ONLINE JOURNAL OF BIOTECHNOLOGY RESEARCH

Editor: Dr.G. Kumar

ISSN: 0975-1734

visit:http://www.mmpublication.com Volume 1 Issue 4 May 2010

MM PUBLICATION, C/O Bharathan Flims,#35, Thanjavur Main Road, Varaganeri, Tiruchirappalli, Tamilnadu,India 620008
editorial@mmpublication.com

ISSN: 0975-1734

Contents lists available at MMPublication

ONLINE JOURNAL OF BIOTECHNOLOGY RESEARCH

journal homepage:http://www.mmpublication.com

Horizontal Gene Transfer-a Genomic Perspective

Aniruddha Chatterjee[1]* and Archana Pan[2]

1. School of Biotechnology, Chemical & Biomedical Engg., VIT University, Vellore-632014, TN, India

2. Assistant Professor, Centre of Bioinformatics, School of Life sciences, Pondicherry Central University, puducherry-14

Articleinfo

Article history:

Received 15 October 2009
Accepted 15 November 2009

Keywords:
HGT; LGT; Evolution; Genome; Phylogenetic; Genome-wide; Chromosome; Nucleotide ;Signature; Phylogenetic tree:

Abstract

Horizontal gene transfer is a phenomenon by which one organism transfers its genetic material to other organism but not to its offspring. Transformation, Transduction and bacterial conjugation is the common method by which gene gets laterally transferred. Horizontal gene transfer is a frequent event between microbes especially bacteria and the transferred genes provides selective benefit to the recipient organism. Horizontal gene transfer also plays a key role in determining the evolutionary history of the organism where the extent of lateral transfer is high. Single bioinformatics or molecular biology tool is incapable of systematically identifying all laterally transferred genes in a single organism. The integration of different available methods can bring out the best possible results.In this article we aimed to review the phenomenon of Horizontal gene transfer as a whole. We also set sights on the methods of analyzing the event *insillico* and overcome the problems associated with it.

INTRODUCTION

Horizontal gene transfer (HGT), also known as Lateral gene transfer (LGT), is a process in which an organism transfers genetic material to another cell that is not its offspring. By contrast, *v*ertical transfer occurs when an organism receives genetic material from its ancestor, e.g. its parent or a species from which it evolved. Most studies on microbial evolution focused primarily on the vertical transmission of genetic information, but there is a recent awareness that horizontal gene transfer is a significant phenomenon in bacterial evolution. Horizontal gene transfer is common among bacteria, even very distantly-related ones. This process is thought to be a significant cause of increased drug resistance; when one bacterial cell acquires resistance, it can quickly transfer the resistance genes to many species. Enteric bacteria appear to exchange genetic material with each other within the gut in which they live. In this article we have discussed the probable mechanisms of Horizontal gene transfer and its implications and also revisited the powerful tools of identifying the phenomenon at genomic level.

Mechanisms for horizontal gene transfer

Transformation

The genetic alteration of a cell resulting from the introduction, uptake and expression of foreign genetic material (DNA or RNA) is called as transformation. This process is relatively common in bacteria, but less common in eukaryotes. The process is known as transformation (Lorencz and Wackernagel, 1994).This process is relatively common in bacteria, but less common in eukaryotes. Transformation is often used to insert novel genes into bacteria for experiments, or for industrial or medical applications. Earlier scientists have opined that the genes would be broken down by the DNA-digesting enzymes in the environment. But recently it has been found that the DNA is protected from enzyme attacks when adsorbed to solid particles. In this way genes from genetically engineered organisms may be transferred to bacteria and other microorganisms. The sources may be:Debris ploughed back from genetically engineered crops. The dead or dying cells are likely to release naked DNA that may survive for many hours. An experiment showed that when adhered to clay soil particles, the DNA survived at an average about 28 hours. Dead cells from the feces or other excretions of genetically engineered fish may release DNA. Naked DNA may survive on the ocean surface for 45-83 hours and for up to 235 hours in the bottom sediment. Feces from genetically engineered domestic animals contain myriads of dead and even alive cells with DNA from the animal. Some DNA may have already been transferred bacteria when in the gut.

