Opinion

Rates of Lateral Gene Transfer in Prokaryotes: High but Why?

Michiel Vos,[1,*] Matthijn C. Hesselman,[2] Tim A. te Beek,[3] Mark W.J. van Passel,[2,4] and Adam Eyre-Walker[5]

Lateral gene transfer is of fundamental importance to the evolution of prokaryote genomes and has important practical consequences, as evidenced by the rapid dissemination of antibiotic resistance and virulence determinants. Relatively little effort has so far been devoted to explicitly quantifying the rate at which accessory genes are taken up and lost, but it is possible that the combined rate of lateral gene transfer and gene loss is higher than that of point mutation. What evolutionary forces underlie the rate of lateral gene transfer are not well understood. We here use theory developed to explain the evolution of mutation rates to address this question and explore its consequences for the study of prokaryote evolution.

The Fluidity of Prokaryote Genomes

It has long been recognized that even closely related strains of bacteria and archaea can greatly differ in gene content [1–3]. The rate and promiscuity with which genes change residence in prokaryote genomes is of such magnitude that many have rejected the concept of prokaryote species (e.g., [4]) or the possibility of reconstructing a bifurcating tree of prokaryote life (e.g., [5]). The realization of the high rates of gene content turnover has led to the paradigm of a core genome of genes present in all members of a taxonomic group and an accessory genome present in only a subset of members (with the total complement of genes in a taxon termed the pan genome [6]). The extent to which related genomes differ in gene content varies for different species, with some having a relatively 'closed' genome (i.e., a core genome that is large compared to the accessory genome) and some species an 'open' or 'flexible' genome (i.e., a relatively small core genome and a large pan genome) [7].

A diverse set of mechanisms underlies the evolution of gene content but they can be grouped into three main classes: gene loss, gene gain through duplication (parology), and gene gain though lateral gene transfer (LGT, xenology) [8,9]. Knowledge on the impact of LGT on bacterial evolution continues to rapidly expand (e.g., [10–13]). In many prokaryotes, LGT is known to be a more important process for determining gene gain than is gene duplication [14–16]. Gene loss can be due to two processes: mutational deletion and the lateral transfer of gene absence by recombination between a strain with and without the gene (likely helped by the presence of homologous flanking sequence [17]), although it is not well understood which of the two mechanisms is most important.

The rate at which genomes accumulate random variation, due to point mutation or wholesale changes in gene content, dictates the speed and mode of evolution. Both mutational changes and changes in gene content provide the raw material for selection to act on, but the rate of each process can be modified through selection too. Apart from the fitness effects of individual LGT events, it is therefore crucial to understand how second-order selection acts on the rate of change itself. In this paper we provide a short summary of the relatively few studies detailing the

Trends

The rate of lateral gene transfer (LGT) in prokaryotes can be high, and is likely to have profound consequences for genome evolution and adaptation. However, the selective forces underlying LGT rates have remained relatively unexplored. Insights from theory developed to explain the evolution of mutation rates can be applied to the evolution of LGT rates.

We identify two main scenarios that could explain the high observed rates of LGT. In the first scenario, LGT rate is optimal while in the second scenario the rate of LGT is suboptimal.

The diverse mechanisms and fitness effects of LGT events make it challenging to address evolutionary costs and benefits. A combination of population genomic and experimental approaches will be needed to establish LGT rates and the distribution of fitness effects of LGT events.

[1]European Centre for Environment and Human Health, University of Exeter Medical School, University of Exeter, Penryn, UK
[2]Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands
[3]Netherlands Bioinformatics Centre, Nijmegen, The Netherlands
[4]Centre for Zoonoses and Environmental Microbiology, Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[5]School of Life Sciences, University of Sussex, Brighton, UK

*Correspondence: m.vos@exeter.ac.uk (M. Vos).

rate and fate of lateral gene transfers. The main part of this paper then discusses how models developed to explain the evolution of the mutation rate can be applied to the evolution of LGT rate, which is likely to be high in many species. Specifically, we discuss whether selection drives optimal rates of LGT or whether selection for lower rates is constrained. Finally, we highlight some of the implications our findings have for our understanding of evolutionary microbiology and identify areas for future research.

