Review

Horizontal gene transfer in evolution: facts and challenges
Luis Boto*

Departamento Biodiversidad y Biología Evolutiva, Museo Nacional Ciencias Naturales, CSIC, C/ José Gutierrez Abascal 2, 28006 Madrid, Spain

The contribution of horizontal gene transfer to evolution has been controversial since it was suggested to be a force driving evolution in the microbial world. In this paper, I review the current standpoint on horizontal gene transfer in evolutionary thinking and discuss how important horizontal gene transfer is in evolution in the broad sense, and particularly in prokaryotic evolution. I review recent literature, asking, first, which processes are involved in the evolutionary success of transferred genes and, secondly, about the extent of horizontal gene transfer towards different evolutionary times. Moreover, I discuss the feasibility of reconstructing ancient phylogenetic relationships in the face of horizontal gene transfer. Finally, I discuss how horizontal gene transfer fits in the current neo-Darwinian evolutionary paradigm and conclude there is a need for a new evolutionary paradigm that includes horizontal gene transfer as well as other mechanisms in the explanation of evolution.

Keywords: evolution; horizontal gene transfer; evolutionary mechanisms

1. THE RISE OF HORIZONTAL GENE TRANSFER

At the beginning, the founders of molecular phylogenetics used molecular information from different proteins and genes to reconstruct phylogenetic relationships between organisms (Zuckerkandl & Pauling 1965). One of these molecular markers, the sequence of 16S RNA genes, was proposed to be a good tool for reconstructing old phylogenetic relationships (Woese et al. 1990). It was soon realized, however, that different genes could lead to phylogenetic incongruence and conflicting phylogenies, mainly in micro-organisms (Hilario & Gogarten 1993; Gupta & Singh 1994; Golding & Gupta 1995; Whitehouse et al. 1998), by grouping together species or species groups that are split by other morphological, physiological or molecular markers.

On this background, the concept of horizontal gene transfer, ‘the non-genealogical transmission of genetic material from one organism to another’ (Goldenfeld & Woese 2007), is a source of new genes and functions to the recipient of the transferred genetic material. In this sense, it is a mechanism that permits the acquisition of evolutionary novelties. But these acquisitions are primarily non-genealogical, questioning, in my opinion, the neo-Darwinian conception of a gradualist process driving the appearance of novel traits and functions.

In this review, I will discuss how horizontal gene transfer fits in with current evolutionary thinking, as well as the challenges that it proposes for the current evolutionary paradigm.
thing: how does horizontal gene transfer challenge the traditional neo-Darwinian view of evolution as a gradual process of variation with descendence? I review below the state of art in relation to these questions.

2. HOW IMPORTANT IS HORIZONTAL GENE TRANSFER IN EVOLUTION OF EARTH’S BIOTA?

Since the emergence of horizontal gene transfer as a way of explaining phylogenetic incongruence using different gene trees, a considerable number of studies have been published about genes that have been acquired by horizontal gene transfer (Gogarten et al. 2002; Lerat et al. 2005), both in Bacteria (Saunders et al. 1999; Ochman et al. 2000) and Archaea (Doolittle & Logsdon 1998; Faguy & Doolittle 1999), as well as in eukaryotes (Andersson 2005). These studies show that the transfer can occur not only among but also between domains in all possible directions: from Bacteria to Archaea (Rest & Mindell 2003), from Archaea to Bacteria (Gophna et al. 2004), from Archaea to Eukarya (Andersson et al. 2003), from Bacteria to Eukarya (Watkins & Gray 2006), from Eukarya to Bacteria (Guljamow et al. 2007) and even within Eukarya (Nedelcu et al. 2008). However, it is in bacterial and archaean evolution that horizontal gene transfer has been more widely documented and accepted.

A recurrent question addressing the importance of horizontal gene transfer in evolution is how many genes in any given organism have been acquired by horizontal gene transfer. It is evident that in Bacteria and Archaea, even the transfer of a single or a few genes can give recipient organisms the opportunity to exert a new function, exploiting new ecological niches (Fournier & Gogarten 2008). However, its importance as an evolutionary mechanism can be limited if only a few horizontal gene transfer events have taken place in the history of life.

