Indian J. Microbiol. (June 2008) 48:228–242

REVIEW

Phylogeny vs genome reshuffling: horizontal gene transfer

Sadhana Lal · Simrita Cheema · Vipin C. Kalia

Received: 15 January 2008 / Accepted: 6 June 2008

Abstract The evolutionary events in organisms can be tracked to the transfer of genetic material. The inheritance of genetic material among closely related organisms is a slow evolutionary process. On the other hand, the movement of genes among distantly related species can account for rapid evolution. The later process has been quite evident in the appearance of antibiotic resistance genes among human and animal pathogens. Phylogenetic trees based on such genes and those involved in metabolic activities reflect the incongruencies in comparison to the 16S rDNA gene, generally used for taxonomic relationships. Such discrepancies in gene inheritance have been termed as horizontal gene transfer (HGT) events. In the post-genomic era, the explosion of known sequences through large-scale sequencing projects has unraveled the weakness of traditional 16S rDNA gene tree based evolutionary model. Various methods to scrutinize HGT events include atypical composition, abnormal sequence similarity, anomalous phylogenetic distribution, unusual phyletic patterns, etc. Since HGT generates greater genetic diversity, it is likely to increase resource use and ecosystem resilience.

Keywords Evolution · Phylogeny · Horizontal gene transfer

S. Lal · S. Cheema · V. C. Kalia (✉)
Microbial Biotechnology and Genomics;
Institute of Genomics and Integrative Biology (CSIR),
Delhi University Campus, Mall Road,
Delhi - 110 007,
India
e-mail: vckalia@igib.res.in, vc_kalia@yahoo.co.in
Tel: +91 / 11 / 27666156,
Fax: +91 / 11 / 2766 7471; 2741 6489

Introduction

Organismal evolution of living beings occurs largely through transfer of genetic material. This process is primarily instrumental in vertical inheritance among closely related organisms. The creative evolutionary movement of genetic information among remotely related species has been traditionally resisted by biologists. The skepticism around this evolutionary route is largely due to lack of experimental explanations and quite a few cases, which failed on closer scrutiny [1]. The most widely accepted forms of horizontal transfer of genetic information have been the evolution of eukaryotic chloroplasts and mitochondria as bacterial endosymbiotes [2, 3]. Protists (microbial eukaryotes) have existed as parasites with relic genomes [4] and anaerobic protists seem to have acquired hydrogenosomes via multiple endosymbioses [5]. It sparked interest in 1944, when scientists began accumulating experimental evidences on the ability of microbes to uptake 'naked' DNA from their environment and incorporate it in to their genomes [6]. The main question for almost a century has been whether these horizontal movements of genes occur at a rate that significantly influences evolution. In the pre-genomic era, it has been difficult issue to deal [7], however, in the post-genomic era, the subject has been reviewed quite vigorously [1, 5, 8–16]. During this period, more and more candidates for the wonderful natural events of horizontal (Lateral) gene transfer (HGT) have been identified. Horizontal gene flow is being viewed as a competitive model for the origin of microbial diversity [17]. In view of the evidences that most archaeal and bacterial genomes and certain eukaryotes contain genes from multiple sources [9], "chimerism" or HGT cannot be dismissed as a trivial issue. Particularly so since different estimates reflect that

2 to 60% of all prokaryotic genes have been affected by HGT [18, 19]. HGT blur taxonomic boundaries and such cases are impossible to predict [20]. These cause profound disagreements. It does not come as a surprise that it is one of the most controversial topics in evolutionary microbiology [19, 21].

The need for HGT

Under challenging environmental conditions, organisms draw from a pre-existing genetic repertoire, which may not be sufficient to ensure their survival. DNA rearrangements often result in novel combinations of existing functionalities. Most natural mutations stem from mutations in the Mismatch Repair System (MRS), which increases mutation rate and confer a hyper-recombination phenotype [22]. Acquired genes play a major role in bacterial diversification and allow rapid exploitation of new environments [20]. Extensive HGT accelerates niche invasion and genes that confer advantages against host defense mechanism may be shared widely among pathogens [23]. Approximately 250 HGT events within *Synechococcus-Prochlorococcus* clade, and >150 in γ-Proteobacterial genus, *Pseudomonas* alone have been recorded [23]. Among these, Proteobacteria and Low G+C Gram-positive best represent long distance HGT [23].

Many bacterial operons confer highly beneficial functions that may be of long-term use to their new hosts: such functions include biosynthesis of amino-acids, co-factors or other metabolites, the degradation of compounds as carbon or nitrogen sources or transport of metabolites. Simultaneous transfer of all genes is required for such functions to be conferred to the host. Organization of these genes in to operons is primarily beneficial to the constituent genes. It is a selfish property since clustered genes can propagate successfully and their co-regulation would be most beneficial [24]. For example, the *phd*/*doc* system of bacteriophage P1 prevents the host death / death on curing, encoding a long-lived toxin (e.g., Doc) and a short-lived antidote (Phd). Deletion or loss of these genes from the cell will allow toxin to outlive its antidote and cause cell death. Hence, cells are addicted to the presence of antidote. The proximal location of toxin- and antidote-encoding genes is selfish in that it allows the effective co-transfer in to naïve genomes – transfer of one gene without the other is useless [24]. Similar situations also hold good for Type II restriction / modification systems, organization of antibiotic resistance gene [24]. Horizontal transfer of selfish operons most probably promotes bacterial diversification [24]. Transfer of operons en block, however, provides a selective advantage by furnishing new metabolic capacities, resulting in simultaneous generation of several aberrant gene pairs. Although co adapted nature of genes can accelerate the co-transfer of genes contributing to a single function, the Selfish operon model predicts that clusters of functionally related genes can colonize naïve genomes or genomes having lost two genes of a single pathway [25]. The *Salmonella typhimurium cob* operon, the organization of the cobalmin (Coenzyme B_{12}) biosynthetic operon and the propanediol degradation (*pdu*) operon, exemplifies the introduction of a new degradative pathway and the biosynthetic pathway. The adjacent location of the *cob* (cobalmin (Coenzyme B_{12}) biosynthetic) and *pdu* operons demonstrates that an organism by horizontal transfer may gain complex functions [26, 27]. Many such introgressed operons have been identified within the *Escherichia coli* and *S. typhimurium* chromosomes [25]. A remarkable similarity between the structure and sequence of the *trp* (tryptophan biosynthesis) operon in *Brevibacterium lactofermentum* and of enteric bacteria can be accounted for by the process involving loss and reacquisition of genes [28].

In a "catastrophic" situation, it will be beneficial for organisms to share their individual capabilities and survive. This might have forced dissemination of gene cluster (operons) involved in the catabolism of xenobiotics in polluted environments [29–33]. Arsenic resistance in microbes through *arsB* gene (conferring arsenite, As [3+], resistance) seems to have evolved in response to changes in redox conditions, by an event of duplication of *arsA* gene [34] and subsequent HGT events [35]. In key biogeochemical cycles, microbial sulfate reduction pathway has been shown to evolve early in Earth's history [36]. Most living sulphate-reducing microbes belong to bacterial groups, with archaea, *Archaeoglobus* species as exceptions. It seems *Archaeoglobus* species have acquired the gene for dissimilatory sulfite reductase (*dsr*), a key enzyme in sulfate reduction, through HGT [37].

