

Chapter 12

Clonally Evolving Pathogenic Bacteria



Sofia Hauck and Martin C. J. Maiden

1 Clonality in Bacteria

A population is a group of organisms that occupy the same niche and that share an evolutionary path, adapting to their environment through the process of natural selection. Members of a population are closely related and united by a recent common ancestor or ancestral population. Despite these similarities, bacterial populations are rarely composed of single clones that are genetically identical. Indeed, genetic variation within a population is a necessary condition for evolution by means of natural selection to occur. A population of identical clones cannot adapt, as all of its members have genomes that are equally suited to any set of conditions (Nielsen 2005).

Natural selection results in changes in the frequency of genotypes according to their fitness, as higher rates of reproduction and survival increase the abundance of the fittest genotypes within the population. Likewise, lower rates of success reduce the frequency of the least fit genotypes, sometimes to the point of elimination. This process therefore reduces the genetic diversity of the population over time, as it can only maintain or decrease diversity. Diversity is not completely reduced by natural selection because it is balanced by the introduction of new variants through mutation and recombination. All genetic variation ultimately arises from mutations, but sexual processes can introduce mutations from other populations via horizontal gene

S. Hauck · M. C. J. Maiden (✉)
Department of Zoology, University of Oxford, Oxford, UK
e-mail: martin.maiden@zoo.ox.ac.uk

© Springer International Publishing AG, part of Springer Nature 2018
P. H. Rampelotto (ed.), *Molecular Mechanisms of Microbial Evolution*, Grand Challenges in Biology and Biotechnology,
https://doi.org/10.1007/978-3-319-69078-0_12

307

transfer (HGT) as well as causing the reassortment of genes within a population (Feil and Spratt 2001).

Sexual recombination in bacteria differs from that of plants and animals in both frequency and range of source genetic material. Bacterial HGT occurs widely and promiscuously, with genetic material sourced throughout the domain, rather than exclusively from other members of the same species (Cohan 2002). Unlike most plants and animals, which recombine as a matter of course during each reproductive cycle, HGT in bacteria occurs at rates and extents that vary widely by species and by ecological situation (Hanage et al. 2006a). Across the bacterial domain, a wide spectrum of recombination rates has been observed, from very high rates that erase most clonal signal to complete clonality (Spratt and Maiden 1999). In a fundamentally non-clonal population, genes are the evolutionary units, as recombination is so frequent that it overwhelms the non-random associations among genetic variants in a genome, collectively named “linkage disequilibrium.” A recombination rate that is ten times higher than the mutation rate may be enough for the linkage caused by shared descent to break down entirely (Cohan 1994). When this is the case, each variant of a particular gene can exist in a range of genomic backgrounds and therefore evolve independently. This is not an hypothetical extreme as it occurs in *Helicobacter pylori*, a human-specific pathogen that causes chronic gastric infections (Suerbaum et al. 1998), which has been found to exchange up to a tenth of its genome in one multi-strain infection over a period of 4 years (Cao et al. 2015).

At the opposite end of the spectrum is full clonality, which is a consequence of limited or no HGT (Tibayrenc and Ayala 2002). HGT can occur occasionally without affecting a clonal population structure, as long as it is infrequent enough to remain beneath the “clonality threshold.” This is the point at which recombination is too rare to act as a homogenising force within the population and prevent the divergence of distinct branches (Tibayrenc and Ayala 2015). In the extreme and unusual case of a fully clonal population, where no recombination can be observed, genes cannot be exchanged among individuals and therefore each genome evolves independently. The phylogeny of a population in these conditions fits the “Russian doll” model, where, with increasing resolution, each branch of the species phylogeny is subdivided into smaller distinct and diverging branches (Tibayrenc and Ayala 2016).

2 The Clonal Paradigm

This article will focus on fully clonal populations, where HGT is entirely absent or sufficiently rare that it has essentially no impact on the evolution, population structure, and ecology of the organism in question during the timeframe under consideration (Shapiro 2016). The first major consequence of limited or no HGT is the inability to repair mutations that arise from time to time. Bacteria can use a variety of mechanisms for repairing DNA damage, including nucleotide or base excision repair and mismatch repair. These mechanisms can lower the rate of mutation by reversing damage as it happens, but without HGT, bacteria cannot

repair mutations that have been copied to both stands of DNA (van der Veen and Tang 2015). Two other major consequences follow from a lack of HGT: the absence of a mechanism whereby different mutations can be combined to make novel variants and the inability to acquire genetic material from other organisms. All three of these affect the diversity of a population. In the absence of HGT, a population is both limited in its ability to generate new genetic diversity and to maintain existing diversity, as it is frequently subject to purges (Namouchi et al. 2012) by means of bottlenecks when population sizes are small or genome-wide sweeps when selection is strong (Smith et al. 2006). This lack of diversity has consequences for the adaptability and long-term viability of a population.

Populations that do not participate in HGT typically exhibit very low genetic diversity, although a lack of diversity is also a feature of relatively young populations, as over time mutations will accumulate even in the absence of HGT. The term “clonal” can describe both of these traits, the genealogical concept of reproduction through clonal descent and the genetic concept of identical clones forming all members of a population. This state of low diversity within the population is referred to as genetic monomorphism. The confusion between these definitions is particularly pronounced because these two traits often, but not necessarily, co-exist in bacteria (Tibayrenc and Ayala 2012).

From the perspective of evolutionary and molecular epidemiological studies, the low diversity of monomorphic bacteria can present problems for the identification of epidemiologically relevant variants for practical purposes. When typing resolution is not sufficient to distinguish similar isolates, the crucial task of tracking pathogen transmission becomes more challenging (Kohl et al. 2014). Improvements in the speed, cost, and accuracy of whole-genome sequencing (WGS) made in the 10 years prior to writing this article ensured that these bacteria could be examined at an unprecedented level of detail, covering virtually all of the existing genetic variation. Consequently, fully clonal, monomorphic pathogen species have been some of the first to be analysed in-depth by this technology, precisely because their low diversity simplifies large-scale genomics studies (Achtman 2008). This research has made it possible to discover the common traits of fully clonal bacteria.

Most known examples of highly clonal pathogens are lineages nested within a more diverse phylogeny of nonpathogenic organisms (Achtman 2008). An example of this is the agent of tuberculosis, *Mycobacterium tuberculosis*, which is perhaps the best studied example of a fully clonal bacterium. It is one of nine lineages within the *Mycobacterium tuberculosis* complex, each of which is host-specific, though some spillover in hosts does exist (Ghodbane et al. 2014). The complex as a whole is considered to be one species, due to its high degree of relatedness (99% or more average nucleotide identity), but individual species names are still used to refer to each subspecies, or ecotype, within it (Djelouadji et al. 2011). There are a number of parallel examples elsewhere in the domain including the following: *Yersinia pestis*, the agent of plague, which is best considered a specific clone of *Yersinia pseudotuberculosis* as it is nested within this more diverse species (Rasmussen et al. 2015);

Bordetella pertussis, which causes whooping cough, is a monophyletic branch of the *B. bronchiseptica* species (Diavatopoulos et al. 2005); and *Bacillus anthracis*, responsible for anthrax, is an equivalent branch of the *B. cereus* phylogeny (Okinaka et al. 2006). Not all clonal populations are given species names, as can be observed with *Salmonella enterica* subspecies *enterica* serovar Typhi, the agent of typhoid fever (Didelot et al. 2011).