Transduction

It is the process in which bacterial DNA is moved from one bacterium to another by a bacterial virus (a bacteriophage, commonly called a phage). There is a class of viruses, called bacteriophages that are specialized at infecting bacteria. They can be described as a package of DNA surrounded by a protein coat. This coating has special properties that make it stick on to bacterial cell walls. When attached, they inject the DNA into the bacteria. Viral DNA has an ability to force the invaded cell to switch over from normal activity to producing millions of copies of the whole virus. Or the virus may insert its DNA into the chromosome of the bacteria. When new viruses are produced in the bacteria, a piece of bacterial DNA may become packaged in the virus coat. The virus will thereby carry this bacterial DNA to other bacteria where it may become a part of their chromosome. This is called transduction. Mostly the bacteriophages attach only to a few kinds of bacteria. But they have a tendency to broaden their host range. Bacteriophage transduction is especially common in water.

Bacterial conjugation

It is a process in which a living bacterial cell transfers genetic material through cell-to-cell contact. It has long been believed that mating or conjugation only occurs between members of the same bacterial species. This has recently been contradicted by research that reported a high degree of cross-species mating (Clewell, 1993). Mating is mediated by a special small piece of DNA, the

* Corresponding author. Address:School of Biotechnology, Chemical & Biomedical Engg., VIT University, Vellore-632014, TN, INDIA. E-mail address:raj_success123@yahoo.com (Aniruddha Chatterjee*).

plasmid that is separate from the bacterial chromosome. Most plasmids are also carriers of antibiotic resistance genes. The plasmids are transferred at mating.Also kinds of 'jumping genes' are transferred separately at mating. They can jump between different positions in the chromosome or between chromosomes and plasmids. Among these jumping genes there are the genes that support efficient transfer of antibiotic resistance genes.

Horizontal gene transfer: the milestones

Before the advent of the genomics era, only a handful of horizontal gene transfer events were documented in the literature (Smith *et al*., 1992). And even though it had been argued that gene acquisition from foreign species could potentially have a great impact on evolution (Syvanen, 1985), it was not until after the genomic sequences of numerous prokaryotic and eukaryotic organisms became publicly available that the traditional tree-based evolutionary model was seriously challenged, considering even the possibility of substantial gene exchange (Doolittle, 1999; Groisman *et al*., 1992; Medigue *et al*.,1991). In particular, it was first observed that some *Escherichia coli* genes exhibit codon frequencies that deviate significantly from those of the majority of its genes (Aravind *et al*., 1998). Also, the genomes of *Aquifex aeolicus* and *Thermotoga maritima*, two hyperthermophilic bacteria, supported the hypothesis of a massive gene transfer from archaeal organisms with which they shared the same lifestyle (Nelson *et al*., 1999).

Genes susceptible to horizontal transfer

Though considerable HGT seem to have occurred within and across prokaryotes and eukaryotes, not all genes are equally likely to be transferred. For HGT to be successful the acquired genes must persist in the host chromosome; the gene would persist only if it provided a selective benefit to the recipient organism (Lawrence, 1999). Essential genes, ubiquitously present in all organisms, like those encoding rRNA operons, are therefore less likely to undergo successful transfer as genomes naïve to their functions are rare. Horizontal transfer of genes under weak or transient selection, in contrast, could greatly benefit lineages that do not contain any functional orthologs of these genes (Lawrence, 1999). In bacteria, the genes whose products catalyze steps in a single pathway and provide for weakly selected functions often tend to assemble into operons (even when their products do not interact physically), probably to facilitate horizontal transfer of such genes into naïve genomes (Lawrence,1996). Upon introgression, an operon provides a novel metabolic function allowing its new host to exploit effectively a novel ecological niche (Koonin *et al*., 1997; Lawrence and Roth, 1998). The physical proximity of the genes in an operon, however, provides no selective benefit to the individual organism it rather enhances the fitness of the operon itself by allowing efficient horizontal transfer of its constituent genes, otherwise susceptible to loss by genetic drift. The operon organization is therefore considered to be a 'selfish' property of the constituent genes (Koonin *et al*., 1997). According to the complexity hypothesis (Jain *et al*., 1999), genes participating in transcription, translation and related process (informational) genes are seldom transferred because they are, in general, members of large, complex systems, products of which are less likely to function successfully in a foreign cytoplasm. Operational (housekeeping) genes, on the other hand, are typically members of small assemblies of few genes and hence can undergo horizontal transfer more frequently.

