

The Conflict Between Horizontal Gene Transfer and the Safeguard of Identity: Origin of Meiotic Sexuality

Nicolas Glansdorff · Ying Xu · Bernard Labedan

Received: 12 June 2009 / Accepted: 8 September 2009 / Published online: 26 September 2009
© Springer Science+Business Media, LLC 2009

Abstract Contrary to a widespread opinion, horizontal gene transfer (HGT) between distantly related microorganisms (such as Bacteria and Archaea) has not been demonstrated to occur on a large scale. Except for transfer of mobile elements between closely related organisms, most alleged HGT events reflect phylogenetic discrepancies that can be explained by a variety of artefacts or by the differential loss of paralogous gene copies either originally present in the Last Universal Common Ancestor (LUCA) to the three Domains (a sophisticated, genetically redundant and promiscuous community of protoeukaryotes), or created by duplications having occurred at later times. Besides, (i) there is no experimental evidence for the facile acquisition of foreign DNA between distant taxa and (ii) important biological constraints operate on the phenotypic success of genetic exchange at several levels, including protein–protein interactions involved in metabolic channelling; stable integration and expression of foreign DNA is, therefore, expected to require strong selection. Explaining phylogenetic discrepancies by artefacts or loss of paralogs does not eliminate difficulties in retracing species genealogy but maintains the picture of a universal

tree of life, HGT between distant organisms being reduced to a trickle. We illustrate our thesis by the phylogenetic analysis of carbamoyltransferases, a family of paralogous proteins. Among higher eukaryotes HGT appears of limited scope except in asexual organisms. We suggest that meiotic sexuality (a hallmark of eukaryotes) emerged in the genetically redundant and protoeukaryotic LUCA as a molecular identity check providing a defence mechanism against the deleterious effects of HGT.

Keywords LUCA · HGT · Paralogy · Phylogenetic inconsistencies · Tree of life · Sex · Meiosis

Introduction

It is more and more fashionable to underline the importance of horizontal gene transfer (HGT) in comparisons of genomes and related evolutionary studies. In nearly most cases, observing a discrepancy in the phylogenetic tree of a particular protein family triggers the reflex to conclude that the anomalous grouping of two species that are taxonomically unrelated is due to HGT.

However, it is not always appreciated how stringently phenotypic constraints may operate on genetic exchange between microorganisms. Efficiency of microbial metabolism, of growth and of cell division requires a high degree of coordination within a complex but highly ordered network of functions. Thanks to partial redundancy and to homeostatic regulatory mechanisms, the latter is relatively robust. Nevertheless, even minor disturbances, not always easy to identify, can prevent this network to function optimally. For example, among recombinants obtained between non-isogenic strains of the same bacterial species,

As this manuscript was under revision, Nicolas Glansdorff died accidentally during a solitary hike in a remote, icy area of Northern Scandinavia. At 71, Nicolas was still impelling new ideas and maintaining the excellence of his scientific contribution. We will miss deeply his great humanity, his humour, and his immense culture.

N. Glansdorff · Y. Xu
JM Wiame Research Institute for Microbiology and Vrije
Universiteit Brussel, 1 ave E. Gryzoon, 1070 Brussels, Belgium

B. Labedan (✉)
Institut de Génétique et Microbiologie, Université Paris Sud,
CNRS UMR 8621, Bâtiment 400, 91405 Orsay Cedex, France
e-mail: bernard.labedan@igmors.u-psud.fr

the growth rate may vary substantially; this is a serious matter because a mere 10% difference in doubling time between two microbial strains dividing in about 60 min will result in a two-fold difference in cell number after only 10 generations. It is thus clear that however robust the cell functional network may be, it is constantly under selection against any departure from the operational balance fitting a given environment. By the same token, it appears unlikely that any substantial amount of DNA could pass from one species to a different one without exerting an effect on survival of the recipient and thus becoming subject to selection. For DNA of foreign origin to efficiently contribute to the survival of a particular organism, it would have to generate an advantageous phenotype. Moreover, only strong selection could bypass the natural restrictions imposed by various mechanisms on the integrity, expression and integration of foreign DNA.

These considerations constitute a serious but often neglected caveat to consider when discussing the role that HGT may have played in the course of evolution. In the following sections, we will (i) discuss the type of evidence usually produced to support this proposal, (ii) consider alternative explanations, (iii) emphasize the necessity of efficient safeguards against genetic promiscuity, (iv) examine evidence to this effect and (v) discuss how Archaea and Bacteria on the one hand and Eukaryotes on the other, appear to have dealt with this necessity. We further suggest that meiotic sexuality—a eukaryotic hallmark—has emerged as a defence mechanism against HGT in the cellular context of a protoeukaryotic Last Universal Common Ancestor (LUCA) after the RNA to DNA genomic transition (see Glansdorff et al. 2008 for an exhaustive definition of LUCA).

Challenging the Evidences Alleged to Support the Occurrence of HGT Among Archaea and Bacteria

There are two main types of discrepancies that HGT is assumed to explain (Ochman et al. 2000): a disparity in DNA sequence (often a difference in G + C content associated with remnants of mobile elements) and a phylogenetic incongruity. In general, the first case corresponds to recent events, whereas the second one is viewed as the result of ancient transfers where differences in nucleotide composition were progressively erased by “amelioration” (Lawrence and Ochman 1997).

Detecting Recent Events

The most direct evidence for a case of HGT is the association of the gene or genes incriminated with mobile elements able to integrate into the host chromosome, such as a plasmid, (pro)phage, transposon, integron or pathogenicity

island (for a recent review, see Zaneveld et al. 2008 and references inside). Actual transfer, either under direct selection or reasonably explained by selection occurs readily between strains of the same species such as *Escherichia coli* and, less frequently, between broadly related organisms such as Proteobacteria or Firmicutes (Dobrindt et al. 2003; Leplae et al. 2006). Such recent HGT events are a well-understood cause of natural variation and have certainly contributed to shape the genomes of microorganisms though they do not appear to bring the existence of a universal tree of life in doubt (see further).

Detecting Ancient Events: Actual HGT or Incongruences Due to Methodological Errors?

In many instances, gene trees were found to conflict with the species tree originally established by comparing SSU rRNA gene sequences; this constitutes the main type of evidence alleged to support HGT when more direct indications are not available (Boucher et al. 2003; Doolittle 1999a, b; Koonin et al. 2001). However, as it has been stated in a seminal paper (Esser et al. 2004), asserting HGT explanation for unexpected branching order implies “the assumptions that the interpretation of individual gene trees is straightforward and that the reconstruction of gene trees is, at the extreme, infallible”.

