Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Microbiology

Trends and barriers to lateral gene transfer in prokaryotes

Ovidiu Popa and Tal Dagan

Gene acquisition by lateral gene transfer (LGT) is an important mechanism for natural variation among prokaryotes. Laboratory experiments show that protein-coding genes can be laterally transferred extremely fast among microbial cells, inherited to most of their descendants, and adapt to a new regulatory regime within a short time. Recent advance in the phylogenetic analysis of microbial genomes using networks approach reveals a substantial impact of LGT during microbial genome evolution. Phylogenomic networks of LGT among prokaryotes reconstructed from completely sequenced genomes uncover barriers to LGT in multiple levels. Here we discuss the kinds of barriers to gene acquisition in nature including physical barriers for gene transfer between cells, genomic barriers for the integration of acquired DNA, and functional barriers for the acquisition of new genes.

Address
Institute of Molecular Evolution, Heinrich-Heine University of Düsseldorf, Universitätstr. 1 40225, Düsseldorf, Germany

Corresponding author: Dagan, Tal (tal.dagan@hhu.de)

Current Opinion in Microbiology 2011, **14**:615–623

This review comes from a themed issue on
Genomics
Edited by Luis Serrano and Rotem Sorek

Available online 17th August 2011

1369-5274/$ – see front matter

DOI 10.1016/j.mib.2011.07.027

Introduction

Prokaryotes possess the unique ability to acquire DNA from the environment, or their neighbors, and incorporate it into their genome in a process called lateral gene transfer (LGT) [1]. Accumulating evidence shows that LGT plays a major role in prokaryote genome evolution [2–4], affecting virtually all genes [5–7], with only few genes that are resistant to it [8]. Lateral gene transfer is crucial to our understanding of microbial evolution; furthermore, as a source of natural variation it facilitates the emergence of novel infectious diseases through the spread of virulence mechanisms (e.g. [9,10]).

The known mechanisms for LGT include transformation, conjugation, transduction, and gene transfer agents. Transformation involves the uptake of naked DNA from the environment [11,12]. Conjugation is the transfer of DNA via plasmids, a process that is mediated by a proteinaceous cell-to-cell junction, forming a tunnel through which the DNA is transferred [13,14]. Transduction is DNA acquisition following a phage infection [12], and gene transfer agents (GTA) are phage-like DNA-vehicles that are produced by a donor cell and released to the environment [15,16] (Figure 1). An additional transfer mechanism – nanotubes – was discovered recently [26••]. These are tubular protrusions composed of membrane components that can bridge between neighboring cells and conduct the transfer of DNA and proteins (Figure 2).