Whole-genome comparisons have uncovered sets of genes that are restricted to organisms that have independently adapted to a common life style such as Archaeal and Bacterial hyperthermophiles [38] or the intracellular pathogens *Rickettsia* and *Chlamydia* [39]. This movement has happened across intergenic boundaries including the Gram-positive and Gram-negative barrier. In fact massive HGT from endosymbiotic bacteria to the nucleus forms the basis of the theory of the origin of eukaryotes [40]. The evolutionary analysis of 31 photosynthetic genes shows that the earliest photosynthetic organisms diverged from a bacterial lineage (Proteobacteria) [41]. HGT and resulting phylogenetic incongruities document the process of gene-transfer-mediated organismal diversification [10]. These characteristics contribute significantly to enhance evolutionary rates, which may be further accelerated by

environmental stimuli. Gene swapping (HGT) is perhaps an important mechanism of biological insurance against sudden geochemical changes and perhaps also shortens time-scale for succession [42].

Agents of HGT

The evolution of microbes occurs mainly through transformation, conjugation and transduction. However, the genetic mechanisms for acquisition and dissemination are enhanced through mobile genetic elements, gene cassettes, transposons, plasmids and bacteriophages [43–45]. The mobilization of resistant determinants has occurred with the aid of transposons, where the gene conferring a selectable phenotype is flanked by two insertion sequences. For example, insertion sequence IS10 flank a tetracycline resistant determinant and a regulatory gene to form transposon Tn10. Similarly, IS50 elements flank an operon that confers resistance to kanamycin, bleomycin and streptomycin to form a transposon Tn5 [46]. These transposons can integrate into the chromosomes of phylogenetically diverse bacterial species. Integrons, on the other hand, incorporate promoterless genes, there by converting them into functional genes [46]. Integrons have been implicated in the acquisition of virulence determinants by the cholera-causing bacterium, *Vibrio cholerae* [47]. Unlike the acquisition of antibiotic resistance, adoption of a pathogenic life style usually involves a fundamental change in a microorganism's ecology. The pathogenic strains of *Streptococcus pneumoniae* [48], *Neisseria meningitidis* [49, 50] and *N. gonorrhoeae* [51] were originally sensitive but have recently acquired resistance to antibiotics. The appearance of multicellular organisms has provided bacteria with new environments, whereby they acquired virulence. Most pathogenicity genes are located in the bacterial chromosome as pathogenicity islands (PAI). These resemble defective bacteriophages and plasmid mediated conjugate systems or compound transposons [52]. The sporadic phylogenetic distribution of the pathogenic organism finds support from the discovery that horizontally acquired PAIs, which are the major contributors to the virulent nature of many pathogenic bacteria [8,52] for example 70 kb PAI-1 of uropathogenic *E. coli* [53], 35 kb LEE island of enteropathogenic *E. coli* [54], 24 kb SH1-3 island of *Shigella flexneri* [55,56] and 17 kb SP1-3 island of *Salmonella enterica* [57].

Detection of Horizontal gene transfer

Evidences for acquisition of genetic material for a character from a donor organism are derived largely from the available sources rather than actual observations. Gene transfer event involves introduction of DNA in a lineage, where the acquired character is limited to the descendants of the recipient and absent from closely related taxa. It results in a restricted phylogenetic distribution. DNA segments gained through HGT often display a scattered phylogenetic distribution among unrelated strains or species [12, 13]. These species-specific DNA regions may show unduly high levels of DNA or protein sequence similarity to genes from otherwise divergent taxa [10]. Most of the methods for detecting HGT include atypical composition, abnormal sequence similarity (i.e. greatest similarity with a gene from distantly related species), anomalous phylogenetic distribution and unusual phyletic patterns. [12, 13, 58, 59].

Atypical composition and codon usage

The genome composition in terms of the bases G and C (G+C content) varies significantly as a function of the position within the codon. The overall G+C content is computed by considering all of the nucleotides in a genome [60]. A comparison of any two homologous DNA regions that display an abrupt change in similarity raises the possibility of a mosaic structure and hence a horizontal transfer [20]. Another criteria for identifying horizontal transfers is to compare codon usage of a gene with the codon bias of the host organism. The codon usage patterns generally differ significantly from organism to organism [61, 62]. Thus, genes whose nucleotide or codon composition are significantly different from the mean for a given genome are considered as probable horizontal acquisitions [60, 63–66]. Codon bias can be used to support more compelling evidences. In *E. coli* and *S. typhimurium* 90% genes are closely related, 10% of the genome in *S. typhimurium* encodes functions not found in *E. coli*. GC contents in these unique genes are frequently significantly lower than the 50% average for the entire genome i.e., remote origins for these regions [67]. The *spa* gene (secretion of Ipa proteins) of *Salmonella* spp. has a GC content of ~30–40%, which is atypical of *S. typhimurium* coding sequences [68]. In addition, its homologues are also not found among taxa closely related to *Salmonella* [68]. The mosaic pattern of GC content among *spa* genes suggest that like the *rfa* and *rfb* operons, (lipopolysaccharide synthesis) the *spa* operon may represent another example of an operon of genes assembled from different chromosomes. *S. typhimurium oad* (oxaloacetate decarboxylase) operon have aberrant GC contents ~65% G+C and 87% G+C in the third codon position. *cat* (chloramphenicol acetyltransferase) operon of *Acinetobacter calcoaceticus* has a mosaic structure indicative of recent assembly from

multiple sources [69]. *rfa* (35–39% GC); *rfb* (31–40% GC); *pdu* (59% GC); *cob* (alamin adenosyltransferase, 59% GC) all show GC contents atypical of the *S. typhimurium* genome`. A large number of *S. enterica* genes, which are absent in other enteric species including *E. coli*, have base compositions that differ significantly from the overall 52% G+C content of the entire chromosome [64, 70]. Within *S. enterica*, certain serovars (with distinct flagellar composition and /or lipopolysaccharide surface antigens) may contain more than a megabase of DNA not present in other serovars, as assessed by a genomic subtraction procedure. The base composition of these serovar-specific sequences suggests that at least half were gained through HGT [71]. The genome of *Thermotaga maritima* has an average G+C content of 46%. Within the rRNA operon (16S-23S-5S), the spacer regions are occupied by isoleucine transfer RNA and alanyl tRNA. These spacer regions have a significantly higher G+C content of 62%. This contrasts with a region of significantly lower G+C content (34%) encoding lipopolysaccharide biosynthesis (LPS) proteins. On the basis of analysis of G+C ratio, G-C skew (G-C/G+C) and asymmetric distribution of oligomers, a characteristic bacterial origin of replication identified in *T. maritima* genome was also observed in the genomes of Archaea *Methanococcus jannaschii* [72] and *Archaeoglobus fulgidus* [73]. Such evidences support HGT between archaea and bacteria [74].

Genome's composition i.e., G+C content was initially proposed as its signature [60]. Due to periodicity of the DNA code, it was implied by the organization of the coding regions into codons. Since G+C content varies significantly as a function of the position within the codon, four discrete G+C content signatures can be identified. The first corresponds to the overall G+C content and is computed by considering all of the nucleotides in a genome. Each of the remaining three signatures are denoted by G+C(k), with k = 1,2,3. Each number corresponds to the value of the G+C content as the latter is determined by considering only those nucleotides occupying the kth position within each codon; unlike the G+C signature which is computed across all genomic positions, only coding regions are used in the computation of G+C(k) [16].

