Bacterial genome evolution within a clonal population: from in vitro investigations to in vivo observations

Beaume, Marie-Emilie; Monina, Nadezda; Schrenzel, Jacques; Francois, Patrice

How to cite

BEAUME, Marie-Emilie et al. Bacterial genome evolution within a clonal population: from in vitro investigations to in vivo observations. In: Future microbiology, 2013, vol. 8, n° 5, p. 661–674. doi: 10.2217/fmb.13.28

This publication URL: https://archive-ouverte.unige.ch/unige:33331

Publication DOI: 10.2217/fmb.13.28

Review
Future Microbiology

Bacterial genome evolution within a clonal population: from *in vitro* investigations to *in vivo* observations

Marie Beaume‡[1], Nadezda Monina‡[1], Jacques Schrenzel[1] & Patrice François*[1]
[1]Genomic Research Laboratory, Infectious Diseases Service, University of Geneva Hospitals, Rue Gabrielle-Perret-Gentil 4, CH-1211 Geneva 14, Switzerland
*Author for correspondence: Tel.: +41 223 729 337 ■ Fax: +41 223 729 830 ■ patrice.francois@genomic.ch
‡Authors contributed equally

Bacteria are faced with a diversity of environmental stresses that include high salt concentrations, heavy metals and pH fluctuations. Adaptation to resist such stresses is a complex phenomenon that involves global pathways and simultaneous acquisition of multiple unrelated properties. During the last 3 years, the development of new technologies in the field of molecular biology has led to numerous fundamental and quantitative *in vitro* and *in vivo* evolutionary studies that have improved our understanding of the principles underlying bacterial adaptations, and helped us develop strategies to cope with the health burden of bacterial virulence. In this review, the authors discuss the evolution of bacteria in the laboratory and in human patients.

In vitro adaptive dynamics

Recent developments in high-throughput sequencing and analysis allow the performance of evolution studies at the nucleotide resolution level [1–5]. Darwin's theory of evolution [6] is widely accepted within the modern scientific community. The Darwinian view relies on genetic drift and natural selection. The best-adapted individuals within a particular environment have a higher capacity to reproduce, thus leading to an over-representation of their genotype among the descendant population. Observing the ways in which evolution shapes species in a natural context is difficult, owing to the long time periods needed for natural selection [6,7]. Scientists circumvent this difficulty by taking advantage of the short generation times of bacteria, which enables the observation of evolution in real time under 'laboratory conditions'. Moreover, because bacteria mostly reproduce asexually, and have much smaller genomes, studies of genomic evolution are simplified.

Winner & loser populations: the importance of epistatic interactions

Competition between two asexual populations leads to the selection of the most adapted population [8]. At the end of the evolutionary race, the populations with poor plasticity can disappear or dramatically decrease in number, becoming 'losers' of the race, whereas the populations able to find solutions for survival dominate and represent 'winners'. Therefore, bacteria not only vary in traits that determine their own fitness at a specific time, but also vary in their capacity to generate the best-adapted descendants.

A long-term evolution experiment conducted by Lenski provides opportunities to address and answer central questions of evolutionary biology [9]. Twelve *Escherichia coli* populations have been cultivated from a common ancestor for more than 20 years (over 56,000 generations). Every 500 generations (75 days), bacterial cultures were frozen as a 'fossil' sample record [9].

In order to determine whether various changes in DNA provided organisms with different capacities for adaptation, the 500th and the 1000th generations were characterized [8]. After 500 generations, two genotypes of *E. coli* that differed by several mutations were identified, representing loser and winner genotypes. Based on evolutionary theory, it was predicted that 'eventual winners' would be fitter than 'eventual losers', but interestingly, competitive-fitness measures indicated that early clones of higher fitness were not those that became fixed. An additional experiment was performed to help understand this phenomenon [8]. Ten isolates from a 500th generation population were grown for an additional 883 generations. 'Eventual winners', although of lower fitness early on, were more able to survive in the long run and had the most potential to evolve within the experimental environment [8].

The first generation of 'losers' acquired an advantageous mutation in the *topA* gene, which

Keywords
■ bacterial adaptation ■ genome evolution ■ genome plasticity ■ metabolism ■ virulence

Future Medicine part of fsg

codes for a protein that helps package DNA [8]. However, this particular mutation prevented the acquisition of a beneficial mutation in *spoT*, a gene that encodes a regulator of global gene expression via synthesis and degradation of (p)ppGpp. Earlier isolates from the winner lineage gained a mutation that was less advantageous in terms of immediate fitness in comparison with loser lineage; however, they eventually produced the most adapted descendants and became fixed in the population [8]. These gene interactions are commonly known as epistasic interactions. Woods *et al.* demonstrated that the difference between the evolutionary potential of each lineage could be explained by epistasis, and that there was 'second-order selection' during the evolution process [8].

The key role of epistatic interactions during evolution was notably highlighted in two other studies: one with long-term laboratory evolution of *E. coli* [10], and one with an engineered strain of *Methylobacterium extorquens* [11]. These studies showed that as beneficial mutations accumulate in one genome, the gain in fitness achieved by each mutation diminishes. This contrasted with the fitness gains exhibited by beneficial mutations when they were in separate genomes. Therefore, the greater the number of beneficial mutations present in one genome, the more damaging epistasis was.

Dynamics of adaptive genome evolution

Long-term evolution experiments enabled Barrick *et al.* to investigate the correlation between adaptative dynamics and genome evolution [2]. Fitness assays were performed and genomes of *E. coli* populations at 2000, 5000, 10,000, 15,000, 20,000 and 40,000 generations were sequenced and analyzed [2]. Gains in measurable fitness decelerated over time, but the acquisition of beneficial mutations was nearly constant for 20,000 generations, indicating genomes were not evolving neutrally. Interestingly, the same lineage evolved an increased mutation rate (hypermutable phenotype) after 20,000 generations, resulting in a substantially increased rate of genome evolution mostly due to the accumulation of neutral mutations [2]. Therefore, the complexity of correlation between genomic and adaptive evolution seems clear and requires additional investigation.

Cooper and Lenski showed that the *E. coli* isolates at 20,000 generations had a 70% increase in fitness relative to their ancestor; however, this was ten-times less than the fitness gain measured at 5000 generations [12]. Although fitness increased linearly until 1000 generations and then plateaued, mutations accumulated linearly until the 20,000 generations [2]. This observation could be explained by a model of stochastic and constant acquisition of neutral mutations during evolution [13]. According to the model, beneficial mutations would be more commonly fixed during the early phase of laboratory evolution; however, over time, their fixation would decelerate. Furthermore, detecting beneficial mutations would become more difficult in later generations because neutral mutations, which had always been accumulating by drift, would be more numerous. However, Barrick *et al.* presented strong evidence that nearly all mutations were beneficial and therefore, the discordance observed between genomic and adaptive evolution cannot be explained by the accumulation of neutral substitutions [2].

Finally, the rate of mutation changed drastically between 20,000 and 40,000 generations due to a single base-pair insertion in the gene encoding the DNA repair protein MutT [2]. This evolved hypermutability is demonstrated by the numbers of mutations at 20,000 versus 40,000 generations: the 20,000th generation accumulated 29 single-nucleotide polymorphisms (SNPs) and 16 deletion/inversion/insertion events, whereas the 40,000th generation contained 627 SNPs and 26 deletion/inversion/insertion events, including numerous neutral mutations. These studies thus illustrate that increases in fitness and genome evolution are complex processes, even within a constant and controlled laboratory environment.