Quantifying the Rate and Fate of Changes in Gene Content

A plethora of comparative genomics studies have demonstrated the rate of prokaryote gene content turnover to be very high. However, systematic attempts to estimate how fast gene turnover occurs over evolutionary time, expressed per genome per generation or relative to point mutation, have been rare. To our knowledge, the first study to explicitly address this question was by Hao and Golding [18] who used a maximum likelihood method to quantify both insertion and deletion rates of genes relative to mutation in a set of *Bacillus cereus* group genomes. According to their published estimates, genes were gained and lost at a rate approximately 4.4 times the rate of nucleotide substitution per site [18]. Similar estimates were found in *Streptococcus* [19] and *Corynebacterium* [20] genomes using the same methodology. A recent estimate of LGT rate in *Pseudomonas syringae* based on stochastic mapping methodology (after corrections necessary for working with genomes that are not sequenced to complete closure) was found to be four orders of magnitude higher than the estimate for *B. cereus* [21]. Individual *P. syringae* lineages could be shown to have acquired thousands of genes in the same period in which a 1% amino acid divergence accrued in the core genome. The rate also appears to be very high in *Escherichia coli*, where strains that have almost no single nucleotide changes in their genes they share can differ substantially in gene content [22].

These and other studies (e.g., [23,24]) have also demonstrated that most gene content changes are short-lived. In most studies this is apparent from a decrease in the rate of gene content change relative to the rate of nucleotide substitution in the deeper branches of the phylogeny compared to the tips. The same pattern is most elegantly demonstrated in the *P. syringae* and *E. coli* studies [21,22] in which the vast majority of individual gene gains are mapped to a single strain. There are three potential explanations for why most gene content changes are transient. First, the vast majority of gene content changes might be neutral (i.e., have no selective effects) and, like mutations, are lost from the population by random **genetic drift** (see Glossary). This seems an unlikely explanation, because under this model the relative rates of gene content change and nucleotide substitution would remain constant across different phylogenetic depths. Second, most gene content changes might be deleterious [6,8,18,22]. There are a variety of reasons to expect that LGT events are likely to have a negative effect on fitness [25]; apart from an added cost of translation, the high coding density of bacterial genomes means that inserted genes are likely to disrupt existing translation. Moreover, newly introduced gene products could be outright toxic through gene dosage effects, interfere with existing cellular interactions, or interfere with each other [26]. Third, accessory gene gains could be beneficial but only very transiently so. However, their removal would still imply that these genes ultimately have a deleterious effect.

The fact that the majority of lateral gene transfers are likely to be deleterious and quickly removed by **purifying selection** (unless this is made impossible by the presence of addiction mechanisms [27]), highlights the necessity of comparing very closely related genomes in order to reliably estimate the LGT rate. As a large number of deleterious events will have gone undetected, the ratio of gene content changes to mutation is likely to be significantly higher than estimated. Using data on the two pairs of most closely related *P. syringae* strains in the Nowell *et al.* [21] study, we divided the estimated number of gene gains by the number of estimated

Glossary

Distribution of fitness effects (DFE): mutations, as well as lateral gene transfer (LGT) events, have fitness effects that can be broadly divided into three categories. First, there are mutations that decrease fitness. Second, there are 'neutral' mutations, which have little or no effect on fitness. Third, there are advantageous mutations, which increase fitness by allowing organisms to adapt to their environment. However, in reality, there is a continuum of selective effects, stretching from those that are strongly deleterious, through weakly deleterious mutations, to neutral mutations and then on to mutations that are mildly or highly adaptive. The distribution of fitness effects refers to the relative frequencies of these types of mutation.

Genetic drift: the change in frequency of a mutation or accessory gene in a population as a result of chance, not selection.

Mutation rate modifier: a gene variant that influences the mutation rate.

Purifying selection: selection against genomic changes that lower fitness.

mutations that have occurred since the pairs of genomes separated. (Mutational divergence was estimated by assuming that synonymous mutations are neutral and by multiplying the number of synonymous substitutions per site by the total number of sites in the genome.) Gene gain was estimated to have occurred at ~20% the genomic rate of point mutation. However, these strains were not particularly closely related, with many hundreds of mutations separating them. Purifying selection therefore can be expected to have removed many deleterious gene gain events, leading this to be a highly conservative estimate of LGT rate. Moreover, single gene gain events usually equate to many nucleotide changes and so our measure of genomic change by LGT is conservative.