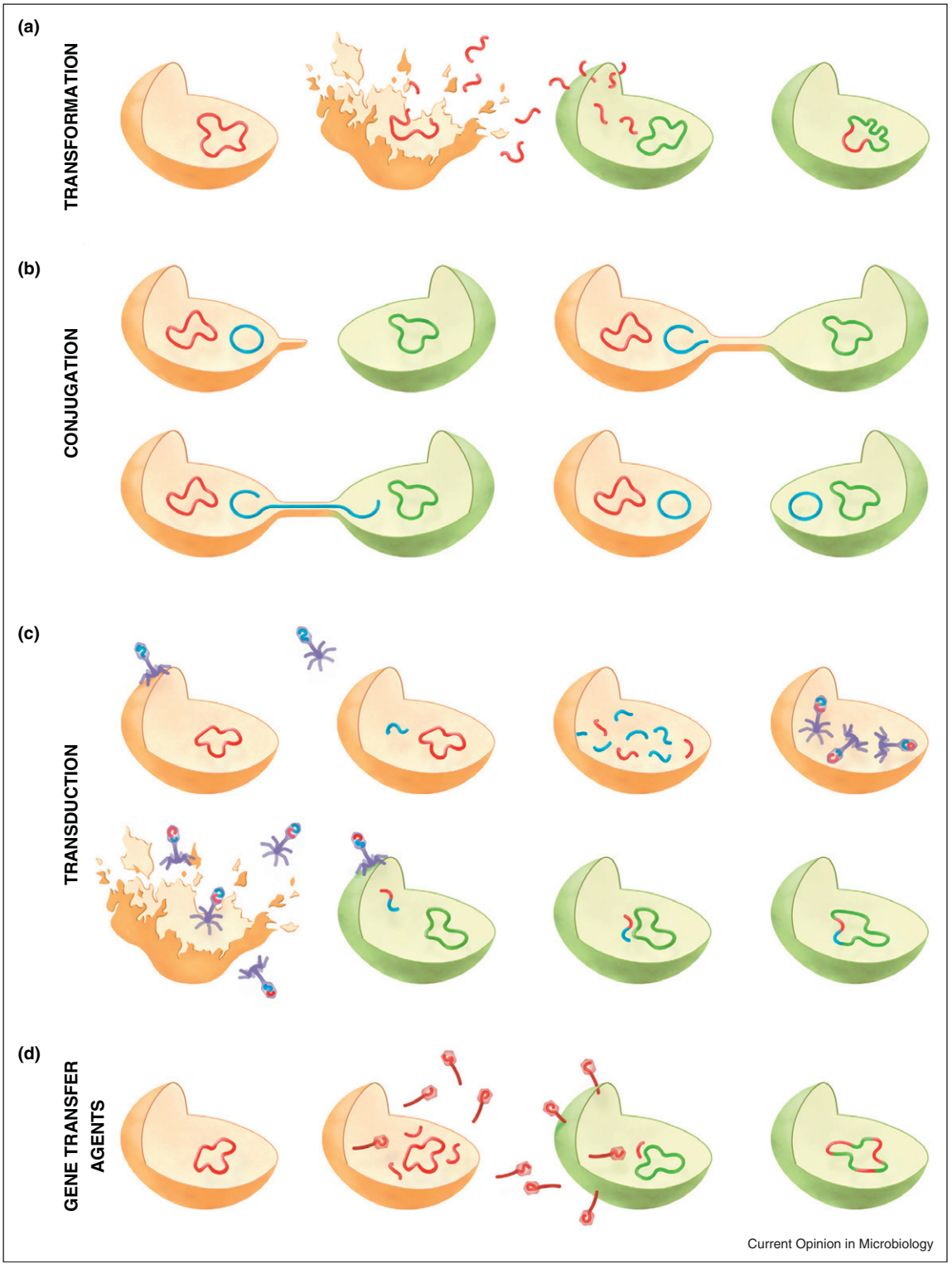
Lateral gene transfer frequency

Several experiments have been conducted in order to quantify the frequency of LGT in nature. For example, Babić *et al.* [27] tested the success rate of gene acquisition by conjugation in *Escherichia coli*. Using a plasmid encoding a gene for fluorescence protein (YFP) they quantified the odds for a successful integration of plasmid genes into the recipient genome. They found that in 96% of the population the YFP gene was integrated into the chromosome and inherited to the next generation. The percolation of an acquired DNA within the population can be extremely fast in *Bacillus subtilis* where the cells are arranged in chains. Tracking the spread of an integrative and conjugative element (ICE) encoding a gene for green fluorescence protein (GFP) under the microscope showed that in 43 (81%) out of 53 cases a recipient cell turned into a donor and transconjugated the ICE to the next cell in line, often within 30 min [28••].

Lateral gene transfer via transduction takes place during a phage infection. Hence gene acquisition by this transfer mechanism depends on the survival of the recipient. In a recent study Kenzaka *et al.* [29] quantified the survival rate of phage infected enteric bacteria as 20% of the population. These surviving bacteria may acquire DNA from previous hosts of the attacking phage. Recent measurements of LGT by gene transfer agents (GTAs) in marine α-proteobacteria revealed that this transfer mechanism is probably the most efficient one. McDaniel *et al.* [30••] measured the frequency of LGT by the acquisition rate of Kanamycin resistance gene. Their results show that gene transfer by GTA is more efficient than transformation or transduction by orders of magnitude.

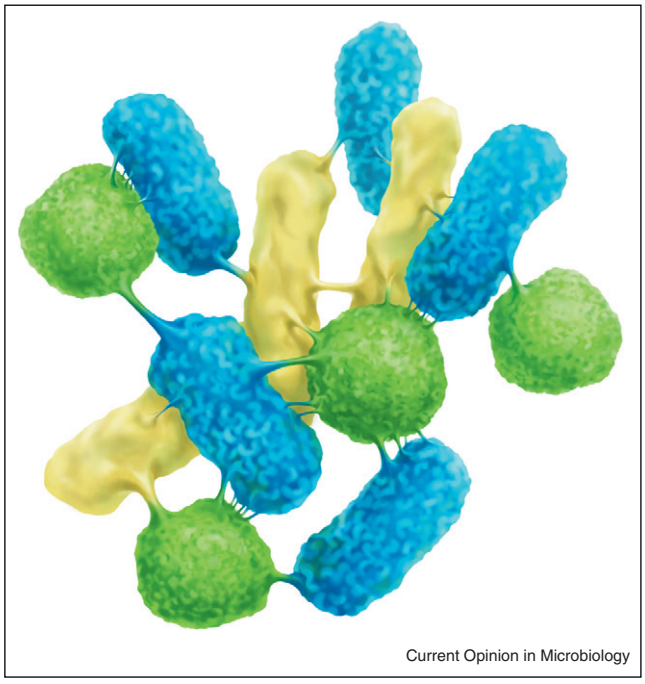
The exact mechanism of DNA transfer via nanotubes is yet unknown. In addition to intracellular molecules, nanotubes conduct also nonconjugative plasmids and even viral particles [26••]. The promiscuity of the nanotubes attachment and their architecture dimensions suggest that they play an important role in all transfer

Figure 1



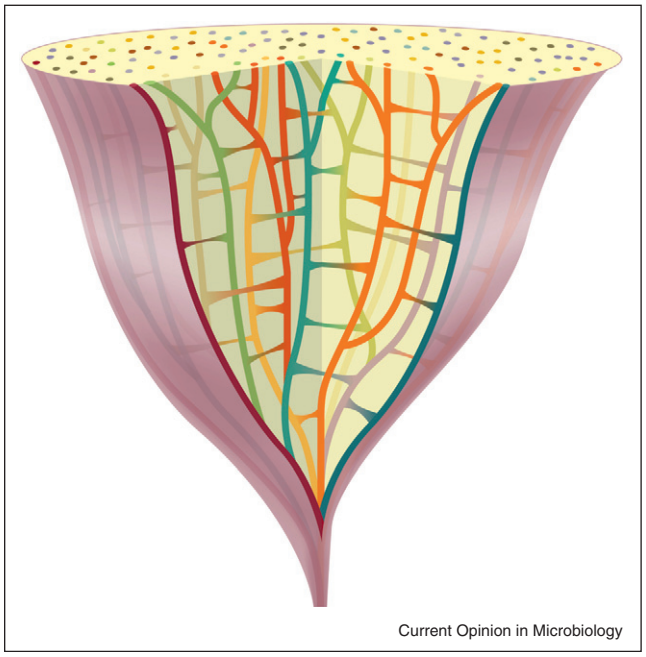
LGT mechanisms. **(a)** The uptake of raw DNA in transformation is enabled during a competence state that involves 20–50 proteins, including the type IV pillus and type II secretion system proteins [11,12]. In some species, an effective transformation requires the presence of uptake signal sequences (USSs; called also DUS: DNA uptake signal). These are specific DNA motifs, about 10 bp long, that are encoded within the recipient genome in a frequency that is much above that expected by chance [17]. Environmental DNA molecules bearing the USS motif are recognized by specific receptors at the cell surface, imported into the cytoplasm, and can then be readily integrated into the recipient chromosomes, usually via homologous recombination [11,12,18,19]. **(b)** During conjugation, plasmids can integrate into the recipient chromosomes by homologous

Figure 2



A schematic representation of cells interconnected by nanotubes. The nanotubes are between 30 and 130 nm wide and up to 1 μm long. The tube dimension is correlated with the distance between the connected cells. Proximal cells are commonly connected by several small nanotubes, while thicker tubes connect distant cells. The rate of transfer via the nanotubes correlates negatively with the size of the transferred substance. Cellular interconnection mediated by nanotubes is not species specific. However, morphology and diameter of the tubes seem to depend on characteristics of the connected cells [26••].