The extent and rate of transfer in bacteria

Availability of complete genome sequences of diverse bacterial species and the ability to recognize horizontally transferred genes solely on the basis of their intrinsic sequence characteristics allows one to estimate the total amount of putative foreign genes in such genomes without resorting to database searching or phylogenetic analyses. Analysis of 19 complete genomes (Ochman *et al*.,2000) has revealed that the amount of laterally acquired sequences vary widely among bacteria ranging from virtually none in certain bacteria with small genome sizes such as *Rickettsia prowazekii*, *Borrelia burgdoferi* and *Mycoplasm genitalium*, to nearly 17% in *Synechocystis* PCC6803 and 18% in *E. coli*. It is likely that the intracellular habitat of some bacterial organisms such as *R. prowazekii* and *M. genitalium* shields these organisms from exposure to potential gene donors, thereby reducing the possibility of acquiring foreign sequences. To assess the rate of effective horizontal transfer in any organism, one must estimate the amount of foreign DNA segments persisting in the host chromosome and maintaining their protein coding potential (maintenance of coding potential implies that nonsense mutations in these genes have been eliminated by natural selection). The persistence of a foreign gene in a bacterial genome can be estimated by exploring the patterns of nucleotide substitutions that have taken place since its arrival (Lawrence and Roth, 1998). At the time of introduction, laterally transferred genes bear the sequence characteristics of the donor genome. But once acquired, they come under the influence of directional mutation pressure appropriate to the recipient genome and eventually start to reflect the DNA composition of their new host.

Frequency of this gene transfer

Horizontal transfer of genes appears to be a common practice between microbes, probably due to the important role of such transfer in shaping the architecture of microbial genomes, conferring novel metabolic capabilities to the recipient genome and enabling an organism to explore new ecological niches. The impact of interspecies gene transfer is radically different from that of the spontaneous mutation. Point mutations can lead only to the subtle refinement and alteration of existing metabolic functions, but horizontal gene transfer has the capability of introducing, immediately upon integration, completely novel physiological traits.

Amelioration of HGT

At the time of introduction, horizontally transferred genes have the base composition and codon usage pattern of the donor genome. But because transferred genes are subject to those mutational processes affecting the recipient genome, the acquired sequences will incur substitutions and eventually come to reflect the DNA composition of the new genome. This process of "amelioration"whereby a sequence adjusts to the base composition and codon usage of the resident genomeis a function of the relative rate of G/C to A/T mutations. Based on substitution rates estimated for *E. coli* and the mutational bias of this species, it is possible to predict the amount of time required after transfer for a transferred gene to fully resemble native DNA (Lawrence and Ochman, 1997).

Impact of horizontally transferred genes

The significance of horizontal gene transfer goes beyond helping interpret phylogenetic congruencies in the evolutionary history of genes. In fact, there is strong evidence that pathogenic bacteria can develop multi-drug resistance simply by acquiring antibiotic resistance genes from other bacteria (Gophna *et al*., 2006; Ortutay *et al*., 2003).