Applying Theory Developed for the Evolution of Mutation Rate to the Evolution of Gene Content Turnover

The fact that most LGT events are deleterious poses the question why LGT seemingly occurs at high rates. Both rates of mutation and gene content change are controlled by intrinsic factors (i. e., those under genetic control of the cell) and extrinsic factors (i.e., those mediated by the environment, including other biological agents). Mutation rate is in part determined by the fidelity of the replication and repair machinery of the cell, and can be elevated directly by extrinsic stress factors such as ultraviolet (UV) radiation and reactive oxygen, or indirectly by stress-induced gene expression changes lowering the efficiency of DNA repair [28]. The LGT component of gene content change is in part determined by the genetic regulation of competence (controlling the uptake of foreign DNA by transformation) and by innate and adaptive immunity systems preventing infection by plasmids and transducing bacteriophage [29], but is also crucially dependent on the concentration and diversity of free DNA and mobile genetic elements in the cell's environment. Although, in the case of LGT, the balance between intrinsic and extrinsic determinants is likely shifted towards the latter, we here argue that there is no clear-cut boundary between different types of genomic change and that LGT can be analysed in the theoretical framework developed for mutation.

It has been well established that the evolution of mutation rate is crucially dependent on three factors: (i) **the distribution of fitness effects (DFE)** of new mutations, (ii) the costs associated with keeping mutation rate under control, and (iii) the presence of recombination. When it is assumed that all mutations are deleterious, selection is expected to evolve to bring the mutation rate down to zero. However, because mutations also provide a source of genetic variation required for adaptation, a more realistic expectation is that rates have evolved to an optimum level. The optimum mutation rate then is a balance between, on one hand, the effects of deleterious mutations and the metabolic cost of limiting the rate at which they occur, and on the other hand, the benefits of occasionally obtaining mutations that aid adaptation to a changing environment [30–35]. Following this reasoning, the more frequent and strongly selected advantageous mutations are, the higher the optimum mutation rate will be.

Theory has demonstrated that recombination rate is a crucial factor determining the evolution of the mutation rate [34,35]. In asexual organisms, the loci determining mutation rate (**mutation rate modifiers**) are genetically linked to the mutations they bring forth. However, in the presence of recombination, mutation rate modifiers become genetically unlinked from the mutations they generate, and so cannot hitchhike to a higher frequency with beneficial mutations, or to a lower frequency when linked to deleterious mutations. In sexual populations, selection will tend to minimise the mutation rate irrespective of what the optimum mutation rate is [34,36] (Figure 1). This rather counterintuitive situation is caused by the fact that deleterious mutations greatly outnumber advantageous mutations. As both types of mutation will be associated with the modifier for approximately the same length of time, selection will favour a minimum mutation rate even when this is not beneficial over the longer term. As a consequence, optimal mutation rates can evolve only in largely asexual populations [34].

	Low cost of fidelity	High cost of fidelity
Asexual	①	②
Sexual	③	④
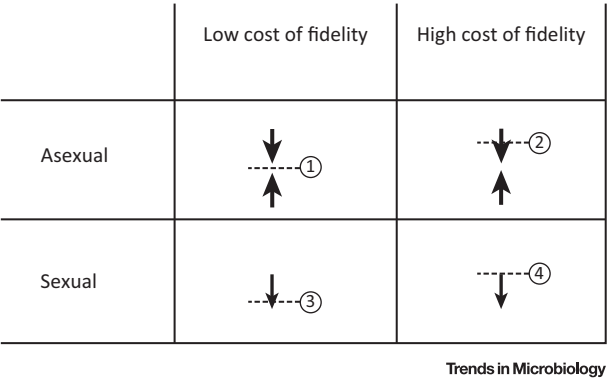
Trends in Microbiology

Figure 1. Four Different Scenarios of Selection on the Rate of Gene Content Turnover in Prokaryotes. In species that are largely asexual, selection could theoretically result in an optimal, nonminimal rate of lateral gene transfer (LGT) (two arrows). When costs of minimizing LGT rate are low, this optimum could be attained in reality (scenario 1), when costs are high, LGT rate will remain higher than the optimum (scenario 2). When recombination results in unlinking the genes that control LGT rate from the accessory genes they bring into the genome, theory predicts that selection will act to minimize LGT rate and this selection is relatively inefficient (small downward arrows). Again, depending on the cost of minimizing LGT rate, realized rates will be closer (scenario 3) or less close (scenario 4) to the optimum rate.