Parallel evolution

The parallel evolution of beneficial mutations/phenotypes has been reported in prokaryotes [14], eukaryotes [15], viruses [16,17] and bacteriophages [18]. Stanek *et al.* studied the repeated appearance of mutations in 12 populations of *E. coli* propagated during 20,000 generations [19]. A beneficial adenine insertion in the protein-binding site upstream of the *glmUS* operon was found in one population [19]. The mutation occurred in the population between the 500th and 1500th generations and conferred a fitness increase of approximately 5% with respect to the ancestor. Interestingly, none of the other 11 populations contained any modification in the same genomic region. Two explanations were proposed: clonal interference, where competition among beneficial mutations of different clones resulted in selection of only one of

them; and epistasis. Although mutations did occur during the evolution of these populations, additional sequencing of 36 genes in 50 isolates and the ancestor revealed very few mutations in these 36 randomly chosen gene regions from 12 populations [20].

Cooper *et al.* identified the mechanism involved in the parallel phenotypic modification associated with ribose catabolism [21]. All 12 populations of *E. coli* lost D-ribose catabolic function after 2000 generations when grown in glucose-minimal medium. Molecular analysis revealed an insertion sequence (IS), located upstream of the ribose operon (*rbs*) that was responsible for the loss of *rbs* functionality. A few years later, Crozat *et al.* proposed DNA topology as a key target for selection, because of its influence on the overall patterns of gene expression regulation [22]. Twelve *E. coli* populations alternated between active growth and nutrient exhaustion, a regimen that is known to influence DNA topology. Indeed, most populations had a change in the level of DNA supercoiling, potentially owing to mutations in *topA* or *fis*, which encode topoisomerase I and a repressor of the *gyrAB* transcription, respectively [23]. While Crozart's study was based on phenotypic observations [22], Woods *et al.* studied genetic parallelism within the same 20,000 generation populations of *E. coli* [24]. Statistical tests proved that mutations common to multiple populations conferred a selective advantage in a particular environment; therefore, natural selection led to parallel changes [24]. Two beneficial mutations appeared in the *pbpA* operon (involved in cell wall biosynthesis) within the first 2000 generations of *E. coli* laboratory evolution. These mutations led to a reduction in the cellular concentrations of PBP2, resulting in a spherical cell shape with increased volume. Following this, Cooper *et al.* used microarrays to compare two *E. coli* isolates. As a result of their parallel evolution, the transcriptomic changes observed in each population displayed the same pattern relative to their common ancestor [25]. Pelosi *et al.* extended this work using proteomics [26] and exceptional parallelism in protein expression profiles was detected. Impressive correlation between two methods allowed identification of parallel changes in global regulatory networks: ppGpp and cAMP-CRP regulons [26]. Moreover, the authors found the mutation in *rbs* and discovered another common change in *malT*, which encodes a transcriptional regulator of the maltose regulon [21]. Most of evolved bacteria did not show an increased cell size [27].

Nusbaum *et al.* also observed parallel evolution in *Vibrio cholerae* grown under high selection pressure [3]. Comparative genomics of five evolved rifampicin-resistant descendants with their antibiotic-susceptible ancestor revealed one difference in the *rpoB* gene, which encodes the β-subunit of RNA polymerase.

Leiberman *et al.* performed another parallel evolutionary study on 112 *Burkholderia dolosa* isolates collected from 14 patients over a period of 16 years [28,29]. They identified a set of genes that appeared to evolve in parallel in the human host: a glycosyltransferase gene involved in production of O-antigen repeats, and a homolog of the gene encoding DNA gyrase subunit A. The latter had nonsynonymous mutations that conferred resistance to ciprofloxacin and had arisen independently in several individuals, indicating strong positive selection [28].

Abundance of the mutator genotypes

Sniegowski *et al.* reported isolates that had undergone spontaneous mutations to the hypermutable phenotype, or 'mutators', during the long-term evolution (initiated in 1988) of *E. coli* in batch cultures [30]. Among the 12 populations, three showed mutation rates 100-fold higher than their ancestor. Additionally, data from Cooper and Lenski revealed that mutators emerged after 20,000 generations in four populations [12]. By using multicopy plasmids bearing the parental allele of seven genes encoding DNA repair proteins, Sniegowski found that mutations in *mutS*, *uvrD* and *mutL* genes conferred the mutator phenotype [30]. In addition, *mutT* was strongly supported in work carried our by Lenski *et al.*, as defects in this gene were frequently observed within the 12 populations of *E. coli* [20]. The mutator phenotype seemed to be stable over generations. However, Taddei *et al.* demonstrated that the mutator genotype could arise from a nonmutator genetic background and could also revert back to a wild-type genotype [31].

Mutators appeared as the consequence of 'second-order selection', as described by Taddei *et al.* [31] and could be coenriched by beneficial mutation selection. The selection of a specific mutation by particular environmental conditions resulted in the conversion of 50% of the population to mutator genotypes. Moreover, three rounds of selection for particular mutations were sufficient to produce a population with 100% of the mutator genotype [32]. Thus, mutations leading to the mutator genotype did not generally confer real advantages in a

particular environment; however, they spread by hitchhiking with the beneficial mutation to which they were genetically linked. Interestingly, Visser *et al.* demonstrated that the high mutation rate of mutators did not lead to the acceleration of adaptive evolution in asexual populations [33].

The prevalence of mutators is also important in natural populations of bacteria. LeClerc and coworkers found 1–2% of the population to be mutators in pathogenic *E. coli* and *Salmonella enterica* [34], and 20% of cells in *Pseudomonas aeruginosa* populations associated with cystic fibrosis (CF) were mutators [35]. The presence of these mutators in natural populations can be potentially related to the rapid evolution of antibiotic resistance.

Specific evolution to the environment: variation in metabolic functions

Adaptation to a particular environment also occurs via specific metabolic changes that may not be beneficial in alternative environments. There are two relevant mechanisms described in the literature: antagonistic pleiotropy corresponds to a mutation that is beneficial in one environment, but detrimental in another; and mutations that accumulate by neutral drift in one environment, and could be advantageous in another [7]. Thus, bacteria have to balance the benefit of optimized growth in one medium with the cost of keeping such mutations in the genome. Cooper and Lenski analyzed the decay of unused catabolic functions among the 12 *E. coli* populations in the 'long-term experiment' [12]. Bacterial growth on 64 substrates was assessed using phenotypic arrays. After 20,000 generations in minimal glucose medium, cells became specialized for growth in this environment. The majority of parallel reductions in catabolic functions arose during the first 2000 generations, before the emergence of any mutator population. Finally, this specialization appeared to be more consistent with antagonistic pleiotropy than the accumulation of mutations [12], similar to the previously reported loss of D-ribose catabolic function [21].

In order to understand how cells adapted to glucose-limited medium, Travisano and Lenski evaluated the relative fitness of these bacterial populations in 11 novel environments, with lactose, mannose or trehalose as unique carbon sources [36]. The 12 populations seemed to cluster in at least six physiologically distinct groups. Performance varied widely between tested environments, especially in media where the bacteria used different transport mechanisms than those that are used for growing in glucose. In the glucose-limited environment, bacteria evolved and acquired specific mechanisms to increase their capacity to use glucose. In media where the uptake mechanism differed from that used for glucose, fitness did not improve appreciably. Using 2000 *E. coli* generations, Travisano *et al.* explored the role of historical contingency of *E. coli* populations [37]. After 2000 generations in glucose-limited medium, *E. coli* populations were grown in maltose-containing medium for an additional 1000 generations. Immediately following this environmental change, most of the cells did not grow significantly. However, eventually all the populations were able to grow on maltose as a sole carbon source. Finally, the authors showed that the adaptive history of each population could dramatically shape future evolution.

