

The Bacterium That Got Infected by a Cow!

Horizontal Gene Transfer and Evolution

Saurabh Dhawan and Tomás J Ryan

This article is a commentary on horizontal gene transfer and how it is modifying our understanding of evolution.

The bizarre observation, made in 1970, that the enzyme *trypsin* of bacterium *S. griseus*, though different from that of the related bacterial species, is similar to the bovine trypsin, led a surprised Brian Hartley to comment, "...the bacterium must have been infected by a cow!" The radical instance of a bacterium picking up a gene from an organism outside its species was later explained as an example of horizontal or lateral gene transfer, a fundamental biological phenomenon that this article will attempt to review.

Let us start from the age-old question of heredity; how does the genetic material perpetuate itself? The expected answer will be, only *vertically* – that is from parent-to-offspring – a dogmatic view central to most of the evolutionary paradigms. Increasing evidence arising from the massive genome projects now tell us that genes could well be transmitted *horizontally*, i.e. from one organism to another of a different species. In contrast to the downward passage of genes between successive generations, horizontal gene transfer (HGT) refers to sideways passage of genes between discrete species.

The Recent Rush

Horizontal gene transfer is by no means a newfound biological phenomenon. The process was observed in bacteria as far back as the 1950's when the fundamental microbial processes of conjugation and transformation were being characterized. However, at that time it was dismissed as being very rare and too insignificant to play a role in evolution. The recent flood of information from the genome sequence data has revealed that horizontal



(left) Saurabh Dhawan is currently pursuing an MS in neuro-cognitive psychology at Ludwig-Maximilians University, Germany. His interests are in life processes underlying evolution, development and human cognitive abilities while his current scientific efforts are focussed on the neural basis of conscious perception.

(right) Tomás J Ryan is currently a PhD student in the Genes to Cognition Group at the Wellcome Trust Sanger Institute, Cambridge. His primary research interests lie in molecular evolution and the genetic basis of neural development, synaptic plasticity and cognition.

Keywords

Horizontal gene transfer, molecular evolution, phylogenetic trees, bacterial evolution, the selfish gene.



gene transfers have occurred in evolutionary history much more often than the scientific community had first appreciated.

Well-documented cases of HGT have come up across domains. In a study involving *E. coli*, the bacterium present in the human gut, investigators at University of Rochester calculated that nearly 20 % of its DNA does not come from its parental lineage. Instead it originated in other microbes and was transferred laterally to *E. coli*. More surprisingly still, cases of HGT have also been found between prokaryotes and eukaryotes; it was HGT of certain genes from bacteria to unicellular yeast that accounts for the ability of certain yeast strains to propagate in the absence of oxygen – the biological process that we later capitalised on in the baking of bread and fermentation of alcohol. HGT from bacteria to multi-cellular organism is of even greater evolutionary interest as it has been shown to facilitate speciation during evolution. For instance, a genome fragment of *Wolbachia*, the bacterial endosymbiont, has been shown to have been translocated to the X chromosome of its host organism, the Adzuki Bean Beetle, during very recent evolutionary time. This event is of significance because a gene transferred here causes cytoplasmic incompatibility¹ and therefore would promote reproductive isolation between different sub-populations of beetles, ultimately leading to speciation. At even higher levels of evolutionary processes, HGT can be observed between different eukaryotes. Though the instances noticed in animals are rare and inconclusive, HGT has been demonstrated between different plant species, specifically between parasitic and host plants, and in both directions.

¹This is a phenomenon where crosses between infected males and uninfected females are non-productive.

Box 1. Outsourcing is Not New

Some bacteria were recently shown to use HGT to outsource their work to viruses! Certain cyanobacterial species are continuously in deficit of photosynthetic protein D1 due to light induced damage. During evolution, a copy of the bacterial gene for this protein was transferred to the genome of *S-PM 2*, a virus which infects them. The virus could thus produce a photosynthetic protein, which though of no direct use, enhances the chances of survival of its host and hence itself, demonstrating the utilisation of horizontal gene transfer during symbiosis.