How do the theoretical expectations for the evolution of mutation rate outlined above apply to LGT rate in prokaryotes? Lateral gene transfer and gene loss events can have a range of fitness effects (see Box 1 for a brief review of laboratory experiments). On average, gene transfers are more likely to decrease fitness than are nonsynonymous mutations (although this is not always true; in the case of antibiotic resistance, plasmids carry a lower fitness cost than do point mutations on the chromosome [37]). However, it is likely that the proportion of changes with strongly beneficial fitness effects is higher for LGT than it is for mutation. LGT events can result in the uptake of whole genes, operons, plasmids, or 'genomic islands' [38], allowing for the immediate gain of new functions which could have strongly beneficial effects when adapting to a novel environment. For example, cryptic prophages in *E. coli* can aid the cell in withstanding a variety of adverse environmental conditions [39]. The DFE of lateral gene transfers thus could potentially be more conducive to the evolution of a higher optimal rate (Figure 2, Key Figure).

Whether bacteria classify as sexual or asexual in the context of the association of modifiers of LGT rate with the rest of the genome has, to our knowledge, not been addressed. It could be argued that LGT itself is a form of recombination, shuffling accessory genes across different genomic backgrounds, and that linkage between LGT modifiers and the accessory genes whose uptake they govern thus cannot be achieved. However, unlike meiotic sex, where mutation rate modifiers are shuffled into a new genome every generation, it is conceivable that a novel gene taken up by a recipient cell stays associated with that genomic background for an appreciable time. This could mean that it is possible that prokaryotes have evolved high LGT rates in order to secure the uptake of rare but highly beneficial genes (Figure 2, Key Figure) [21]. The concomitant influx of many more deleterious genes then is counteracted by purifying selection either on the level of individual genes through gene loss or, more likely, on the level of genomes through cell death.

An alternative explanation of the observed data is that the LGT rate is actually higher than the theoretical optimum (Figure 1). The classic explanation of why mutation rates cannot evolve to be lower based on the physiological costs associated with optimizing replication and repair fidelity [34] could theoretically also apply to LGT. For instance, the mechanisms of resistance to mobile genetic agents, such as transducing phage, are diverse and physiologically complex [29] and the cost of complete protection against infection could be prohibitively high or simply impossible to achieve [40]. However, this argument is less likely to hold for LGT mediated by transformation,

Box 1. Fitness Effects of Changes in Gene Content: Experimental Evidence

Comparative genomics studies are suggestive of adaptive benefits of many of the observed changes in gene content in free-living bacteria (e.g., [20,52,53]); however, it is obvious that such approaches are biased towards changes that have been maintained over evolutionary time by selection [54]. Controlled experiments can circumvent this bias.

Gene Gain by Lateral Gene Transfer (LGT)

In an original approach dating from the era when it was still common to sequence prokaryote genomes to completion, Sorek *et al.* used raw sequencing data of a large number of genome projects to identify which genes are under-represented in clone libraries (necessitating subsequent noncloning-based resequencing) [55]. It could be shown that specific gene families systematically failed to be cloned into the *Escherichia coli* host. When a subset of these genes was artificially expressed in *E. coli*, the majority inhibited growth, indicating toxic effects. A bioinformatics survey could show that these same genes were, on average, less widely distributed among species, confirming the generality of the boundaries to their lateral transfer. This study demonstrates that a minority of introduced genes can have direct negative effects on fitness. The actual number of genes resisting LGT must be higher when it is taken into account that the genes under study here were physically forced into the host by electroporation and were present on plasmids, not the chromosome.