The species *E. coli* is partly defined by its inability to metabolize citrate (Cit^-) under toxic conditions. However, one of the 12 *E. coli* populations in the 'long-term experiment' evolved this ability, starting from generation 31,500 [38]. The glucose-limited medium these populations evolved in also contained an abundance of citrate. The authors suggested that this phenotype did not evolve sooner because more than one mutation was necessary. The Cit^+ clones increased dramatically in population size, but did not eliminate the Cit^- population, which was a superior competitor for limited glucose. The authors performed additional replicate experiments to distinguish whether evolution of the ability to utilize citrate required rare mutation(s), or was dependent on more common mutations occurring in a certain genetic background. Statistical analysis rejected the hypothesis of a rare mutation. Indeed, three independent experiments showed the same tendency for the Cit^+ acquisition. Thus, authors concluded that 'chance' was not sufficient to explain this adaptation and suggested that their results supported the hypothesis of historical contingency. In addition to the evolution of citrate catabolism, three other strains evolved a dependence on citrate for optimal growth in glucose after 50,000 generations [39]. Another laboratory evolution study by Dekel and Alon showed that after approximately 250 generations, *E. coli* grown in an abundance of lactose evolved increased expression of β-galactosidase, the enzyme responsible for lactose cleavage [40].

Bacterial resistance to viral infections or antibiotics could also shape the balance between the costs and benefits of mutations, producing

unique tradeoffs. Meyer *et al.* assessed changes in resistance against bacteriophage infection in *E. coli* populations that grew for 45,000 generations in the absence of viruses [41]. When exposed to three viruses, two phenotypes were observed: the ancestral state of partial T6* (T6 mutant with altered host range) resistance was lost; and resistance to λ evolved, despite no contact with λ during evolution. The strong parallelism between the six bacterial cell lines suggested that this evolution was not the consequence of a random drift. Authors suggested that resistance to the λ virus is associated with pleiotropic effects of a beneficial mutation linked to a metabolic process [41]. In addition, loss of phage T6* resistance strongly supports the hypothesis of phenotypic costs that favor the loss of resistance when not necessary. However, resistance to viral infection does not always have a fitness cost (e.g., phage T5 infection described by Lenski and Levin [42]).

Recently it was shown that vancomycin-intermediate *S. aureus* lost its resistance capacity once vancomycin pressure disappeared [43]. The hetero-vancomycin-resistant *S. aureus* population evolved to a vancomycin-susceptible population after 2 weeks of growth in a drug-free medium. On the contrary, two different classes of antibiotic (vancomycin and β-lactam derivatives) could trigger the evolution of vancomycin-susceptible *S. aureus* towards a hetero-vancomycin-resistant *S. aureus* population. Overall, it was shown that a small fraction of individuals ($<10^{-9}$) in a susceptible population was not affected by the antimicrobial treatment [44], suggesting that resistant variants already existed and would be selected by the antibiotic pressure, without any *de novo* acquisition of mutations.

The evolutionary importance of *rpoS* expression level (global transcriptional regulator) was reported in several studies [45–47]. It was demonstrated that strains expressing more *rpoS* were more resistant to stress but less efficient in their metabolic functions [48]. Conversely, bacteria with low *rpoS* expression levels displayed broad metabolism, but low resistance to environmental stresses.

Biofilms are known for their clinical relevance. The formation and evolution of biofilms in *Acinetobacter* spp. and *Pseudomonas putida* was studied by Hansen *et al.* [49]. Mutations in one species led to a special adaptive interaction between the two species within the biofilm. This new association produced biofilms that were more stable and efficient in producing biomass when compared with ancestral biofilms. Whereas results from the studies discussed earlier suggested that the adaptation of one population resulted in the elimination of another, Hansen *et al.* showed that bacterial coevolution could result in a functional gain.

In vivo evolution of bacterial genomes during infection

Numerous *in vitro* studies on bacterial evolution have illuminated important concepts of evolution, including the impact of selection pressure, random drift and bacterial aging. However, laboratory evolution is limited to relatively small population sizes and simplified laboratory conditions [50]. Therefore, whether *in vitro* results are applicable to the *in vivo* adaptation of bacteria to their hosts remains to be demonstrated.

Strategies used by bacteria to evolve in a particular environment & consequences

Acquisition & deletion of genomic elements

Mobile genetic elements

Genome plasticity facilitates the adaptation of pathogens in response to clinical challenges and plays a key role in the emergence of new epidemic pathogens (Figure 1). Approximately 160 new bacterial diseases have been discovered during the last 70 years [51]. Some of these pathogens caused severe infections in environments that differed from the niche they were originally associated with. For example, *Burkholderia cepacia* complex demonstrates a phenotypic duality: it provides agricultural benefits, but can also cause detrimental human infections [52]. For decades *S. aureus* was considered to be a nosocomial pathogen. Currently, it is widespread in the community, affecting healthy people who have no contact with the hospital environment and leading to major outbreaks worldwide [53].

The main mechanisms of horizontal gene transfer are transformation (acquisition of free linear or circular DNA from the environment), conjugation (DNA transfer between two organisms) and bacteriophage transduction. Recently, it was proposed to refer to all mobile genetic elements as the 'mobilome' [54].

It was reported by Ochman *et al.* that mobile genetic elements (MGEs) account for <1% (*Mycoplasma*) to approximately 12.8% (*E. coli*) of the total genome [55], while in *S. aureus*, it is approximately 20% [56]. Therefore, the amount of MGEs and their size varies between species. Whole-genome comparisons of five sets of *Helicobacter pylori* strains sequentially isolated over 3–16 years from four chronically infected

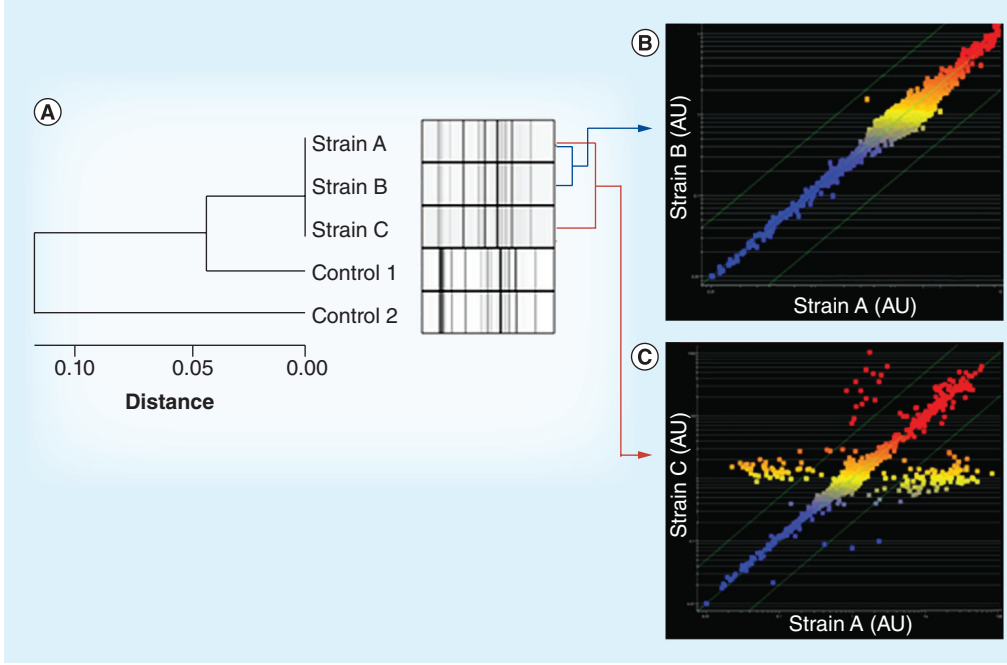

Figure 1. Genomic alterations during the course of an infection. (A) Illustrative example of genotyping by using multilocus variable-number tandem-repeat analysis showing the relatedness between *Staphylococcus aureus* isolates from the same patient [121]. The scale shows the distance between isolates resulting from the analysis of the band pattern. A distance <0.2 is found for related isolates and <0.1 for closely related isolates. **(B)** Scatterplot resulting from comparison of two 'clonal' isolates (using multilocus variable-number tandem-repeat analysis), by microarray. Fluorescence levels (arbitrary units) recorded for a pair of 'clonal' strains are plotted on the x- (strain A) and y-axis (strain B) and show a homogeneous repartition. **(C)** Scatterplot obtained for another pair of 'clonal' isolates showing spot distribution outside cutoff lines (in blue), suggesting acquisition (upper-left quadrant of the plot) and loss of genes (bottom-right quadrant). The alignment of the plots on the diagonal proves the presence of the corresponding genome regions in both strains. No significant changes are detected in **(B)**, whereas **(C)** shows important evolution events.