In the absence of wholly sexual reproduction, which is used to define animal species, and the widespread occurrence of HGT in bacteria from a variety of sources, precise definitions of bacterial species remain difficult, contingent, and to a degree controversial (Achtman and Wagner 2008). For pathogenic bacteria, the importance of accurately circumscribing disease-causing microbes has led to pathogenicity being used as the defining trait of a species. However, pathogenicity, like many other phenotypic traits, correlates poorly with genotype, as is becoming increasingly clear as genotyping methods improve. A species name that may not be “accurate,” that is, not defining a distinct monophyletic cluster of genotypic and phenotypic traits, may nonetheless still be used in cases where distinguishing dangerous strains from harmless genetic relatives can be a life-or-death concern. This concept of a *nomen periculosum* is one example of the many complications in the field of bacterial taxonomy (Stackebrandt et al. 2002).

The complex nature of phenotypic and genotypic clusters in bacteria, both in defining clear clusters by either set of criteria or in correlating genotype to phenotype, has meant that a universal bacterial species concept is yet to be achieved. While the search for a unifying definition continues, “species” remains a term without one particular meaning, largely used as a label of convenience (Bapteste and Boucher 2009). Lineages are given names “dependent on the level of divergence thought appropriate” and only “where this serves a useful purpose” (Lan and Reeves, 2001). Clonal populations present particular problems in defining species clusters, as the absence of recombination means that each lineage is “irreversibly set” on divergence from other members of its species (Hanage et al. 2006b). For this reason, “species,” “clone,” “serovar,” or any other terms for a specific phylogenetic group of bacteria are used here as per the convention for that species.

Despite HGT not being a mechanism in the maintenance of clonal lineages, it is often key to the evolutionary events that have led to them becoming distinct from their parental populations. HGT events can be catalysts for the expansion of a population to a new niche and form critical junctions in the evolutionary path of a species (Wiedenbeck and Cohan 2011). A clear example of this has been observed in the history of *Bacillus anthracis*, the agent of anthrax. It diverged from *Bacillus cereus* after acquiring the pXO1 and pXO2 plasmids, which includes the anthrax toxin genes implicated in the high fatality rate of the disease (Kolstø et al. 2009). After this acquisition, which occurred perhaps 20,000 years ago, *B. anthracis*'s unique pathobiology has caused it to specialise into a novel ecological niche (Keim et al. 2009). The evolutionary history of *S. enterica* serovar Typhi displays a similar pattern. *S. enterica* serovar Typhi lost the promiscuous HGT that is characteristic of the *S. enterica* species, changed niche as it transformed from a

gastroenteric infection to an invasive one and underwent functional gene loss and genome degradation in the process (Holt et al. 2008).

Irreversible specialisation is the only possibility for fully clonal populations, due to the loss of adaptability that comes with restricted HGT (Moran and Wernegreen 2000). Parasitism itself is a form of specialisation, and the extreme host specificity and obligate pathogenicity evident in many fully clonal species is an even more extreme version of this specialization (Shapiro 2016). While the possibility of a discovery bias caused by the overrepresentation of pathogens in bacterial research cannot be discounted, it appears that all persistent fully clonal bacteria are pathogenic (Achtman 2012).

Many of the most deadly bacterial pathogens are species with low genetic diversity that fit the fully clonal paradigm (Namouchi et al. 2012). For example, the etiological agent of plague, *Y. pestis*, was responsible for the Black Death pandemic that killed nearly a third of Europeans in the fourteenth century, and this organism remains the cause of small but often fatal outbreaks in the modern era (Gage and Kosoy 2005). Meanwhile, *M. tuberculosis* is the most deadly bacterial pathogen in the world, infecting 100 million new patients and causing 1.5 million deaths a year, and appears to have coevolved with humans (Niemann et al. 2016). *S. enterica* serovar Typhi causes 200,000 deaths a year and may have been the cause of the plague of Athens in 430 B.C. (Galan 2016).

These species are genetically monomorphic despite their wide geographic range and long history as human pathogens, stretching over tens of thousands of years. While this is not substantial stretch of time in evolutionary terms, it is enough for analyses of mutation rate to be made across species and for phylogeographic signals to emerge within species (Comas et al. 2013). A comparison between modern *Y. pestis* strains and DNA recovered from a Black Death victim showed that few genetic changes occurred in the interceding 660 years, with only 10 cases of altered gene order and only 97 chromosomal sites and 6 plasmid single-nucleotide sites differing from modern strains (Bos et al. 2011). For *M. tuberculosis*, whole-genome alignments find that even the most distant strains that can be isolated from humans differ at most in 1800 single nucleotide polymorphisms or SNPs (Coscolla and Gagneux 2014). In contrast, over 15,000 SNP differences have been found between isolates of *Staphylococcus aureus* (Planet et al. 2017), and *Campylobacter jejuni* isolates can differ by 13,000 SNPs even within one outbreak (Moffatt et al. 2016).

Lifestyle stages with slow or no growth, common to many clonal pathogenic bacteria, may further exacerbate their low genetic diversity, by prolonging generation times and slowing evolutionary rates (Ohta 2011). *M. tuberculosis* can enter a granuloma phase in infected individuals, resulting in essentially asymptomatic carriage (Ford et al. 2011). *Bacillus anthracis* can survive as a spore in dry soil for decades, potentially centuries, in which time it does not replicate and therefore does not accumulate mutations (Rasko et al. 2011). This does not seem to be a necessary condition, however, as *Bordetella pertussis* is an obligate human pathogen with no chronic carrier state (Park et al. 2012). The source-sink evolutionary dynamics caused by carrier stages in their lifestyles may also reduce variability by stabilising the genome, as adaptations that are advantageous in the infectious phase confer no

fitness benefits for invasive disease. This may be the case in *Yersinia pestis*, which constantly circulates in rodent populations between its infrequent outbreaks in humans (Ayyadurai et al. 2008) and in *S. enterica* serovar Typhi, whose reservoir is human asymptomatic carriers that may shed the bacteria for decades over their lifetimes (Holt et al. 2008).

3 Limiting the Rate of Recombination

To date, pathogens with restricted host ranges are among the best studied clonally evolving populations. A host represents a highly controlled and relatively uniform environment, and the restriction to a given host range deepens the extreme level of specialization these populations have undergone. As specialization increases, newly introduced genes are more likely to disturb existing favorable interactions in the genome, an effect named recombinational load (Michod et al. 2008). Because of this, it is advantageous for clonally evolving bacteria to actively suppress their rate of HGT so as to reduce this load. The realised rate is at a maximum equivalent to the potential rate, which is determined by the available mechanisms for HGT, but which in practice is restricted by ecological conditions of the recipient bacteria (Yahara et al. 2016).