A related variation of the G+C(k) content idea is the Codon Adaptation Index (CAI) [75], where only coding regions are used in its computation, unlike the G+C signature which is computed across all genomic positions [16]. CAI measures the degree of correlation between a given gene's codon usage and the codon usage that is deduced by considering only highly expressed genes from the organism under consideration [76]. Karlin *et al* [77] found that the codon biases observed in ribosomal proteins deviate the most from the biases of the average gene such as in *E. coli*. On the basis of this observation, they defined 'alien' genes as those genes whose codon bias was high in comparison to the one observed in ribosomal proteins and also exceeded a critical level when compared to that of the average gene.

Abnormal sequence similarity

A gene sequence from a particular organism shows the strongest similarity to a homolog from a distant taxon forms the basis for horizontal transfer hypothesis. For example, TM0005 of *T. maritima* (a bacteria) is most similar to archaeal genes: of these, 451 (24%) have best-hits in Archaea. These may be the thermophilic genes. Majority of housekeeping functions are most similar to orthologues in eubacterial species. In contrast, 49% of transporters and 42% of conserved hypothetical proteins are most similar to archaeal genes [74]. Wide variation has been observed in the size and organization of 19 genomes analysed, and the amount of horizontally acquired DNA represented - as those open reading frames (ORFs) whose sequence characteristics depart from the prevalent features of their resident genome - ranges from virtually none in some organisms with small genome sizes, such as *Rickettsia prowazekii, Borrelia burgdorferi* and *Mycoplasma genitalium*, to nearly 17% in *Synechocystis* PCC6803 [60].

Unusual Phyletic Patterns

Phyletic patterns represent the distribution of clusters of orthologous groups (COGs) across genomes [78, 79]. These are useful for tracking the evolutionary events such as vertical gene inheritance, gene loss and horizontal transfer. These can also be used as a post-homology method of predicting protein function and provide clues to the prevailing trends in genome evolution [80]. This was based on the assumption that genes/COGs encoding functionally linked proteins are co-inherited (simultaneously present or simultaneously absent) in the same subsets of genomes. Phyletic patterns are coded as strings of bits, standing for presences or absences of homologs in different genomes. The Hamming distance of 3 bits or less between phyletic patterns is a useful similarity threshold for detecting functionally linked genes [81]. Generally speaking, phyletic patterns are binary vectors in species space and distance between them can be measured in different ways: primarily through a graph in which the number of automatically identifiable, biologically relevant clusters are maximized. These unusual, rare patterns are of particular interest, suggesting the possibility of differential gene loss and horizontal transfer of genes

[78, 79]. Of the COGs represented in all archaea, 16 so far have no members from other domains of life and comprise a unique archaeal genomic signature, whereas 61 are exclusively archaeo-eukaryotic. This archaeo-eukaryotic component shows that the process of evolution has been more complex than simple vertical inheritance and has involved extensive HGT between archaea and bacteria, at least outside the core gene set [82-84]. Phylogenetic reconstruction might be more reliable method for identifying HGTs [85].

Anomalous phylogenetic distribution

Another method that has been used to determine whether HGT has occurred involves inferring evolutionary trees for many genes in many genomes. It takes into account the presence or absence of a gene (or gene family) on a phylogenetic tree. Consistent gene presence in a clade indicates that the corresponding gene was present in the ancestor of that clade, whereas occasional absence of a gene might result from gene loss. A fragmented distribution of a gene family across very distantly related species is indicative of horizontal gene transfer (HGT) events [59]. And these phylogenies of individual genes or proteins are not in complete accord with ribosomal RNA (rRNA) phylogeny [86]. Thus it is important that at least two criteria are matched for constructing phylogenetic trees: 1) establish orthology relationships between genes; and 2) translate gene presence-absence data into a tree structure. These targets can be achieved by various routes: i) defining orthologus as intergenomic best hits (BeTs) and converting it in to intergenomic distances; using these to build hierarchical classification trees [87-89] ii) using BLAST E-values to establish homology for computing intergenomic distances and construct trees [90] iii) using intergenomic FASTA z-scores followed by single linkage clustering to identify orthologous groups, then apply PARSIMONY ANALYSIS to reconstruct trees [91,92]; iv) use clusters of orthologous groups of proteins (COGs) [79] and build either parsimony [93] or least-squares trees [94] which are good for small set of gene sequences [57,95,96]; or apply distance based method, for larger set of gene sequences [97]; v) use the co-efficient of co-occurrence of genomes in COGs for calculating intergenomic distances and construct neighbour-joining trees [98]; vi) compute the ratio of orthologs (identified as reciprocal BeTs) to the number of genes in the smaller genome and construct least-squares trees [99]; vii) build dendograms on the basis of the predicted protein fold composition of genomes [94,100]. Comparisons of 16S rRNA sequences have revolutionized the understanding of the diversity and phylogenetic relationships of all organisms [101]. Therefore, to better understand the extent of mosaicism within species, it is important to compare and contrast the relationships of other molecules with those of 16S rRNA.

Although the trees produced by these different approaches may not be directly comparable, however, several major trends are recognizable. The major factor, which determines tree topology, appears to be the magnitude of gene loss during evolution. The major lineages and plasticity of prokaryotic genomes can be assigned to selective pressure during the adaptation to the environment e.g., parasitism in Proteobacteria. Removal of parasites from the analysis [91,92] or normalization of the intergenomic distances [88,89] results in reasonable phylogenetic reconstructions. Deletion of obviously foreign (horizontally transferred) genes from the analysis strengthens the ribosomal gene trees [102]. Ribosomal gene-based analysis was also substantiated by those constructed on the basis of gene content in the entire genomes [91]. On the other hand, a combination of any of these methods and a proper calculation of evolutionary distances results in correct trees.

Tree of Life

Molecular systematics has become the primary way to determine evolutionary relationships among microorganisms because morphological and other phenotypic characters are either absent or change too rapidly to be useful for phylogenetic inference [103, 104]. Living organisms were categorized as eubacteria and eukaryotes, until Woese [105, 106] gave a new dimension to this concept. According to rRNA trees, the Archaea has two major kingdoms: Crenarchaeota and Euryarchaeota [103]. The kingdom Crenarchaeota generally consists of hyperthermophiles or thermoacidophiles where as the kingdom Euryarchaeota spans a broader ecological range to encompass hyperthermophiles, methanogens, halophiles and even thermophilic methanogens. Archaea have certain metabolic regimes, which differ greatly from those operating in Bacteria and eukaryotes [107–109]. Evolutionary trends derived from single-gene trees of metabolic and biosynthetic enzymes provide no clear support for any particular grouping of domains. On the other hand, gene-trees based on informational enzymes, tend to more solidly support the sisterhood of the Archaea and eukaryotes. Many potential candidates have been chosen to generate phylogenetic trees similar to those generated from 16S rRNA, for example HSP70 (eukaryotic chaperonin), GroEL (bacterial chaperonin), EF-TU (Translation Elongation factor Tu), ATPase-ß-subunit, 23S rRNA, RNA polymerases and RecA (Recombinase A)

[101]. Prokaryotic and eukaryotic evolutionary pattern has been drawn largely on the basis of some highly conserved gene portions of 16S rRNA [100, 101]. The accumulating database of 16S rRNA sequences, which now includes over 471,792 annotated (16S rRNA) sequences, provides an extra incentive to focus on this molecule [110]. Although 16S rRNAs have been stably transferred between species, other genes may not have been similarly transmitted [111]. Incidentally, even ribosomal RNA shows evidence of HGT over certain portions [112]. A combination of cladistic analysis and traditional phylogenetics has the potential of making a convincing case for HGT [113].