Different computational methods have been used to infer from sequence comparisons whether a particular gene appears to obey vertical transmission throughout any one of the three Domains or, on the contrary, to have been transferred horizontally (Beiko et al. 2005 and references therein; Beiko and Ragan 2009). These phylogenetic methods are liable to various types of artefacts that rapidly become problematic when more and more divergent sequences are being compared (Beiko and Ragan 2009). Such is the case for a vast number of HGT alleged to have occurred between representatives of different Domains.

Indeed, several methodological errors (Than et al. 2007; Beiko and Ragan 2009) can lead to incongruent trees without the need to invoke HGT. (i) With increasing sequence divergence, it becomes more and more difficult to correctly identify genuine orthologs (for a review, see Kuzniar et al. 2008). Accordingly, families of orthologs are often contaminated with hidden paralogs, leading to gene trees that do not fit the organismal tree (Kurland et al. 2003). (ii) Lineage sorting (random genetic drift) is another phenomenon that can lead to mistaken inference of HGT particularly when analyzing closely related organisms, where coalescent effects may not be ignored when reconciling gene trees (Maddison, 1997; Than et al., 2007; Beiko and Ragan, 2009). (iii) Incorrect multiple sequence alignments can lead to false phylogenetic inferences especially when non-homologous residues are mistakenly aligned.

Moreover, it has been demonstrated that the number of aligned positions may be critical. Many “HGT” may arise as a consequence of stochastic errors increasing with decreasing alignment length (Swofford et al. 2001), and with increasing ratios of terminal to internal branch lengths (Philippe et al. 2005).

Barriers to DNA Integration

Besides neglecting the statistical errors summarized above, many supporters of the prevalence of HGT are also underestimating crucial biological parameters. For horizontally transferred DNA to become successful, a number of barriers must be overcome. For instance, “transformation proficiency does not necessarily translate into correspondingly high rates of interspecific gene transfer: although species of *Haemophilus* are naturally competent, the *H. influenzae* Rd genome bears little foreign DNA beyond two large prophages” (Ochman et al. 2000). Phage transduction is even more restricted due to very high selectivity of bacteriophage receptors and the efficiency of host restriction endonucleases as a barrier against foreign invading DNA (for reviews and references inside, see Arber 2000; Matic et al. 1996; Murray 2002). Likewise, mobilizable plasmids require a complex cell apparatus in both donor and recipient to be transferred by conjugation (for a review and references inside, see Sørensen et al. 2005).

Besides these constraints on uptake there are other efficient barriers on DNA integration and expression (see, for instance Navarre et al. 2006; Dorman 2007). Moreover, successful transfer requires compatibilities of transcription machineries (especially in the case of Bacteria and Archaea), efficient translation (Taoka et al. 2004) and, as emphasized *supra*, phenotypic success, i.e. the formation of properly regulated and efficient protein complexes, since low dosage impedes complex formation (Deutschbauer et al. 2005), and imbalance of complex subunits can lead to harmful protein aggregation (Papp et al. 2003). To these difficulties Kurland (2005) added the problem of patchiness: genes horizontally transferred in one patch of a taxon would be missing in other patches; only strong selection would ensure survival and predominance of the horizontally transferred gene in the descendants of the taxon.

Certain authors emphasized that differences appear to exist between lineages concerning the propensity to produce the phylogenetic inconsistencies they interpret as HGT; Boucher and Baptiste (2009) even proposed recently to distinguish between “open” and “closed lineages”. The fact that these lineages appear interspersed along the branches of any prokaryotic tree raises the question of how much statistical or sampling artefacts may be responsible for such a distinction. For example, taxa bearing different

species names such as many *Enterobacteriaceae* related to *E. coli*, might appear as an open lineage because of the amount of knowledge accumulated on the mobile elements of these closely related organisms, whereas other ones, less intensely studied, would appear genetically more isolated. At any rate, before elaborating on the evolutionary consequences of alleged differences in susceptibility to HGT and on the extent of the phenomenon—in particular denying the existence of a universal tree of life (Doolittle and Baptiste, 2007)—it should be investigated whether alternative explanations are not being neglected. Indeed, “in cases where equally if not more plausible mechanisms exist, extraordinary events such as horizontal gene transfer do not provide the best explanation” (Salzberg et al. 2001), the more so that direct experimental evidence for widespread interdomain HGT has not been forthcoming.

Lack of Experimental Evidence for Widespread, Interdomain HGT

Evidence is not proof, yet even a cursory survey of the literature (in particular the first sentence of many papers) shows that the mere existence of phylogenetic inconsistencies is accepted by many as a demonstration of HGT occurrence and of the prominent role it would have played in evolution. Rarely has the non-existence of actual (i.e. experimental) evidence appeared to matter so little in the interpretation of a biological phenomenon. Yet, if HGT could occur so readily between members of distant phyla, such as Archaea and Bacteria, or either of these and Eukarya, why is not the literature replete with actual demonstrations? The answer is probably more circumstantial than scientific: such experiments might provide negative, unpublishable results; it is more rewarding to continue publishing sequence comparisons and systematically conclude from phylogenetic inconsistencies in favour of HGT, the more so that it would fall in a context already well-disposed to that type of interpretation. We should hope to witness a reversal of this tendency.

A Major Alternative to Present-Day HGT-Based Interpretations

The Last Universal Common Ancestor (LUCA) Was a Complex Organism

The existence of a LUCA was one of Darwin’s original ideas (Padian 2008). In recent years, the profile of this hypothetical precursor has undergone a marked transformation: it was conceived originally as the simple cellular forerunner of progressively more complex entities, but from several contributions stressing different or complementary aspects (Castresana 2001; Delaye et al. 2004;

Forterre 1995; Glansdorff 2000, Glansdorff et al. 2008, 2009; Kandler 1994; Kurland et al. 2006; Labedan et al. 1999, 2004; Ouzounis et al. 2006; Poole et al. 1999; Wang et al. 2007; Woese 1998, 2002) the LUCA would rather appear to have been an already complex organism, a member of a genetically promiscuous community of protoeukaryotic cells, rich in paralogous gene copies, genetically redundant for every essential gene and ready to undergo reductive evolution toward the simpler archaeal and bacterial common ancestors while continuing its evolution toward the eukaryotic state. Genetic redundancy probably was a safeguard against still imperfect mechanisms for the partition of genetic material between daughter cells (Glansdorff et al. 2008) and provided material required for the development of basic functional innovations, largely by the “patchwork” mechanism of metabolic evolution (Jensen 1976). This protoeukaryotic LUCA announced many of the traits that we consider today hallmarks of the eukaryotes. In a recent publication (Glansdorff et al. 2008), we have gathered and discussed the multifaceted evidence that favours such a conclusion rather than the alternative of eukaryogenesis by the merger of a bacterium and an archaeon.