In a more recent experimental study, Knöppel *et al.* [56] inserted *Bacteroides*, *Proteus*, and human intestinal phage genes into a neutral site in a *Salmonella* genome, followed by growth rate measurements of the resulting mutants in competition with the ancestral clone. Fitness effects were generally not significantly different from the control and were only in a minority of cases slightly negative. It must be noted that the effect of gene disruption was not considered in this study and that even fluorescence-activated cell sorting (FACS)-based competition assays, as used here, lack the power to detect detect small, but nevertheless evolutionarily significant, fitness effects [56,57]. Most artificially introduced nonsynonymous substitutions would probably not have yielded significant fitness effects either, even though it is well established that the majority of such mutations are deleterious and removed over longer time scales [58]. However, this study does indicate that it is likely that a sizeable proportion of foreign genes could be taken up and maintained for many generations and serve as the raw material of adaptation.

Gene loss

High rates of observed gene loss can be explained by two general processes. In the first, genetic drift through frequent population bottlenecks can cause deleterious loss of genes. This process has been experimentally observed in *Salmonella* forced to undergo single-cell bottlenecks during laboratory evolution [59] and can be deduced from genome data for endosymbionts and obligate pathogens. In a second process, referred to as genome stream lining, gene loss is advantageous when it rids genomes of unused and metabolically burdensome genes. Again, in *Salmonella*, Koskiniemi *et al.* could demonstrate that a large portion of randomly introduced deletions conferred significant fitness benefits [60]. Furthermore, during prolonged laboratory evolution, spontaneous deletions that increased fitness were observed. In another evolution experiment using *Methylobacterium*, Lee and Marx found that portions of a megaplasmid representing up to 10% of the genome were lost over the course of adaptation [61]. In a separate mutation-accumulation experiment, in which selection was absent, no such deletions were observed. Both studies could verify the adaptive benefits of deletions by reconstructing them in the ancestral strain and observing increased growth rate.

Key Figure

The Hypothetical Distribution of Fitness Effects for Point Mutations and Lateral Gene Transfer (LGT) Events
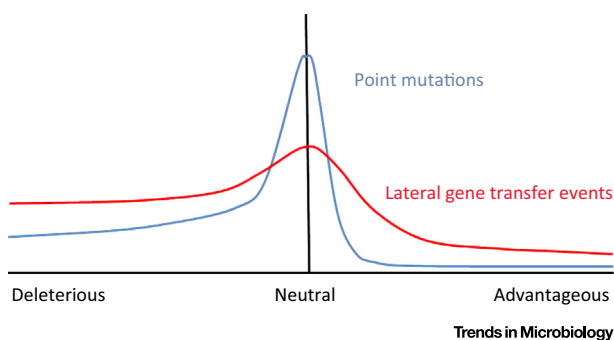
Figure 2. We might imagine that LGT events are generally subject to stronger positive and negative selection, and hence that the distribution for LGT events is more dispersed than the distribution for single-nucleotide mutations.

which is under control of the cell itself [41]. Inactivating the ability to transform would lower LGT rate and is expected to come at a direct fitness gain, and so the presence of this mechanism [42] would argue against the scenario where selection to reduce the mutation rate is constrained (but note that this reasoning ignores any potential benefits of transformation other than recombination, i.e., DNA repair or nutrition [43,44]).

Another explanation of how selection to lower mutation or LGT rate can be constrained comes in the form of the drift-barrier hypothesis [30,31]. This hypothesis is based on the fact that the fitness gains of achieving lower rates of change will become incrementally smaller, up to a point where selection is unable to drive down rates of genomic change any further and rates of change are solely influenced by drift [30,31]. The drift-barrier hypothesis predicts that species with high effective population size N_e and therefore more efficient selection (the selection coefficient s equals the inverse of N_e) have the lowest mutation rates. Although mutation rate indeed shows a negative correlation with N_e, [30], it has been argued that observed genomic mutation rates are higher than the theoretical minimum set by drift [35]. Regardless of the above disagreement, the possibility that rates of gene content change could be high relative to point mutation in free-living bacteria with large effective population size (e.g., $N_e \sim 10^7$ for *E. coli* [45]) suggests that selection does not act to minimize either LGT or gene loss or both (i.e., is not at the drift limit).