DNA can be transferred between bacteria by three mechanisms, all of which transfer a small amount of DNA from a donor to a recipient: transduction, conjugation, and transformation. Briefly, transduction involves infectious virions (most often bacteriophages) as vectors, while in conjugation the vectors are self-replicating plasmids or other mobile genetic elements. Bacteria can protect themselves against these parasitic elements by means of endonucleases, enzymes that cleave foreign DNA. With these enzymes, invading DNA sequences can be destroyed before or after they recombine with the bacterial chromosome and consequently prevent cases of successful HGT from being passed on to subsequent generations (Cohan 2002). Restriction endonucleases recognise and cleave DNA at specific “target” sequences, while the more recently discovered CRISPR-Cas system detects the clustered regularly interspaced short palindromic repeats (CRISPR) after which it is named. The high specificity of the targets in both restriction and CRISPR systems is key to their ability to protect bacteria from “foreign” DNA without damaging their own genomes.

In contrast, transformation depends on a recipient with functional and active competency machinery to complete the uptake of DNA from the environment. This requires a set of highly specialised loci that are often spread throughout the genome and which are conserved across a species (Croucher et al. 2016). Because expressing this machinery carries a high cost, competence is typically regulated so that it is transiently expressed only under specific circumstances. When the loss of one of the genes occurs, relaxed selection on the others ensures that they are quickly lost. These losses often signal a point of no return for the lineages in which they

occur, resulting in their extinction, albeit in the long-term perspective of hundreds of millions of years (Redfield et al. 2006).

Even in the presence of the possibility of HGT, for any of these mechanisms to succeed at facilitating transfer between two bacteria, the donor and recipient must be in close proximity to one another, so that the DNA may be passed from one to the other. Spatial isolation may therefore be a barrier to gene flow (Polz et al. 2013) and is the basis for the “starving sex hypothesis” that posits clonality as a passive process, occurring due to the absence of opportunity for mating (Tibayrenc and Ayala 2016). This may in fact be the case in many pathogenic bacteria, for which transmission events often mean only a small founding population infects each new host (Bergstrom et al. 1999).

The limitation of small founding populations can prevent the impact of recombination in yet another way: if the genetic distance between host and recipient is too small, the exchange of DNA is inconsequential and indeed unobservable. After DNA is exchanged, homologous recombination is a necessary step for genetic material to be incorporated into the genome and subsequently heritable. This requires that the donor DNA be paired with a chromosomal sequence with which it has some sequence identity, at least on its flanking sequences. If the donor sequence and host sequence are identical, homologous recombination may indeed be occurring, but the like-for-like exchange leaves no trace in a case of “invisible sex” (Tibayrenc and Ayala 2015). If both donor and recipient bacteria are descendants of a recent transmission event, they may be too closely related to recombine in ways that can be observed. This may in fact be the most common form of HGT within bacteria, a mechanism that is advantageous due to its utility in DNA repair (Feil and Spratt 2001).

With the high resolution possible by WGS analyses, data from multiple samples of a single patient have shown that multi-clonal infections are more common than previously thought (Votintseva et al. 2014). For *M. tuberculosis*, DNA recovered from human remains indicates that multiple strain infections were commonplace in eighteenth century Europe (Kay et al. 2015). It is also likely that any colonising bacteria are not alone in their niche, as every area of the human body has an associated microbiome (Huttenhower et al. 2012); however, physical proximity to distantly related bacteria may not lead to HGT as the frequency of transformation is reduced exponentially as sequence divergence increases. Therefore, the rate of successful HGT is reduced as the genetic distance between a host and a recipient increases and the homology between their genome sequences weakens (Cohan 1994). Genetic distance between bacteria can also lower the efficacy of plasmids and bacteriophages as vectors of HGT as their host ranges are limited to similar bacteria (Cohan 2002).

In this way, recombination may not occur, even when potential mechanisms are present and the donor and recipient bacteria are physically close, if the bacteria are either too distantly related (no HGT will occur) or too closely related (HGT is non-observable). HGT is a major force in the evolution of the bacteria domain as a whole, with estimates that up to 20% of genes across the domain have been recently mobilised by this method. Bacteria rapidly adapt to changes in their environment

using this method for gene flow to swap the genes in their accessory genome. In a fully clonal population, HGT is not available as a mechanism for increasing genetic diversity within a species, and so mutation remains as the only option (Feil and Spratt 2001).

4 Mutation and the Evolution of Clonal Pathogens

Mutations arise from DNA damage or copying errors and are therefore random with respect to gene function. Some gene features, such as homopolymeric tracts, may locally raise the rate of mutation and therefore be found primarily in genes where this effect is beneficial, but even mutations such as this occur randomly (Orsi et al. 2010). According to Ohta's nearly neutral theory model, the majority of mutations have little or no effect on phenotype, with a small selective coefficient if any, and out of these, most are slightly deleterious. A minority of mutations will be fatal, and an even smaller number will confer a strong positive effect on fitness (Ohta 1992). Obligate pathogens, by the nature of the specialization inherent in their ecology, can be particularly likely to incur fitness costs from most mutations (Achtman and Wagner 2008). The consequence of this is that point mutations, by and large, generate mutations that negatively impact the fitness of the bacterium in which they arise.

Hyper-mutating bacteria, those that are defective in DNA proof-reading mechanisms, can develop new traits, such as antibiotic resistance, at faster rates, but this comes at the cost of their overall fitness of the organism as a whole (Woodford and Ellington 2007). This is because deleterious mutations cannot be easily removed from the genome in the absence of HGT. In strictly clonal organisms, with no HGT, linkage disequilibrium is so strong that the entire genome is the evolutionary unit and natural selection can only remove a deleterious mutation through a strong purifying selective pressure that causes a genome-wide sweep. By this mechanism, however, other deleterious mutations can become fixed in a population by hitchhiking with beneficial mutations during selective sweeps, as the sudden decrease in population size causes their mild negative effect to be counteracted by the overwhelmingly positive effect acting upon the allele selected for by the sweep (Fay and Wu 2000). Population bottlenecks, like sweeps, can remove or fix mutations; however, in this case, due to the small effective population size, it is random genetic drift, rather than natural selection, which is the predominant force. Consequently, such bottlenecks often fix deleterious mutations that actually reduce fitness.

Distinguishing between sweeps and bottlenecks, especially in retrospect, can be difficult, due to their similar effects on population structure, in that they both cause a sudden and severe reduction in the effective population size (Smith et al. 2006). The survival of a particular variant through a bottleneck may not be entirely decided by random genetic drift when more virulent genotypes lead to higher bacterial load within the host and a consequently higher likelihood of transmission to a new host.

This may explain the bias of clonal lineages to high disease-causing capacity and greater severity or virulence (Coscolla and Gagneux 2014).