Before LUCA Crystallization, a Community of Progenotes Was Actively Exchanging Their Genetic Material

According to Woese (2002) and Vetsigian et al. (2006), the very emergence of the communal progenote, the development of a universal genetic code and of LUCA’s genome rested on continuous and unrestricted HGT between cells as yet devoid of restriction enzymes and cell walls; this contrasts with the more vertical evolution, which members of either of the three Domains engaged after the “crystallization” (Woese 1998, 2002) of the LUCA into Archaea, Bacteria and the Eukaryotic lines. From this point onwards, we must indeed assume that HGT was drastically reduced in order to understand the perpetuation and further differentiation of well-distinct archetypes; the question is, how much? In the present context this means: must all bonafide phylogenetic inconsistencies be attributed to HGT, after elimination of all possible artefacts? The answer is an emphatic no, as discussed below.

The LUCA as a Putative Source of Ancient Paralogs Generating Phylogenetic Inconsistencies

A genetically redundant and promiscuous LUCA population would have been a reservoir of multiple gene copies expected to segregate in an unpredictable way in any of the three emerging Domains. Such a situation predicts the emergence of ancient phylogenetic inconsistencies by a simple mechanism: differential loss of paralogs (Labadan

et al. 1999, 2004; Glansdorff et al. 2008). Moreover, considering that gene duplications occurred again and again during billions of years at frequencies that do not appear to differ in prokaryotes and eukaryotes (Lynch and Conery 2003), it would be surprising if this type of event was not the basis of a large number of phylogenetic inconsistencies; see for example, the recurrent paralogy observed in the evolution of archaeal chaperonins (Archibald et al. 2000) and family B DNA polymerases (Edgell et al. 1998). We insist that our model allows overcoming the logical contradiction stated by Dagan and Martin (2007) that “although differential gene loss can account for patchy distributions in individual instances, it cannot be invoked to account for all such patterns, because the inferred size of ancestral genomes would become unrealistically large”. Their statement is based on the so-called “genome of Eden” concept (Doolittle et al. 2003), which postulates that the LUCA was simple because primitive (Doolittle 1999b), a view that also underestimates the incidence of gene duplication and the existence of biological barriers to HGT. Therefore, contrary to the assumption of Dagan and Martin (2007), we do not have to envision the need of “incremental allowance of LGT to solve the genome-of-Eden problem”.

It can be difficult to distinguish the loss of a paralog from an acquisition by HGT without evidence for a specific mechanism, such as the presence of sequences signalling the possible intervention of a mobile element. Furthermore, when HGT is assumed to have occurred a very long time ago, close to the branching point between the phyla investigated (see for example the comparative analysis of some of the genes involved in isoprenoid biosynthesis by Boucher et al. 2004), the interpretation of the discrepancy comes very close indeed to the hypothesis of differential loss of ancestral paralogs. Kunin et al. (2005) have attempted to evaluate the roles, respectively, played by gene loss and by HGT in genomic evolution; they concluded that HGT had been over-evaluated but their analysis did not address the specific problem of individual paralogies since they focused on gene families; therefore, how many events they classified as HGT could still be due to loss of paralogs remains to be determined. It is however noticeable how their approach reduces the frequency of HGT alleged between distant taxa to “thin vines” on a robust tree of life. Makarova et al. (2005), while studying paralogous genes in eukaryotes, made a distinction between true paralogs and “pseudo-paralogs” the latter being detected by their similarity to a prokaryotic homolog; however, this systematically eliminates from consideration possible paralogs predating the emergence of the three Domains from the LUCA. Last but not least, recent comparisons between prokaryotes and eukaryotes (in particular higher eukaryotes, where putative HGT instances are not

frequent, Soria-Carrasco and Castresana 2008; Galtier and Daubin 2008), bring an important contribution to the debate by showing that paralogy problems and phylogenetic artefacts strongly affect phylogenies across the three Domains in a similar way; this minimizes the alleged impact of HGT in prokaryotes and strengthens the concept of a universal tree of life.

Relationships Between HGT, Core and Pan-Genome, and Ecological Forces

It was reported that HGT (i.e. phylogenetic discrepancies interpreted in this way) appears more frequent between phylogenetically and/or ecologically related taxa (Gogarten et al. 2002; Jain et al. 2003; Comas et al. 2006). This seems at first sight reasonable and it is probable that most of the real cases of HGT belong to this category, but it is subject to a caveat: differential loss of paralogs will also create phylogenetic inconsistencies among the descendants of an ancestral cell line adapting to related environments, so that, again, the distinction can become difficult. In fact, phylogenetic inconsistencies affecting taxa that are ecologically unrelated could be best accounted for by differential loss of paralogs, whereas inconsistencies found among taxa that are both ecologically and phylogenetically related could in principle be explained by either HGT or loss of paralogs. Furthermore, the suggestion that the transferability of genes seems to depend on their functions (Jain et al. 2003; Nakamura et al. 2004) is subject to a similar caveat; just as putative horizontally transferred genes, novel genes created by duplication would be constrained by a pre-existing metabolic organization. We would not expect finding diverging paralogs affecting some of the core cellular functions (such as genes involved in information processing) as readily as more peripheral functions. It should be stressed here that asserting the occurrence of a duplication creating paralogs by looking for a genetic redundancy in a putative ancestor is of little diagnostic value since all organisms we can look at are modern and may have lost genes as well.

The possible role of ecological forces in different biogeographical conditions has been recently examined. A very recent and seminal paper (Reno et al. 2009) has studied gain and loss of genetic material in genomes of seven *Sulfolobus islandicus* strains living in three different geographical locations. The biogeographical structures of each corresponding pan-genome show a “spatial partition of the variable gene pool between distinct geothermal regions with local adaptation and dramatically slow gene flows”. The role of HGT in defining the distribution of the variable genes is strictly limited to a “recent strain-specific integration of mobile elements”. Interestingly, these results contrast to previous suggestion that environmental

differences rather than geographical isolation drive differences in gene content (Grogan et al. 2008). Such data underlines the biological barriers to free diffusion of mobile elements from divergent species among distantly located taxa, contradicting the current model proposed by HGT advocates.