Concluding Remarks

Although there is a marked lack of available data on rates of LGT [46], analysis of closely related genomes from a variety of species suggests that prokaryote genomes could potentially change more rapidly due to lateral gene transfers than due to point mutation. Understanding the (in)ability of selection to drive LGT rate has important implications for our understanding of a range of topics in evolutionary microbiology, such as adaptation to novel niches (Box 2). It has become commonplace to estimate the size of the core and accessory genomes of various bacterial species based on samples of sequenced genomes (e.g., [6,7]). The finding that many accessory genes are deleterious and transient means that the accessory genome does not constitute a purely adaptive gene pool. As a consequence, it is likely that the accessory genome will continue to increase as more strains are sampled not because accessory genes are advantageous, but because there is essentially an infinite supply of deleterious genes available for uptake. Conversely, core genome size can be underestimated because at least some cases of gene loss are deleterious.

The selective pressures underlying LGT rate have remained largely unexplored and are not straightforward to tease out, especially as they are governed by multiple distinct mechanisms which are under varying levels of evolutionary constraint [43]. To address the many outstanding

Box 2. Lateral Gene Transfer (LGT) and Adaptation

LGT has a crucial impact on ecological diversification and, over longer time scales, the evolution into more or less distinct species [17,53,62]. Ecological adaptation is expected to occur especially rapidly when clones find themselves in a novel niche filled with many previously not encountered species from which niche-relevant genes could be obtained [11,17,63]. It could be expected that in times of rapid adaptation, the LGT rate evolves to be higher (or is upregulated) whereas in times of relative ecological stasis, LGT rates evolve to be lower. It has recently been demonstrated that natural variation in transformation is under the control of multiple genes [64]. Multiple loci will form a larger target for mutation and LGT, potentially allowing more efficient fine-tuning of a single optimum gene turnover rate (but note that when the rate of LGT modifier turnover is too high it precludes selection for optimal rates). Alternatively, lineages could alternate between high and low rates over short evolutionary time scales, akin to mutators periodically invading populations with low mutation rate [65]. Great natural variation in the degree of transformation among clones in a variety of species has led to the hypothesis that competent species continually give rise to noncompetent lineages which can have a short-term advantage under conditions where little adaptation is required [47]. The finding that CRISPR-Cas loci preventing mobile genetic element-mediated LGT events (e.g., [66]) can be gained or lost over short evolutionary time scales (e.g., [67]) could be consistent with this scenario.

Outstanding Questions

What is the rate of lateral gene transfer (LGT) per genome, per generation, for different prokaryote taxa?

What is the relative contribution of the different genetic mechanisms responsible for LGT, and how readily could the action of these mechanisms themselves be changed through gene gain or loss?

How important is selection against individual newly introduced genes (with disruptive or toxic effects) relative to selection against increasing genome size *per se* (increased costs of replication and expression)?

Are rates of homologous recombination and LGT (nonhomologous recombination) positively correlated?

Loss of competence genes has been observed to be followed by loss of CRISPR-Cas loci [47]. How do different genetic mechanisms responsible for LGT interact?

The uptake of genes is not random with regard to genomic position [48]. Does LGT take place in genomic hotspots, and how would such linkage influence LGT rate evolution?

Can environmental cues that result in physiological upregulation of LGT processes reveal the selective benefits of these processes?

How important are compensatory mechanisms such as 'xenogeneic silencing' [49] in reducing the cost of deleterious LGT events relative to selection against novel genotypes?

Testing for selection in comparative genomics studies generally relies on traditional population genetics approaches based on mutation in homologous sequences. However, solely employing core-genome-based tests will give an incomplete and biased view of adaptive evolution. Can tests for selection that are based on accessory gene content be developed (see e.g., [50,51])?

questions (see Outstanding Questions), comparative genomics approaches could be used to identify differences in LGT rates between species and correlate these with ecological and genetic characteristics. Perhaps more promising are controlled experiments (Box 1) quantifying rates of gene gain and loss, possibly in the form of LGT accumulation experiments (analogous to mutation accumulation experiments). Understanding whether high LGT rates are primarily due to fighting a losing battle against continuous invasion by selfish genetic elements, or due to selection to obtain 'rare gems' amidst the rubble that is the community metagenome, will be of fundamental importance to the study of prokaryote evolution.