patients revealed many recombinant clusters of polymorphisms [57]. Recombinant regions had a mean length of 394 bp and were nonrandomly distributed across chromosomal regions up to 20 kbp. The members of the *hop* family exhibited a high recombination frequency in all isolates. Selection of these recombinant genomes during chronic infections was associated with the role of Hop proteins in adhesion to human gastric epithelium [57]. Apparently, these mechanisms play an important role in the genomic plasticity of bacteria [58,59].

The emergence of pathogens from commensal populations is of a great importance [60,61]. For instance, it was found that *mecA* of methicillin-resistant pathogenic staphylococcal strains originated within the genus *Staphylococcus* from a *mecA* homolog identified in *Staphylococcus sciuri*, where its function was unrelated to β-lactam resistance [60]. In another instance *Shigella* evolved from *E. coli* to become pathogenic by deleting the *cadA* gene, the key determinant for switching from a commensal to pathogenic lifestyle [61].

Insertions & deletions without involvement of identified horizontal transfer

Airway infections in patients with CF offer opportunities to study the development and persistence of bacterial infections. One of the most abundant bacteria in CF is the opportunistic pathogen *P. aeruginosa*, which can reach 10^7–10^9 cells/ml of airway mucus secretion [62]. Smith *et al.* performed comparative genomic analysis of *P. aeruginosa* collected from a single CF patient over 8 years [63]. Multilocus sequence typing confirmed the clonality of these isolates. Analysis revealed numerous loss-of-function mutations owing to premature translation stop or a shift in the reading frame, affecting virulence genes and their regulators as well as genes coding for transport proteins and antibiotic resistance determinants. Genes encoding multidrug efflux pumps appeared to be a common target for mutations in chronic infections, resulting in increased resistance to tobramycin, gentamicin, aminoglycosides and amikacin antibiotics with respect to the parental strain [63]. Moreover, phenotypic changes included alterations in

biofilm formation, the loss of twitching motility, a decrease in pyoverdine production and the modulation of serotype-specific antigenicity. Selection of a population with mutations in, or reduced expression of, virulence factors ensured the long-term survival of *P. aeruginosa* by enabling cells to remain undetected by the immune system [63].

Asymptomatic *H. pylori* infections represent a conundrum. These infections can increase the risk of developing severe pathologies such as gastric cancer [64], but can also protect the host against several diseases [65,66]. In order to correlate pathology evolution with bacterial genome modifications, Oh *et al.* compared the genomes of two *H. pylori* isolates collected from a patient initially diagnosed with chronic atrophic gastritis that developed into gastric adenocarcinoma 4 years later [67]. Comparative whole-genome sequencing analysis of the two isolates revealed nine genes coding for D-alanine:D-alanine ligase A, metalloprotease, methionyl-tRNA formyltransferase, putative ribonuclease N, type I restriction enzyme M protein and four hypothetical proteins acquired during evolution from chronic gastritis to cancer. Six genes were lost in the cancer-associated isolate; these were mostly genes coding for proteins implicated in DNA repair system (e.g., endonuclease III, uracil-DNA glycosylase and transcription-repair coupling factor) [67]. The loss of DNA repair genes resulted in an increased mutation rate, effectively helping the bacteria diversify and escape the host immune system. Moreover, a decrease in DNA repair probably favored the accumulation of additional mutations. Finally, the authors suggested that the genes gained via horizontal transfer probably originated from a member of the gastric microbiota. However, Kraft *et al.* found that the number of gene gains and losses was dwarfed by the larger number of genetic changes that occurred through homologous recombination [68].

Duplications & recombinations

Recombination between two homologous IS can lead to chromosomal deletions, inversions or translocations depending on the IS orientation. IS could also be acquired by horizontal transfer, and result in significant genomic rearrangement [69–73]. Two theories were developed to explain the persistence of IS elements in bacterial genomes: ISs are genomic parasites that increase in number by replicative transposition and lead to an increased rate of deleterious mutations by their genomic insertion; or IS are able to generate beneficial mutations that increase bacterial fitness [74]. For example, in *E. coli*, these beneficial mutations affect genes involved in cell wall synthesis (*pbpA* encoding the penicillin-binding protein 2 and *rodA*) and in central metabolism (*pykF* encoding the pyruvate kinase I and *rbs* implicated in ribose utilization) [74]. It was demonstrated by Dhar and McKinney that transposon-mediated mutations led to various phenotypes and contributed to adaptation of microbial populations to their respective environments [75].

Phage mobilization & conversion

Bacteriophages are responsible for genome alterations leading to functional modifications. Goerke *et al.* investigated *S. aureus* genome plasticity in chronic lung infections and healthy nares colonization [76]. In CF patients, *S. aureus* generates chronic lung infections requiring antibiotic treatment. Comparisons of *S. aureus* isolated from healthy carriers and CF patients revealed that genome alterations were more frequent in CF patients. Antibiotic exposure of CF patients exerted strong selection pressures on *S. aureus* and forced bacterial genomes to evolve. Among the altered *S. aureus* strains, at least 42% corresponded to phage mobilization [76]. Most alterations were due to modification of the β-toxin production by phage insertion and the rest corresponded to phage deletion.

A few years later, Boyle-Vavra *et al.* investigated the involvement of phages in *S. aureus* resistance. Two *S. aureus* isolates, one susceptible and one resistant to daptomycin, were collected [77]. Interestingly, the *hlb* gene (encoding β-hemolysin) was disrupted by a prophage in the daptomycin-susceptible strain, whereas this gene lacked the prophage and was intact in the resistant isolate [77]. Evidently, phages play an important role in bacterial pathogenicity.

Bacteriophages acquired by horizontal transfer can also impart new functions to their bacterial hosts. For example, *S. aureus* gained the ability to produce Panton–Valentine leukocidin or enterotoxins [78,79]. It was shown that the loss of prophages by *S. aureus* strain Newman led to its inability to cause disease [80]. Another study revealed that among three *S. aureus* isolates collected over 26 months from the airways of a CF patient, a subpopulation with variation in phage content led to a reduction of metabolic costs for the whole infected population [81].

Prophage and other MGEs can induce mobility of some pathogenicity islands [82] or interact with plasmids [83]. However, most bacteria that already carry a particular type of phage appear to be resistant to infection by other phages [84].

Nubel and colleagues demonstrated the involvement of phages in mechanisms promoting the virulence and spread of pathogenic bacteria [85].