Evidence for Loss of Ancestral Paralogs

In the above, we have shown that a large proportion of phylogenetic inconsistencies, when they are not statistical artefacts, could in principle be attributed to other causes than HGT, differential loss of paralogs, ancient or recent, being the most obvious one. Accordingly, we are now providing a case study where structural characteristics of enzymes displaying a polyphyletic tree provide specific insight. This analysis also emphasizes the importance of protein–protein interactions for the integration of a foreign gene in the cellular network. It could become a model to evaluate loss of paralogs versus HGT whenever comparable evidence could be obtained (Barba et al. in preparation).

Our hypothesis of differential loss of paralogous copies was presented as the conclusion of a comparative analysis of the genes involved in the metabolism of carbamoyl-phosphate (Labedan et al. 1999; 2004), a key component in the biosynthesis of arginine and pyrimidines. On the basis of several hundreds of supplementary sequences, we can confirm these conclusions that are summarized below and will be published in detail elsewhere (Barba et al. in preparation).

Aspartate carbamoyltransferases (ATCases) occur in different structural classes according to the mode of association of the catalytic PyrB subunit with other polypeptides (either dihydroorotase (PyrC) in class A or the PyrI regulatory subunit in class B) or its occurrence as a free trimeric protein (class C). The PyrB phylogenetic tree is not congruent with the SSU rRNA tree; an almost complete correlation was nevertheless observed between this polyphyletic pattern and the different structural classes of ATCase; this correlation is important because it confers biological significance to the *pyrB* tree, providing a kind of internal control against the artefacts of tree construction that plague many other analyses. Most importantly, the tree becomes coherent when we consider (i) that the LUCA was genetically redundant and contained at least two copies of the ATCase catalytic gene, (ii) gene losses occurred independently in different lineages. This pattern of differential extinctions of paralogous copies explains the data in a more straightforward way than does HGT. It can also explain similar evolution of the paralogous ornithine carbamoyltransferase (OTCase) into different families that also descend from an ancestral gene that was already duplicated in

LUCA. Therefore, the primordial gene duplication of an ancestral carbamoyltransferase occurred in ancestors of LUCA (Labedan et al. 1999).

A HGT-based interpretation of the ATCase polyphyletic pattern would not only assume the occurrence of many ad hoc, independent, events (many of them counterintuitive ecologically) but also the systematic replacement of a resident gene by an often distantly related homolog whereas, even in the case of selection pressure arising from accidental loss, retrieval by a gene from a cell of the same species or from a closely related organism would be by far more likely. Besides, the differences between Archaea and Bacteria regarding promoter structure and transcription machinery would preclude efficient expression of genes transferred from another Domain.

Another major concern about interpreting the carbamoyltransferase tree by multiple HGT is the uniform carbamoyltransferase pattern found among Archaea. Indeed, if HGT is as frequent between Archaea and Bacteria as assumed in many discussions, why do we not find any archaeon with another ATCase than a class B one and why is the ATCase I group (which comprises the related A and C classes) confined to Bacteria? Current ideas on the propensity of HGT to swap genes around fail to explain such a pattern. Rather it would appear that the emergence of Archaea proceeded through a bottleneck (perhaps related to their appearance by thermoreduction, Xu and Glansdorff 2002; Glansdorff et al. 2008) selecting only one ATCase paralog, whereas the reductive evolution leading to Bacteria would have been less restrictive.

Still another feature uncovered with carbamoyltransferases but of possible general significance would disfavour HGT. Carbamoylphosphate synthetases and carbamoyltransferases interact physically to channel the unstable carbamoylphosphate (CP) molecule (Massant et al. 2002; Massant and Glansdorff 2005). It is very likely that the cognate protein interactions are stereospecific; therefore, the replacement by HGT of an ATCase of a particular class by another one would be discriminated against (see supra our discussion of the difficulty of horizontally transferred genes to integrate an interaction network, as also underlined by Papp et al. (2003), Deutschbauer et al. (2005) and Lercher and Pal (2008)). This type of constraint may act even on the replacement of enzymes of one and the same structural class; in that case, it may be weaker but not necessarily negligible. In fact, it was shown that replacing the yeast inactive dihydroorotase (DHOase) domain of the multifunctional CAD protein by the active DHOase domain of the mammalian CAD considerably impairs CP channelling and that all other chimeric constructions alter it significantly (Serre et al. 1998). Metabolic channelling is a phenomenon of general significance (Ovádi and Sreere

2000); HGT would easily disturb specific interactions between proteins operating in the same metabolic channel.

In keeping with the present analysis, Lercher and Pal (2008) pointed out that among enterobacteria, susceptibility to HGT appears negatively correlated with the degree of integration of the corresponding protein in the cellular interaction network, full integration requiring millions of years. Granting that at least some of the genes they identified were true cases of HGT (a reasonable assumption given the rather high degree of relatedness of the organisms investigated), this suggests that they were cotransferred with other, not identified genes under selection pressure.

Such integration constraints are reminiscent of a well-known concept about isolating mechanisms in eukaryotes: Mayr (1954) suggested that genomes are made of *coadapted gene complexes*, which resisted changes and that the selective value of a single allele depended greatly upon the overall genetic environment: According to Mayr (cited by Provine 2004) “Such a well-integrated, coadapted gene complex constitutes an evolutionary unit in spite of its intrinsic variability. Any disharmonious gene or gene combination which attempts to become incorporated in such a gene-complex will be discriminated against by selection” (Mayr 1954, p. 165).

Conclusions Regarding HGT Among Bacteria and Archaea

The study of carbamoyltransferases shows that it is possible to retrace gene evolution with a working hypothesis that is more parsimonious—in terms of the number and nature of events postulated—than multiple HGT between distant organisms. Most importantly, the differential loss of paralogs is not just an alternative explanation but is a *prediction* based on current ideas about the complexity and genetic redundancy of the LUCA. Many instances of polyphyletic patterns can be interpreted in that way or by the loss of more recently duplicated genes, especially when housekeeping, ubiquitous or at least widespread proteins are concerned, for example EF-TU (Ke et al. 2000), the so-called “promiscuous” aminoacyl tRNA synthetases (Doolittle 1999a, b; Woese et al. 2000), proteins involved in isoprenoid biosynthesis (Boucher et al. 2003), family B DNA polymerases (Edgell et al. 1998) and chaperonins (Archibald et al. 2000). Other examples of gene paralogy were already alleged to have originated in the LUCA: histidine biosynthetic genes (Alifano et al. 1996); glutamate dehydrogenase genes (Benachenhou and Baldacci 1991; Benachenhou-Lahfa et al. 1993); genes involved in bioenergetic processes (Castresana 2001) α and β ATPases (Gogarten et al. 1989); aldehyde dehydrogenases (Habenicht et al. 1994); EF-TU and EF-G (Iwabe et al. 1989).

There is therefore no reason to consider the carbamoyl-transferases as an isolated case.