Genome-wide sweeps bear the cost of high genetic load (Nielsen 2005), which can be mitigated if the effective population size is large, or by increasing the rate of HGT, which weakens linkage disequilibrium (Parkhill et al. 2003). As neither of these compensatory options is possible in fully clonal pathogens, purifying selection pressure is relaxed and ineffective in removing deleterious mutations. While strongly deleterious mutations, such as those that cause cell death, are removed immediately from a population, the smaller the negative impact on fitness that a mutation has, the slower will be the speed at which it is removed from the population. This time lag is especially marked when the pathogen population is expanding to a new niche, as is the case in a recently emerged or emerging pathogen species. The rate of non-synonymous mutations remains constant relative to the time of divergence while the synonymous mutation rate decreases (Rocha et al. 2006).

The small effective population sizes of specialised pathogens leave them particularly vulnerable to random genetic drift. Transmission events are in effect population bottlenecks, sudden decreases in population size that lead to only a minority of the population surviving to reproduce further, purging diversity and barring variants from future generations. Mutations that increase in prevalence due to genetic drift are the basis for one of the most widely used methods of pathogen classification, multilocus sequence typing (MLST). This approach indexes housekeeping genes that are key to the survival of bacteria and form part of the core genome that is typically present in all members of a species. Their sequences can be investigated for neutral (or nearly neutral) mutations that reveal phylogenetic relationships (Maiden et al. 1998, 2013).

Deleterious mutations may be fixed by such mechanisms in a clonally evolving population, as a consequence of population bottlenecks or through hitchhiking when genome-wide selective sweeps occur. Combined with relaxed purifying selective pressure, deleterious mutations can accumulate in a process known as Muller's ratchet (Gordo and Charlesworth 2000). During population bottlenecks, small population sizes weaken purifying selection so that random drift becomes the predominant evolutionary force and deleterious mutations can become fixed by chance. As these mutations are effectively random and can cause genes to lose their function, purifying selective pressure on them is further relaxed, causing gene loss and genome degradation. This has been observed in the emergence of *B. pertussis*, an obligate human pathogen, from the generalist *B. bronchiseptica*, which has a much broader-host specificity and can infect a variety of mammals and birds (Gross et al. 2010). The novel host specificity of *B. pertussis* has been accompanied by and perhaps driven by gene loss. In this case, gene loss followed host restriction, most likely due to the removal of genes that are no longer needed for growth and survival in the environment between host infections (Cummings et al. 2004). This can give the bacteria a fitness advantage by streamlining its genome, removing parts which may have associated costs but no longer incur fitness benefits, but reduction in genome size may instead be a consequence of genome degradation due to Muller's ratchet (Mooi 2010). In the case of *B. pertussis*, the reduction in genome size has been

accompanied by a massive increase in insertion sequence copy number and a consequent loss of genome structure, which suggest that this has not simply been a streamlining process and that some element of degradative evolution has also occurred (Bart et al. 2014).

A similar situation is observed in *M. tuberculosis*, where gene loss is associated with increased virulence and transmission (Djelouadji et al. 2011). *Mycobacterium leprae*, the agent of leprosy, is an even more extreme example from within the same genus. It is a fully clonal organism, marked by gene decay so extreme that half the genes in its chromosome have become inactive since the bottleneck that formed the emergence of the species (Cole et al. 2001). Homologous recombination, through its origin as a tool for DNA repair, is also an effective way to excise parasitic DNA from the genome. Species that do not participate in HGT are therefore left vulnerable to insertion elements and other forms of parasitic DNA, which may accelerate the degradation of the genome (Croucher et al. 2016). The pattern forms a general syndrome of genome evolution, where host restriction is followed by pseudogene formation, gene loss, and increasing numbers of insertion elements (Moran and Plague 2004). This syndrome can even be observed in the parallel evolution of independent clonal lineages. *S. enterica* serovars Paratyphi A and Typhi shared genes by HGT in the recent past, but their subsequent evolution shows evidence of convergence. Both serovars have evolved to a greater degree of host restriction and to cause systemic disease through the loss of function in many of the same genes (Holt et al. 2009).

5 Examples of Adaptive Evolution in Clonal Pathogens

As we have seen, strictly clonal pathogens have limited means by which they can adapt to changes in their environment, being constrained in their ability to generate new variants and to reassort these variants by HGT, even within the same species. Nevertheless, when sufficiently strong, selective pressures can impact the evolution of these pathogens and lead to adaptation based on single mutations that arise in the population. Most often this is observed following positive periodic selection, where a strong selective pressure, following, for example, a change to the environment, causes a genome-wide sweep that removes any isolates that do not have the allele or alleles most advantageous to the new environment (Shapiro 2016). In pathogenic bacteria, such events are typically associated with changes caused by medical interventions including the introduction of specific treatments.

B. pertussis has been a significant cause of whooping cough morbidity and mortality in children worldwide, but while the implementation of effective vaccines successfully reduced the disease burden, the bacterium may still be circulating in the population as a milder infection in adults (Guiso 2014). Whole-cell vaccines for *B. pertussis* are highly effective but often cause mild side effects and occasionally severe ones. For this reason they were replaced by acellular vaccines in the late 1980s and early 1990s, after which a worldwide selective sweep was observed in the

circulating *B. pertussis* strains. The previously dominant *ptxP1* organisms were replaced with strains carrying the *ptxP3* allele, which is associated with an increase in the expression of pertussis toxin. The resulting change in population composition caused phenotypic consequences beyond the pertussis toxin genes, as other mutations hitchhiked with the *ptxP3* variant (de Gouw et al. 2014).

The limited amount of genetic diversity present in clonally evolving populations may belie their adaptive potential. The few existing points of variability in a genetically monomorphic population can have a disproportionately high impact on their biology (Reiling et al. 2013). In *M. tuberculosis*, whole-genome sequencing has revealed that the species can adapt when faced with sufficiently strong selection pressure, as is the case with the emergence of antibiotic resistance in the face of treatment. The difficulty in achieving therapeutic doses of antibiotics in the intracellular environment of *M. tuberculosis* compounds the problem of resistance in this species, as therapy with a single antibiotic is rarely effective and the use of combination therapies accelerates the emergence of multidrug-resistant strains (Moreno-Gamez et al. 2015). Despite the need for each resistance-conferring mutation to occur sequentially within each lineage, resistance is commonplace, and multidrug resistance is a growing concern for tuberculosis treatment worldwide (Hershberg et al. 2008).

Hypermutability can enhance the ability of clonal lineages to adapt rapidly and may therefore be actively selected in clonal pathogen populations. This appears to have been the case for the W-Beijing lineage of *M. tuberculosis*, a genotype which is associated with high virulence, in terms of the rate of activation of tuberculosis disease in patients (Merker et al. 2015). This phenotype enhances the spread of tuberculosis in large dense urbanised human populations but would be less advantageous in less dense human populations where the chronic latent stage, rather than the acute symptomatic stage, would be favoured. The mutations promoting virulence in the W-Beijing lineage have probably played a role in the pandemic spread of this lineage during the twentieth century (Hanekom et al. 2011). The genetic basis for this increased pathogenicity is variable throughout the lineage, with the common factor appearing to be the relaxation of purifying selection on the genes that control replication and DNA repair, leading to a higher rate of mutations. This increase in genomic variability provides an evolutionary advantage under stressful conditions and can compensate for the loss of adaptability caused by the lack of horizontal gene transfer (Vultros Dos et al. 2008).