Save or abandon the tree

The ribosomal RNA (rRNA) sequence based phylogenetic analysis resulted in a "standard model" of evolutionary (species) tree [103, 104, 114, 115]. However, in the present genome era, the validity of rRNA based species tree is being questioned as incongruencies between topologies and phylogenetic taxonomy are recorded rather frequently [13, 19–21, 38, 58, 66, 86]. Genome comparisons indicate that HGT and differential gene loss are major evolutionary phenomena that involve a large fraction of genes in prokaryotes. The magnitude of these happenings casts doubt on the "Tree of Life'. The present trend of constructing evolutionary tree is more of a genome-scale gene trees rather than a "complete" demonstration of evolutionary process. The ring like representation provides a broad phylogenetic framework for testing theories for the origin and evolution of eukaryotic genomes [15]. This scheme based on analyzing whole genome sequences thus is a radical departure from conventional thinking. It allows HGT to be used in assessing genome based phylogeny and it recovers the connection between prokaryotes and eukaryote genomes [14].

Microbes frequently involved in HGT

Extent of HGT

The availability of complete genomic sequences provides an opportunity to quantify the amount of horizontally transferred sequences among diverse microbes, which varies from virtually none in some organisms (with small genome size) such as *R. prowazekii*, *B. burgdoferi* and *Mycoplasma genitalium* to nearly 17% in *Synechocystis* PCC6803 genome [60]. Recent evidences suggest that HGT is a major and continuing force in archaeal and bacterial evolution e.g., 18% ORFs in *E. coli* genome [116] were introduced since this species diverged from the *Salmonella* genomes supports HGT. *Methanococcus jannaschii* has informational genes for translation, transcription, replication and protein secretion being much more similar to eukaroytes than bacteria. The operational genes, responsible for the eukaryote on the other hand were most closely related to their counterparts found in *E. coli* and *Synechocystis* [117].

Methanogen, *A. fulgidus* [73] bears many genes for fatty acid metabolism that are unknown in other archaeal genomes. Tryptophan biosynthesis pathway *A. fulgidus* seems very closely related to eubacterium *Bacillus subtilis* even though large distances separate these two on the 16S rRNA tree. The hyperthermophilic Eubacteria *Aquifex aeolicus* and *T. maritima* each contain a large number of genes that are most similar in their protein sequences and in some cases, in their arrangements, to homologues in thermophilic archaea. Twenty four per cent of *Thermotaga*'s 1,877 ORFs [74] and about 16% of *Aquifex*'s 1,512 ORFs display their highest match to an Archaeal protein. On the other hand, mesophilic organisms such as *E. coli*, *B. subtilis* and *Synechocystis* have much lower proportion of genes that are most similar to Archaeal homologues [38]. Significant transfer of genes from plants to bacterium, *Deinococcus radiodurans* [118] illustrates the scope of HGT. Table 1 illustrates certain microbes that are frequently involved in HGT.

Transcription and regulatory activities

DNA dependent RNA polymerases (DdRp) that catalyze gene transcription in all cells are complex enzymes which consist of 5–15 subunits. Four of these subunits (α, β′ and ω) constitute the structural core and is conserved in all cells. Comparative analysis of the domain architectures of β-β′ subunits (which form active sites) and σ [70] of DdRp and their phylogenetic analysis revealed a fundamental split among bacteria. A striking deviation was observed in *Aquifex*, whose DdRp clustered with Proteobacteria, Chlamydia, Spirochaetes, *Cytophaga-Chlorobium* and *Planctomycetes* where as its ribosomal proteins group with *Thermotoga*. Although evolution of DdRp appeared to be generally dominated by vertical inheritance, HGT of some or all of the subunits, might have played a role in displacement of ancestral genes in several lineages, such as *Aquifex*, *Thermotoga* and *Fusobacterium* [12, 119].

Phosphofructokinase (PFK) a key regulator enzyme of glycolysis has an intricate evolutionary history. Amino acids essential for ATP-PFK and PPi-phosphofructokinase (PFP) at 104 and 124 position [120]. Several duplication and HGT flux the phylogeny of PFK [121]. Conservation

Table 1 Organisms involved in horizontal gene transfer

Organisms / Phylogenetic similarity	Character acquired through HGT	References
Anabaena sp., *Synechocystis* sp., *Thermosynechococcus elongatus* and Eukaryota.	Phosphoglycerate kinase and phosphopyruvate hydratase (Carbohydrate metabolism)	[158]
Aquifex aeolicus and *Thermotoga maritima*	Large number of genes that are most similar in their protein	[117]
Archaeoglobus fulgidus	Fatty acid metabolism; Tryptophan biosynthesis pathway closely related to eubacterium *Bacillus subtilis*	[73, 117]
Bacillus subtilis and *B. halodurans*	Prophage like regions	[117]
Bacteria and fungi	ß-glucuronidase (*gus*) genes	[159]
Bacteria and archaea	*Fts*Z, a cell division protein	[160, 161]
Bacteria and rumen Ciliates	Expressed genes	[162]
Chlamydia trachomatis	Polymorphic membrane protein C gene	[163]
Clostridium acetobutylicum	3-Isopropylmalate dehydratase, small subunit 3-Isopropylmalate dehydratase, large subunit 2-Isopropylmalate synthase	[164]
Cyanobacteria and α-proteobacteria (*Rickettsia* and *Ehrlichia*) are nearest to chloroplast and mitochondria	HSP60, a major bacterial antigenic protein	[165, 166]
Cyanobacteria and γ-proteobacteria	ArsC gene (arsenate reductase)	[167]
Cyanobacteria and chloroplast genome of *Euglena myxocylindracea*	*psbA* intron	[168]
Deinococcus radiodurans	dTDP-4- dehydrorhamnose epimerase (Lipopolysaccharide biosynthesis)	[118, 164]
Escherichia coli	70 kb PAI-1 and 35 kb LEE island	[57, 58]
E. coli, B. halodurans, and archaea	Subunits of acyl-CoA: acetate CoA transferase	[113]
Eukaryotes and archaea	Large RNAP subunits	[169, 170]
Eukaryotes and bacteria	Mg^{2+}- or Mn^{2+}-dependent protein phosphatases (PPMs)	[171–173]
Flowering plants and *Gnetum*	Mitochondrial nad1 intron 2 and adjacent exons b and c	[174]
Haloarcula vallismortis and eukaryotes, bacterial homologs	Isoforms of GAPDH	[175]
Methanobacterium thermoformicium and *Neisseria gonorhoeae*	Type II restriction enzyme systems	[176]
Methanococcus jannaschii	Informational genes resistance to eukaryotes than bacteria. Operational genes closely related to *E. coli* and *Synechocystis*.	[117]
Mycoplasma and ε-proteobacteria	RuvB clusters	[164]
Neisseria meningitidis and *N. gonorrhoeae*	Resistance to antibiotics	[49–51]
Plant-like host organism and *Chlamydia*	High proportion of proteins	[177]
Rickettsia conorii and *R. prowazekii*	Preprotein translocase subunit SecE (Ribosomal proteins); Undecaprenyl pyrophosphate Synthase (Lipid metabolism)	[164]
Salmonella enterica	17 kb SP1-3 island and Unique genes	[57, 165]
Shigella flexneri	24 kb SH1-3 island	[55, 56]
Streptococcus pneumoniae	Resistance to antibiotics	[48]
Sulfolobus and eukaryotes	Amido- synthetase domains (*carB* genes)	[178]
Ureaplasma urealyticum	F0F1-type ATPase (α-subunit, δ-subunit, ε-subunit)	[164]
Vibrio cholerae	Virulence determinants	[47]
Bacillus cereus ATCC14579 (Firmicutes) and *Chlorogloeopsis fritschii* (Cyanobacteria)	*pha*C gene involved in polyhydroxyalkanoate biosynthesis	[147]

Desulfitobacterium hafniense (Firmicutes) and *N. meningitidis* MC58 (β-Proteobacteria)	*pha*B gene involved in polyhydroxyalkanoate biosynthesis	[147]
Sinorhizobium meliloti (α-Proteobacteria) and *Z. ramigera* (β-Proteobacteria)	*pha*A gene involved in polyhydroxyalkanoate biosynthesis	[147]

of 2 distally related PFK enzymes in a single species could be explained in terms of adaptability, each copy allowing the use of either phospho- donor or the other, potentially enhancing the fitness of the species [120].