Tools for identifying horizontally transferred genes at the genomic level

There is currently no single bioinformatics tool capable of systematically identifying all laterally acquired genes in an entire genome (Sheila and Terry, 2007). Available methods for identifying horizontal transfer generally rely on finding anomalies in either nucleotide composition or phylogenetic relationships with orthologous proteins. Nucleotide content and phylogenetic relatedness methods have the advantage of being independent of each other, but often give completely different results. There is no 'gold standard' to determine which, if either is correct, but it has been suggested that different methodologies may be detecting lateral transfer events of different relative ages (Lawrence and Ochman, 2002). A variety of strategies have been used to predict horizontal gene transfer using nucleotide composition of coding sequences. Early methods flagged genes with atypical G + C content; later methods evaluate codon usage patterns as predictors of horizontal transfer (Lawrence and Ochman, 1997; Sharp and Li, 1987). A variety of so-called 'genomic signature' models have been proposed, using nucleotide patterns of varying lengths and codon position. These models have been analyzed individually and in various combinations, using sliding windows, Bayesian classifiers, Markov models, and support vector machines (Dalevi *et al*., 2006).One limitation of nucleotide signature methods is that they can suggest that a particular gene is atypical, but provide no information as to where it might have originated. To discover this information, and to verify the validity of positive candidates, signature-based methods rely on subsequent validation by phylogenetic methods. These crosschecks have revealed many clear examples of both false positive and false negative predictions in the literature

The fundamental source of error in predictions based on genomic signature methods is the assumption that a single, unique pattern can be applied to an organism's entire genome (Gophna *et al*., 2006). This assumption fails in cases where individual proteins require specialized, atypical amino acid sequences to support their biological function, causing their nucleotide composition to deviate substantially from the 'average' consensus for a particular organism. Ribosomal proteins, a well known example of this situation, must often be manually removed from lists of horizontal transfer candidates generated by nucleotide-based identification methods (Ortutay *et al*., 2003).The assumption of genomic uniformity is also incorrect in the case of eukaryotes that have historically acquired a large number of sequences through horizontal transfer from an internal symbiont, or an organelle like mitochondrion or chloroplast. For example, the number of genes believed to have migrated from chloroplast to nucleus represents a substantial portion of the typical plant genome (Guindon and Perriere, 2001) Phylogenetic methods seek to identify horizontal transfer candidates by comparison to a baseline phylogenetic tree (or set of trees) for the host organism. Baseline trees are usually constructed using ribosomal RNA and/or a set of well-conserved, well-characterized protein sequences (Tsirigos and Rigoutsos, 2005). Each potential horizontal transfer candidate protein is then evaluated by building a new phylogenetic tree, based on its individual sequence, and comparing this tree to the overall baseline for the organism. More recently, a number of automated tree building methods have used statistical approaches to identify trees for individual genes that do not fit a consensus tree profile (Martin, 2003).

CONCLUSION:

Although phylogenetic trees are generally considered the best available technique for determining the occurrence and direction of horizontal transfer, they have a number of known limitations. Analysts must choose appropriate algorithms, out groups, and computational parameters to adjust for variability in evolutionary distance and mutation rates for individual data sets. Results may be inconclusive unless a sufficient number and diversity of orthologous sequences are available for the test sequence. In some cases, a single set of input data may support multiple different tree topologies, with no one solution clearly superior to the others. Building trees is especially challenging in cases where the component sequences are derived from organisms at widely varying evolutionary distances.Perhaps the biggest drawback to using tree-based methods for identifying horizontal transfer candidates is that these methods are very computationally expensive and time consuming; it is currently impractical to perform them on large numbers of genomes, or to update results frequently as new information is added to underlying sequence databases. Even a relatively small prokaryotic genome requires building and analyzing thousands of individual phylogenetic trees. To manage this computational complexity, many authors exploring horizontal transfer events have been forced to limit their calculations to one or a few candidate sequences at a time.More recently, semi-automated methods have become available for building multiple phylogenetic trees at once (MacLeod *et al*., 2005). These methods are suitable for application to whole genomes, and include screening routines to identify trees containing potential horizontal transfer candidates. However, to achieve reasonable sensitivity without an unacceptable false positive rate, these methods still require each candidate tree identified by the automated screening process to be manually evaluated. One notable exception has been the work of Koonin *et al*, who searched for horizontal transfer in 31 bacterial and archaeal genomes by a combination of BLAST searches with semi-automated and manual screening techniques.A new approach to rapid, genome-wide identification and ranking of horizontal transfer candidate proteins is also recently presented by the name of Dark horse algorithm (Sheila and Terry, 2007). The method is quantitative, reproducible, and computationally undemanding. It can be combined with genomic signature and/or phylogenetic tree-building procedures to improve accuracy and efficiency.