Today, it is possible to address this question with respect to the evolution of Bacteria and Archaea using the abundant number of available completed genome sequences. However, in order to achieve this goal, it is necessary to previously identify what genes of an organism have been transferred. In this way, several criteria have been proposed and discussed (Ragan 2001, 2002; Lawrence & Ochman 2002) that permit this identification with more or less trouble: mainly, bias in codon usage and different base composition in relation to others genes in the genome and phylogenetic incongruence.

Criteria based on codon usage bias and differential base composition have undergone several criticisms (Koski et al. 2001; Kuo & Ochman 2009). The foreign genes are ameliorated after many generations (Marri & Goldberg 2008) because both sets of genes, resident and transferred, support the same mutational biases once they are part of the same genome. In this way, these criteria could allow the identification of recent gene transfer events but are inefficient in detecting events that occurred a long time ago.

In relation to phylogenetic incongruence, the problem is that it is difficult to differentiate between horizontal gene transfer and gene loss in a phylogenetic tree (Ragan 2001, 2002; Lawrence & Ochman 2002; Snell et al. 2002; Zhaxybayeva et al. 2007). A careful comparison of the different phylogenetic methods employed along with the use of a test of compatibility between trees can provide some clues about the process involved (Gogarten & Townsend 2005).

Despite these caveats, a lot of work has been carried out in the last few years to try to deal with the importance of horizontal gene transfer in Bacteria and Archaea evolution (Faguy & Doolittle 1999; Nelson et al. 1999; Ochman et al. 2000; Ochman 2001; Gogarten et al. 2002; Kurland et al. 2003; Philippe & Douady 2003; Pål et al. 2005a; Choi & Kim 2007; Koonin & Wolf 2008), and the attained results are controversial depending on whether the emphasis is placed on gene transfer or gene loss (Kurland et al. 2003). However, the emergent picture is that horizontal gene transfer plays a larger role in microbial evolution than previously thought.

Today it is estimated that between 1.6 and 32.6 per cent of the genes of each microbial genome have been acquired by horizontal gene transfer (Koonin et al. 2001). Moreover, a recent study by Dagan et al. (2008) using a network analysis of shared genes in 181 sequenced prokaryotic genomes shows that this increases dramatically to 81 ± 15 per cent if the cumulative impact of horizontal gene transfer towards lineages is considered, which reinforces the importance of this mechanism in microbial evolution.

On the other hand, it has been assumed that the role of horizontal gene transfer in eukaryotic evolution is less important compared with its relevance in prokaryotic evolution. Evidently, the lower number of sequenced eukaryotic genomes compared with that of bacterial and archaean genomes makes it difficult to estimate how many genes have been acquired by horizontal gene transfer in eukaryotes. However, we are only starting to realize that horizontal gene transfer is not a negligible force in modulating eukaryotic genome evolution.

Presently, we already know that horizontal gene transfer is important in unicellular eukaryote evolution (Huang et al. 2004; Andersson 2005), and it is widely accepted that eukaryotic nuclear genomes contain several genes of microbial origin that have been transferred from mitochondrial and plastid ancestors via ancient endosymbiotic events (Keeling & Palmer 2008; Lane & Archibald 2008).

But a recent study also highlights the role of horizontal gene transfer modulating the evolution of other eukaryotic genomes in a different way: by acquisition of novel genes from living endosymbionts (Hotopp et al. 2007).

In addition, it has been shown that the acquisition of foreign genes from bacteria and other eukaryotes is also important in fungi (Richards et al. 2006) and plant evolution (Martin et al. 1993; Huang & Gogarten 2008), and has contributed to the evolution of bdelloid rotifers (Gladyshchev et al. 2008). On the other hand, horizontal gene transfer in eukaryotic evolution would be even more important if we consider animal and plant hybridization a massive horizontal gene transfer event, as well as the proposed role of hybridization in evolution (Seehausen 2004).

In the following section, I will focus the discussion on the impact of horizontal gene transfer in prokaryotic evolution.
3. HORIZONTAL GENE TRANSFER AND MICROBIAL EVOLUTION

The impact of horizontal gene transfer in microbial evolution (Bacteria and Archaea) is dependent on the number of genes that have been transferred to and successfully maintained in microbial genomes, but it is also dependent on the extent of the phenomena, in both evolutionary time framework (considering both recent and ancient events) and phylogenetic distance between organisms.