Microevolution: the particular case of the SNPs

SNPs & the emergence of antibiotic resistance

To explore evolution of multidrug resistance in *S. aureus* in response to antimicrobial therapy, Mwangi *et al.* compared genome isolates from the bloodstream of a patient undergoing chemotherapy [4]. Despite treatment with vancomycin and β-lactam antibiotics, the patient died from complicated endocarditis within 12 weeks. The initial vancomycin-susceptible isolate and the evolved vancomycin-resistant isolate differed by 35 nucleotides at 31 different loci [4]. A nonsynonymous mutation in the *vraR* operon may have been selected by the glycopeptide therapy. Surprisingly, vancomycin-resistant isolates became more resistant to daptomycin (100-fold) even though this antibiotic was not used in the treatment [4]. Resistance to daptomycin was due to a mutation in *rpoC*, encoding the RNA polymerase β′-subunit, or in the *yyc* gene cluster, which encodes a two-component system response regulator identified in the cell wall regulon [4]. Mutational alteration of these global regulators and two-component systems (the EnvZ/OmpR two-component signal transduction system in *E. coli*) was frequent and had important effects on evolution [86]. Additionally, comparative analysis of daptomycin-susceptible and -resistant isolates of *S. aureus* by Boyle-Vavra and coworkers revealed a point mutation in the multiple peptide resistance factor, *mprF*, encoding lysyl-phosphatidylglycerol transferase, which prevented cell disruption by a decreased daptomycin interaction with the cell surface [77].

To identify mutations conferring daptomycin resistance, the genome of laboratory-derived daptomycin-resistant *S. aureus* was compared with that of its daptomycin-susceptible parent. Single point mutations occurred in *mprF* (encoding lysyl-phosphatidylglycerol synthetase), *yycG* (encoding histidine kinase), and *rpoB* and *rpoC* (encoding subunits of RNA polymerase). Sequence analysis of the same set of genes from clinical isolates also found an association between mutations in *mprF* and *yycG* and increased resistance to daptomycin in the hospital setting [87].

SNPs & pathogen invasion & persistence

Emergence of the small colony variant phenotype (small colony size, reduced growth rate and appearance of auxotrophy) in Staphylococci is associated with host cell invasion and persistence, and antimicrobial susceptibility [88]. *S. aureus* isolates obtained from a patient with recurrent bacteremia demonstrated significant resistance to antibiotics and small colony variant phenotype [89]. Comparison between the parental and evolved resistant strains showed that four point mutations were sufficient to produce persistence during chronic infection and ensure survival to multiple antimicrobial chemotherapies (linezolid, rifampicin and ciprofloxacin). A particular substitution in *relA* reduced RelA hydrolase activity and triggered accumulation of the intracellular signaling molecule (p)ppGpp, giving rise to the bacterial stringent response. Additionally, the evolved strain harbored other mutations in *rpoB* and *parC* (encoding topoisomerase IV) leading to a reduction in antimicrobial susceptibility and the promotion of bacterial persistence [89].

SNPs & emergence of mutator phenotypes

Each CF patient is typically infected by a *P. aeruginosa* strain that persists throughout their life [90]. Cramer *et al.* reported genomic microevolution in the two most common clonal complexes (C and PA14) that were isolated from airways of CF patients over a period of more than 20 years [91]. The PA14 lineage had 15 nucleotide changes and a large deletion. No mutations occurred in clone C during the first 3 years in CF lungs, but 959 polymorphisms were observed 15 years later [91]. More than 98% of these SNPs were transitions. This high number of mutations was likely due to a mutation in the *mutL* gene, resulting in an increased mutation rate. Hypermutable strains seemed to be frequent in *P. aeruginosa*, as well as in *S. aureus* and *Haemophilus influenzae* [92,93]. Mutations in the DNA mismatch repair system occurred more frequently in bacteria isolated from patients with chronic diseases, compared with those with acute infections [94]. Moreover, it was shown that mutators lost virulence factors, but not resistance to antibiotics [95].

Mutation rate & *in vivo* parallelism

According to the CF study by Mena *et al.*, *P. aeruginosa* nonmutator lineages gained 0.25 mutations/year, whereas mutator strains acquired more than three mutations/year of infection [96]. Comparing *P. aeruginosa* clonal isolates of CF airways, Yang *et al.* found that adaptation was rapid in the first 20,000 generations then reached a plateau between 50,000 and

200,000 generations [97]. *P. aeruginosa* populations were highly genetically homogeneous, despite the complexity and variability of their surrounding ecosystem [97]. However, several studies reported that the spread of new favorable alleles increased the genetic diversity. An experimental coevolution study between the nematode *Caenorhabditis elegans* hosting the pathogen *Bacillus thuringiensis* during 48 host generations revealed that, for both organisms, genetic diversity increased during coevolution [98,99]. Recently, Young *et al.* studied evolutionary dynamics of *S. aureus* by comparing strains from healthy nasal carriage with isogenic isolates causing fatal bloodstream infection [100]. Authors studied clonal *S. aureus* strains of a patient who developed disease 15 months after joining the study. Results revealed that the number of mutations in nasally carried staphylococci increased slowly with 2.7 mutations/megabase/year [100]. This result is consistent with the rates of spontaneous mutations for bacterial genomes: 0.003 mutations per genome per replication, as calculated by Drake *et al.* [101]. In conclusion, only eight mutations were associated with the progression from asymptomatic to bloodstream infection. Interestingly, 50% of these mutations were premature stop codons causing early termination of translation [102].

Another question raised in several studies was related to the reproducibility of evolution [5,57,103]. An evolutionary experiment in *P. aeruginosa* by Huse *et al.* (39,000 generations were achieved) confirmed parallel evolution in CF lungs [103]. However, the mutation rate strongly varied between isolates. For example, mutation rates of clonal *H. pylori* strains isolated from four chronically infected patients from the same geographical area varied from 0.5 to 6.5 × 10^{-5}/site/year [57]. Mutation rates in *Mycobacterium tuberculosis* isolates from sick macaques during latency and reactivated states of disease were 2–3 × 10^{-10} mutations/bp/generation [5].

In conclusion, *in vivo* ecosystems constitute complex and highly dynamic environments composed of numerous niches. The infecting bacterium should rather be considered as a population that is not uniformly exposed to an identical stress, but to numerous microenvironments that could trigger and select for a variety of adaptive responses [104].

Coevolution

Recently, the Black Queen hypothesis was proposed by Morris and colleagues [105,106] based on the inability of the ocean-dwelling bacterium *Prochlorococcus*, which exclusively grows in the presence of other species, to grow in the laboratory. According to this hypothesis, organisms can lose essential functions when there is another species present to perform that function. Another example would be host–bacteria coevolution, where bacteria encounter challenges such as the host immune system, accessing nutrients, coping with drug treatments and the competing with other microbiota.

After decades of asymptomatic carriage, the colonizing *S. aureus* strain can evolve to a highly virulent organism causing severe disease and even death [107]. According to Wertheim *et al.*, nasal carriage of staphylococci increased the risk of auto-infection by a factor of three for hospitalized patients; however, the mortality rate was higher among the noncarriers with bacteremia [107,108]. It was reported that more than 80% of *S. aureus* strains isolated from blood of bacteremic patients were nearly identical to those in the anterior nares of the same patients [109], suggesting origin from nasal colonies. The immune response is one of the strongest evolutionary forces that drives adaptation in infectious bacteria. For example, evolution from acute to chronic *P. aeruginosa* infection was associated with a reduction of immunogenic factors such as secreted virulence factors, O-antigen expression and bacterial motility [62,63,110]. Along with the immune response, drug therapy leads to an important selection pressure for invading bacteria. Recently, Zdziarski *et al.* reported that the host can influence bacterial genome evolution and, therefore, personalize its microbial flora [111–113]. Comparison of genomes of an asymptomatic *E. coli* bacterium and its descendants after therapeutic bladder colonization was consistent with adaptive evolution specific to each host environment [111]. Often, commensal flora compete with pathogens and can protect the host from colonization by pathogens and their associated damages [114]. However, it was also demonstrated that pathogens could obtain genomic material from commensal cells that enabled colonization and persistence [115]. Moreover, some interactions between commensals and pathogens led to the synergistic enhancement of virulence [116]. Niche competition also occurs between conspecific pathogens, where ancestral and evolved populations compete for resources [45] and where horizontal gene transfers could arise [76]. Thus, clonal competition often results in expansion and persistence of epidemic clones [117]. Cases of interspecies pathogen interference have been also described. An inverse

relationship between *S. aureus* and *Streptococcus pneumoniae* colonization was identified in the nasopharyngeal carriage [118,119]. These studies show that polymicrobial community dynamics can be an important determinant of infection. Additionally, bacteria themselves are a host for bacteriophages and need to continuously evolve to escape phage infection [120].