How Does the Gene Hop?

How do these genes cross all boundaries and hop between species? There are three major mechanisms for this in prokaryotes. In *transformation*, a cell picks up genetic material directly from its environment. Alternatively, in *conjugation*, donor and recipient bacteria come in physical contact so that DNA exchange occurs directly between the two cells. Lastly, as a part of their life cycle, certain *viruses* that infect bacteria (known as bacteriophages or phages) inject their DNA into bacteria. Once inside, the viral genes replicate to make new copies, which then get repackaged to form the virus. At times however, the new viruses carry in their package some bacterial DNA instead of their own. This foreign DNA in the virus head can then be transferred to the whole spectrum of species the virus infects, a process known as *transduction*.

The above three mechanisms only account for horizontal gene transfers between bacteria and do not explain how bacterial genes are evident in vertebrate and plant genomes. Due to the increasing complexity of higher eukaryotes, more convoluted mechanisms of HGT must be present to allow prokaryote to eukaryote transfer. However, the details of HGT mechanisms are trivial in comparison to evolutionary significance of the HGT events themselves and hence will not be discussed here.

Detecting a Swap

The emergence of HGT into the lime light has been an indirect consequence of the numerous large scale genome sequencing projects. A laterally acquired gene can be loosely distinguished from the parental genes in the following ways:

A Scattered Family: Since the transfer event was unique to the recipient, the swapped gene will be limited only to itself and will be absent from its close relatives. This gene can then be compared to distant genomes to search for the donor. As in the earlier example, the *S. griseus* trypsin was found to be similar to that of the cow but different from that of closely related bacteria



and hence identified as a laterally transferred gene. This is the premise used in applying systematic phylogenetic analysis, where trees are constructed to illustrate the evolutionary relationships of similar genes across different species.

Mobility Marks: In a swapped gene, DNA sequences adjacent to it will contain vestigial marks of the vehicle sequences that had caused their transfer.

Odd Man Out: DNA sequences of a particular species show certain unique properties; for example, their overall G+C (types of bases in DNA) content is similar. Gene segments that have come laterally from a different source retain the sequence properties of their donor genome and can thus be distinguished from the original DNA.

Shifts in the Paradigms

HGT, a process now central to the concept of evolution was neglected through much of the development of contemporary evolutionary theory. Its recent acceptance as a common genetic mechanism has put fresh challenges before the existing paradigms of evolution. A number of interesting new problems and approaches that have thus emerged are discussed below; though, they remain hotly debated.

The Open Gene Market: A New look at the Bacterial World

The post World War II period saw the widespread use of antibiotics and many of the bacterial diseases were thought to have been conquered. However, in 1960's the same infectious bacteria reappeared in many countries causing a large number of deaths. This time they were equipped with antibiotic resistance genes and hence were untreatable by the four or five general antibiotics earlier used successfully against them. The resistance genes originally absent in these infectious bacteria were picked up laterally from other bacterial genera. The selective pressure caused by the massive use of antibiotics accelerated this transfer.



For the infectious bacteria, it was a war of life or death against the antibiotics, so they borrowed the weapons from their relatives and survived.

It became evident that bacteria were not discrete species living in genetic isolation like the eukaryotes. Rather they lived in an open gene market, with access to the common global reserve of all bacterial genes, from which they could ‘buy’ (laterally acquire) genes if and when both necessity (selective pressure) and chance acted together. Each group of bacteria existing in a particular ecological niche would thus harness a no-vel combination of different genes available to it in the market.

To add substance to these metaphors, Howard Ochman of the University of Arizona and his colleagues suggest that unlike eukaryotic evolution, bacterial speciation might be ‘majorly driven by a high rate of horizontal transfers’. New genes from far away confer new advantageous skills, enabling bacteria to adapt to new environments and hence survive the hostile conditions.