In this section, I will review our current knowledge about the type of genes involved in transfer events and the way in which these genes are maintained post-transfer. This discussion can shed light on two important questions concerning the impact of horizontal gene transfer in prokaryotic evolution. Has horizontal gene transfer been equally prevalent throughout evolution? Is horizontal gene transfer more prevalent between closely related organisms than between those that are distantly related? Finally, I will review the impact that horizontal gene transfer has on the reconstruction of phylogenetic relationships between Bacteria and Archaea, as well as on the species concept in the prokaryotic world.

(a) Gene transfer and the fate of transferred genes

Horizontal gene transfer results both from the successful transfer of genetic material (mediated by processes such as conjugation, transduction or transfection and ulterior recombination) and from the survival of the transferred genetic material throughout the generations. The presence of certain physical barriers to transfer, as well as different selective forces over the transferred genes, may explain observed differences in the type of genes involved in horizontal gene transfer.

In 1999, Jain et al. proposed the complexity hypothesis (Jain et al. 1999) to explain the observed differences in horizontal gene transfer susceptibility between genes. This hypothesis proposes that the so-called informational genes (involved in DNA replication, transcription and translation, and whose products participate in multiple molecular interactions) are less prone to horizontal gene transfer than operational genes (involved in cell maintenance and whose products have few interactions with other molecules).

This hypothesis gained support from the characterization by Bayesian inference of recently acquired genes in prokaryotic genomes (Nakamura et al. 2004), which has shown that the fraction of transferred genes is biased towards genes involved in DNA binding, pathogenicity and cell surface functions, all of them included among the functions of operational genes. This work also shows, however, that not all operational genes are participating equally in horizontal gene transfer events.

In addition, a recent study (Sorek et al. 2007) searched for genes that cannot be cloned in Escherichia coli in all of the completed bacterial and archaeal genomes. The authors found that, despite the fact that there are not any a priori barriers to cloning any gene in E. coli (see below), some informational genes represent a substantial fraction of unclonable genes. However, there are some informational genes that can also be cloned.

Another recent study (Hao & Golding 2008), which used maximum-likelihood models that incorporate rate variation to evaluate the contribution of gene insertion and deletion among 173 completed bacterial genomes, shows that according to the complexity hypothesis, informational genes are less prone to be involved in horizontal gene transfer than operational genes, but the difference between informational and operational genes is a small fraction of the overall observed variation.

Taking in all, these studies suggest that the functional split-up of genes into those that are informational and those that are operational is not an absolute predictive tool for transferred genes.

A complementary approach to understand whether there are differences between genes, with regard to their involvement in transfer events, is to determine what processes are involved in post-transfer gene maintenance. It is accepted that the maintenance of a transferred gene is associated with positive selection (Gogarten et al. 2002; Pál et al. 2005b). In this way, genes having a useful function are preserved while useless genes are removed.

Several recent studies shed interesting results with regard to the maintenance of transferred genes (see review by Kuo & Ochman 2009). For example, it has been shown that the integration of a single transferred gene into regulatory interaction networks is very slow (Lercher & Pál 2008) in the case of genes providing the receptor with new functions, and it is dependent on the number of partners for the gene product in the regulatory network, according to the complexity hypothesis. Moreover, products of these genes are integrated in the periphery of the corresponding regulatory network. On the other hand, transferred genes codifying for products with few partners are more easily deleted from genomes (Rocha 2008).

Recent work also shows that the transfer of complete operons enables the integration and maintenance processes, allowing a rapid gain of function and facilitating the coordinate regulation of the new genes in the receptor (Price et al. 2008).

Taking in all, these results support that the involvement of their products in multiple molecular interactions (complexity) is a more important constraint to transfer and maintenance of genes in the prokaryotic world than the functional class to which the transferred genes belong.

(b) Incidence of horizontal gene transfer over evolutionary time

The studies discussed above suggest that the evolutionary distance between organisms can be another important constraint with regard to transfer because genes transferred between organisms separated a long time ago were found to participate in very different regulatory networks. On the other hand, I have previously discussed that ancient transfer events are difficult to detect because of the amelioration process that affects the evolution of foreign genes in the receptor genome (Almeida et al. 2008; Marri & Golding 2008; Kuo & Ochman 2009).