Figure 3



A phylogenetic tree of microbial genomes. Tree branches correspond to genomes and branch colors represent different lineages. Horizontal connections between the branches correspond to LGT events. Bifurcating phylogenetic tree models cannot account for the lateral component of microbial genome evolution [3].

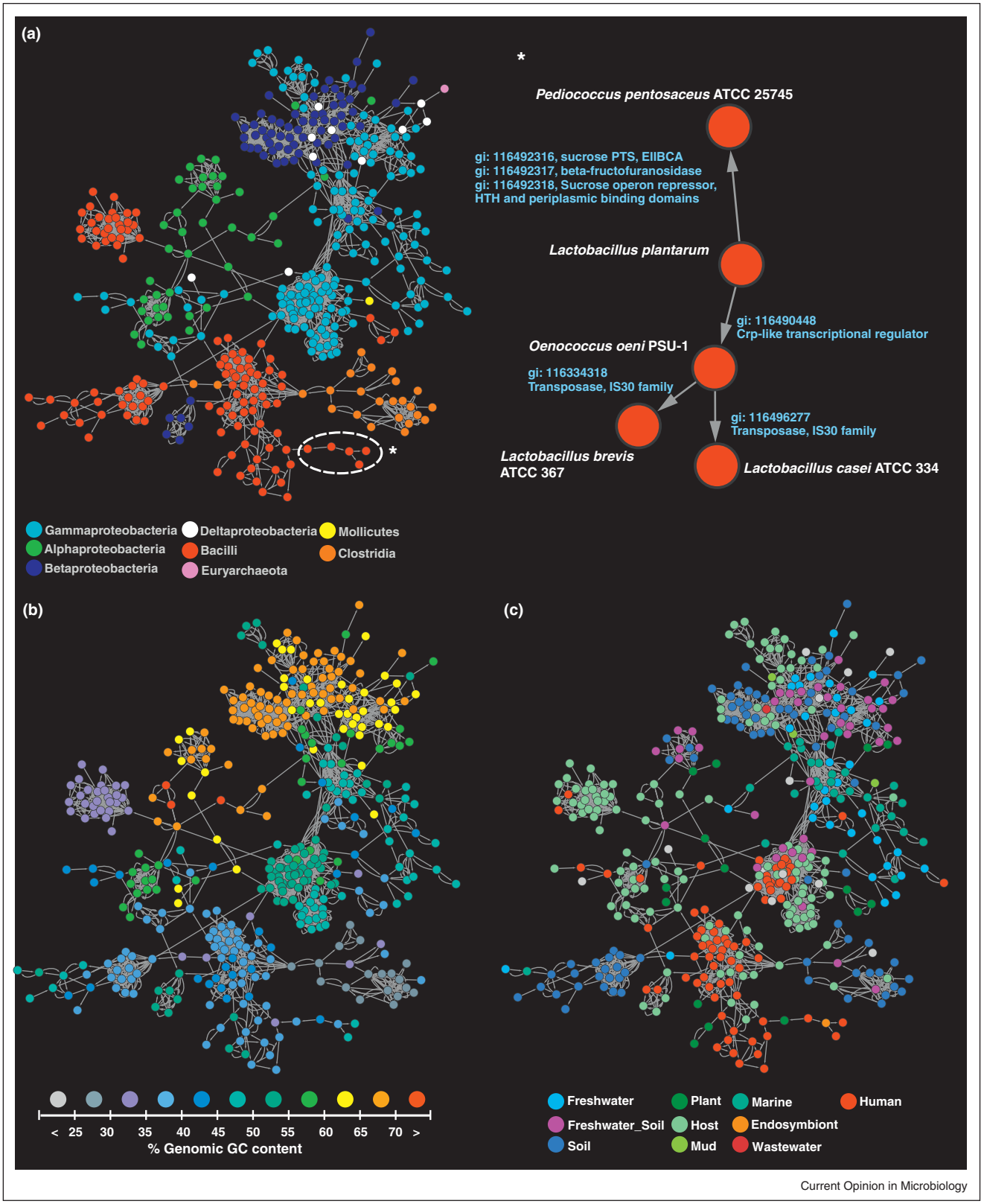
mechanisms, by enabling the propagation of acquired DNA within the population.

We know that LGT occurs in the laboratory, the issue is how often it occurs in the wild and how important it is during evolution. Phylogenetic reconstruction of microbial genes reveals that LGT plays a major role in shaping prokaryotic genomes [5–7,31•,32]. In a pioneering study, Lawrence and Ochman [33] identified all *E. coli* genes that were acquired since its divergence from the *Salmonella* lineage by their aberrant codon usage. They estimated that 755 (18%) of the 4288 genes in *E. coli* strain MG1655 were laterally acquired over a time period of about 14 million years (Myr) and estimated the LGT rate as 16 kb/1 Myr per lineage [33]. Using gene distribution patterns across 329 proteobacterial genomes, Kloesges *et al.* [32] recently estimated that at least 75% of the protein families have been affected by LGT during evolution. Gene transfer rate in those families is on average 1.9 events per protein family per lifespan [32]. Similar estimates were found in phylogenetic analyses of broader taxonomic samples [5–7].

The impact of LGT during genome evolution can be estimated either by the proportion of recently transferred genes whose unusual base composition and codon usage still bears the marks of acquired DNA [33–35] or by phylogenetic analysis of individual genes including recent and ancient LGTs alike (e.g. [36–39]). A survey of genes having aberrant nucleotide composition within proteobacterial genomes revealed that $21 \pm 9\%$ of the genes in those genomes comprises recent acquisitions

(Figure 1 Legend Continued) recombination that may entail insertion sequences (ISs) or other sequences conserved between plasmid and recipient chromosomes that carry the minimal sequence similarity required for homologous recombination [14,20,21]. **(c)** Phages recognize possible hosts by specific receptors found on the cell surface. Many phages include in their genomes chunks of DNA taken coincidentally from previous hosts. These are transferred to the new host during the integration of the phage genome into the host chromosomes. DNA integration into the host chromosome is generally mediated by the phage-encoded enzymes that specifically integrate the phage into the chromosome of the infected recipient [12]. **(d)** DNA stored in GTAs is imported into the recipient in a generalized transduction process mediated by a cellular *RecA* recombination system [22]. GTAs, unlike phages, are linked to transfer of genomic DNA only and GTA-induced cell lysis was not observed [23]. The mechanism of DNA packing and capsule release from the cell is still unknown. GTA systems have been documented not only in oceanic α-proteobacteria, but also in few archaebacteria and some spirochaetes [16,24,25].