RecX is a bacterial regulatory protein which acts as an anti-recombinase to quell inappropriate recombinational repair during normal DNA metabolism [122, 123]. It is found in bacteria and some plants (*Arabidopsis thaliana* & *Oryza sativa*) but not in archaea. Plant Rec-X like protein was found to have evolved by a gene fusion event between N-terminal domain of unknown origin and RecX domain within plant cells. Gene fusion events probably occurred in plant cells [124].

Microtubules of the eukaryotic cytoskeleton are built by αβ-tublin heterodimers. These are thought to be unique to eukaryotes, however their homologues FtsZ can be found in bacteria. BtubA and BtubB of free living bacterium, *Prosthecobacter*, shows higher sequence homology to eukaryotic tubulin than to FtsZ and were perhaps transferred from a eukaryotic cell by HGT [125].

Niche enhancing abilities

Genome degradation and exploitation of host cellular processes by *Rickettsia*, *Orientia* and *Wolbachia* genomes show a few genes potentially acquired by HGTs such as the paralogous gene family coding for proteins with ankyrin repeat domains. Variably present gene may reflect host-adaptation processes and give hints to the differences in lifestyle and host preferences [126]. Reduction in *Helicobacter pylori* and *H. acinonychis* genome sizes compared to *H. hepaticus* has been suggested as a result of a host adaptation process following a host jump [127]. The presence of a large number of genome islands indicates a large flexible gene pool for this group of bacteria. The most prominent of these PAIs is related to cytotoxin-associated gene (*cag*). The *cag* PAI encodes a bacterial type IV secretion system, a bacterial virulence factor for acute infection. Type IV secretion systems are widespread among other bacteria including the genera *Agrobacterium, Bartonella* and *Bordetella* [128,129]. These systems reflect of potential HGT enhancement mechanisms.

Many Gram-negative bacteria, pathogens and symbionts of animals and plants, have developed secretion systems, termed type III secretion systems (TTSS) that mediate elaborate interactions with their hosts. Genes encoding TTSS are predominantly located on unstable genetic element - plasmids or PAIs: PAI-1 and PAI-2 in *S. enterica* serovar *Typhimurium*, LEE on enteropathogenic *E. coli*, hrp-PAI in *Pseudomonas syringae* and plasmid of S. *flexneri, Yersinia enterocolitica* and *Ralstonia solanacearum*. Thus TTSS could have been acquired by one or more HGT events [130].

Pseudomonas is a unique bacterium with abilities to infect plants and animals [131]. The large *P. aeruginosa* PA14, pathogenicity island (PAPI-1) is a cluster of 108 genes that encode for a number of virulence features and play an important role in the evolution of *P. aeruginosa* by expanding its natural habitat from soil & water to animals & human infections. The genes *int* (integrase) and *soj* (chromosome partitioning protein), located at the opposite ends of PAPI-1, stabilize the circular form and are required for inter-strain transfer [132].

Pox viruses (*Pox viridae*) are a family of double stranded DNA viruses with no RNA stage. *Orthopoxvirus* (OPV) are highly invasive and virulent. OPVs are among the most dreaded pathogens on earth. Poxvirus are possible vector for horizontal transfer of the retroposons from reptiles (advanced snakes) to mammals [133], since up to 50% of the poxvirus gene families show some evidence of HGT from other hosts [134].

Metabolic activities

Salinibacter ruber is a remarkable bacterium whose phenotype resembles that of hyperhalophilic Archaea (Halo archaea). The genome sequence supports the convergence of its four rhodopsins: one of which resembles bacterial proteorhodopsin and three are of haloarcheal type [85]. To asses the nature and number of potential HGTs in *S. ruber* genome, "competitive matching" query method was employed [135]. In this method, any ORF that has a match only to genomes outside the Bacteroides/ Chlorobi group or a match at least 0.05 normalized BLAST score units better outside this group was counted as a potential HGT; 1470 ORFs showed such discordant similarities by this BLAST-based method [85]. NosZ of the *nos* cluster in *Salinibacter* encodes a nitrous oxide reductase similar to that found in a variety of proteobacteria and in *Haloarcula marismortui*. Similarly, *cco*O and *cco*N, components of a cbb3-type

cytochrome oxidase were acquired likely by HGT from a β-proteobacteria. Although *Salinibacter* does not posses a *cco*P gene, which is thought to be an essential subunit of cbb3-type cytochrome oxidase, a gene flanking *cco*ON, appears to have been acquired through HGT from a β-proteobacterium and may be a substitute for *Cco*P function [85].

Phylogenetic distribution distinguishes two structurally distinct classes of MoFe nitrogenase and is not consistent with organismal phylogeny. The patchy distribution of genes has been attributed by some to HGT while others explained the uneven distribution to loss of function in certain limneages [136].

Aerobic degradation of Hexachlorocyclohexane (HCH) is carried out mostly by strains of *Sphingomonas paucimobilis* and *Rhodanobacter lindaniclasticus*. The primary enzyme in γ-HCH degradation is HCH dehydrochlorinase encoded by the *linA* gene. In addition to other evidences, the G+C content of all the *linA* genes reported so far are lower than those of *linB, linC, linD* and *linE* genes, suggesting that *linA* gene might have been acquired by HGT, through the involvement of IS*6100* [33]. New pathway(s) for the degradation of β- and δ-HCH exist in *Sphingobium indicum* B90A, which appear to be independent of that involving α- and γ-HCH [32].