Future perspective

- By considering the entire genome rather than targeted genes, whole genome sequencing will contribute to more in depth analyses of evolution inside the bacterial genome;
- Progress and advances in metagenomic tools will lead to analyses of genome evolution in the context of the surrounding bacterial ecosystem;
- Appreciating the way in which genomic changes affect transcriptomic and proteomic variations will enable estimations of global adaptation in particular scenarios;
- Single cell genomics and transcriptomics will allow the evaluation of genomic change at the level of individual organisms, rather than the population;
- A better understanding of pathogen genome evolution will come from evaluating competition and interaction between biotic and abiotic components of an ecosystem (e.g., the host immune system);
- Despite remarkable progress in the genomics of bacteria, the mechanisms of virulence acquisition, whether through genome reduction or horizontal gene transfer remain unclear for many pathogens;
- Advances in functional genomics will provide understanding of bacterial antibiotic resistance mechanisms not limited to acquisition of specific resistance genes;
- The complexity of problems posed by evolution is enormous and it seems clear that integrative approaches are required to limit pathogen burden, unify cellular and molecular biology, for statistical analysis and for physical modeling;
- Advances in high-throughput sequencing technologies has led to an exponentially growing number of sequenced bacterial genomes, providing diversity of bacterial strains within numerous species. However, the success will depend on the reliability of analysis and the accuracy of queries.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

New approaches for genome evolution studies

- Recently developed transcriptomic, genomic and proteomic tools have helped guide our understanding of the mechanisms underlying bacterial adaptation in different environments. However, many important questions are still unresolved.

Microorganisms in the exploration of genome evolution & adaptation

- Microorganisms offer a powerful tool for testing evolutionary hypotheses in the laboratory in order to explore fundamental principles of evolution. *In vivo* studies are becoming more common as the biases associated with artificial *in vitro* settings become apparent.

Antibiotic selection pressures & emergence of resistance

- Acquisition of antibiotic resistance by bacteria exacerbates major human health challenges and therefore represents a pivotal issue for modern society. Identifying the genetic basis of adaptation progressively allows improvement of therapies, the design of new preventive measures and will enable personalized treatment of resistant bacterial infections.
- In addition, the environment is rapidly changing. The widespread use of antibiotics and antibacterial sanitizers is progressively contributing to the evolution of microbial communities, by selecting for resistant bacteria that can survive current conditions.

References

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1. Shendure J, Ji H. Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145 (2008).
2. Barrick JE, Yu DS, Yoon SH *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461, 1243–1247 (2009).

▪ **Relative fitness improvement and genomic mutation acquisition during 20,000 *Escherichia coli* generations.**

3. Nusbaum C, Ohsumi TK, Gomez J *et al.* Sensitive, specific polymorphism discovery in bacteria using massively parallel sequencing. *Nat. Methods* 6, 67–69 (2009).
4. Mwangi MM, Wu SW, Zhou Y *et al.* Tracking the *in vivo* evolution of multidrug

resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc. Natl Acad. Sci. USA* 104, 9451–9456 (2007).

▪ **Contributes to the elucidation of the complex pathways of bacterial genome evolution during infection and acquisition of drug resistance.**

5. Ford CB, Lin PL, Chase MR *et al.* Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat. Genet.* 43, 482–486 (2011).
6. Darwin C. *On the Origin of the Species by Means of Natural Selection: or, the Preservation of Favoured Races in the Struggle for Life*. John Murray, UK (1859).

▪▪ **Revolution of the evolution theory.**

7. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4, 457–469 (2003).

▪▪ **First and longest long-term *in vitro* evolution study.**

8. Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Second-order selection for evolvability in a large *Escherichia coli* population. *Science* 331, 1433–1436 (2011).
9. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* 138, 1315–1341 (1991).
10. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332, 1193–1196 (2011).
11. Chou HH, Chiu HC, Delaney NF, Segrè D, Marx CJ. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332, 1190–1192 (2011).
12. Cooper VS, Lenski RE. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407, 736–739 (2000).
13. Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, UK (1985).
14. Nakatsu CH, Korona R, Lenski RE, de Bruijn FJ, Marsh TL, Forney LJ. Parallel and divergent genotypic evolution in experimental populations of *Ralstonia sp. J. Bacteriol.* 180, 4325–4331 (1998).
15. Losos J, Jackman T, Larson A, Queiroz K, Rodriguez-Schettino L. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279, 2115–2118 (1998).
16. Crandall KA, Kelsey CR, Imamichi H, Lane HC, Salzman NP. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate ratio to detect selection. *Mol. Biol. Evol.* 16, 372–382 (1999).
17. Reid SD, Herbelin CJ, Bumbaugh AC, Selander RK, Whittam TS. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406, 64–67 (2000).
18. Cunningham CW, Jeng K, Husti J *et al.* Parallel molecular evolution of deletions and nonsense mutations in bacteriophage T7. *Mol. Biol. Evol.* 14, 113–116 (1997).
19. Stanek MT, Cooper TF, Lenski RE. Identification and dynamics of a beneficial mutation in a long-term evolution experiment with *Escherichia coli*. *BMC Evol. Biol.* 9, 302 (2009).
20. Lenski RE, Winkworth CL, Riley MA. Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. *J. Mol. Evol.* 56, 498–508 (2003).
21. Cooper VS, Schneider D, Blot M, Lenski RE. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J. Bacteriol.* 183, 2834–2841 (2001).
22. Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D. Long-term experimental evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics* 169, 523–532 (2005).
23. Schneider R, Travers A, Kutateladze T, Muskhelishvili G. A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol. Microbiol.* 34, 953–964 (1999).
24. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 103, 9107–9112 (2006).
25. Cooper TF, Rozen DE, Lenski RE. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 100, 1072–1077 (2003).
26. Pelosi L, Kuhn L, Guetta D *et al.* Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* 173, 1851–1869 (2006).
27. Lenski RE, Mongold JA. Cell size, shape, and fitness in evolving populations of bacteria. In: *Scaling in Biology*. Oxford University Press, UK, 221–235 (2000).
28. Lieberman TD, Michel JB, Aingaran M *et al.* Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat. Genet.* 43, 1275–1280 (2011).
29. Lenski RE. Chance and necessity in the evolution of a bacterial pathogen. *Nat. Genet.* 43, 1174–1176 (2011).
30. Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387, 703–705 (1997).
31. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution. *Nature* 387, 700–702 (1997).
32. Mao EF, Lane L, Lee J, Miller JH. Proliferation of mutators in a cell population. *J. Bacteriol.* 179, 417–422 (1997).
33. Arjan JA, Visser M, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. Diminishing returns from mutation supply rate in asexual populations. *Science* 283, 404–406 (1999).
34. LeClerc JE, Li B, Payne WL, Cebula TA. High mutation frequencies among *Escherichia coli* and *Salmonella pathogens*. *Science* 274, 1208–1211 (1996).
35. Oliver A, Canton R, Campo P, Baquero F, Blazquez J. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* 288, 1251–1253 (2000).
36. Travisano M, Lenski RE. Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. *Genetics* 143, 15–26 (1996).
37. Travisano M, Mongold JA, Bennett AF, Lenski RE. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267, 87–90 (1995).
38. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 105, 7899–7906 (2008).
39. Leiby N, Harcombe WR, Marx CJ. Multiple long-term, experimentally-evolved populations of *Escherichia coli* acquire dependence upon citrate as an iron chelator for optimal growth on glucose. *BMC Evol. Biol.* 12, 151 (2012).
40. Dekel E, Alon U. Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436, 588–592 (2005).
41. Meyer JR, Agrawal AA, Quick RT, Dobias DT, Schneider D, Lenski RE. Parallel changes in host resistance to viral infection during 45,000 generations of relaxed selection. *Evolution* 64, 3024–3034 (2010).