Figure 4



A directed network of LGT [49]. The nodes correspond to contemporary or ancestral species that are connected by directed edges of LGT. The edges point from the LGT donor to the recipient. **(a)** Node color corresponds to species taxonomic classification (see legend at the bottom). A cluster of

[32]. Gene distribution patterns across the same species sample suggest that, on average, 74 ± 11% of the genes in each genome have been laterally transferred at least once during evolution [32].

Phylogenetic reconstruction of microbial genome evolution

Lateral gene transfer during microbial genome evolution poses an acute problem to standard phylogenetic reconstruction methodology. Species phylogeny is customarily represented by using bifurcating phylogenetic trees. However, the tree model applies only to the reconstruction of vertical inheritance where genetic material is transferred from ancestral species to their descendants. More realistic models of prokaryotic genome evolution have to embrace lateral gene transfer in addition to vertical inheritance (Figure 3) [2,3,40,41•,42], yet such methods are still scarce. Network models are an alternative to bifurcating trees because they permit the reconstruction and depiction of reticulated evolutionary events such as recombination, gene fusions, and lateral gene transfer [43]. A network is composed of nodes (or vertices) connected by edges corresponding to entities connected by pairwise relations [44,45]. In a phylogenomic network the nodes are completely sequenced genomes and the edges correspond to phylogenetic relations between the genomes that they connect. Phylogenomic networks can be reconstructed from shared gene content (e.g. [6,31•,46]), shared sequence similarity (e.g. [47,48]), or phylogenetic trees [36,49].

The directed network of lateral gene transfer (dLGT) is a phylogenomic network recently developed in order to study the lateral component of recent microbial genome evolution [49]. The nodes in this network correspond to species or their ancestors. The edges represent recent lateral gene transfer events between the species that they connect and they are directed from the donor to the recipient in the LGT event (Figure 4A). Reconstructing a network of recent LGTs has two main advantages: first, the phylogenetic reconstruction of recent transfers is based on a comparison of relatively conserved gene sequences, which is less susceptible to phylogenetic artifacts [53]. Second, bacterial genomes are highly dynamic and may change considerably over time [54,55]; focusing on recent LGTs allows the coupling of the information regarding LGT and current cellular characteristics of donors and recipients. Using the dLGT network one can study trends in – and barriers to – LGT during microbial evolution. In what follows we present some of the insights that this approach permits.

Donor–recipient similarity barrier

Most of the detected LGT in the dLGT network occur between closely related species from the same taxonomic group (Figure 4A) [49]. A graphical representation of the network with species colored by their genomic GC content reveals that clusters of densely connected donors and recipients are very similar in their genomic GC content (Figure 4B). Furthermore, the difference in genomic GC content between donors and recipients is <5% for most (86%) of connected pairs [49]. This suggests that there exists a biological barrier for gene acquisition from donors of dissimilar genomic GC content. Indeed, one such mechanism was discovered in *Salmonella typhimurium* where a histone-like protein (*H-NS*) functions as a transcriptional repressor of GC-poor ORFs [56]. A comparison between the GC content of genes silenced by the *H-NS* repressor (46.8%) and the overall genomic GC content of *S. typhimurium* LT2 (52.2%) reveals that this mechanism is highly sensitive to foreign DNA with lower GC content than that of the genome [56]. However, it is apparent from the dLGT network that some LGT does occur between donors and recipients having difference GC content (Figure 4B). A possible bypass for *H-NS* silencing is provided by the plasmid encoded protein *sfh*, which has been shown to suppress the activity of *H-NS*, enabling the expression of GC-poor ORFs within the genome of *S. typhimurium* [57]. The *sfh* bearing plasmid was isolated from several enteric species, and its DNA sequence is GC-poor. Hence the *sfh* gene allows this plasmid to be transferred among enteric bacteria by escaping from the transcription suppression of the *H-NS* protein [57,58].

To further study the effect of donor–recipient genome sequence similarity on LGT frequency we calculated the similarity between the genomes of connected donors and recipients as the total length of all identical ≥20 bp genomic segments between the two genomes, divided by the recipient genome size. Using this measure we found that the dLGT network is enriched for connected donors and recipients having similar genome sequences [49]. Furthermore, donor–recipient genome sequence similarity and LGT frequency are positively correlated ($r_s = 0.55$, $P \ll 0.01$) [49]. This suggests that LGT is more frequent among closely related species, having similar

(Figure 4 Legend Continued) connected Bacilli (marked with a star) is enlarged to exemplify the network underlying data. Species names are shown next to the nodes. Gene identifier and protein annotation of detected recent gene transfers are noted next to the corresponding edge. The lateral acquisition of genes for sucrose utilization in *Pediococcus pentosaceus* from *Lactobacillus plantarum* has been suggested before [50]. *Oenococcus oeni* (strain PSU-1) – that is associated with malolactic fermentation (MLF) in wine – is connected with three different *Lactobacillus* species as donor and recipient. Bon *et al.* [51] showed recently that gene acquisition from various donors, especially lactic acid Bacilli, contributes to genome plasticity in this species and suggested LGT as a mechanism to enhance *O. oeni* tolerance for harsh wine conditions. **(b)** Node color corresponds to the genomic GC content calculated as the proportion of Guanine and Cytosine (GC) nucleotides within the genome (see scale at the bottom). **(c)** Nodes in the network are colored by habitat (the ten main habitats are listed at the bottom). Habitat annotation was extracted from the GOLD database (ver. 12/2010) [52]. Ancestral nodes are colored by the habitat of their descendants if it is homogeneous or gray otherwise.

genomes, while LGT between distantly related species is more rare [49]. This observation has been recently supported by a study of LGT using simulated genome evolution [59] and through the study of tyrosil-tRNA synthetase phylogeny [60]. High donor–recipient genome similarity in Gram-negative bacteria from the Neisseriales or Pasteurellales orders could be due to frequent gene acquisition by transformation (Figure 1) leading to a high frequency of uptake signal sequences (USSs) in the genomes of both donor and recipient.