It should be stressed that HGT does not have to be as rampant as it would appear from the present literature to have played an important, qualitative role in prokaryotic evolution: granting the force of selection, it is possible to understand how single genes or operons—catabolic operons, resistance traits, energetically useful genes, such as those for proteorhodopsin (Frigaard et al. 2006)—may have been transferred from one organism to a distant one. Even complex adaptations, such as the progressive emergence of thermophily may have included HGT at some critical step by the transfer of pleiotropic traits such as reverse gyrase (Forterre et al. 2000).

What the available evidence argues against, however, is the notion of a ready opportunity for the facile acquisition of foreign DNA. It is clear that less biased interpretations of polyphyletic patterns, as well as experimental evidence for distant HGT (presently lacking) are required to obtain a more balanced appreciation of the evolutionary forces at work. From the phylogenetic point of view, replacing HGT by loss of paralogs does not eliminate practical difficulties in retracing the genealogy of species; however, it maintains the picture of a tree where gene filiation remains mostly vertical.

By contrast, HGT readily explains the mosaicism of bacteriophage genomes; in that particular case, the extent of the phenomenon clearly imposes a reticulate type of classification (Lima-Mendez et al. 2008) reflecting their natural history.

Eukaryotes, HGT and the Scope of Genetic Innovation

Among higher eukaryotes, animals are not expected to be prone to HGT because of the necessity for the laterally transferred DNA to gain access to the germ line and to undergo sufficient expansion in order not to become lost. Moreover, alleged transfers from bacteria to mammals have been reconsidered and explained more parsimoniously by gene loss from an ancestor (Salzberg et al. 2001). Likewise, the acquisition of cellulase genes by termites from a bacterial source (a textbook example) has been shown to be erroneous (Davison and Blaxter 2005). Transfer of bacterial genes by a maternally transmitted bacterial parasite has been reported however (Dunning Hotopp et al. 2007) but the majority of the transferred genes have been pseudogenized (Blaxter 2007; Nikoh et al. 2008). Thus, it is unknown whether these transfers represent innovation in animal or plant evolution or are just another flavour of nuclear mitochondrial fragments like elements (Richly and Leister 2004) without evolutionary significance (Blaxter 2007).

Of a more evolutionary value are a few convincing occurrences of endosymbiotic gene transfer (EGT, reviewed in Timmis et al. 2004). For instance, nuclear genes encoding for chloroplast proteins have been transferred from endosymbiotic plastids present in alga after these alga have been ingested by the ascoglossan sea slug *Elysia crispata* (Pierce et al. 2003). Here, the distinction between gene origins via EGT versus HGT is crucial and can be argued if it is possible to determine whether the transferred genes can be traced back to a unique source and are found in most if not all related taxonomic lineages versus sporadic gene origin in particular lineages and from multiple different sources, respectively (Moustafa et al. 2009). Such a distinction has been made recently in the case of diatoms, showing that a large proportion of the algal green genes detected in diatom is due to an ancient endosymbiosis that occurred in the common ancestor of chromalveolates (Moustafa et al. 2009).

Some authors made a distinction between HGT and transfers mediated by viruses (DeFilippis and Villarreal 2001); however viruses and transposons are precisely the kind of vectors we would expect to mediate occasional HGT so that the question is rather semantic. Transposons may exhibit narrow or wide host ranges, such as, respectively, the P elements of *Drosophila* or the AhT elements (those discovered by McClintock 1951) that can spread among animals and plants (Silva et al. 2004; Emelyanov et al. 2006; Pace et al. 2008). The possibility of rather distant transfers therefore exist in principle but the actual frequency seems altogether low, even in plants where germ lines are more exposed than in animals (Diao et al. 2006). It should be stressed here than we are focusing on possible transfers by potential vectors among groups of organisms that may appear widely different from an anthropomorphic point of view (such as diverse tetrapods, or fishes and mammals) but have a high degree of molecular compatibility, something that became already clear in early experiments on heterologous cell fusions (Harris and Watkins 1965; Harris 1970); the situation is very different when we consider distant groups of prokaryotes, such as phyla belonging to different Domains.

This raises the question of the alleged role of HGT in the acquisition of genetic innovation. In simple organisms such as Archaea, Bacteria, and some protists, expansion of metabolic capacities is an important evolutionary challenge. Genes conferring new metabolic properties could have been transferred horizontally under selection. A polyphyletic pattern of genes involved in a particular pathway is however by itself not a proof of HGT (see supra) and it is hard to see how it could be taken as an indication of metabolic innovation. Moreover, the intrinsic potential of microorganisms for functional innovation is already considerable as attested by many experimental studies. One of the main avenues is

that of promiscuous enzymes in both microorganisms (McLoughlin and Copley 2008) and complex eukaryotes (Azam et al. 2008); their evolvability is estimated to be near *inexhaustible* (for recent reviews, see Khersonsky et al. 2006; Tokuriki and Tawfik 2009).

The most important eukaryotic innovations are of another nature, however, such as modifications of the anatomy and physiology and occurrence of new behavioural possibilities (Arthur 1997). Yet, chimpanzees and man, for example despite striking differences in mental development and body posture, have extremely similar DNA sequences. Hence, the notion that rearrangements of resident genes and other mutations with regulatory effects (including the intervention of McClintock's "controlling elements", i.e. transposons) may be the main sources of innovation in higher organisms (Wray 2007); actual HGT might be of minor or no importance. It could therefore be the case that higher eukaryotes not only do not indulge in a lot of HGT but did not need to import foreign material in order to acquire their most remarkable properties. Nevertheless the question arises: if they are exposed to foreign DNA in various ways, especially "selfish DNA" (Dawkins, 1976), how did they restrain its impact?