Several studies suggest that gene transfer could effectively be more frequent for short and intermediate evolutionary distances but uncommon between organisms that are separated by large evolutionary time frames (Ochman et al. 2000; Brugger et al. 2002; Nakamura et al. 2004; Ge et al. 2005; Choi & Kim 2007;
Dagan et al. (2008). A recent study (Wagner & De la Chaux 2008) has analysed the evolution of 2091 insertion sequences in 438 completely sequenced prokaryotic genomes and found only 30 cases of presumed transfer events among distantly related clades. Twenty-three of these events seemed to be ancient while only seven were recent.

However, instances of gene transfer between Archaea and Bacteria have been described (Rest & Mindell 2003; Gophna et al. 2004), which shows that horizontal gene transfer could affect evolution in the prokaryotic world along the different evolutionary times.

A recent study (Kanhere & Vingron 2009) compares the distance between orthologues and the intergenomic distances to try to detect ancient transfer events in prokaryotic genomes. The authors found that 118 of the 171 gene transfer events were between Archaea and Bacteria, and they correspond mainly to metabolic genes. Seventy-four per cent of these events were transfers from Bacteria to Archaea and the remaining 26 per cent were transfers from Archaea to Bacteria. Only 53 genes were gene transfer events between bacteria phyla and corresponded mainly to genes involved in translation.

Despite the fact that this approach is limited to protein families that evolve at a constant rate, the study shows interesting trends within the interdomain gene transfer. It seems that the most transfers between Bacteria and Archaea have taken place in the Bacteria to Archaea direction.

On the other hand, studies by Zhaxybayeva et al. (2006, 2009) show that the intra-phylum versus inter-phylum gene exchange is different among different bacteria lineages: in Cyanobacteria, intra-phylum gene change seems to be more important than inter-phylum exchange (Zhaxybayeva et al. 2006). In Thermotogales, however, and in particular in Firmicutes (Zhaxybayeva et al. 2009), the inter-phylum exchange is dominant over intra-phylum gene transfer. In addition, the proposed multiple gene exchange between ε-proteobacteria and Aquificales seems another example of inter-phylum exchange dominance (Boussau et al. 2008).

Taking in all, these studies suggest that despite the fact that gene transfer can be more frequent between closely related organisms, it may also take place between distantly related organisms, contributing to evolution of Archaea and Bacteria.

On the other hand, the fact that recent transfer events can be more easily detected adds a bias to the study of gene transfer in prokaryotic evolution, confounding the real impact of ancient gene transfer events.

(c) Prokaryotic phylogenetic relationships, bacterial species concept and horizontal gene transfer
Reconstructing the phylogenetic relationships between Bacteria and Archaea, morphological characters are of limited use (Bohannon 2008) compared with metabolic and molecular markers. However, horizontal gene transfer challenges in many cases the correct reconstruction of these relationships, confounding the phylogenetic signal present in these markers. Some authors (Doolittle 1999; Martin 1999; Doolittle & Bapteste 2007) question whether it is possible to reconstruct an accurate phylogenetic tree for the microbial world at all, considering the existence of horizontal gene transfer events. Others support the idea that some core genes are never transferred (Wolf et al. 2002; Brown 2003), thus maintaining a true phylogenetic signal that enables the reconstruction of a microbial phylogenetic tree. Finally, others (Kurland et al. 2003; Kurland 2005) consider that the existence of many barriers to gene transfer between organisms lowers the impact of horizontal gene transfer in phylogenetic reconstruction.

A recent study (Sorek et al. 2007) dealing with this topic searched in all the completed bacterial and archaeal genomes for genes that cannot be cloned in E. coli as a proxy to the study of barriers against horizontal gene transfer (see §3(a)). Their results suggest that there are no absolute barriers to gene transfer because genes in all the families considered can be cloned in E. coli from at least one of the genomes.

In addition, the results of a network analysis of shared genes (Dagan et al. 2008) agree with the idea that horizontal gene transfer leaves no gene family untouched.

Supporting the idea that horizontal gene transfer challenges the reconstruction of phylogenetic relationships among prokaryotes, another study claims that less than 0.7 per cent of the prokaryotic genes may be considered core genes (Bapteste et al. 2008), making the construction of a phylogenetic tree unsustainable.