The distribution of shared genes across genomes has a strong phylogenetic signal [7] hence there must be some restrictions to DNA acquisition among prokaryotes. Barriers to LGT between distantly related species (having dissimilar genomes) are still poorly understood but are thought to depend on the transfer mechanism. During transformation and conjugation the acquired DNA is commonly integrated into the genome by homologous recombination [12], which requires high similarity between recombining sequences [61,62]. Genes encoded within plasmids that are transferred by conjugation may also be integrated into the recipient chromosome by transposases [13,14] whose function is independent of donor–recipient sequence similarity. However, recent advances in studying the function of the microbial anti-phage CRISPR system (for review see [63•]) revealed that this system also identifies and degrades foreign plasmids in addition to phages [64]. Hence the CRISPR system may function as a barrier for conjugative gene transfer by blocking non-self plasmids [63•]. During transduction, the acquired DNA is integrated into the recipient genome by the phage enzymes whose function is independent of donor–recipient genome similarity [12]. Barriers to phage-mediated gene transfer between distantly related species could be related to the frequency of phages whose host range is species-specific. Such phages indeed exist in marine environments; a test of host specificity for 44 clonal cyanophages revealed that 25 phages were *Prochlorococcus*-specific and 7 were *Synechococcus*-specific, while the remaining 12 (27%) could infect both species [65]. A recent report of antagonistic coevolution between *Pseudomonas fluorescens* and its parasitic phage [66] reveals that species-specific phages exist in terrestrial environments as well.

Ecological barrier

The physical distance between the donor and recipient in the LGT event depends upon the LGT mechanism. In transformation the distance between the donor and recipient depends upon the raw DNA stability within the environment [67]. Conjugation requires that the donor and recipient will be close enough for the formation of the conjugation tunnel. Transduction is considered as the longest range LGT mechanism because it entails phage mobility [67]. This suggests that most transfers should occur within habitats. A graphical representation of the dLGT network with species colored according to their habitat (Figure 4C) reveals that several clusters of the highly connected donors and recipients are variable in their habitat classification, yet, most (74%) of the detected LGT in the network occur between donors and recipients residing in the same habitat. A network of shared transposases among 774 microbial genomes supplies further support for the rarity of inter-habitat gene transfers [68•].

Phylogenomic analyses of microbial genomes support this notion [31•,32]. For example, Halary *et al.* [31•] reconstructed a network of shared protein families among various genetic entities including microbial chromosomes, plasmids, and phage genomes. A comparison of network properties between plasmids and phage genomes revealed that plasmids are more frequently connected within the network in comparison to phages. From this they concluded that conjugation is more frequent than transduction in nature [31•].

Functional barrier

Once imported and integrated into the genome, acquired genes still have to adapt within the genome in order to be retained during evolution. Microbes tend to delete non-functional or otherwise unneeded DNA from their genomes [69,70]. Therefore, the fixation of the acquired DNA within the genome is highly dependent on its functionality or utility to the recipient under selectable environmental conditions [54,71,72]. In order to be expressed, the gene has to be either inserted near a recognized promotor, bring one with it, or be acquired together with the corresponding regulator. Hence, acquired genes that are inserted within existing regulatory circuits [73,74] or have a promotor of similar GC content as the recipient genomes [8], have a higher probability to be retained by the recipient.

Gene encoding by suboptimal codons that do not fit the tRNA pool of the recipient has been considered a barrier to LGT [75,76]. However, two recent experimental studies show that the impact of codon usage on the expression and retention of acquired genes might have been overrated. Kudla *et al.* [77] compared the expression level of 154 synthetic gene copies encoding for GFP that varied randomly in synonymous sites (i.e. in their codon usage). The synthetic genes were cloned (i.e. transferred) into *E. coli* and the expression level of GFP was measured by the fluorescence level of the cultures. The result showed that codon usage and fluorescence level of genes are not correlated within the recipient ($r = 0.02$) [77]. Hence the expression level of acquired genes within a recipient cell immediately after the acquisition is independent of their codon usage. In another study, Amorós-Moya *et al.* [78•] compared the fitness of three *E. coli* cultures into which they cloned a chloramphenicol resistance gene encoded by three different codon usage regimes: optimal, GC-rich, and AT-rich. Their results showed that cultures encoding for

the suboptimal gene variants were 10–20 times more sensitive to the antibiotics than those that encoded for the optimal variant. However, within 358 generations of experimental evolution (roughly 54 days) under antibiotic selection, these differences vanished. Interestingly, the compensating mutations were restricted to *in cis* substitutions within the gene promotor or *in trans* substitutions in the host genome, with no substitutions in the gene coding sequence [78•]. In this experiment the new gene acquisition, even of one that is encoded by suboptimal codons, is highly advantageous because the selection regime acts on bacteria that have either low or no antibiotic resistance (i.e. its a matter of life and death). Notably, the evolution of elevated protein expression level in this experiment took place without nucleotide substitutions leading to optimal codons.

Most laterally transferred genes perform metabolic functions while the transfer of genes performing information processing (including replication, transcription, and translation) is rare [49,79,80]. According to the complexity hypothesis [79], the scarcity of lateral transfer of information processing genes is attributed to their role in complex structures. Proteins that function in a complex structure, for example ribosomal proteins, are adapted to their common function. An LGT event that leads to replacement of such a gene with a less adapted homolog will result in a 'squeaking wheel' within the complex and reduced fitness of the recipient [79]. In a recent study, Cohen *et al.* [81] tested the relative impact of functional category and the number of interacting partners on LGT frequency. Their result showed that the complexity hypothesis still passes a reality test in the genomic era. However, LGT barriers owing to multiple interacting partners are not restricted to information processing genes only, but may be observed in all functional categories [81]. Acquisition frequency of metabolic genes depends on their role within the cellular metabolic network [71]. A study of the laterally acquired genes within the *E. coli* metabolic network showed that LGT is more frequent among enzymes involved in peripheral reactions (uptake and metabolism of nutrients) in comparison to those involved in central reactions (biomass production) [71].