Origin of Meiotic Sexuality as an Anti-HGT Mechanism: A Molecular Identity Check

Bdelloid rotifers were reported to contain bacterial, fungal and plant genes that are clustered in telomeric regions with mobile elements (Gladyshev et al. 2008). A basic difference between these lower Metazoa and higher eukaryotes is sexuality, absent in rotifers for a very long time (Mark Welch and Meselson 2000). We feel this correlation between the secondary (the ancestral rotifer was mictic) loss of sexuality and the presence of putatively horizontally transferred genes may be significant. Although various forms of sexuality have been observed in the three Domains (such as DNA conjugation in various bacteria and archaea), we suggest that meiotic sexuality, which is essentially an identity check for the genetic material, originated as a mechanism to restrict HGT by triggering elimination of discrepancies or, in the grossest cases, by interfering with the correct pairing of chromosomes. When chromosomes pair, a number of mechanisms (repair, conversion, recombination) are triggered allowing the elimination of deleterious differences from the descendance. We assume meiosis to have emerged in a protoeukaryote descendant of the LUCA, after the RNA to DNA transition, in a cellular context where genetic redundancy was already the rule (Glansdorff et al. 2008), thus setting the stage for a mechanism of chromosomal identity check. The origin of meiotic sexuality should indeed be sought in an immediate benefit rather than in the future advantages of genetic

recombination, something natural selection could not have foreseen. Later on, other mechanisms, such as gene silencing, may have concurred in setting a barrier to the spread of infective DNA in eukaryotes (Kurland 2005). Though the origin of sex (i.e. amphimixis) is usually considered an unsolved mystery (Charlesworth 2006; Hadany and Beker 2007), we found a recent suggestion by O'Dea (2006) that is clearly in keeping with our hypothesis since it postulates that sex emerged as a mechanism solving "intragenomic conflicts" by permitting recombinational elimination of "worthless DNA"; O'Dea's suggestion however does not refer to HGT as a possible *primum movens* nor to a protoeukaryotic descendant of LUCA as the seat of this innovation.

Conclusions

Current reasons for believing in the widespread occurrence of HGT between distant taxa (such as those belonging to different Domains) originate in great part from phylogenetic inconsistencies emerging from comparisons of gene sequences, in the noticeable absence of direct experimental evidence. A large proportion of alleged HGT may be due to "hidden paralogy", in particular to haphazard segregation in the common ancestors of the three Domains of redundant gene copies already present in the LUCA. Whereas selection may have stabilized some distant, inter-domain transfers, in general, HGT would have been selected against as soon as it threatened the cellular and metabolic integrity of the host cell. Indeed, the integration of a foreign gene in a pre-existing interaction network appears to be slow and inefficient (Lercher and Pal 2008) and probably requires cotransfer with a selected gene. If it is already difficult to believe that a metabolic or regulatory gene could be readily exchanged between two relatively close organisms, the probability of such an event occurring between Domains is approaching a null value.

We propose the hypothesis that eukaryotes evolved sex (i.e. meiosis) as an identity check limiting the impact of foreign DNA on cell survival. Indeed, if HGT may in some cases have provided material for a useful innovation, most frequently it must have appeared as an invasion against which it proved necessary to invent defence mechanisms. Sex would be one of them, specific of eukaryotes, having emerged in a descendant of the protoeukaryotic LUCA (Glansdorff et al. 2008; Kurland et al. 2006). By a curious twist, sex would appear as a safeguard against promiscuity.

Acknowledgements We thank Dr. Massimo Di Giulio for his invitation to the late Nicolas Glansdorff to write a manuscript for this special issue and the two reviewers for their positive comments and constructive suggestions. BL lab is supported by the CNRS (UMR 8621), the Agence Nationale de la Recherche (ANR-05-MMSA-0009

MDMS_NV_10), and the Université Paris-Sud (PPF Bioinformatique et Biomathématiques).

References

- Alifano P, Fani R, Lio P, Lazcano A, Bazzicalupo M, Stella Carlomagno M, Bruni CB (1996) Histidine biosynthetic pathway and genes: structure, regulation and evolution. *Microbiol Rev* 60:44–69
- Arber W (2000) Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol Rev* 24:1–7
- Archibald JM, Logsdon JM Jr, Doolittle WF (2000) Origin and evolution of eukaryotic chaperonins: phylogenetic evidence for ancient duplications in CCT genes. *Mol Biol Evol* 17:1456–1466
- Arthur W (1997) The origin of animal body plans. Cambridge University Press, Cambridge
- Azam S, Jouvét N, Jilani A, Vongsamphanh R, Yang X, Yang S, Ramotar D (2008) Human glyceraldehyde-3-phosphate dehydrogenase plays a direct role in reactivating oxidized forms of the DNA repair enzyme APE1. *J Biol Chem* 283:30632–30641
- Beiko RG, Ragan MA (2009) Untangling hybrid phylogenetic signals: horizontal gene transfer and artifacts of phylogenetic reconstruction. *Methods Mol Biol* 532:241–256
- Beiko RG, Harlow TJ, Ragan MA (2005) Highways of gene sharing in prokaryotes. *Proc Natl Acad Sci USA* 102:14332–14337
- Benachenhou N, Baldacci G (1991) The gene for a halophilic glutamate dehydrogenase: sequence, transcription, and phylogenetic analysis. *Mol Gen Genet* 230:345–352
- Benachenhou-Lahfa N, Forterre P, Labedan B (1993) Evolution of glutamate dehydrogenase genes: evidence for two paralogous protein families and unusual branching patterns of the archaeobacteria in the universal tree of life. *J Mol Evol* 36:335–346
- Blaxter M (2007) Symbiont genes in host genomes: fragments with a future? *Cell Host Microbe* 2:211–213
- Boucher Y, Bapteste E (2009) Revisiting the concept of lineage in prokaryotes: a phylogenetic perspective. *Bioessays* 31:526–536
- Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau ME, Nesbø CL, Case RJ, Doolittle WF (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet* 37:283–328
- Boucher Y, Kamekura M, Doolittle WF (2004) Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol Microbiol* 52:469–497
- Castresana J (2001) Comparative genomics and bioenergetics. *Biochim Biophys Acta* 1506:147–162
- Charlesworth B (2006) The evolutionary biology of sex. *Curr Biol* 16:R693–R695
- Comas I, Moya A, Azad RK, Lawrence JG, Gonzales-Candela F (2006) The evolutionary origin of Xanthomonadales genomes and the nature of the horizontal gene transfer process. *Mol Biol Evol* 23:2049–2057
- Dagan T, Martin W (2007) Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc Natl Acad Sci USA* 104:870–875
- Davison A, Blaxter M (2005) Ancient origin of glycosyl hydrolase family 9 cellulase genes. *Mol Biol Evol* 22:1273–1284
- Dawkins R (1976) The selfish gene. Oxford University Press, Oxford
- DeFilippis V, Villarreal LP (2001) Lateral gene transfer or viral colonization? *Science* 293:1048
- Delaye L, Becerra A, Lazcano A (2004) The nature of the last common ancestor. In: de Pouplana LR (ed) The genetic code and the origin of life. Landes Bioscience/Eureka.com/Kluwer Academic Plenum publishers, Georgetown, TX/New York, NY
- Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, Nislow C, Giaever G (2005) Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* 169:1915–1925
- Diao X, Freeling M, Lisch D (2006) Horizontal transfer of a plant transposon. *PLoS Biol* 4:e5
- Dobrindt U, Agerer F, Michaelis K, Janka A, Buchrieser C, Samuelson M, Svanborg C, Gottschalk G, Karch H, Hacker J (2003) Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J Bacteriol* 185:1831–1840
- Doolittle WF (1999a) Phylogenetic classification and the universal tree. *Science* 284:2124–2128
- Doolittle WF (1999b) Lateral genomics. *Trends Cell Biol* 9:M5–M8
- Doolittle WF, Bapteste E (2007) Pattern pluralism and the Tree of Life hypothesis. *Proc Natl Acad Sci USA* 104:2043–2049
- Doolittle WF, Boucher Y, Nesbo CL, Andersson JO, Roger AJ (2003) How big is the iceberg of which organellar genes in nuclear genomes are but the tip? *Philos Trans R Soc Lond B* 358:39–57
- Dorman CJ (2007) H-NS, the genome sentinel. *Nat Rev Microbiol* 5:157–161
- Dunning Hotopp JC, Clark ME, Oliveira DCSG, Foster JM, Fischer P, Munoz Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317:1753–1756
- Edgell DR, Malik SB, Doolittle WF (1998) Evidence of independent gene duplications during the evolution of archaeal and eukaryotic family B DNA polymerases. *Mol Biol Evol* 15:1207–1217
- Emelyanov A, Gao Y, Naqvi NI, Parinov S (2006) Trans-kingdom transposition of the maize *Dissociation* element. *Genetics* 174:1095–1104
- Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius-Dietrich G, Henze K, Kretschmann E, Richly E, Leister D, Bryant D, Steel MA, Lockhart PJ, Penny D, Martin W (2004) A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol* 21:1643–1660
- Forterre P (1995) Thermoreduction, a hypothesis for the origin of prokaryotes. *CR Acad Sci III* 318:415–422
- Forterre P, Bouthier de la Tour C, Philippe H, Duguet M (2000) Reverse gyrase from hyperthermophiles: probable transfer of a thermoadaptation trait from Archaea to bacteria. *Trends Genet* 16:152–154
- Frigaard NU, Martinez A, Mincer TJ, DeLong EF (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* 439:847–850
- Galtier N, Daubin V (2008) Dealing with incongruence in phylogenomic analyses. *Philos Trans R Soc Lond B Biol Sci* 363:4023–4029
- Gladyshev EA, Meselson M, Arhipova IR (2008) Massive horizontal gene transfer in bdelloid rotifers. *Science* 320:1210–1213
- Glansdorff N (2000) About the last common ancestor, the universal life-tree and lateral gene transfer: a reappraisal. *Mol Microbiol* 38:177–185
- Glansdorff N, Xu Y, Labedan B (2008) The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol Direct* 3:29
- Glansdorff N, Xu Y, Labedan B (2009) The origin of life and the last universal common ancestor: do we need a change of perspective? *Res Microbiol*. doi:10.1016/j.resmic.2009.05.003
- Gogarten JP, Kibak H, Dittrich P, Taiz L, Bowman EJ, Bowman BJ, Manolson MF, Poole RJ, Date T, Oshima T (1989) Evolution of the vacuolar H⁺-ATPase: implications for the origin of eukaryotes. *Proc Natl Acad Sci USA* 86:6661–6665