Recently, the pan-genome concept—initially developed to determine how many genomes should be sequenced from any given bacterial species in order to get an accurate representation of the whole gene repertoire (Tettelin et al. 2005, 2008; Hogg et al. 2007) and to define the complete set of genes present in a prokaryotic group—has been applied to the study of the complete set of genes present in sequenced bacterial genomes (Lapierre & Gogarten 2009). In this interesting study, the authors search for the presence of homologue genes in 573 completed genomes using BLAST and conclude that only 8 per cent of the genes in a typical bacterial genome (approx. 250 gene families in all genomes) are present in 99 per cent of the sampled genomes and therefore can be considered to be core genes that are part of the extended core genome or set of shared genes.

Lapierre & Gogarten identify two other gene categories: the so-called character genes (shared by a group of organisms) present in a subset of genomes (64 per cent of the genes in a typical bacterial genome and 7900 gene families in all genomes) and the so-called accessory genes present in only one or in only a few of the genomes (28 per cent of the genes in a bacterial genome and an infinite number of gene families in all genomes). The authors also suggest that character gene evolution is mainly based on mutation, gene duplication and horizontal gene transfer, while horizontal gene transfer and gene losses are involved in the evolutionary history of accessory genes.

The final conclusion of this study is that the bacterial ‘pan-genome’ (the set of all genes present in bacteria) is of infinite size, demonstrating the plasticity of the genome evolution in prokaryotes.

Disregarding the fact that the methodology used can lead to underestimations or overestimations of particular gene categories, this study strongly underlines the impact that horizontal gene transfer has had on the
evolutionary history of prokaryotes and provides us with important clues towards understanding the evolution of prokaryotic genomes.

Bearing in mind the points discussed above, some authors support the point that it is impossible to reconstruct a tree of life, considering the pervasiveness of horizontal gene transfer in modulating Bacteria, Archaea and even Eukaryota genome evolution. The metaphors of a web of life (Doolittle 1999) and a ring of life (Rivera & Lake 2004) have therefore been proposed as an alternative.

The web of life tries to represent phylogenetic relationships as a tree with many interbranch connections. On the other hand, the ring of life attempts to represent the three life domains as being connected in a single ring, considering that the eukaryotic domain has been originated from a hybridization event between representatives from the Bacteria and Archaea domains (proposed by several authors but hotly debated; Embley & Martin 2006) as well as the widespread gene exchange occurring between both domains.

Anyway, several tools have been proposed to deal with horizontal gene transfer in phylogenetic reconstruction, which encompass phylogenetic networks, supermatrices and supertrees (Huson & Bryant 2006; Galtier & Daubin 2008). The problem is that these tools are not free of criticisms (Rannala & Yang 2008). Several authors even propose that horizontal gene transfer may be used as a tool in phylogenetic reconstruction, providing sinapomorphies that facilitate the reconstruction of problematic relationships (Huang & Gogarten 2006; Keeling & Palmer 2008).

Wagner & De la Chaux’s (2008) study discussed above, which shows that horizontal gene transfer can be an infrequent event between distant clades, provides new arguments to those who defend the reliability of phylogenetic inferences in the microbial world. If Wagner and De la Chaux’s results are true for other non-mobile DNAs, the chance of obtaining a reliable phylogenetic reconstruction in the microbial world will increase.

In addition, a new method for phylogeny reconstruction in prokaryotes based on genome organization has recently been proposed (Merkl & Wiezer 2009) that seems robust in prokaryotes based on genome organization has recently been proposed (Merkl & Wiezer 2009) despite the fact that environmental cues may be used to define bacterial species (Cohan & Perry 2007; Ward et al. 2008). Other authors (Gevers et al. 2005) propose a sequential approach, using rRNA sequences to define prokaryotic genera, as well as a multi-locus sequence analysis to define species in the genera, and using different sets of genes for each genus or prokaryotic group. The most radical standpoints (Bapteste & Boucher 2008) propose a new systematic based on the redefinition of natural groups and evolutionary units considering the prevalence of horizontal gene transfer in the microbial world.

In this new frame, evolutionary units can be considered at different levels (composite evolutionary units), and the history of currently established entities can be split up into the different histories of their evolutionary units (transferred genes, vectors, etc.). This new microbial systematic proposes the construction of a taxonomically interactive database that encompasses overlapping groups, as well as the multiplication of names and taxa, once biological units involved in the evolutionary process have been identified. In this sense, the concept of open and closed lineages (Boucher & Bapteste 2009) discussed above can be very useful.