Organismal Adaptation and the Selfish Gene

In his much acclaimed and controversial book *The Selfish Gene* Richard Dawkins presents an alternative perspective of life on earth. According to his arguments, all biological entities are merely ‘vehicles’ for DNA molecules to replicate themselves. To put it simply, genes became organized and packaged into complex organisms, for instance humans, because doing this ensured their own spread and survival. He called the genetic material as ‘the selfish gene’, where selfishness indicates that the unit of life is primarily concerned with its own replication. “A chicken is DNA’s way of making more DNA”, he remarks.

HGT seems to give a whole new dimension to this theory. A gene that can hop across genomes appears to be the most selfish of them all, for it is not restricted to perpetuate only in its parental lineage and can maintain its presence in a number of distinct species. This ‘infidelity’ (lateral transfer into another genome) will preserve the gene, even if its source species eventually

It became evident that bacteria were not discrete species living in genetic isolation like the eukaryotes.

Unlike eukaryotic evolution, bacterial speciation might be ‘majorly driven by a high rate of horizontal transfers’. New genes from far away confer new advantageous skills, enabling bacteria to adapt to new environments and hence survive the hostile conditions.



Horizontally transferred genes and their associated 'gene vehicles' act as evolutionary agents in their own right, surviving not only because of being advantageous to the host genome, but because they are especially transferable between genomes.

becomes extinct. Such genes and their associated 'gene vehicles' act as evolutionary agents in their own right, surviving not only because of being advantageous to the host genome, but because they are especially transferable between genomes.

The situation however is not so simple. HGT, in the respect that it allows for the promiscuous spread and persistence of individual genes, does seem to support Dawkins' universe. However, another often overlooked property of HGT strongly argues against the reasoning of the selfish gene and the objectivity of genes in general. In HGT events, the transferred gene is maintained in the host genome only if it confers some useful function to the host. If redundant, the transferred gene will eventually be removed by evolutionary gene loss. Moreover, in order to be functionally useful to the host, it must mutate and undergo the same process of positive Darwinian selection as the vertically transferred genes would. This is evident in cases where the transferred gene may be entirely truncated in the recipient, a notable feature of host-parasite HGT in immune evasion.

Considering the above argument, the selfish gene does not seem to be so accurate a notion. A gene may be accepted only if it is of use to the host and for this it must allow itself to be altered without compromise, according to the needs of the host. The organism is still the object of necessity in evolution, not the genetic material. Everything that happens at the genetic level is allowed by and/or is caused by the organism itself and for its own benefit. Those working on molecular evolution often quote Theodosius Dobzhansky who said, "Nothing makes sense in biology except in the light of evolution". We hold here that nothing makes sense in molecular evolution except in the light of biology.

The organism is still the object of necessity in evolution, not the genetic material.

The Fall of the Trees

Biologists show a greater fervour for phylogenetic trees than do environmentalists for the real ones. Traditionally they have been using tree models, called phylogenetic trees to depict



evolutionary relationships between different species. (See *Figure 1a*). Today, genome sequence data is used to infer such relationships. The logic is that, since an organism acquires all its genes vertically from its parents, two species showing a similar gene sequence must have got them from the same parent. Thus, genomes are compared for the degree of similarity between different genes and the information generated is then used to build family trees. (See *Figure 1b*).

However, there are many who object to this approach of tree building. The gene being compared for making a family tree need not necessarily have come from a common parent. Now we know that it might as well have been picked up from a stranger organism. The possibility that different genes in the same organism might have different parentage makes it seem unreasonable to equate the history of a gene to that of the organism. Gene swapping has thus blurred the very notion of a universal tree of life by putting in horizontal bars connecting its vertical branches. This prompted some scientists to dismiss all gene based evolutionary trees. W F Doolittle of the University of Dalhousie, Canada offers an alternative – “A single universal phylogenetic tree is not the best way to depict relationships between all living and extinct species. Instead a web or net-like model (See *Figure 1c*) with diverging and intersecting branches reflecting the importance of HGT might provide a more appropriate visual metaphor”.