42. Lenski RE, Levin BR. Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *Am. Nat.* 125, 585–602 (1985).
43. Cui L, Iwamoto A, Lian JQ *et al.* Novel mechanism of antibiotic resistance originating in vancomycin-intermediate *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 50, 428–438 (2006).
44. Canton R, Morosini MI. Emergence and spread of antibiotic resistance following exposure to antibiotics. *FEMS Microbiol. Rev.* 35, 977–991 (2011).
45. Zambrano MM, Siegele DA, Almiron M, Tormo A, Kolter R. Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science* 259, 1757–1760 (1993).
46. Notley-McRobb L, King T, Ferenci T. *rpoS* mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. *J. Bacteriol.* 184, 806–811 (2002).
47. Ferenci T. What is driving the acquisition of *mutS* and *rpoS* polymorphisms in *Escherichia coli*? *Trends Microbiol.* 11, 457–461 (2003).
48. King T, Ishihama A, Kori A, Ferenci T. A regulatory trade-off as a source of strain variation in the species *Escherichia coli*. *J. Bacteriol.* 186, 5614–5620 (2004).
49. Hansen SK, Rainey PB, Haagensen JA, Molin S. Evolution of species interactions in a biofilm community. *Nature* 445, 533–536 (2007).
50. Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. Experimental evolution. *Trends Ecol. Evol.* 27, 547–560 (2012).
51. Jones KE, Patel NG, Levy MA *et al.* Global trends in emerging infectious diseases. *Nature* 451, 990–993 (2008).
52. Govan JR, Balandreau J, Vandamme P. *Burkholderia cepacia* – friend and foe. *ASM News* 66, 124–125 (2000).
53. Kazakova SV, Hageman JC, Matava M *et al.* A clone of methicillin-resistant *Staphylococcus aureus* among professional football players. *N. Engl. J. Med.* 352, 468–475 (2005).
54. Siefert JL. Defining the mobilome. *Methods Mol. Biol.* 532, 13–27 (2009).
55. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304 (2000).
56. Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc. Natl Acad. Sci. USA* 98, 8821–8826 (2001).
57. Kennemann L, Didelot X, Aebischer T *et al.* *Helicobacter pylori* genome evolution during human infection. *Proc. Natl Acad. Sci. USA* 108, 5033–5038 (2011).
58. Treangen TJ, Abraham AL, Touchon M, Rocha EPC. Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiol. Rev.* 33, 539–571 (2009).
59. Lindroos H, Vinnere O, Mira A, Repsilber D, Naslund K, Andersson SGE. Genome rearrangements, deletions, and amplifications in the natural population of *Bartonella henselae*. *J. Bacteriol.* 188, 7426–7439 (2006).
60. Wu S, Piscitelli C, de Lencastre H, Tomasz A. Tracking the evolutionary origin of the methicillin resistance gene: cloning and sequencing of a homologue of *mecA* from a methicillin susceptible strain of *Staphylococcus sciuri*. *Microb. Drug Resist.* 2, 435–441 (1996).
61. Maurelli AT, Fernandez RE, Bloch CA, Rode CK, Fasano A. 'Black holes' and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 95, 3943–3948 (1998).
62. Yang L, Jelsbak L, Molin S. Microbial ecology and adaptation in cystic fibrosis airways. *Environ. Microbiol.* 13, 1682–1689 (2011).

▪ **Provides insights into the genomic adaptation of an opportunistic pathogen progressing from acute to chronic infections.**

63. Smith EE, Buckley DG, Wu Z *et al.* Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl Acad. Sci. USA* 103, 8487–8492 (2006).
64. Huang JQ, Zheng GF, Sumanac K, Irvine EJ, Hunt RH. Meta-analysis of the relationship between *cagA* seropositivity and gastric cancer. *Gastroenterology* 125, 1636–1644 (2003).
65. Blaser MJ. Hypothesis: the changing relationships of *Helicobacter pylori* and humans: implications for health and disease. *J. Infect. Dis.* 179, 1523–1530 (1999).
66. Ye W, Held M, Lagergren J *et al. Helicobacter pylori* infection and gastric atrophy: risk of adenocarcinoma and squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *J. Natl Cancer Inst.* 96, 388–396 (2004).
67. Oh JD, Kling-Backhed H, Giannakis M *et al.* The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. *Proc. Natl Acad. Sci. USA* 103, 9999–10004 (2006).
68. Kraft C, Stack A, Josenhans C *et al.* Genomic changes during chronic *Helicobacter pylori* infection. *J. Bacteriol.* 188, 249–254 (2006).
69. Louarn JM, Bouché JP, Legendre F, Louarn J, Patte J. Characterization and properties of very large inversions of the *E. coli* chromosome along the origin-to-terminus axis. *Mol. Gen. Genet.* 201, 467–476 (1985).
70. Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M. Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. *Genetics* 156, 477–488 (2000).
71. Menard A, de Los Santos PE, Graindorge A, Cournoyer B. Architecture of *Burkholderia cepacia* complex sigma70 gene family: evidence of alternative primary and clade-specific factors, and genomic instability. *BMC Genomics* 8, 308 (2007).
72. Ziebuhr W, Dietrich K, Trautmann M, Wilhelm M. Chromosomal rearrangements affecting biofilm production and antibiotic resistance in a *Staphylococcus epidermidis* strain causing shunt-associated ventriculitis. *Int. J. Med. Microbiol.* 290, 115–120 (2000).
73. Song H, Hwang J, Yi H *et al.* The early stage of bacterial genome-reductive evolution in the host. *PLoS Pathog.* 6, e1000922 (2010).
74. Schneider D, Lenski RE. Dynamics of insertion sequence elements during experimental evolution of bacteria. *Res. Microbiol.* 155, 319–327 (2004).
75. Dhar N, McKinney JD. *Mycobacterium tuberculosis* persistence mutants identified by screening in isoniazid-treated mice. *Proc. Natl Acad. Sci. USA* 107, 12275–12280 (2010).
76. Goerke C, Papenberg S, Dasbach S *et al.* Increased frequency of genomic alterations in *Staphylococcus aureus* during chronic infection is in part due to phage mobilization. *J. Infect. Dis.* 189, 724–734 (2004).
77. Boyle-Vavra S, Jones M, Gourley BL *et al.* Comparative genome sequencing of an isogenic pair of USA800 clinical methicillin-resistant *Staphylococcus aureus* isolates obtained before and after daptomycin treatment failure. *Antimicrob. Agents Chemother.* 55, 2018–2025 (2011).
78. Betley MJ, Mekalanos JJ. Staphylococcal enterotoxin A is encoded by phage. *Science* 229, 185–187 (1985).

