanalysis. *Am. Anthropologist*, 95, 51-72.
- TORRONI (A.), SEMINO (O.), SCOZZARI (R.), SIRUGO (G.), SPEDINI (G.), ABBAS (N.), FELLOUS (M.), et al. 1990. Y chromosome DNA polymorphisms in human populations: differences between Caucasoids and Africans detected by 49a and 49f probes. *Annals of hum. genetics*, 54, 287-296.
- UNDERHILL (P.A.), JIN (L.), ZEMANS (R.), OEFNER (P.J.), CAVALLI-SFORZA (L.L.). 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc. of the Nat. Academy of Sciences of the USA*, 93, 196-200.
- WHITFIELD (L.S.), SULSTON (J.E.), GOODFELLOW (P.N.). 1995. Sequence variation of the human Y chromosome. *Nature*, 378, 379-380.

Les échantillons sont: Alb1: Albanais; Alb2: Albanais de Calabre; Alg: Algériens; Apu: Pouilles; Ash1: Ashkenazes; Ash2: Ashkenazes; Ash3: Ashkenazes; Aus: Aborigènes Australiens; Bam: Bamileke (Cameroun); Bas: Basques d'Espagne; Cal: Calabrais; Cam: Camerounais; Cat: Catalans; Chi: Chinois; Cze: Tchécoslovaques; Dam: Dama; Egy: Egyptiens; Eng: Anglais; Eth: Ethiopiens; Fre: Français; Ful: Peuls (Sénégal); Gab: Gabonais; Gre: Grecs; Her: Herero; Him: Himba; Ital: Italiens; Ita2: Italiens; Ita3: Italiens; Leb: Libanais; Lem: Lemba; Man: Malinkés (Sénégal); Mex: Amérindiens du Mexique; NAS: Sepharades d'Afrique du Nord; NES: Sepharades du Proche Orient; NoS1: Sardes du Nord; NoS2: Sardes du Nord; Ped: Péd; Pol: Polynésiens; Por: Portugais; Sek: Sekele San; Sep: Sepharades; Sic: Siciliens; SoS1: Sardes du Sud; SoS2: Sardes du Sud; Sot: Sotho; Spa: Espagnols; Swa: Swazi; Tso: Tsonga; Tsu: Tsumkwe San; Tsw: Tswana; Tun: Tunisiens; Tur1: Turques d'Istanbul; Tur2: Turques de Konya; Ven: Venda; Wol: Wolofs (Sénégal); Xho: Xhosa; Zai: Zaïrois; Zul: Zulus.

Contribution à l'étude de la reconnaissance et de la perception des visages: une approche morphométrique

Le visage humain est la partie du corps que nous utilisons le plus pour reconnaître nos semblables. Il joue un rôle central dans la communication en véhiculant via le langage et les expressions faciales une multitude d'informations différentes. Parmi tous les objets que nous sommes amenés à reconnaître chaque jour, la reconnaissance de l'objet «visage» est un processus très élaboré, et semble être plus performante que celle des autres classes d'objets. En effet, nous sommes capables de distinguer rapidement parmi des milliers de visages différents, alors que ces stimuli diffèrent relativement peu les uns des autres sur le plan structurel¹. De plus, l'acquisition de ces capacités n'implique pas d'apprentissage formel particulier, et a lieu de manière quasi-automatique au cours de l'enfance et du début de l'adolescence. La

1. Un visage familier est reconnu en approximativement 0,5 secondes, malgré le grand nombre d'instances de visages que nous stockons, et la grande similitude structurelle entre eux. Les individus d'âge adulte peuvent enregistrer de nombreux nouveaux visages à partir de représentations photographiques qu'ils ont vu pendant un temps très court (50 visages pendant 5 secondes chacune, par exemple). Les performances de reconnaissance seront supérieures à 90% (Carey 1992).