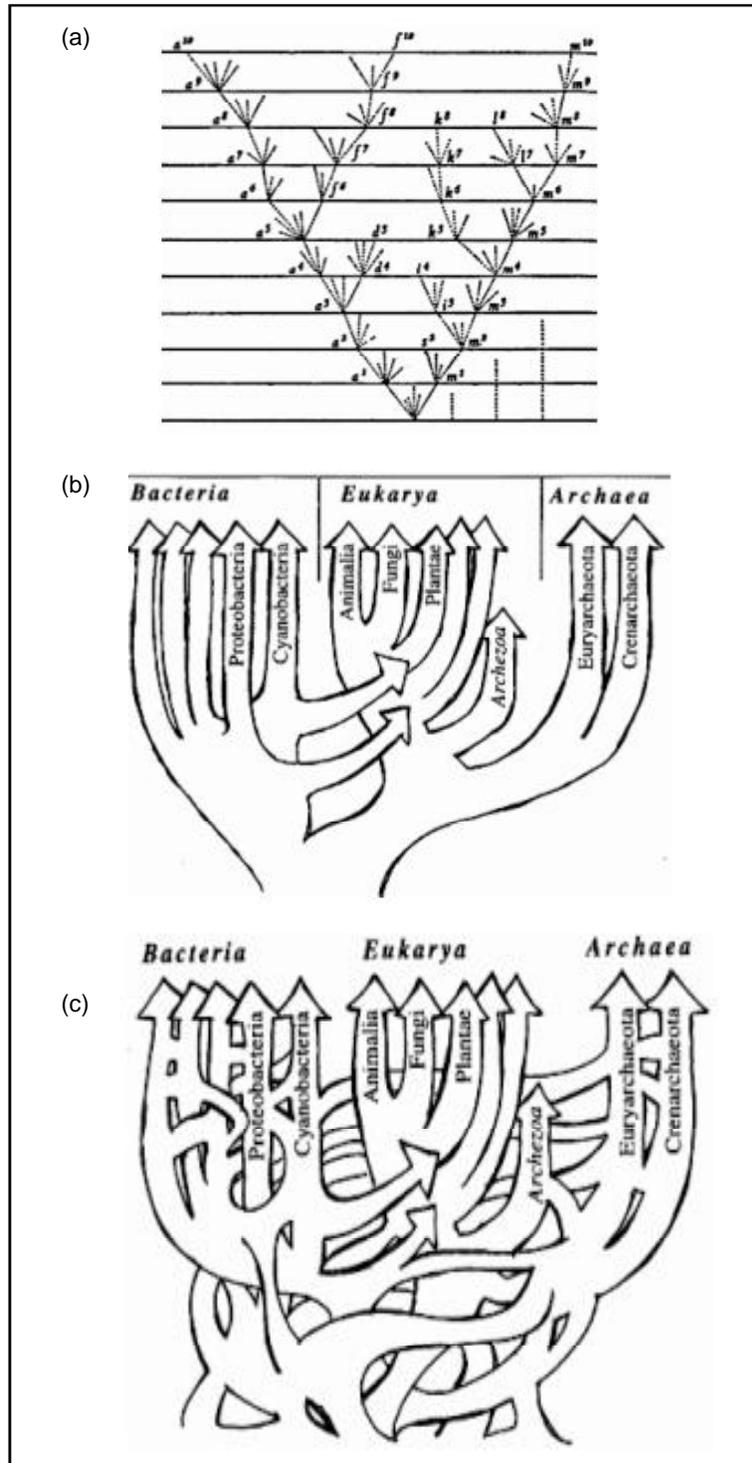
Box 2. Bugs in Our Genome

We ourselves have not been left untouched. A report in the much-acclaimed Human Genome Draft published in *Nature* suggests that we possess in our genome a striking 223 bacterial genes, which were acquired laterally from bacteria during the course of vertebrate evolution. These genes are widespread among bacteria but among eukaryotes are present only in vertebrates. It is possible that these genes were originally present in all lineages and got selectively lost in eukaryotic lineages other than vertebrates. However, their absence through genomes as distant as that of yeast, worms and flies makes this explanation less plausible. It is more likely that these genes entered the vertebrate lineage by horizontal transfer from bacteria at some point during primate, vertebrate, chordate, or even metazoan evolution, so a common ancestor of human beings and related vertebrates could have acquired new genes this way.



Figure 1. Transformation of phylogenetic trees. (a) A phylogenetic tree based on the morphological characters of different organisms. The above figure is a part of the only graphical figure in the *Origin of Species!* Darwin first used it to represent the divergence of variants within a species, showing successively more difference in a single lineage (a1 through a10) and splitting into multiple lineages (m, s, i, and so forth). (b) A modern phylogenetic tree based on comparison of rRNA gene sequences assuming only vertical transfer. It represents the relationships between the three main domains of life viz. Archaea, Bacteria and Eukaryotes. (c) The recently proposed phylogenetic-web model in which vertical branches of a gene-based tree, get connected by lateral transfer events.

Modified from: Doolittle 1999 (Figures 1-3).



Drama, Fiction and the Hazards of Genetic Engineering

The recent BBC drama, *Fields of Gold*, which portrayed the health risk of GM (genetically modified) food, suggested that the antibiotic resistance genes could jump across laterally from GM food to different microbes, leading to the outbreak of a 'superbug'. This is another aspect of HGT, as a sensationalist topic for the tabloids and an outcry for the environmental activists pointing out the hazards of GM food. HGT has even entered science fiction books as the stuff of doomsday prophecies.