S. enterica serovar Typhi has also adapted to the pressure imposed by antibiotic use, with 15 mutations that confer fluoroquinolone resistance being observed at the *gyrA* gene in just a decade. This accumulation of adaptations as a consequence of strong positive selection is in stark contrast to the evidence of neutral evolution in the remainder of *S. typhi* genome: genetic drift is the predominant force throughout the genome, with no observed difference in the selective forces in genes associated with housekeeping functions or pathogenicity (Roumagnac et al. 2006).

6 Consequences of Clonality

As we have discussed, clonal population structures can lead to specialization and vice versa, leading to the evolution of a highly specialised bacterial clonal pathogen, which has lost much of its ancestors' metabolic capacity. The ecological and evolutionary consequences of the interaction between clonality and specialisation are a continuous cycle of small effective population sizes that relax purifying pressure and increase the dominance of genetic drift, causing deleterious mutations which may arise to accumulate in the genome. These lead to genome degradation and genetic isolation, which in turn reduces the genetic diversity of a population through sweeps and bottlenecks. Specialising to a specific niche, under these circumstances, can become the only possible evolutionary path that is available to the organism in question, however short-lived that path may ultimately be.

Sexual reproduction is costly for the organisms that engage in it. They must bear the cost of searching and competing for a mate, of producing male offspring, and of sharing only half their genes with any offspring. Why sexual reproduction is widely maintained despite these costs is an “evolutionary puzzle” (Lehtonen et al. 2012). In vertebrates, clonality is always a short-lived strategy that inevitably leads to extinction, though it can be rescued by even occasional recombination (Avisé 2015). In bacteria, parasexual processes, mediated by gene transfer and recombination mechanisms, may have evolved primarily as a DNA repair mechanism, belying their far-reaching evolutionary consequences as a method of genetic exchange. The ability of HGT to generate diversity and allow genes, rather than entire genomes, to become the evolutionary units is critical for the long-term success of bacterial species. As with vertebrates, even occasional recombination seems to prevent the complementary processes of clonality and specialisation, by markedly improving the viability of populations over that of a strictly clonal lifestyle (Tibayrenc and Ayala 2012).

When considering the time scales over which clonality can emerge as a stable trait of a given highly specialised pathogen, it is necessary to “distinguish between short term emergence of clonal complexes and the more long-term evolutionary history of a bacterial population” (Feil and Spratt 2001). The concept of clonal lineages emerging from a context with a high frequency of recombination, due to small population sizes that lead to genetic isolation, predates genome sequencing (Levin 1981). Expanding upon this theory led to the epidemic clone model, which posits that even in populations with rampant recombination, successful clones will occasionally cause epidemic waves (Maynard Smith et al. 1993). In this model, a freely recombining population, comprising many variants that participate in HGT, occasionally gives rise to a clonal lineage that successfully emerges as a transmissible pathogen (Turner and Feil 2007).

Are the clonal lineages that survive over millennia fundamentally different from those that survive only until the next genotype in clonal replacement takes over? It is possible that the various pandemics of *Vibrio cholerae*, the cause of the diarrheal disease cholera, are an intermediate between short-lived epidemic clones and more

long-lived clonal species such as *M. tuberculosis*. Sublineages that define each epidemic are observed over the space of a few decades but are replaced by successive waves of new sublineages. These lineages can also develop into hypermutators, again with the pattern of fast adaptive evolution to a specific pressure matched by a decrease in overall fitness (Didelot et al. 2015).

In *Salmonella enterica*, two time scales of evolution have been proposed, based on studies of the clonal *S. typhi* serovar. Evolution to antibiotic resistance occurs in clinically observable time, affecting transmission and infections, with a much slower process of genetic drift fixing neutral mutations gradually, resulting in “two distinct epidemiological dynamics” being present in one population structure (Roumagnac et al. 2006). This two-level evolution may explain, at least in part, the predominant nature of clonal evolution being one of stability over thousands of years and marked only by the slow accumulation of mostly neutral mutations but overlaid with the rapid adaptation in response to the strong selection pressures that are the consequence of antibiotic use or vaccine implementation.

Intriguingly, in another serovar of this species, *S. enterica* serovar Typhimurium, a novel clone appears to currently be emerging as a highly specialised pathogen in observable time. The ST313 sublineage of *S. typhimurium* causes invasive disease rather than the gastroenteritis typical of this serovar. Further, the disease that it causes is associated with much higher mortality, with up to half of adult cases resulting in death. The ST313 clone has, through a HGT event, gained a virulence-associated plasmid that bears many antibiotic resistance genes, and this seems to have enabled a change in its ecology that has resulted in a shift in its pathology. The change to an invasive disease niche has subsequently led to genome degradation, including loss of many of the same genes that are now absent from *S. typhi* (Kingsley et al. 2009), in an apparent example of convergent evolution.

Strictly clonal, obligate pathogens of one or a few host species can be seen as a quirk of evolution. They have emerged as specialised clones of large, more adaptable, and diverse populations, but their invasion of this novel niche, with its consequent process of specialization, comes with the associated costs of loss of adaptability and perhaps leading to a process of degradative evolution that inevitably leads to extinction, even if the host species survives. This process of extinction may take very many host generations and, for a number of human pathogens, this seems to have been ongoing for many millennia, sufficient for global distribution of these seemingly evolutionarily doomed pathogens. Given the high virulence that these pathogens can attain, a fuller understanding of the evolutionary path taken by these bacteria will aid us in containing the long-standing threats that they pose and that may be posed by novel pathogens that take the same evolutionary path as a consequence of changes in human ecology. In this respect, it is worth noting as a final comment, that our own species is a relative newcomer in terms of worldwide distribution and ecological success (indeed we are unique among the family Hominidae in this respect), and it is therefore unsurprising that the obligate clonal pathogens of humans are also of recent date, with some of them, a notable example being *M. tuberculosis* evolving alongside *Homo sapiens*, with the fate of both species inexorably linked (Soares et al. 2012) (Fig. 12.1).