Exploiting natural HGT for engineering biosynthetic pathways

On the top of the organizational and societal agenda are themes such as environmental pollution and bioenergy for sustainable development. Biofuels such as bioethanol, biodiesel, bio-oil, hydrogen (H_2), methane, etc. have the potential to provide a sustainable energy system for the society [137]. The whole issue of bioenergy revolves around the feed and the organisms with abilities to produce them [138]. Biological wastes of plant and animal origins as feed have a great potential to provide organic matter for microbial action. These wastes are available in large quantities and from diverse sources [137]. Anaerobic digestion is one of the most efficient approaches to completely degrade and stabilize biowastes. This multiple step bioconversion process involves hydrolytic bacteria, hydrogen producer, polyhydroxyalkanoates (PHA, bioplastic) producers, secondary metabolite producers, metabolic enzyme producers and ultimately methanogens [139–142]. We need robust microbes for efficient and economical treatment of wastes. These organisms should not only be capable of producing the desired product such as H_2 and PHA but also be able to survive and metabolize the wastes which are otherwise labeled as pollutants. In the last few decades, in spite of immense efforts to search efficient H_2 and PHA producers, there has been little progress through conventional methods [138, 143]. In this post genomic era, the availability of 940 bacterial, 48 archaeal and 162 eukaryotic sequenced genomes (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi), has infact "flooded" us with a large amount of data. With the use of bioinformatics tools, it has been possible to detect novel organisms with properties such as H_2-, PHA-, antibiotic- production [144–146]. These studies have also revealed those organisms, which are at present categorized as "non"-producers, because they lack a gene or two of the biosynthetic pathway. Comparative genomic analysis of 253 sequenced genomes for *phaA*, *phaB*, and *phaC* genes of PHA biosynthesis has revealed incongruencies in the phylogenetic trees for these genes. Further analysis of these phylogenetic discrepancies for parameters such as G+C content, CAI and Chi-square tests lead to the suggestion that HGT may be a major contributor for the evolution of this stress induced metabolic pathway [147] (Table 1). It has been possible to trace the potential donors for transforming the non-producers into producers through genome shuffling [147]. This technique can be extended to other metabolic pathways and the process of HGT can exploited to generate naturally engineered organisms.

Conclusions

New discoveries about novel microorganisms influence our thinking process. Intragenomic plasticity and inter-species horizontal mobility of operons are thought to be important facets of archaeal, bacterial and eukaryotic genome evolution. It has led to the view that HGT is a "major force", rather than an interesting but anecdotal event [88]. An increasing quantum of evidence of HGT suggests the importance of these phenomena for evolution. It is well-known among bacteria, its occurrence has become recognized among higher plants and animals in the last decade or so. The scope for HGT is essentially the entire biosphere. Here, bacteria and viruses act as genes reservoirs and aid in their trafficking. However, the prevalence and importance of HGT in the evolution of multicellular eukaryotes remain unclear [148]. Horizontal transfer of genes from bacteria to some fungi, especially the yeast *Saccharomyces cerevisiae*, has been well documented [149]. The role of *Wolbachia*. an endosymbiont as important potential source of genetic material has been reported in adzuki bean beetle [150] and in arthropods and filarial nematodes [151]. On the other hand, evidences exist where mitochondrial genes were reported to be horizontally transferred to parasites of the Rafflesiaceae plant family from their hosts [152, 153] and from eukary-

otic chloroplasts to the mitochondria of the bean *Phaseolus* [154].

Different studies suggest that evolution occurred in a communal, a loosely knit, diverse conglomeration of primitive cells that evolved as a unit. Subsequently it developed into several distinct communities: bacteria, archaea and eukaryotes). By swapping genes freely, these diverse organisms shared various of their talents with their contemporaries [155]. In order to meet the ultimate challenge of comprehending the ecosystem as a whole, we need to understand the principles underlying the assembly of biogeochemical systems and the flow of materials and energy [42]. Increasingly powerful search tools and further developments in theoretical and experimental approaches will be needed to help expand our understanding of the role of HGT in ecosystem dynamics and evolutionary processes. In view of the significance of HGT, a new line of comprehensive evolution coupling vertical and HGTs has been drawn [156]. A polyphasic approach is expected to make significant clarity in the science of bacterial evolution [157]. In conclusion, we agree that HGT generates greater genetic diversity which is likely to increase resource use and ecosystem resilience [42].

Acknowledgements

We are thankful to Director, Institute of Genomics and Integrative Biology, CSIR for providing the necessary funds, facilities and moral support.

References

1. Syvanen M (1994) Horizontal Gene Transfer: Evidence and possible consequences. Annu Rev Genet 28:237–261
2. Lopez-Garc P & Moreira D (1999) Metabolic symbiosis at the origin of eukaryotes. Trends Biochem Sci 24:88–93
3. Dutta C & Pan A (2002) Horizontal gene transfer and bacterial diversity; J Biosci (Suppl 1) 27:27–33
4. He CY, Striepen B, Pletcher CH, Murray JM & Roos DS (2001) Targeting and processing of nuclear-encoded apicoplast proteins in plastid segregation mutants of *Toxoplasma gondii*. J Biol Chem 276:28436–28442
5. Martin W & Mueller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41
6. Avery OT, MacLeod CM & McCarty M (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* **149:**297–326
7. Smith MW, Feng DF & Doolittle RF (1992) Evolution by acquisition: the case for horizontal gene transfers. Trends Biochem Sci 17:489–493
8. Groisman EA & Ochman H (1996) Pathogenicity islands: bacterial evolution in quantum leaps. Cell 87:791–794
9. Brown JR & Doolittle WF (1997) *Archaea* and the prokaryote-to-eukaryote transition. Microbiol Mol Biol Rev 61: 456–502
10. Doolittle WF (1999) Phylogenetic classification and the universal tree. Science 284:2124–2128
11. Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA 97:8392–8396
12. Koonin EV, Makarova KS & Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709–742
13. Philippe H & Douady CJ (2003) Horizontal gene transfer and phylogenetics. Curr Opin Microbiol 6:498–505
14. Martin W & Embley TM (2004) Early evolution comes full circle. Nature 431:134–137
15. Rivera MC & Lake JA (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. Nature 431: 152–155
16. Tsirigos A & Rigoutsos I (2005) A new computational method for the detection of horizontal gene transfer events. Nucleic Acids Res 33:922–933
17. Charlebois RL, Beiko RG & Ragan MA (2003) Branching out. Nature 421:217
18. Lerat E, Daubin V, Ochman H & Moran NA (2005) Evolutionary origins of genomic repertoires in bacteria. PLoS Biol 3:e130
19. Choi I-G & Kim S-H (2007) Global extent of horizontal gene transfer. Proc Natl Acad Sci USA 104:4489-4494
20. Ochman H, Herat E & Daubin V (2005) Examining bacterial species under the specter of gene transfer and exchange. Proc Natl Acad Sci USA 102:6595–6599
21. 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
22. Aertsen A & Michiels CW (2005) Diversify or die: generation of diversity in response to stress. Crit Rev Microbiol 31: 69–78
23. Beiko RG, Harlow TJ & Ragan MA (2005) Highways of gene sharing in prokaryotes. Proc Natl Acad Sci USA 102: 14332–14337
24. Lawrence JG (1999) Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. Curr Opin Genet Dev 9:642–648
25. Lawrence JG & Roth JR (1996) Selfish operons: Horizontal transfer may drive the evolution of gene clusters. Genetics 143:1843–1860
26. Lawrence JG & Roth JR (1995) The cobalamin (coenzyme B_{12}) biosynthetic genes of *Escherichia coli.* J Bacteriol 177: 6371–6380
27. Lawrence JG & Roth JR (1996) Evolution of coenzyme B_{12} synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. Genetics 142:11–24
28. Matsui K, Sano K & Ohtsubo E (1986) Complete nucleotide and deduced amino acid sequences of the *Brevibacterium luctofermentum* tryptophan operon. Nucleic Acids Res 14: 10113–10114
29. Williams PA & JR Sayers (1994) The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas.* Biodegradation 5:195–217