Genetic engineering often involves the direct transfer of genes from one organism into another, usually of unrelated species. Known as transgenic technology, this is nothing but HGT aided by human intervention rather than nature. In nature however, gene transfer across species is limited by a number of barriers, which is why natural gene swappings are of such limited extent. In order to overcome the natural limitations, genetic engineers have made a huge variety of artificial gene vehicles designed to cross all such barriers. Opponents of GM foods suspect that once introduced in a crop, these vehicles may well jump out again and reinsert at another site or to the genome of another organism. This could theoretically lead to catastrophic gene combinations that neither existed in nature nor were desired by man. Consequences might include creation of novel infectious viruses or spread of antibiotic resistance genes to bacteria making their infections untreatable. Mae-Won Ho of the Institute of Science in Society, UK cautions, "There is an urgent need to establish effective regulatory oversight to prevent the escape and release of these dangerous constructs into the environment, and to consider whether some of the most dangerous experiments should be allowed to continue at all".

Although the threat does exist, from an academic scientist's point of view, it is far too exaggerated. Michael Syvanen, co-author of a book on HGT at the University of California, Davis states, "If and when experiments prove that horizontal transfer from GM foods occur, some vectors might indeed create an

The antibiotic resistance genes could jump across laterally from GM food to different microbes, leading to the outbreak of a 'superbug'.



How can the occurrence of gene swapping be reconciled with the fact that discreteness of species has been maintained?

unacceptable risk. However there are no confirmed studies yet, of this type of gene transmission occurring in field conditions". The US Food and Drug Administration (FDA) also maintains that HGT from plants to bacteria does not look feasible and for now has dismissed this as an evidence against GM foods, though the FDA trials are on. The benefits of GM food need to be cautiously weighed against the risks of unregulated gene transfer. They are both too important for either to be ignored.