79. Kaneko J, Kimura T, Narita S, Tomita T, Kamio Y. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage phiPVL carrying Panton–Valentine leukocidin genes. *Gene* 215, 57–67 (1998).
80. Bae T, Baba T, Hiramatsu K, Schneewind O. Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. *Mol. Microbiol.* 62, 1035–1047 (2006).
81. McAdam PR, Holmes A, Templeton KE, Fitzgerald JR. Adaptive evolution of *Staphylococcus aureus* during chronic endobronchial infection of a cystic fibrosis patient. *PLoS ONE* 6, e24301 (2011).
82. Malachowa N, Deleo FR. Mobile genetic elements of *Staphylococcus aureus*. *Cell. Mol. Life Sci.* 67, 3057–3071 (2010).
83. Jensen SO, Lyon BR. Genetics of antimicrobial resistance in *Staphylococcus aureus*. *Future Microbiol.* 4, 565–582 (2009).
84. Lindsay JA, Holden MTG. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol.* 12, 378–385 (2004).
85. Nubel U, Dordel J, Kurt K *et al.* A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. *PLoS Pathog.* 6, e1000855 (2010).
86. Giraud A, Arous S, De Paepe M *et al.* Dissecting the genetic components of adaptation of *Escherichia coli* to the mouse gut. *PLoS Genet.* 4, e2 (2008).
87. Friedman L, Alder JD, Silverman JA. Genetic changes that correlate with reduced susceptibility to daptomycin in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 50, 2137–2145 (2006).
88. Proctor RA, von Eiff C, Kahl BC *et al.* Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. *Nat. Rev. Microbiol.* 4, 295–305 (2006).
89. Gao W, Chua K, Davies JK *et al.* Two novel point mutations in clinical *Staphylococcus aureus* reduce linezolid susceptibility and switch on the stringent response to promote persistent infection. *PLoS Pathog.* 6, e1000944 (2010).
90. Romling U, Fiedler B, Bosshammer J *et al.* Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic fibrosis. *J. Infect. Dis.* 170, 1616–1621 (1994).
91. Cramer N, Klockgether J, Wrasman K, Schmidt M, Davenport CF, Tummler B. Microevolution of the major common *Pseudomonas aeruginosa* clones C and PA14 in cystic fibrosis lungs. *Environ. Microbiol.* 13, 1690–1704 (2011).
92. Prunier AL, Malbruny B, Laurans M, Brouard J, Duhamel J, Leclercq R. High rate of macrolide resistance in *Staphylococcus aureus* strains from patients with cystic fibrosis reveals high proportions of hypermutable strains. *J. Infect. Dis.* 187, 1709–1716 (2003).
93. Roman F, Canton R, Pérez-Vazquez M, Baquero F, Campos J. Dynamics of long-term colonization of respiratory tract by *Haemophilus influenzae* in cystic fibrosis patients shows a marked increase in hypermutable strains. *J. Clin. Microbiol.* 42, 1450–1459 (2004).
94. Oliver A, Mena A. Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin. Microbiol. Infect.* 16, 798–808 (2010).
95. Hogardt M, Hoboth C, Schmoldt S, Henke C, Bader L, Heesemann J. Stage-specific adaptation of hypermutable *Pseudomonas aeruginosa* isolates during chronic pulmonary infection in patients with cystic fibrosis. *J. Infect. Dis.* 195, 70–80 (2007).
96. Mena A, Smith EE, Burns JL *et al.* Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *J. Bacteriol.* 190, 7910–7917 (2008).
97. Yang L, Jelsbak L, Marvig RL *et al.* Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl Acad. Sci. USA* 108, 7481–7486 (2011).

▪ **Provides a better understanding of the evolutionary dynamics and mutation rates of an opportunistic pathogen in a natural environment.**

98. Brockhurst MA, Rainey PB, Buckling A. The effect of spatial heterogeneity and parasites on the evolution of host diversity. *Proc. Biol. Sci.* 271, 107–111 (2004).
99. Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proc. Natl Acad. Sci. USA* 107, 7359–7364 (2010).
100. Young BC, Golubchik T, Batty EM *et al.* Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. *Proc. Natl Acad. Sci. USA* 109, 4550–4555 (2012).
101. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics* 148, 1667–1686 (1998).
102. Yang J, Tauschek M, Robins-Browne RM. Control of bacterial virulence by AraC-like regulators that respond to chemical signals. *Trends Microbiol.* 19, 128–135 (2011).
103. Huse HK, Kwon T, Zlosnik JEA, Speert DP, Marcotte EM, Whiteley M. Parallel evolution in *Pseudomonas aeruginosa* over 39,000 generations *in vivo*. *mBio* 1, e00199-10 (2010).
104. Hoboth C, Hoffmann R, Eichner A *et al.* Dynamics of adaptive microevolution of hypermutable *Pseudomonas aeruginosa* during chronic pulmonary infection in patients with cystic fibrosis. *J. Infect. Dis.* 200, 118–130 (2009).
105. Morris JJ, Lenski RE, Zinser ER. The Black Queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio* 3, e00036-12 (2012).
106. Sachs JL, Hollowell AC. The origins of cooperative bacterial communities. *mBio* 3, e00099-12 (2012).
107. van Belkum A, Melles DC, Nouwen J *et al.* Co-evolutionary aspects of human colonisation and infection by *Staphylococcus aureus*. *Infect. Genet. Evol.* 9, 32–47 (2009).
108. Wertheim HF, Vos MC, Ott A *et al.* Risk and outcome of nosocomial *Staphylococcus aureus* bacteraemia in nasal carriers versus non-carriers. *Lancet* 364, 703–705 (2004).
109. von Eiff C, Becker K, Machka K, Stammer H, Peters G. Nasal carriage as a source of *Staphylococcus aureus* bacteremia. Study Group. *N. Engl. J. Med.* 344, 11–16 (2001).
110. Bianconi I, Milani A, Cigana C *et al.* Positive signature-tagged mutagenesis in *Pseudomonas aeruginosa*: tracking patho-adaptive mutations promoting airways chronic infection. *PLoS Pathog.* 7, e1001270 (2011).
111. Zdziarski J, Brzuszkiewicz E, Wullt B *et al.* Host imprints on bacterial genomes – rapid, divergent evolution in individual patients. *PLoS Pathog.* 6, e1001078 (2010).
112. Harris SR, Feil EJ, Holden MTG *et al.* Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474 (2010).
113. Devi SM, Ahmed I, Khan AA *et al.* Genomes of *Helicobacter pylori* from native Peruvians suggest admixture of ancestral and modern lineages and reveal a western type *cag*-pathogenicity island. *BMC Genomics* 7, 191 (2006).
114. Shinefield HR, Wilsey JD, Ribble JC, Boris M, Eichenwald HF, Dittmar CI. Interactions of staphylococcal colonization. Influence of normal nasal flora and antimicrobials on inoculated *Staphylococcus aureus* strain 502A. *Am. J. Dis. Child.* 111, 11–21 (1966).
115. Stecher B, Robbiani R, Walker AW *et al.* *Salmonella enterica* serovar *typhimurium* exploits inflammation to compete with the intestinal microbiota. *PLoS Biol.* 5, 2177–2189 (2007).

116. Sibley CD, Duan K, Fischer C *et al.* Discerning the complexity of community interactions using a *Drosophila* model of polymicrobial infections. *PLoS Pathog.* 4, e1000184 (2008).

117. Jelsbak L, Johansen HK, Frost AL *et al.* Molecular epidemiology and dynamics of *Pseudomonas aeruginosa* populations in lungs of cystic fibrosis patients. *Infect. Immun.* 75, 2214–2224 (2007).

118. McNally LM, Jeena PM, Gajee K *et al.* Lack of association between the nasopharyngeal carriage of *Streptococcus pneumoniae* and *Staphylococcus aureus* in HIV-1-infected South African children. *J. Infect. Dis.* 194, 385–390 (2006).

119. Bogaert D, van Belkum A, Sluijter M *et al.* Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *Lancet* 363, 1871–1872 (2004).

120. Gomez P, Buckling A. Bacteria–phage antagonistic coevolution in soil. *Science* 332, 106–109 (2011).

121. Francois P, Huyghe A, Charbonnier Y *et al.* Use of an automated multiple-locus, variable-number tandem repeat-based method for rapid and high-throughput genotyping of *Staphylococcus aureus* isolates. *J. Clin. Microbiol.* 43, 3346–3355 (2005).