- Gogarten JP, Doolittle WF, Lawrence JG (2002) Prokaryote evolution in light of gene transfer. *Mol Biol Evol* 19:2226–2238
- Grogan DW, Ozarzak MA, Bernander R (2008) Variation in gene content among geographically diverse *Sulfolobus* isolates. *Environ Microbiol* 10:137–146
- Habenicht A, Hellman H, Cerff R (1994) Nonphosphorylating GADPH of higher plants is a member of the aldehyde dehydrogenase superfamily with no sequence homology to phosphorylating GADPH. *J Mol Biol* 237:165–171
- Hadany L, Beker T (2007) Sexual selection and the evolution of obligatory sex. *BMC Evol Biol* 7:245
- Harris H (1970) Cell fusion. Harvard University Press, Cambridge
- Harris H, Watkins JF (1965) Hybrid cells derived from mouse and man: artificial heterokaryons of mammalian cells from different species. *Nature* 205:640–646
- Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc Natl Acad Sci USA* 86:9355–9359
- Jain R, Rivera MC, Moore JE, Lake JA (2003) Horizontal gene transfer accelerates genome innovation and evolution. *Mol Biol Evol* 20:1598–1602
- Jensen RA (1976) Enzyme recruitment in evolution of new function. *Annu Rev Microbiol* 30:409–425
- Kandler O (1994) The early diversification of life. Nobel symposium no. 84, Early life on Earth. Columbia University Press, New York, NY, pp 152–160
- Ke D, Boissinot M, Huletsky A, Picard FJ, Frenette J, Ouellette M, Roy PH, Bergeron MG (2000) Evidence for horizontal gene transfer in evolution of elongation factor Tu in enterococci. *J Bacteriol* 182:6913–6920
- Khersonsky O, Roodveldt C, Tawfik DS (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr Opin Chem Biol* 10:498–508
- Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* 55:709–742
- Kunin V, Goldovsky L, Darzentas N, Ouzounis CA (2005) The net of life: reconstructing the microbial phylogenetic network. *Genome Res* 15:954–959
- Kurland CG (2005) What tangled web: barriers to rampant horizontal gene transfer. *Bioessays* 27:741–747
- Kurland CG, Canback B, Berg OG (2003) Horizontal gene transfer: a critical view. *Proc Natl Acad Sci USA* 100:9658–9662
- Kurland CG, Collins LJ, Penny D (2006) Genomics and the irreducible nature of eukaryote cells. *Science* 312:1011–1014
- Kuzniar A, van Ham RC, Pongor S, Leunissen JA (2008) The quest for orthologs: finding the corresponding gene across genomes. *Trends Genet* 24:539–551
- Labadan B, Boyen A, Baetens M, Charlier D, Chen P, Cunin R, Durbecq V, Glansdorff N, Herve G, Legrain C, Liang Z, Purcarea C, Roovers M, Sanchez R, Toong TL, Van de Casteele M, Van Vliet F, Xu Y, Zhang YF (1999) The evolutionary history of carbamoyltransferases: a complex set of paralogous genes was already present in the last universal common ancestor. *J Mol Evol* 49:461–473
- Labadan B, Xu Y, Naumoff DG, Glansdorff N (2004) Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase. *Mol Biol Evol* 21:364–373
- Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* 44:383–397
- Leplae R, Lima-Mendez G, Toussaint (2006) A first global analysis of plasmid encoded proteins in the ACLAME database. *FEMS Microbiol Rev* 30:980–994
- Lercher MJ, Pal C (2008) Integration of horizontally transferred genes into regulatory interaction networks takes many million years. *Mol Biol Evol* 25:559–567
- Lima-Mendez G, Van Helden J, Toussaint A, Leplae R (2008) Reticulate representation of evolutionary and functional relationships between phage genomes. *Mol Biol Evol* 25:762–777
- Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302:1401–1404
- Maddison WP (1997) Gene trees in species trees. *Syst Biol* 46:523–546
- Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV (2005) Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res* 33:4626–4638
- Mark Welch D, Meselson M (2000) Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* 288:1211–1215
- Massant J, Glansdorff N (2005) New experimental approaches for investigating interactions between *Pyrococcus furiosus* carbamate kinase and carbamoyltransferases, enzymes involved in the channeling of thermolabile carbamoyl phosphate. *Archaea* 1:365–373
- Massant J, Verstreken P, Durbecq V, Kholti A, Legrain C, Beekmans S, Cornelis P, Glansdorff N (2002) Metabolic channeling of carbamoyl phosphate, a thermolabile intermediate: evidence for physical interaction between carbamate kinase-like carbamoyl-phosphate synthetase and ornithine carbamoyltransferase from the hyperthermophile *Pyrococcus furiosus*. *J Biol Chem* 277:18517–18522
- Matic I, Taddei F, Radman M (1996) Genetic barriers among bacteria. *Trends Microbiol* 4:69–72
- Mayr E (1954) Change of genetic environment and evolution. In: Huxley J, Hardy AC, Ford EB (eds) *Evolution as a process*. Allen & Unwin, London, pp 157–180
- McClintock B (1951) Chromosome organization and genic expression. *Cold Spring Harb Symp Quant Biol* 16:13–47
- McLoughlin SY, Copley SD (2008) A compromise required by gene sharing enables survival: implications for evolution of new enzyme activities. *Proc Natl Acad Sci USA* 105:13497–13502
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 324:1724–1726
- Murray NE (2002) Immigration control of DNA in bacteria: self versus non-self. *Microbiology* 148:3–20
- Nakamura Y, Itoh T, Matsuda H, Gojobori T (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat Genet* 36:760–766
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–238
- Nikoh N, Tanaka K, Shibata F, Kondo N, Hizume M, Shimada M, Fukatsu T (2008) Wolbachia genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. *Genome Res* 18:272–280
- O’Dea JD (2006) Did conflict between chromosomes drive the evolution of sex? *Calodema* 8:33–34
- Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304
- Ouzounis CA, Kunin V, Darzentas N, Goldovsky L (2006) A minimal estimate for the gene content of the last universal common ancestor-exobiology from a terrestrial perspective. *Res Microbiol* 157:57–68
- Ovádi J, Srere PA (2000) Macromolecular compartmentation and channeling. *Int Rev Cytol* 192:255–280