30. Wyndham RC, Cashore AE, Nakatsu CH & Peel MC (1994) Catabolic transposons. Biodegradation 5:323–342
31. Van der Meer JR (1997) Evolution of novel metabolic pathways for the degradation of chloroaromatic compounds. Antonie van Leeuwenhoek J Microbiol Serol 71:159–178
32. Sharma P, Raina V, Kumari R, Malhotra S, Dogra C, Kumari H, Kohler H-PE, Buser H-R, Holliger C & Lal R (2006) Haloalkane dehalogenase LinB is responsible for beta- and delta-hexachlorocyclohexane transformation in *Sphingobium indicum* B90A. Appl Environ Microbiol 72(9): 5720–5727
33. Dogra C, Raina V, Pal R, Suar M, Lal S, Gartemann K-H, Holliger C, van der Meer JR & Lal R (2004) Organization of *lin* genes and IS6100 among different strains of hexachlorocyclohexane-degrading *Sphingomonas paucimobilis*: evidence for horizontal gene transfer. J Bacteriol 186(8): 2225–2235
34. Rosen BP (1999) Families of arsenic transporters. Trends Microbiol 7:207–212
35. Gihring TM, Bond PL, Peters S & Banfield JF (2003) Arsenic resistance in the archaeon *Ferroplasma acidarmanus*: new insights into the structure and evolution of arsenic genes. Extremophiles 7:123–130
36. Shen Y, Buick R & Canfield DE (2001) Isotopic evidence for microbial sulphate reduction in the early Archaean era. Nature 410:77–81
37. Klein M, Friedrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H, Blackall LL, Stahl DA & Wagner M (2001) Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. J Bacteriol 183:6028–6035
38. Aravind L, Tatusov RL, Wolf YI, Walker DR & Koonin EV (1998) Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. Trends Genet 14: 442–444
39. Wolf YI, Aravind L & Koonin EV (1998) *Rickettsiae and Chlamydia*: evidence of horizontal gene transfer and gene exchange. Trends Genet 14:442–444
40. De la Cruz F & Davies J (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol 8:128–133
41. Xiong J, Fischer WM, Inoue K, Nakahara M & Bauer CE (2000) Molecular evidence for the early evolution of photosynthesis. Science 289:1724–1730
42. Macalady J & Banfield JF (2003) Molecular geomicrobiology: genes and geochemical cycling. Earth and Planetary Science Letters 209:1–17
43. Di Gioia D, Peel M, Fava F & Wyndham RC (1998) Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America. Appl Environ Microbiol 64:1940–1946
44. Campbell A, Mrazek J & Karlin S (1999) Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. Proc Natl Acad Sci USA 96:9184–9189
45. Dubnau D (1999) DNA uptake in bacteria. Annu Rev Microbiol 53:217–244
46. Hall RM & Collis CM (1995) Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. Mol Microbiol 15:593–600
47. Mazel D, Dychinco B, Webb VA & Davies J (1998) A distinctive class of integron in the *Vibrio cholerae* genome. Science 280:605–608
48. Dowson CG, Coffey TJ, Kell C & Whiley RA (1993) Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. pneumomae*. Mol Microbiol 9:635–643
49. Lujan R, Zhang QY, Saez Nieto JA, Jones DM & Spratt BG (1991) Penicillin–resistant isolates of *Neisseria lactamica* produce altered forms of penicillin–binding protein 2 that arose by interspecies horizontal gene transfer. Antimicrob Agents Chemother 35:300–304
50. Bowler LD, Zhang QY, Riou JY & Spratt BG (1994) Interspecies recombination between the *penA* genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: Natural events and laboratory simulation. J Bacteriol 176: 333–337
51. Spratt BG, Bowler LD, Zhang QY, Zhou J & Smith JM (1992) Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria* species. J Mol Evol 34:115–125
52. Hacker J, Blum-Oehler G, Mühldorfer I & Tschäpe H (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol 23: 1089–1097
53. Blum G, Ott M, Lischewski A, Ritter A, Imrich H, Tschape H & Hacker J (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. Infect Immun 62:606–614
54. McDaniel TK & Kaper JB (1997) A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Mol Microbiol 23:399–407
55. Moss JE, Cardozo TJ, Zychlinsky A & Groisman EA (1999) The selC-associated SHI-2 pathogenicity island of *Shigella flexneri.* Mol Microbiol 33:74–83
56. Vokes SA, Reeves SA, Torres AG & Payne SM (1999) The aerobactin iron transport system genes in *Shigella flexneri* are present within a pathogenicity island. Mol Microbiol 33: 63–73
57. Blanc-Potard AB & Groisman EA (1997) The *Salmonella selC* locus contains a pathogenicity island mediating intramacrophage survival. EMBO J 16:5376–5385
58. Eisen JA (2000) Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. Curr Opin Genet Dev 10:606–611
59. Ragan MA (2001) Detection of lateral gene transfer among microbial genomes. Curr Opin Genet Dev 11:620–626
60. Ochman H, Lawrence JG & Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304
61. Grantham R, Gautier C, Gouy M, Mercier R & Pavé A (1980) Codon catalog usage and the genome hypothesis. Nucleic Acids Res 8:197–c
62. Ellis J, Morrison DA & Kalinna B (1995) Comparison of the patterns of codon usage and bias between *Brugia*, *Echinococcus*, *Onchocerca* and *Schistosoma* species. Parasitol Res 81:388–393