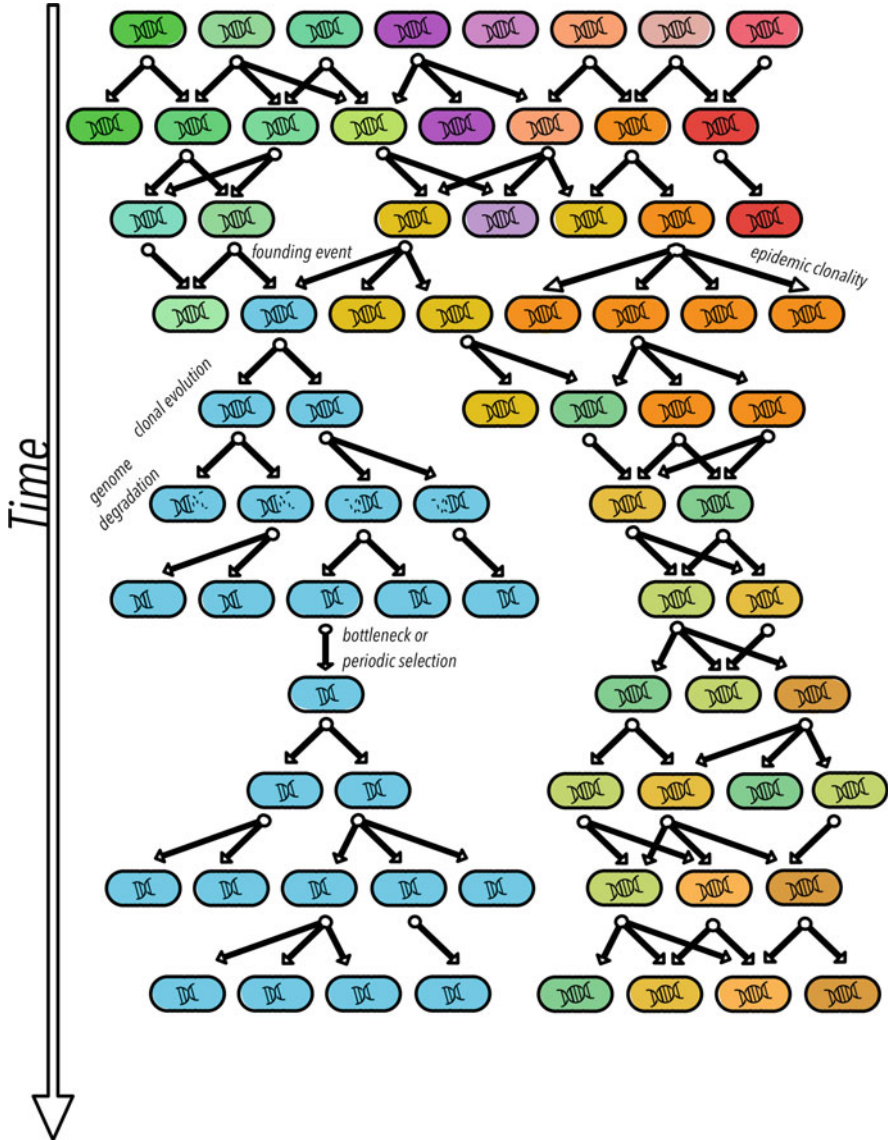


Fig. 12.1 Illustration of bacterial reproduction and evolution, including epidemic clonality and fully clonal evolution. Each oval represents one bacterium and its genome, and each row represents one sampling time period, so that the bacteria in a row are descendants of those in the row above, separated by many generations. At first the population is evolving as is typical of most bacteria, with a mix of clonal descent and horizontal gene transfer. Therefore most bacteria contain genetic material from more than one ancestor, though not all bacteria contribute genetically to further generations. Between the third and fourth time points, one bacterium is particularly successful and rapidly increases in abundance, in an example of *epidemic clonality*. This is a temporary trait, however, and by the fifth time point onward, the evolutionary processes occurring in its descendants have returned to the patterns typical of the ancestral population. A different bacterium within the

References

- Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu Rev Microbiol* 62:53–70. <https://doi.org/10.1146/annurev.micro.62.081307.162832>
- Achtman M (2012) Insights from genomic comparisons of genetically monomorphic bacterial pathogens. 367:860–867. <https://doi.org/10.1098/rstb.2011.0303>
- Achtman M, Wagner M (2008) Microbial diversity and the genetic nature of microbial species. *Nat Rev Microbiol* 6:431–440. <https://doi.org/10.1038/nrmicro1872>
- Avisé JC (2015) Evolutionary perspectives on clonal reproduction in vertebrate animals. *Proc Natl Acad Sci U S A* 112:8867–8873. <https://doi.org/10.1073/pnas.1501820112>
- Ayyadurai S, Houhamdi L, Lepidi H et al (2008) Long-term persistence of virulent *Yersinia pestis* in soil. *Microbiol Immunol* 154:2865–2871. <https://doi.org/10.1099/mic.0.2007/016154-0>
- Baptiste E, Boucher Y (2009) Epistemological impacts of horizontal gene transfer on classification in microbiology. *Methods Mol Biol* 532:55–72. https://doi.org/10.1007/978-1-60327-853-9_4
- Bart MJ, Harris SR, Advani A et al (2014) Global population structure and evolution of *Bordetella pertussis* and their relationship with vaccination. *MBio* 5:e01074–e01074–14. <https://doi.org/10.1128/mBio.01074-14>
- Bergstrom CT, McElhany P, Real LA (1999) Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. *Proc Natl Acad Sci U S A* 96:5095–5100
- Bos KI, Schuenemann VJ, Golding GB et al (2011) A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* 478:506–510. <https://doi.org/10.1038/nature10549>
- Cao Q, Didelot X, Wu Z et al (2015) Progressive genomic convergence of two *Helicobacter pylori* strains during mixed infection of a patient with chronic gastritis. *Gut* 64:554–561. <https://doi.org/10.1136/gutjnl-2014-307345>
- Cohan FM (1994) Genetic exchange and evolutionary divergence in prokaryotes. *Trends Ecol Evol* 9:175–180. [https://doi.org/10.1016/0169-5347\(94\)90081-7](https://doi.org/10.1016/0169-5347(94)90081-7)
- Cohan FM (2002) Sexual isolation and speciation in bacteria. *Genetica* 116:359–370. <https://doi.org/10.1023/A:1021232409545>
- Cole ST, Eiglmeier K, Parkhill J et al (2001) Massive gene decay in the leprosy bacillus. *Nature* 409:1007–1011. <https://doi.org/10.1038/35059006>
- Comas I, Coscolla M, Luo T et al (2013) Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet* 45:1176–1182. <https://doi.org/10.1038/ng.2744>
- Coscolla M, Gagneux S (2014) Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin Immunol* 26:431–444. <https://doi.org/10.1016/j.smim.2014.09.012>
- Croucher NJ, Mostowy R, Wymant C et al (2016) Horizontal DNA transfer mechanisms of bacteria as weapons of intragenomic conflict. *PLoS Biol* 14:e1002394. <https://doi.org/10.1371/journal.pbio.1002394>
- Cummings CA, Brinig MM, Lepp PW (2004) *Bordetella* species are distinguished by patterns of substantial gene loss and host adaptation. *J Bacteriol*. <https://doi.org/10.1128/JB.186.5.1484>



Fig. 12.1 (continued) population becomes reproductively isolated in the fourth time point, after an HGT event that led to its emergence (the *founding event*). The descendants of this bacterium undergo *clonal evolution*, with no HGT observable anywhere in the lineage, so that every bacterium is descended from solely one bacterium in a previous time point. This lineage undergoes *genome degradation* leading to gene loss and eventually to a smaller genome. A *bottleneck or periodic selection* event dramatically reduces the population size of the clonally evolving lineage between the seventh and eighth time points, reducing the diversity of the population and fixing the gene loss in the genome