- Pace JK, Gilbert C, Clark MS, Feschotte C (2008) Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* 105:17023–17028
- Padian K (2008) Darwin's enduring legacy. *Nature* 451:632–634
- Papp B, Pal C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. *Nature* 424:194–197
- Philippe H, Delsuc F, Brinkmann H, Lartillot N (2005) Phylogenomics. *Annu Rev Ecol Evol Syst* 36:541–562
- Pierce SK, Massey SE, Hanten JJ, Curtis NE (2003) Horizontal transfer of functional nuclear genes between multicellular organisms. *Biol Bull* 204:237–240
- Poole AM, Jeffares D, Penny D (1999) Early evolution: prokaryotes, the new kids on the block. *Bioessays* 21:880–889
- Provine WB (2004) Ernst Mayr: genetics and speciation. *Genetics* 167:1041–1046
- Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc Natl Acad Sci USA* 106:8605–8610
- Richly E, Leister D (2004) NUMTs in sequenced eukaryotic genomes. *Mol Biol Evol* 21:1081–1084
- Salzberg SL, White O, Peterson J, Eisen JA (2001) Microbial genes in the human genome: lateral transfer or gene loss? *Science* 292:1903–1906
- Serre V, Guy H, Liu X, Penverne B, Hervé G, Evans D (1998) Allosteric regulation and substrate channeling in multifunctional pyrimidine biosynthetic complexes: analysis of isolated domains and yeast-mammalian chimeric proteins. *J Mol Biol* 281:363–377
- Silva JC, Loreto EL, Clark JB (2004) Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol* 6:57–71
- Sørensen SJ, Bailey M, Hansen LH, Kroer N, Wuertz S (2005) Studying plasmid horizontal transfer in situ: a critical review. *Nat Rev Microbiol* 3:700–710
- Soria-Carrasco V, Castresana J (2008) Estimation of phylogenetic inconsistencies in the three domains of life. *Mol Biol Evol* 25:2319–2329
- Swofford DL, Waddell PJ, Huelsenbeck JP, Foster PG, Lewis PO, Rogers JS (2001) Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst Biol* 50:525–539
- Taoka M, Yamauchi Y, Shinkawa T, Kaji H, Motohashi W, Nakayama H, Takahashi N, Isobe T (2004) Only a small subset of the horizontally transferred chromosomal genes in *Escherichia coli* are translated into proteins. *Mol Cell Proteomics* 3:780–787
- Than C, Ruths D, Innan H, Nakhleh L (2007) Confounding factors in HGT detection: statistical error, coalescent effects, multiple solutions. *J Comput Biol* 14:517–535
- Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5:123–135
- Tokuriki N, Tawfik DS (2009) Protein dynamism and evolvability. *Science* 324:203–207
- Vetsigian K, Woese C, Goldenfeld N (2006) Collective evolution and the genetic code. *Proc Natl Acad Sci USA* 103:10696–10701
- Wang M, Yafremava LS, Caetano-Anolles D, Mittenhal JE, Caetano-Anolles G (2007) Reductive evolution of architectural repertoires in proteomes and the birth of the tripartite world. *Genome Res* 17:1572–1585
- Woese C (1998) The universal ancestor. *Proc Natl Acad Sci USA* 95:6854–6859
- Woese C (2002) On the evolution of cells. *Proc Natl Acad Sci USA* 99:8742–8747
- Woese CR, Olsen GJ, Ibba M, Söll D (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* 64:202–236
- Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* 8:206–215
- Xu Y, Glansdorff N (2002) Was our ancestor a hyperthermophilic prokaryote? *Comp Biochem Physiol A* 133:677–688
- Zaneveld JR, Nemergut DR, Knight R (2008) Are all horizontal gene transfers created equal? Prospects for mechanism-based studies of HGT patterns. *Microbiology* 154:1–15