While waiting for new tools that will allow us to include horizontal gene transfer in phylogenetic reconstructions, the unanswered question still remains of whether horizontal gene transfer prevents us from using the tree paradigm in microbial evolution, as well as the species concept in the microbial world.

4. HORIZONTAL GENE TRANSFER IN CURRENT EVOLUTIONARY THINKING

Despite some recent attempts to integrate horizontal gene transfer with the neo-Darwinian paradigm of evolution (Arber 2008), horizontal gene transfer involves the exchange of genetic material between different organisms in a single generation.

Evidently, once transferred, natural selection can determine which genes spread throughout populations. But the acquisition of novel genes in a single generation, in my point of view, is far from the gradualist or slow-change principle proposed by the neo-Darwinians (Dobzhansky 1937; Mayr 1993), and it has a certain Lamarckian taste (Goldenfeld & Woes 2007).

As has been suggested (Feder 2007), single-nucleotide modifications usually only modify existing genes and functions. The acquisition of new genes and functions requires other mechanisms. It is in this sense that horizontal gene transfer, along with gene or genome duplication, hybridization and other mechanisms of gene acquisition (Ohno 1970; Rivera & Lake 2004; Seehausen 2004; Feder 2007), can be considered an evolutionary mechanism that challenges the neo-Darwinian paradigm (O’Malley & Boucher 2005; Koonin 2009).

I am aware that this is a disputed claim because many authors consider horizontal gene transfer a process of rapid change, however gradualist (the question of what gradualism means in evolution deserves a more extended
discussion and goes beyond the scope of this review). It is also true that not all authors agree with the relative importance of horizontal gene transfer in evolution, as has been discussed above.

But it is also true that authors are moving towards a continuous sense of acceptance (from strong rejecters and moderate rejecters to moderate accepters and strong supporters) in the light of new evidences (O’Malley & Boucher 2005), and this process is determining the shifting in the current paradigm.

It is in this sense that I think that horizontal gene transfer, along with the consideration of development as an internal evolutionary force, as proposed by evo-devo (Gould 1977; Alberch 1982), epigenetic changes (Jablonka & Lamb 2005; Bird 2007) and other emerging concepts such as evolvability (Wagner 2008), needs to be integrated in a new synthesis or paradigm, which will explain both eukaryotic and prokaryotic evolution. This new synthesis (without denying the role of natural selection) needs to incorporate the emerging evolutionary knowledge (Dean & Thornton 2007; Pigliucci 2007; Carroll 2008; Koonin 2009), including mechanisms other than single-point mutations and gradual variability.

Only in this way will it be possible to come to a more robust evolutionary theory, which will be able to overcome the caveats of the neo-Darwinian theory or modern synthesis, especially those that derive from ‘ad hoc’ adaptationist explanations for the new knowledge.

Obviously, this proposal is not a new one—Darwin himself shows a pluralistic spirit about other forces involved in evolution (Darwin 1859)—and different authors have proposed a similar integration of the different approaches, like evo-devo and comparative genomics among others. In this sense, my position is that not all in evolution is black or white. Selection and neutral variation, phylogeny and development, gradualism and innovation, vertical and horizontal inheritance: every one of these is a piece of an intricate puzzle, and it is thus necessary to piece them together to achieve a coherent understanding of evolution.

5. CONCLUSIONS

Horizontal gene transfer is an important force modulating evolution in the prokaryotic world and the evolution of particular eukaryotes. Although gene exchange is easier in closely related organisms, horizontal gene transfer occurred between both domains in the evolution of Archaea and Bacteria. However, it is a disputed point whether horizontal gene transfer precludes the reconstruction of phylogenetic relationships in the microbial world. In any case, horizontal gene transfer is not a canonical or typical evolutionary mechanism. Thus, I agree with other authors that there is a need for a new paradigm in evolution that includes horizontal gene transfer among other neo-Darwinian and non-neo-Darwinian mechanisms.

The author is thankful to A. G. Sáez, A. García Valdecasas and I. Doadrio for the critical reading of the manuscript and for their very helpful suggestions, and to Elena Bulmer for the review of the English text. I am also indebted to two anonymous referees for suggestions that substantially improved the manuscript.

REFERENCES