Author Contributions

M.V. and A.E-W. conceived the paper, performed analyses and wrote the manuscript, M.C.H., T.A.tB. and M.W.J.VP. performed analyses, helped with data interpretation and manuscript revision.

References

1. Lan, R. and Reeves, P.R. (2000) Intraspecies variation in bacterial genomes: the need for a species genome concept. *Trends Microbiol.* 8, 396–401

2. Young, J.P.W. *et al.* (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* 7, R34

3. van Passel, M.W. *et al.* (2008) The emergence and fate of horizontally acquired genes in *Escherichia coli*. *PLoS Comput. Biol.* 4, e1000059

4. Doolittle, W.F. (2012) Population genomics: how bacterial species form and why they don't exist. *Curr. Biol.* 22, R451–R453

5. Koonin, E.V. and Wolf, Y.I. (2009) The fundamental units, processes and patterns of evolution, and the tree of life conundrum. *Biol. Direct* 4, 33

6. Tettelin, H. *et al.* (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. U.S.A.* 102, 13950–13955

7. Kislyuk, A.O. *et al.* (2011) Genomic fluidity: an integrative view of gene diversity within microbial populations. *BMC Genomics* 12, 32

8. Kuo, C.H. and Ochman, H. (2009) The fate of new bacterial genes. *FEMS Microbiol. Rev.* 33, 38–43

9. Rocha, E.P. (2008) Evolutionary patterns in prokaryotic genomes. *Curr. Opin. Microbiol.* 11, 454–460

10. Wiedenbeck, J. and Cohan, F.M. (2011) Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol. Rev.* 35, 957–976

11. Smillie, C.S. *et al.* (2011) Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 480, 241–244

12. Kloesges, T. *et al.* (2011) Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Mol. Biol. Evol.* 28, 1057–1074

13. McDaniel, L.D. *et al.* (2010) High frequency of horizontal gene transfer in the oceans. *Science* 330, 50

14. Treangen, T.J. and Rocha, E.P. (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* 7, e1001284

15. Dagan, T. *et al.* (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10039–10044

16. Pal, C. *et al.* (2005) Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* 37, 1372–1375

17. Polz, M.F. *et al.* (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* 29, 170–175

18. Hao, W. and Golding, G.B. (2006) The fate of laterally transferred genes: life in the fast lane to adaptation or death. *Genome Res.* 16, 636–643

19. Marri, P.R. *et al.* (2006) Gene gain and gene loss in *Streptococcus*: is it driven by habitat? *Mol. Biol. Evol.* 23, 2379–2391

20. Marri, P.R. *et al.* (2007) The role of laterally transferred genes in adaptive evolution. *BMC Evol. Biol.* 7, S8

21. Nowell, R.W. *et al.* (2014) The extent of genome flux and its role in the differentiation of bacterial lineages. *Genome Biol. Evol.* 6, 1514–1529

22. Touchon, M. *et al.* (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* 5, e1000344

23. Didelot, X. *et al.* (2012) Impact of homologous and non-homologous recombination in the genomic evolution of *Escherichia coli*. *BMC Genomics* 13, 256

24. Hao, W. and Golding, G.B. (2010) Inferring bacterial genome flux while considering truncated genes. *Genetics* 186, 411–426

25. Baltrus, D.A. (2013) Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* 28, 489–495

26. Millan, A.S. *et al.* (2015) Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nat. Commun.* Published online April 21, 2015. http://dx.doi.org/10.1038/ncomms7845

27. Engelberg-Kulka, H. and Glaser, G. (1999) Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu. Rev. Microbiol.* 53, 43–70

28. MacLean, R.C. *et al.* (2013) Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat. Rev. Genet.* 14, 221–227

29. Westra, E.R. *et al.* (2012) The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annu. Rev. Gen.* 46, 311–339

30. Sung, W. *et al.* (2012) Drift-barrier hypothesis and mutation-rate evolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18488–18492