REFERENCE

Aravind,L., Tatusov,R.L, Wolf,Y.I, Walker,D.R. and Koonin,E.V. (1998) Trends Genet., 14, 442444

Dalevi, D., Dubhashi, D., Hermansson, M.,(2006) Bioinformatics, 22:517-522

Doolittle,W.F. (1999) Science, 284, 212429

Ed Clewell DB, Bacterial Conjugation, (1993), Plenum Press New York

Gophna, U,, Charlebois, R,L,, Doolittle, W,F,, (2006) Trends Microbiol, 14:64-69

Groisman,E.A., Saier,M.H.,Jr and Ochman,H. (1992) EMBO J., 11, 13091316

Guindon, S,, Perriere, G,, (2001) Mol Biol Evol, 18:1838-1840

Jain, R,, Rivera, M, C, and Lake, J ,A ,1999 Proc. Natl. Acad. Sci. USA 96 38013806

Koonin,E.V., Mushegian,A.R., Galperin,M.Y. and Walker,D.R. (1997) Mol. Microbiol., 25, 619637

Koski, L.,B., Morton, RA., Golding GB,(2001) Mol Biol Evol, 18:404-412

Lawrence, J. G. & Roth, J. R. (1996) Genetics 143, 18431860

Lawrence, J. G. & Ochman, H. (1997) J. Mol. Evol. 44, 383397

Lawrence, J. G and Roth J R (1998) (eds) M Syvanen and C I Kado (London: Chapman and Hall) 208225

Lawrence, J. G (1999), Curr Opin Microbiol., 2(5) 519-23

Lawrence, J. G, Ochman, H., (2002) Trends Microbiol, 10:1-4

Lorencz MG, Wackernagel W., (1994) Microbial Reviews, 58: 563-602

MacLeod D, Charlebois RL, Doolittle F, Bapteste E,(2005) BMC Evol Biol, 5:27

Martin W, (2003) Proc Natl Acad Sci USA, 100:8612-8614

Medigue,C., Rouxel,T., Vigier,P., Henaut,A. and Danchin,A. (1991) J. Mol. Biol., 222, 851856

Nelson,K.E., Clayton,R.A., Gill,S.R., Gwinn,M.L., Dodson,R.J. et al.(1999) Nature, 399, 323329

Ochman H, Lawrence J G and Groisman E A (2000); Nature (London) 405 299304

Ortutay, C., Gaspari, Z., Toth, G., Jager, E., Vida, G., Orosz, L., Vellai, T.,(2003) J Mol Evol, 57:672-680

Sharp PM., Li WH. (1987) Nucleic Acids Res, 15:1281-1295

Sheila Podell and Terry Gaasterland., (2007) Genome Biology, Volume 8, Issue 2, Article R16

Smith,M.W., Feng,D.F. and Doolittle,R.F. (1992) Trends Biochem. Sci., 17, 489493

Syvanen,M., (1985) J. Theor. Biol., 112, 333343

Tsirigos, A., Rigoutsos, I., (2005) Nucleic Acids Res, 33:3699-3707