Another functional barrier to LGT is protein dosage [8]. A genome sequencing project that includes the preparation of Fosmid libraries may be considered as a large-scale experiment in LGT into *E. coli* [8]. Genomic fragments whose cloning into *E. coli* is lethal are suspects for encoding proteins whose acquisition in *E. coli* is extremely disadvantageous [8]. An extensive dataset of lethal fragments collected during genome sequencing projects of 79 diverse species showed that these fragments typically encode for single copy genes. The integration of an additional gene copy into the *E. coli* genome resulted in an elevated protein production that was lethal to the cell [8].

Conclusions

Experimental work shows that gene acquisition by LGT among prokaryotes is frequent and that the percolation of acquired DNA among populations and across generations is rapid. Phylogenomic analyses reveal that LGT has a substantial impact on long-term genome evolution, supplying a mechanism for natural variation that is specific for the prokaryotic domains and allows their adaptation in dynamic environments. Prokaryote genome evolution comprises thus vertical (tree-like) and lateral (network-like) components. At the same time, different types of barriers to LGT on the genomic, species, and habitat levels are becoming increasingly apparent.

The recent discovery of nanotubes [26••] and the highlight of GTA transfer efficiency [30••] along with their high frequency in natural habitats [25] show that there is still much to discover about LGT. Some of the open questions are: what is the role of nanotubes in LGT? What are the pathways for DNA transfer via the tubes? How are GTAs produced in donor cells? How is the GTA cargo-DNA received in recipient cells? Understanding the mechanisms for gene acquisition in prokaryotes will enrich our understanding of prokaryote genetics in the wild. Moreover, many lab protocols are inspired from what we see in nature. Examples for utilizing LGT in basic research include cell transformation and genome sequencing. More research in this field might contribute to novel developments in synthetic biology. Understanding the barriers to LGT and the way they are bypassed in nature, will improve our ability to manipulate microbial cells in the laboratory for research and industrial purposes.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research and a European Research Council grant NETWORKORIGINS to W. Martin. We are thankful to W. Martin, M. Lercher, K. Stucken, and G. Landan for their help in refining the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Dickerson RE: **Evolution and gene transfer in purple photosynthetic bacteria**. *Nature* 1980, **283**:210-212.
2. Doolittle WF: **Phylogenetic classification and the universal tree**. *Science* 1999, **284**:2124-2129.
3. Martin W: **Mosaic bacterial chromosomes: a challenge en route to a tree of genomes**. *Bioessays* 1999, **21**:99-104.
4. Ochman H, Lawrence JG, Groisman EA: **Lateral gene transfer and the nature of bacterial innovation**. *Nature* 2000, **405**:299-304.
5. Mirkin BG, Fenner TI, Galperin MY, Koonin EV: **Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes**. *BMC Evol Biol* 2003, **3**:2.

6. Kunin V, Goldovsky L, Darzentas N, Ouzounis CA: **The net of life: reconstructing the microbial phylogenetic network**. *Genome Res* 2005, **15**:954-959.

7. Dagan T, Martin W: **Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution**. *Proc Natl Acad Sci USA* 2007, **104**:870-875.

8. Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM: **Genome-wide experimental determination of barriers to horizontal gene transfer**. *Science* 2007, **318**:1449-1452.

9. Weigel LM, Clewell DB, Gill SR, Clark NC, McDougal LK, Flannagan SE, Kolonay JF, Shetty I, Killgore GE, Tenover FC: **Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus***. *Science* 2003, **302**:1569-1571.

10. Fondi M, Fani R: **The horizontal flow of the plasmid resistome: clues from inter-generic similarity networks**. *Environ Microbiol* 2010, **12**:3228-3242.

11. Chen I, Christie PJ, Dubnau D: **The ins and outs of DNA transfer in bacteria**. *Science* 2005, **310**:1456-1460.

12. Thomas CM, Nielsen KM: **Mechanisms of, and barriers to, horizontal gene transfer between bacteria**. *Nat Rev Microbiol* 2005, **3**:711-721.

13. Norman A, Hansen LH, Sørensen SJ: **Conjugative plasmids: vessels of the communal gene pool**. *Philos Trans R Soc B* 2009, **364**:2275-2289.

14. Wozniak RAF, Waldor MK: **Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow**. *Nat Rev Microbiol* 2010, **8**:552-563.

15. Solioz M, Yen HC, Marris B: **Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata***. *J Bacteriol* 1975, **123**:651-657.

16. Lang AS, Beatty JT: **Importance of widespread gene transfer agent genes in alpha-proteobacteria**. *Trends Microbiol* 2007, **15**:54-62.

17. Smith HO, Tomb JF, Dougherty BA, Fleischmann RD, Venter JC: **Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome**. *Science* 1995, **269**:538-540.

18. Wang Y, Goodman SD, Redfield RJ, Chen C: **Natural transformation and DNA uptake signal sequences in *Actinobacillus actinomycetemcomitans***. *J Bacteriol* 2002, **184**:3442-3449.

19. Snyder LA, McGowan S, Rogers M, Duro E, O'Farrell E, Saunders NJ: **The repertoire of minimal mobile elements in the *Neisseria* species and evidence that these are involved in horizontal gene transfer in other bacteria**. *Mol Biol Evol* 2007, **24**:2802-2815.

20. Weinert LA, Welch JJ, Jiggins FM: **Conjugation genes are common throughout the genus *Rickettsia* and are transmitted horizontally**. *Proc Biol Sci* 2009, **276**:3619-3627.

21. Bahl MI, Hansen LH, Sørensen SJ: **Persistence mechanisms of conjugative plasmids**. *Methods Mol Biol* 2009, **532**:73-102.

22. Genthner FJ, Wall JD: **Isolation of a recombination-deficient mutant of *Rhodopseudomonas capsulata***. *J Bacteriol* 1984, **160**:971-975.