Reality Check

As with any other scientific idea in its infancy, HGT has its share of doubts and ambiguities. *For example, how can the occurrence of gene swapping be reconciled with the fact that discreteness of species has been maintained?* The answer lies in classification of genes as either informational or operational. There is a current proposal that each organism has a core of informational genes that are responsible for critical functions and are not transferred laterally. Such genes undergo only traditional evolutionary processes and hence maintain the identity of the species. On the other hand, operational genes, those involved in housekeeping activities, are known to undergo rampant HGT. They confer new adaptive functions to the organism but dilute the species discreteness only to a small extent.

A laterally acquired gene can be compared to a new actor who suddenly hops onto the stage (host genome), while the play is on, and has to stay and find a role for himself in it.

What types of genes can swap successfully? James A Lake of the University of California, Los Angeles and his colleagues proposed an answer, which they call as 'the complexity hypotheses'. Genes whose products interact with many other molecules are less likely to function successfully in a foreign host cell than are genes whose products have limited contact with other molecules. Since a transferred gene will be maintained in the host genome only if it can confer some useful function to the host, the second type of genes has a better chance at being swapped successfully. The complexity hypothesis can be easily simplified with an analogy. Proteins in a cell work in concert with each other in what is akin to a multi-character play. A laterally acquired gene can be compared to a new actor who suddenly



hops onto the stage (host genome), while the play is on, and has to stay and find a role for himself in it. The lesser the number of characters in the play he has to learn to deal with, the easier it is for him to adjust and stay (be maintained within the genome). Otherwise, he risks getting kicked off the stage by the stage-master (evolutionary gene loss) for his redundancy (not contributing to any of the cellular functions). This is why the informational genes, that code for RNA's and regulatory proteins, which have to interact with a large number of other proteins to function, are less prone to lateral transfer.

Towards a New Synthesis

The variations generated by processes operating solely on parental genes seem far too slow in giving new functions to the organism. They cannot solely explain the extraordinary ability of life forms to adapt especially in rapidly changing environments. On the other hand, new genes from far away can impart new tempo and new modes in evolution in just one go, enabling the organism to respond quickly to environmental changes.

A new evolutionary model is hence required, especially for prokaryotes, in which both dimensions of genetic acquisition concurrently drive evolution. In this model, natural selection, vertical transfer of genes, gene duplication and exon-shuffling (shuffling of the coding segments within the gene), remain the major driving forces producing bulk evolutionary currency. Yet there is enough flexibility to incorporate inputs and outcomes of HGT as and when they arise. The new understanding will require us to redefine many evolutionary dogmas viz. parental model of gene transmission, species boundaries among prokaryotes and phylogenetic trees, which in turn throw up a huge creative challenge to rewrite the story of evolution, laterally or otherwise!

Suggested Reading

- [1] H Ochman, J G Lawrence, E A Groisman, Lateral Gene Transfer and the Nature of Bacterial Innovation, *Nature*, Vol.405, pp.299–304, 2000.
- [2] W F Doolittle, Lateral Genomics, *Trends in Genetics*, Vol.15, pp.M5–M8, 1999.
- [3] W F Doolittle, Phylogenetic Classification and the Universal Tree, *Science*, Vol.284, pp.2124–2128, 1999.
- [4] J Davies, Inactivation of Antibiotics and the Dissemination of Resistance Genes, *Science*, Vol.264, pp.375–382, 1994.
- [5] S Sonea, A bacterial way of life, *Nature*, Vol.331, pp.216, 1998.
- [6] F Bushman, *Lateral DNA Transfer: Mechanisms and Consequences*, 1st Ed., Cold Spring Harbor Laboratory Press, New York, 2002.
- [7] M Syvanen and C I Kado, *Horizontal Gene Transfer*, 2nd Edition, Academic Press, San Diego, 2002.
- [8] A K Bachhawat, Comparative Genomics, *Resonance*, Vol.11, No.8, pp.22–40, 2006.

Address for Correspondence

Saurabh Dhawan
Christoph-Probst-Strasse
16/103, 80805 Munich
Germany
Email: srdhawan@gmail.com

Tomas John Ryan
Email: tjr37@cam.ac.uk