- de Gouw D, Diavatopoulos D, Heuvelman K et al (2014) Differentially expressed genes in *Bordetella pertussis* strains belonging to a lineage which recently spread globally. *PLoS One* 9:e84523. <https://doi.org/10.1371/journal.pone.0084523>
- Diavatopoulos DA, Cummings CA, Schouls LM, Brinig MM, Relman DA, Mooi FR (2005) *Bordetella pertussis*, the causative agent of whooping cough, evolved from a distinct, human-associated lineage of *B. bronchiseptica*. *PLoS Pathog* 1(4):e45
- Didelot X, Bowden R, Street T et al (2011) Recombination and population structure in *Salmonella enterica*. *PLoS Genet* 7:e1002191–e1002191. <https://doi.org/10.1371/journal.pgen.1002191>
- Didelot X, Pang B, Zhou Z et al (2015) The role of China in the global spread of the current cholera pandemic. *PLoS Genet* 11:e1005072. <https://doi.org/10.1371/journal.pgen.1005072>
- Djelouadji Z, Raoult D, Drancourt M (2011) Palaeogenomics of *Mycobacterium tuberculosis*: epidemic bursts with a degrading genome. *Lancet Infect Dis* 11:641–650. [https://doi.org/10.1016/S1473-3099\(11\)70093-7](https://doi.org/10.1016/S1473-3099(11)70093-7)
- Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics* 155:1405–1413
- Feil EJ, Spratt BG (2001) Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* 55:561–590. <https://doi.org/10.1146/annurev.micro.55.1.561>
- Ford CB, Lin PL, Chase MR et al (2011) Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet* 43:482–486. <https://doi.org/10.1038/ng.811>
- Gage KL, Kosoy MY (2005) Natural history of plague: perspectives from more than a century of research. *Annu Rev Entomol* 50:505–528. <https://doi.org/10.1146/annurev.ento.50.071803.130337>
- Galan JE (2016) Typhoid toxin provides a window into typhoid fever and the biology of *Salmonella typhi*. *Proc Natl Acad Sci U S A* 113:6338–6344. <https://doi.org/10.1073/pnas.1606335113>
- Ghodbane R, Mba Medie F, Lepidi H et al (2014) Long-term survival of tuberculosis complex mycobacteria in soil. *Microbiol Rev* 160:496–501. <https://doi.org/10.1099/mic.0.073379-0>
- Gordo I, Charlesworth B (2000) The degeneration of asexual haploid populations and the speed of Muller’s ratchet. *Genetics* 154:1379–1387. [https://doi.org/10.1016/S0960-9822\(02\)00448-7](https://doi.org/10.1016/S0960-9822(02)00448-7)
- Gross R, Keidel K, Schmitt K (2010) Resemblance and divergence: the “new” members of the genus *Bordetella*. *Med Microbiol Immunol* 199:155–163. <https://doi.org/10.1007/s00430-010-0148-z>
- Guiso N (2014) *Bordetella pertussis*: why is it still circulating? *J Infect* 68(Suppl 1):S119–S124. <https://doi.org/10.1016/j.jinf.2013.09.022>
- Hanage WP, Fraser C, Spratt BG (2006a) The impact of homologous recombination on the generation of diversity in bacteria. *J Theor Biol* 239:210–219. <https://doi.org/10.1016/j.jtbi.2005.08.035>
- Hanage WP, Fraser C, Spratt BG (2006b) Sequences, sequence clusters and bacterial species. *Philos Trans R Soc B* 361:1917–1927. <https://doi.org/10.1098/rstb.2006.1917>
- Hanekom M, Gey Van Pittius NC, McEvoy C et al (2011) *Mycobacterium tuberculosis* Beijing genotype: a template for success. *Tuberculosis* 91:510–523. <https://doi.org/10.1016/j.tube.2011.07.005>
- Hershberg R, Lipatov M, Small PM et al (2008) High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol* 6:e311. <https://doi.org/10.1371/journal.pbio.0060311>
- Holt KE, Parkhill J, Mazzoni CJ et al (2008) High-throughput sequencing provides insights into genome variation and evolution in *Salmonella typhi*. *Nat Genet* 40:987–993. <https://doi.org/10.1038/ng.195>
- Holt KE, Teo YY, Li H et al (2009) Detecting SNPs and estimating allele frequencies in clonal bacterial populations by sequencing pooled DNA. *Vaccine* 25:2074–2075. <https://doi.org/10.1093/bioinformatics/btp344>
- Huttenhower C, Gevers D, Knight R et al (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–214. <https://doi.org/10.1038/nature11234>

- Kay GL, Sergeant MJ, Zhou Z et al (2015) Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* 6:6717. <https://doi.org/10.1038/ncomms7717>
- Keim P, Gruendike JM, Klevytska AM et al (2009) The genome and variation of *Bacillus anthracis*. *Mol Aspects Med* 30:397–405. <https://doi.org/10.1016/j.mam.2009.08.005>
- Kingsley RA, Msefula CL, Thomson NR et al (2009) Epidemic multiple drug resistant *Salmonella typhimurium* causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res* 19:2279–2287. <https://doi.org/10.1101/gr.091017.109>
- Kohl TA, Diel R, Harmsen D et al (2014) Whole genome based *Mycobacterium tuberculosis* surveillance: a standardized, portable and expandable approach. *J Clin Microbiol* 52:2479–2486. <https://doi.org/10.1128/JCM.00567-14>
- Kolstø A-B, Tourasse NJ, Økstad OA (2009) What sets *Bacillus anthracis* apart from other *Bacillus* species? *Annu Rev Microbiol* 63:451–476. <https://doi.org/10.1146/annurev.micro.091208.073255>
- Lan R, Reeves PR (2001) When does a clone deserve a name? A perspective on bacterial species based on population genetics. *Trends Microbiol* 9:419–424
- Lehtonen J, Jennions MD, Kokko H (2012) The many costs of sex. *Trends Ecol Evol* 27:172–178. <https://doi.org/10.1016/j.tree.2011.09.016>
- Levin BR (1981) Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. *Genetics* 99:1–23
- Maiden MCJ, Bygraves JA, Feil E et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95:3140–3145. <https://doi.org/10.1073/pnas.95.6.3140>
- Maiden MCJ, van Rensburg MJJ, Bray JE et al (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol* 11:728–736. <https://doi.org/10.1038/nrmicro3093>
- Maynard Smith J, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? *Proc Natl Acad Sci U S A* 90:4384–4388. <https://doi.org/10.1902/jop.2011.110063>
- Merker M, Blin C, Mona S et al (2015) Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat Genet* 47:242–249. <https://doi.org/10.1038/ng.3195>
- Michod RE, Bernstein H, Nedelcu AM (2008) Adaptive value of sex in microbial pathogens. *Infect Genet Evol* 8:267–285. <https://doi.org/10.1016/j.meegid.2008.01.002>
- Moffatt CRM, Greig A, Valcanis M et al (2016) A large outbreak of *Campylobacter jejuni* infection in a university college caused by chicken liver pâté, Australia, 2013. *Epidemiol Infect* 144:2971–2978. <https://doi.org/10.1017/S0950268816001187>
- Mooi F (2010) *Bordetella pertussis* and vaccination: the persistence of a genetically monomorphic pathogen. *Infect Genet Evol* 10:36–49. <https://doi.org/10.1016/j.meegid.2009.10.007>
- Moran NA, Plague GR (2004) Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* 14:627–633. <https://doi.org/10.1016/j.gde.2004.09.003>
- Moran NA, Wernegreen JJ (2000) Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol Evol* 15:321–326
- Moreno-Gamez S, Hilla AL, Rosenbloom DIS et al (2015) Imperfect drug penetration leads to spatial monotherapy and rapid evolution of multidrug resistance. *Proc Natl Acad Sci U S A* 112: E2874–E2883. <https://doi.org/10.1073/pnas.1424184112>
- Namouchi A, Didelot X, Schöck U et al (2012) After the bottleneck: Genome-wide diversification of the *Mycobacterium tuberculosis* complex by mutation, recombination, and natural selection. *Genome Res* 22:721–734. <https://doi.org/10.1101/gr.129544.111>
- Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197–218. <https://doi.org/10.1146/annurev.genet.39.073003.112420>
- Niemann S, Merker M, Kohl T, Supply P (2016) Impact of genetic diversity on the biology of *Mycobacterium tuberculosis* complex strains. *Microbiol Spectr*. <https://doi.org/10.1128/microbiolspec.TB2-0022-2016>
- Ohta T (1992) The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst* 23:263–286. <https://doi.org/10.2307/2097289>