23. Marrs B: **Genetic recombination in *Rhodopseudomonas capsulata***. *Proc Natl Acad Sci USA* 1974, **71**:971-973.

24. Berglund EC, Frank AC, Calteau A, Vinnere Pettersson O, Granberg F, Eriksson AS, Näslund K, Holmberg M, Lindroos H, Andersson SG: **Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii***. *PLoS Genet* 2009, **5**:e1000546.

25. Zhao Y, Wang K, Budinoff C, Buchan A, Lang A, Jiao N, Chen F: **Gene transfer agent (GTA) genes reveal diverse and dynamic *Roseobacter* and *Rhodobacter* populations in the Chesapeake Bay**. *ISME J* 2009, **3**:364-373.

26. •• Dubey GP, Ben-Yehuda S: **Intercellular nanotubes mediate bacterial communication**. *Cell* 2011, **144**:590-600.
Using electron microscopy the authors revealed the existence of tubular extensions bridging between adjacent cells, serving as a route for exchange of cytoplasmic molecules. This study revealed a previously unrecognized type of bacterial communication.

27. Babić A, Lindner AB, Vulić M, Stewart EJ, Radman M: **Direct visualization of horizontal gene transfer**. *Science* 2008, **319**:1533-1536.

28. •• Babić A, Berkmen MB, Lee CA, Grossman AD: **Efficient gene transfer in bacterial cell chains**. *MBio* 2011, **2**:e00027–11.
Using fluorescence microscopy the authors visualized in real time sequential conjugation events of an integrative and conjugative element (ICE) that encodes for a GFP. This study shows that ICEs can propagate rapidly among bacteria that grow in chains.

29. Kenzaka T, Tani K, Nasu M: **High-frequency phage-mediated gene transfer in freshwater environments determined at single-cell level**. *ISME J* 2010, **4**:648-659.

30. •• McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH: **High frequency of horizontal gene transfer in the oceans**. *Science* 2010, **330**:50.
The authors quantified the frequency of GTA-mediated LGT by cloning a streptomycin kinase gene into donor bacteria encoding for a GTA cassette and measuring the spread of Kanamycin resistance in the population. This study highlights GTAs as the most efficient LGT mechanism currently known in marine environment.

31. • Halary S, Leigh JW, Cheaib B, Lopez P, Bapteste E: **Network analyses structure genetic diversity in independent genetic worlds**. *Proc Natl Acad Sci USA* 2010, **107**:127-132.
Using a network of shared gene content among cellular genomes, plasmids and phages, the authors chart the highways and byways for DNA flow in nature. This study highlights the central role of plasmids as mediators of LGT.

32. Kloesges T, Popa O, Martin W, Dagan T: **Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths**. *Mol Biol Evol* 2011, **28**:1057-1074.

33. Lawrence JG, Ochman H: **Molecular archaeology of the Escherichia coli genome**. *Proc Natl Acad Sci USA* 1998, **95**:9413-9417.

34. Garcia-Vallve S, Romeu A, Palau J: **Horizontal gene transfer in bacterial and archaeal complete genomes**. *Genome Res* 2000, **10**:1719-1725.

35. Nakamura Y, Itoh T, Matsuda H, Gojobori T: **Biased biological functions of horizontally transferred genes in prokaryotic genomes**. *Nat Genet* 2004, **36**:760-766.

36. Beiko RG, Harlow TJ, Ragan MA: **Highways of gene sharing in prokaryotes**. *Proc Natl Acad Sci USA* 2005, **102**:14332-14337.

37. Zhaxybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT: **Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events**. *Genome Res* 2006, **16**:1099-1108.

38. Puigbò P, Wolf YI, Koonin EV: **The tree and net components of prokaryote evolution**. *Genome Biol Evol* 2010, **2**:745-756.

39. Chan Cheong Xin, Beiko Robert G, Ragan Mark A: **Lateral transfer of genes and gene fragments in *Staphylococcus* extends beyond mobile elements**. *J Bacteriol* 2011 doi: 10.1128/JB.01524-10.

40. Bapteste E, O'Malley MA, Beiko RG, Ereshefsky M, Gogarten JP, Franklin-Hall L, Lapointe F-J, Dupré J, Dagan T, Boucher Y, Martin W: **Prokaryotic evolution and the tree of life are two different things**. *Biol Direct* 2009, **4**:34.

41. • Ragan MA: **Trees and networks before and after Darwin**. *Biol Direct* 2009, **4**:43.
In this review the author presents an extensive historical survey of metaphors and models to describe natural systems and genealogical relations over the past 500 years.

42. Swithers KS, Gogarten JP, Fournier GP: **Trees in the web of life**. *J Biol* 2009, **8**:54.

43. Huson DH, Scornavacca CA: **A survey of combinatorial methods for phylogenetic networks**. *Genome Biol Evol* 2011, **3**:23-35.

44. Strogatz SH: **Exploring complex networks**. *Nature* 2001, **410**:268-276.

45. Newman MEJ: *Networks: An Introduction*. Oxford University Press; 2010.

46. Dagan T, Artzy-Randrup Y, Martin W: **Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution**. *Proc Natl Acad Sci USA* 2008, **105**:10039-10044.

47. Lima-Mendez G, Van Helden J, Toussaint A, Leplae R: **Reticulate representation of evolutionary and functional relationships between phage genomes**. *Mol Biol Evol* 2008, **25**:762-777.

48. Fondi M, Bacci G, Brilli M, Papaleo MC, Mengoni A, Vaneechoutte M, Dijkshoorn L, Fani R: **Exploring the evolutionary dynamics of plasmids: the *Acinetobacter* pan-plasmidome**. *BMC Evol Biol* 2010, **10**:59.

49. Popa O, Hazkani-Covo E, Landan G, Martin W, Dagan T: **Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes**. *Genome Res* 2011, **21**:599-609.

50. Naumov DG, Livshits VA: **Molecular structure of the locus for sucrose utilization by *Lactobacillus plantarum*: comparison with *Pediococcus pentosaceus***. *Mol Biol* 2001, **35**:19-27.