31. Lynch, M. (2010) Evolution of the mutation rate. *Trends Genet.* 26, 345–352

32. Drake, J.W. *et al.* (1998) Rates of spontaneous mutation. *Genetics* 148, 1667–1686

33. Kondrashov, A.S. (1995) Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? *J. Theor. Biol.* 175, 583–594

34. Sniegowski, P.D. *et al.* (2000) The evolution of mutation rates: separating causes from consequences. *Bioessays* 22, 1057–1066

35. Martincorena, I. and Luscombe, N.M. (2013) Non-random mutation: The evolution of targeted hypermutation and hypomutation. *Bioessays* 35, 123–130

36. Leigh, E.G., Jr (1970) Natural selection and mutability. *Am. Nat.* 301–305

37. Vogwill, T. and MacLean, R.C. (2015) The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evol. Appl.* 8, 284–295

38. Langille, M.G. *et al.* (2010) Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.* 8, 373–382

39. Wang, X. *et al.* (2010) Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147

40. Vogan, A.A. and Higgs, P.G. (2011) The advantages and disadvantages of horizontal gene transfer and the emergence of the first species. *Biol. Direct* 6, 1

41. Seitz, P. and Blokesch, M. (2012) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol. Rev.* 37, 336–363

42. Johnston, C. *et al.* (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* 12, 181–196

43. Vos, M. (2009) Why do bacteria engage in homologous recombination? *Trends Microbiol.* 17, 226–232

44. Redfield, R.J. (2001) Do bacteria have sex? *Nat. Rev. Genet.* 2, 634–639

45. Charlesworth, J. and Eyre-Walker, A. (2006) The rate of adaptive evolution in enteric bacteria. *Mol. Biol. Evol.* 23, 1348–1356

46. Nielsen, K.M. *et al.* (2014) Detecting rare gene transfer events in bacterial populations. *Front. Microbiol.* 4, 1–12

47. Jorth, P. and Whiteley, M. (2012) An evolutionary link between natural transformation and CRISPR adaptive immunity. *Mbio* 3, e00309–e00312

48. Ray, J.L. *et al.* (2009) Sexual isolation in *Acinetobacter baylyi* is locus-specific and varies 10,000-fold over the genome. *Genetics* 182, 1165–1181

49. Navarre, W.W. *et al.* (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313, 236–238

50. Baumdicker, F. *et al.* (2012) The infinitely many genes model for the distributed genome of bacteria. *Genome Biol. Evol.* 4, 443–456

51. Collins, R.E. and Higgs, P.G. (2012) Testing the infinitely many genes model for the evolution of the bacterial core genome and pangenome. *Mol. Biol. Evol.* 29, 3413–3425

52. Tian, C.F. *et al.* (2012) Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8629–8634

53. Shapiro, B.J. *et al.* (2012) Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48–51

54. Gogarten, J.P. and Townsend, J.P. (2005) Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* 3, 679–687

55. Sorek, R. *et al.* (2007) Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318, 1449–1452

56. Knöppel, A. *et al.* (2014) Minor fitness costs in an experimental model of horizontal gene transfer in bacteria. *Mol. Biol. Evol.* 31, 1220–1227

57. Eyre-Walker, A. and Keightley, P.D. (2007) The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8, 610–618

58. Rocha, E.P. *et al.* (2005) Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J. Theor. Biol.* 239, 226–235

59. Nilsson, A. *et al.* (2005) Bacterial genome size reduction by experimental evolution. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12112–12116

60. Koskiniemi, S. *et al.* (2012) Selection-driven gene loss in bacteria. *PLoS Genet.* 8, e1002787

61. Lee, M-C. and Marx, C.J. (2012) Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet.* 8, e1002651

62. Shapiro, B.J. and Polz, M.F. (2014) Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol.* 22, 235–247

63. Popa, O. *et al.* (2011) Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. *Genome Res.* 21, 599–609

64. Mell, J.C. *et al.* (2014) Extensive cotransformation of natural variation into chromosomes of naturally competent *Haemophilus influenzae*. *G3 (Bethesda)* 4, 717–731

65. Taddei, F. *et al.* (1997) Role of mutator alleles in adaptive evolution. *Nature* 387, 700–702

66. Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845

67. Díez-Villaseñor, C. *et al.* (2010) Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 156, 1351–1361