- Ohta T (2011) Near-neutrality, robustness, and epigenetics. *Genome Biol Evol* 3:1034–1038. <https://doi.org/10.1093/gbe/evr012>
- Okinaka R, Pearson T, Keim P (2006) Anthrax, but not *Bacillus anthracis*? *PLoS Pathog* 2:e122
- Orsi RH, Bowen BM, Wiedmann M (2010) Homopolymeric tracts represent a general regulatory mechanism in prokaryotes. *BMC Genomics* 11(1):102
- Park J, Zhang Y, Buboltz AM et al (2012) Comparative genomics of the classical *Bordetella* subspecies: the evolution and exchange of virulence-associated diversity amongst closely related pathogens. *BMC Genomics* 13:545–545. <https://doi.org/10.1186/1471-2164-13-545>
- Parkhill J, Sebaihia M, Preston A et al (2003) Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet* 35:32–40. <https://doi.org/10.1038/ng1227>
- Planet PJ, Narechania A, Chen L et al (2017) Architecture of a Species: phylogenomics of *Staphylococcus aureus*. *Trends Microbiol* 25:153–166. <https://doi.org/10.1016/j.tim.2016.09.009>
- Polz MF, Alm EJ, Hanage WP (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet* 29:170–175. <https://doi.org/10.1016/j.tig.2012.12.006>
- Rasko DA, Worsham PL, Abshire TG et al (2011) *Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation. *Proc Natl Acad Sci U S A* 108:5027–5032. <https://doi.org/10.1073/pnas.1016657108>
- Rasmussen S, Allentoft ME, Nielsen K et al (2015) Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell* 163:571–582. <https://doi.org/10.1016/j.cell.2015.10.009>
- Redfield RJ, Findlay WA, Bosse J et al (2006) Evolution of competence and DNA uptake specificity in the Pasteurellaceae. *BMC Evol Biol* 6:82. <https://doi.org/10.1186/1471-2148-6-82>
- Reiling N, Homolka S, Walter K et al (2013) Clade-specific virulence patterns of *Mycobacterium tuberculosis* complex strains in human primary macrophages and aerogenically infected mice. *MBio* 4:e00250–e00213. <https://doi.org/10.1128/mBio.00250-13>
- Rocha E, Maynard Smith J, Hurst LD et al (2006) Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J Theor Biol* 239:226–235. <https://doi.org/10.1016/j.jtbi.2005.08.037>
- Roumagnac P, Weill F-X, Dolecek C et al (2006) Evolutionary history of *Salmonella typhi*. *Science* 314:1301–1304. <https://doi.org/10.1126/science.1134933>
- Shapiro BJ (2016) How clonal are bacteria over time? *Curr Opin Microbiol* 31:116–123. <https://doi.org/10.1016/j.mib.2016.03.013>
- Smith NH, Gordon SV, de la Rúa-Domenech R et al (2006) Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat Rev Microbiol* 4:670–681. <https://doi.org/10.1038/nrmicro1472>
- Soares P, Alshamali F, Pereira JB et al (2012) The expansion of mtDNA Haplogroup L3 within and out of Africa. *Mol Biol Evol* 29:915–927. <https://doi.org/10.1093/molbev/msr245>
- Spratt BG, Maiden MCJ (1999) Bacterial population genetics, evolution and epidemiology. *Philos Trans R Soc B* 354:701–710. <https://doi.org/10.1098/rstb.1999.0423>
- Stackebrandt E, Frederiksen W, Garrity GM et al (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 52:1043–1047. <https://doi.org/10.1099/00207713-52-3-1043>
- Suerbaum S, Smith JM, Bapumia K et al (1998) Free recombination within *Helicobacter pylori*. *Proc Natl Acad Sci U S A* 95:12619–12624
- Tibayrenc M, Ayala FJ (2002) The clonal theory of parasitic protozoa: 12 years on. *Trends Parasitol* 18:405–410
- Tibayrenc M, Ayala FJ (2012) Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proc Natl Acad Sci U S A* 109:E3305–E3313. <https://doi.org/10.1073/pnas.1212452109>
- Tibayrenc M, Ayala FJ (2015) How clonal are *Neisseria* species? The epidemic clonality model revisited. *Proc Natl Acad Sci U S A* 112:8909–8913. <https://doi.org/10.1073/pnas.1502900112>

- Tibayrenc M, Ayala FJ (2016) Is predominant clonal evolution a common evolutionary adaptation to parasitism in pathogenic parasitic protozoa, fungi, bacteria, and viruses? *Adv Parasitol* 97:243–325
- Turner KME, Feil EJ (2007) The secret life of the multilocus sequence type. *Int J Antimicrob Agents* 29:129–135
- van der Veen S, Tang CM (2015) The BER necessities: the repair of DNA damage in human-adapted bacterial pathogens. *Nat Rev Microbiol* 13:83–94. <https://doi.org/10.1038/nrmicro3391>
- Votintseva AA, Miller RR, Fung R et al (2014) Multiple-strain colonization in nasal carriers of *Staphylococcus aureus*. *J Clin Microbiol* 52:1192–1200. <https://doi.org/10.1128/JCM.03254-13>
- Vultos Dos T, Mestre O, Rauzier J et al (2008) Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. *PLoS One* 3:e1538–e1538. <https://doi.org/10.1371/journal.pone.0001538>
- Wiedenbeck J, Cohan FM (2011) Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev* 35:957–976. <https://doi.org/10.1111/j.1574-6976.2011.00292.x>
- Woodford N, Ellington MJ (2007) The emergence of antibiotic resistance by mutation. *Clin Microbiol Infect* 13:5–18. <https://doi.org/10.1111/j.1469-0691.2006.01492.x>
- Yahara K, Didelot X, Jolley KA et al (2016) The landscape of realized homologous recombination in pathogenic bacteria. *Mol Biol Evol* 33:456–471. <https://doi.org/10.1093/molbev/msv237>