51. Bon E, Delaherche A, Bilhère E, De Daruvar A, Lonvaud-Funel A, Le Marrec C: ***Oenococcus oeni* genome plasticity is associated with fitness**. *Environ Microbiol* 2009, **75**:2079-2090.

52. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC: **The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata**. *Nucleic Acids Res* 2010, **38D**:D346-D354.

53. Landan G, Graur D: **Heads or tails: a simple reliability check for multiple sequence alignments**. *Mol Biol Evol* 2007, **24**:1380-1383.

54. Hao W, Golding GB: **The fate of laterally transferred genes: life in the fast lane to adaptation or death**. *Genome Res* 2006, **16**:636-643.

55. van Passel MW, Marri PR, Ochman H: **The emergence and fate of horizontally acquired genes in *Escherichia coli***. *PLoS Comput Biol* 2008, **4**:e1000059;
Kuo C-H, Ochman H: **The extinction dynamics of bacterial pseudogenes**. *PLoS Genet* 2010, **6**:e1001050.

56. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC: **Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella***. *Science* 2006, **313**:236-238.

57. Doyle M, Fookes M, Ivens A, Mangan MW, Wain J, Dorman CJ: **An H-NS-like stealth protein aids horizontal DNA transmission in bacteria**. *Science* 2007, **315**:251-252.

58. Dillon SC, Cameron AD, Hokamp K, Lucchini S, Hinton JC, Dorman CJ: **Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella typhimurium* identifies a plasmid-encoded transcription silencing mechanism**. *Mol Microbiol* 2010, **76**:1250-1265.

59. Puigbò P, Wolf YI, Koonin EV: **Search for a 'Tree of Life' in the thicket of the phylogenetic forest**. *J Biol* 2009, **8**:59.

60. Andam CP, Williams D, Gogarten JP: **Biased gene transfer mimics patterns created through shared ancestry**. *Proc Natl Acad Sci USA* 2010, **107**:10679-10684.

61. Majewski J, Cohan FM: **DNA sequence similarity requirements for interspecific recombination in Bacillus**. *Genetics* 1999, **153**:1525-1533.

62. Lovett ST, Hurley RL, Sutera VA Jr, Aubuchon RH, Lebedeva MA: **Crossing over between regions of limited homology in *Escherichia coli*: RecA-dependent and RecA independent pathways**. *Genetics* 2002, **160**:851-859.

63. • Marraffini LA, Sontheimer EJ: **CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea**. *Nat Rev Genet* 2010, **11**:181-190.
CRISPRs (clustered regularly interspaced short palindromic repeats) are a more recently discovered inherited prokaryotic immune system that protects cells from bacteriophages and conjugative plasmids. In this review the authors elucidate the mechanisms of CRISPR interference and its role in microbial physiology and evolution.

64. Marraffini LA, Sontheimer EJ: **CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA**. *Science* 2008, **322**:1843-1845.

65. Sullivan MB, Waterbury JB, Chisholm SW: **Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus***. *Nature* 2003, **424**:1047-1051.

66. Gómez P, Buckling A: **Bacteria-phage antagonistic coevolution in soil**. *Science* 2011, **332**:106-109.

67. Majewski J: **Sexual isolation in bacteria**. *FEMS Microbiol Lett* 2001, **199**:161-169.

68. • Hooper SD, Mavromatis K, Kyrpides NC: **Microbial co-habitation and lateral gene transfer: what transposases can tell us**. *Genome Biol* 2009, **10**:R45.
Using a network of shared transposases the authors reveal that the majority of LGTs occur between closely related species residing in the same habitat. This study demonstrates the utility of genomic data for research in microbial ecology.

69. Moran NA, McLaughlin HJ, Sorek R: **The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria**. *Science* 2009, **323**:379-382.

70. Burke GR, Moran NA: **Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids**. *Genome Biol Evol* 2011, **3**:195-208.

71. Pál C, Papp B, Lercher MJ: **Adaptive evolution of bacterial metabolic networks by horizontal gene transfer**. *Nat Genet* 2005, **37**:1372-1375.

72. Zhaxybayeva O, Doolittle WF: **Lateral gene transfer**. *Curr Biol* 2011, **21**:R242-R246.

73. Davids W, Zhang Z: **The impact of horizontal gene transfer in shaping operons and protein interaction networks – direct evidence of preferential attachment**. *BMC Evol Biol* 2008, **8**:23.

74. Dorman CJ: **Regulatory integration of horizontally-transferred genes in bacteria**. *Front Biosci* 2009, **14**:4103-4112.

75. Medrano-Soto A, Moreno-Hagelsieb G, Vinuesa P, Christen JA, Collado-Vides J: **Successful lateral transfer requires codon usage compatibility between foreign genes and recipient genomes**. *Mol Biol Evol* 2004, **21**:1884-1894.

76. Tuller T, Girshovich Y, Sella Y, Kreimer A, Freilich S, Kupiec M, Gophna U, Ruppin E: **Association between translation efficiency and horizontal gene transfer within microbial communities**. *Nucleic Acids Res* 2011, **39**:4743-4755.

77. Kudla G, Murray AW, Tollervey D, Plotkin JB: **Coding-sequence determinants of gene expression in *Escherichia coli***. *Science* 2009, **324**:255-258.

78. • Amorós-Moya D, BedhommeS, Hermannand M, Bravo IG: **Evolution in regulatory regions rapidly compensates the cost of nonoptimal codon usage**. *Mol Biol Evol* 2010, **27**:2141-2151.
Using experimental evolution approach the authors demonstrate that non-optimal codon usage is a minor obstacle for the fixation of laterally acquired genes. The regulation of gene expression level in this experiment was independent of codon adaptation.

79. Jain R, Rivera MC, Lake JA: **Horizontal gene transfer among genomes: the complexity hypothesis**. *Proc Natl Acad Sci USA* 1999, **96**:3801-3806.

80. Coscollá M, Comas I, González-Candelas F: **Quantifying nonvertical inheritance in the evolution of *Legionella pneumophila***. *Mol Biol Evol* 2011, **28**:985-1001.

81. Cohen O, Gophna U, Pupko T: **The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer**. *Mol Biol Evol* 2011, **28**:1481-1489.