63. Medigue C, Rouxel T, Vigier P, Henaut A & Danchin A (1991) Evidence for horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol 222:851–856
64. Lawrence JG & Ochman H (1997) Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol 44: 383–397
65. Mrazek J & Karlin S (1999) Detecting alien genes in bacterial genomes. Ann NY Acad Sci 870:314–329
66. Garcia-Vallve S, Romeu A & Palau J (2000) Horizontal gene transfer in bacterial and archaeal complete genome. Genome Res 10:1719–1725
67. Groisman EA, Sturmoski MA, Solomon FR, Lin R & Ochman H (1993) Molecular, functional, and evolutionary analysis of sequences specific to *Salmonella*. Proc Natl Acad Sci USA 90:1033–1037
68. Groisman EA & Ochman H (1993) Cognate gene clusters govern invasion of host epithelial cells by *Salmonella typhimurium* and *Shigella flexneri*. EMBO J 12:3779–3787
69. Shanley MS, Harrison A, Parales RE, Kowalchuk G, Mitchell DJ & Ornston LN (1994) Unusual G+C content and codon usage in *catZJF*, a segment of the *ben-cat* supra-operonic cluster in the *Acinetobacter caloaceticus* chromosome. Gene 138:59–65
70. Groisman EA, Saier Jr MH & Ochman H (1992) Horizontal transfer of a phosphatase gene as evidence for the mosaic structure of the *Salmonella* genome. EMBO J 11: 1309–1316
71. Lan R & Reeves PR (1996) Gene transfer is a major factor in bacterial evolution. Mol Biol Evol 13:47–55
72. Bult CJ, White O, Olsen GJ, Zhou L, Fleishmann RD, Sutton GG, Blake JA, Fitzgerald LM, Clayton RA, Gocayne JD, Kertavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kaine BP, Borodovsky M, Klenk HP, Fraser CM, Smith HO, Woese CR & Ventor JC (1996) Complete genome sequence of the methanogenic Archaeon, *Methanococcus jannaschii*. Science 273:1058–1073
73. Klenk H-P, Clayton RA, Tomb J-F, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea KP, Bowman C, Fujii C, Garland SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR & Venter JC (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus.* Nature 390:364–370
74. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg, J, Sutton, GG, Fleischmann RD, Eisen JA, White O, Salzberg SL, Smith HO, Venter JC & Fraser CM (1999) Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritime.* Nature 399:323–329
75. Karlin S & Burge C (1995) Dinucleotide relative abun-dance extremes: a genomic signature. Trends Genet 11:283–290
76. Sharp PM & Li W-H (1987) The codon adaptation index- a measure of directional synonymous codon usage bias and its potential applications. Nucleic Acids Res 15:1281–1295
77. Karlin S, Mrázek J & Campbell AM (1998) Codon usages in different gene classes of the *Escherichia coli* genome. Mol Microbiol 29:1341–1355
78. Tatusov RL, Koonin EV & Lipman DJ (1997) A genomic perspective on protein families. Science 278:631–637
79. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND & Koonin EV (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res 29:22–28
80. Glazko GV & Mushegian AR (2004) Detection of evolutionarily stable fragments of cellular pathways by hierarchical clustering of phyletic patterns. Genome Biol 5:R32
81. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D & Yeaster TO (1999) Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc Natl Acad Sci USA 96:4285–4288
82. Koonin EV, Mushegian AR, Galperin MY & Walker DR (1997) Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the Archaea. Mol Microbiol 25:619–637
83. Doolittle WF & Logsdon JM Jr (1998) Archaeal genomics: do archaea have a mixed heritage? Curr Biol 8: R209–R211
84. Jain R, Rivera MC & Lake JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci USA 96:3801–3806
85. Mongodin EF, Nelson KE, Daugherty S, DeBoy RT, Wister J, Khouri H, Weidman J, Walsh DA, Papke RT, Sanchez Perez G, Sharma AK, Nesbø CL, MacLeod D, Bapteste E, Doolittle WF, Charlebois RL, Legault B & Rodriguez-Valera F (2005) The genome of *Salinibacter ruber*: Convergence and gene exchange among hyperhalophilic bacteria and archaea. Proc Natl Acad Sci USA 102: 18147–18152
86. Kurland CG, Canback B & Berg OG (2003) Horizontal gene transfer: A critical view. Proc Natl Acad Sci USA 100: 9658–9662
87. Huynen MA, Snel B & Bork P (1999) Lateral gene transfer, genome surveys, and the phylogeny of prokaryotes. Science 286:1443a
88. Snel B, Bork P & Huynen MA (1999) Genome phylogeny based on gene content. Nat Genet 21:108–110
89. Korbel JO, Snel B, Huynen MA & Bork P (2002) SHOT: a web server for the construction of genome phylogenies. Trends Genet 18:158–162
90. Tekaia F, Lazcano A & Dujon B (1999) The genomic tree as revealed from whole proteome comparisons. Genome Res 9:550–557
91. Fitz–Gibbon ST & House CH (1999) Whole genome-based phylogenetic analysis of free-living microorganisms. Nucleic Acids Res 27:4218–4222

92. House CH & Fitz-Gibbon ST (2002) Using homolog groups to create a whole-genomic tree of free-living organisms: an update. J Mol Evol 54:539-547
93. Wolf YI, Rogozin IB, Grishin NV, Tatusov RL & Koonin EV (2001) Genome trees constructed using five different approaches suggest new major bacterial clades. BMC Evol Biol 1:8
94. Lin J & Gerstein M (2000) Whole–genome trees based on the occurrence of folds and orthologs: implications for comparing genomes on different levels. Genome Res 10: 808–818
95. Beck JT, Zhao S & Wang CC (1994) Cloning, sequencing, and structural analysis of the DNA encoding inosine monophosphate dehydrogenase (EC1.1.1.205) from *Tritrichomonas foetus*. Exp Parasitol 78:101–112
96. Berg OG & Kurland CG (2002) Evolution of microbial genomes: sequence acquisition and loss. Mol Biol Evol 19: 2265–2276
97. Mount DM (2001) Bioinformatics Sequence and Genome Analysis, New York Cold Spring Harbour Laboratory Press p. 247–248
98. Natale DA, Shankavaram UT, Galperin MY, Wolf YI, Aravind L & Koonin EV (2000) Towards understanding the first genome sequence of a crenarchaeon by genome annotation using clusters of orthologous groups of proteins (COGs). Genome Biol 1 research0009.1–0009.19
99. Clarke GDP, Beiko RG, Ragan MA & Charlebois RL (2002) Inferring genome trees by using a filter to eliminate phylogenetically discordant sequences and a distance matrix based on mean normalized BLASTP scores. J Bacteriol 184:2072–2080
100. Wolf YI, Brenner SE, Bash PA & Koonin EV (1999) Distribution of protein folds in the three superkingdoms of life. Genome Res 9:17–26
101. Eisen JA (1995) The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. J Mol Evol 41:1105–1123
102. Brown JW, Daniels CJ & Reeve JN (1989) Gene structure, organization and expression in archaebacteria. Crit Rev Microbiol 16:287–338
103. Woese CR (1987) Bacterial evolution. Microbiol Rev 51: 221–271
104. Woese CR, Kandler O & Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579
105. Fox GE, Magrum LJ, Balch WE, Wolfe RS & Woese CR (1977) Classification of methanogenic bacteria by 16S ribosomal RNA characterization. Proc Natl Acad Sci USA. 74:4537–4541
106. Woese CR & Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090
107. Danson MJ (1988) Archaebacteria: the comparative enzymology of their central metabolic pathways. Adv Microb Physiol 29:165–231
108. Danson MJ (1989) Central metabolism of the archaebacteria: An overview. Can J Microbiol 35:58–64
109. Schönheit P & Schäfer T (1995) Metabolism of hyperthermophiles. World J Microbiol Biotechnol 11:26–75
110. Schloss PD & Handelsman J (2004) Status of the Microbial Census. Microbiol Mol Biol Rev 68:686–691
111. Pace NR, Olsen OJ & Woese CR (1986) Ribosomal RNA phylogeny and the primary lines of evolutionary descent. Cell 45:325–326
112. Yap WH, Zhong Z & Wang Y (1999) Distinct Types of rRNA Operons exist in the genome of the Actinomycete *Thermomonospora chromogena* and evidence for horizontal transfer of an entire rRNA operon. J Bacteriol 181: 5201–5209
113. Yanai I, Wolf YI & Koonin EV (2002) Evolution of gene fusions: horizontal transfer versus independent events. Genome Biol 3:research0024.1–0024.13
114. Olsen GJ, Woese CR & Overbeek R (1994) The winds of (evolutionary) change: breathing new life into microbiology. J Bacteriol 176:1–6
115. Doolittle RF & Handy J (1998) Evolutionary anomalies among the aminoacyl-tRNA synthetases. Curr Opin Genet Dev 8:630–636
116. Lawrence JG & Ochman H (1998) Molecular archaeology of the *Escherichia coli* genome. Proc Natl Acad Sci USA 95:9413–9417
117. Rivera MC, Jain R, Moore JE & Lake JA (1998) Genomic evidence for two functionally distinct gene classes. Proc Natl Acad Sci USA 95:6239–6244
118. Makarova KS, Aravind L, Wolf YI, Tatusov RL, Minton KW, Koonin EV & Daly MJ (2001) Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. Microbiol Mol Biol Rev 65:44–79
119. Iyer LM, Koonin, EV & Aravind L (2004) Evolution of bacterial RNA polymerase: implications for large scale phylogeny, domain accretion and horizontal gene transfer. Gene 335:73–88
120. Bapteste E, Moreira D & Philippe H (2003) Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. Gene 318: 185–191
121. Müller M, Lee JA, Gordon P, Gaasterland T & Sensen CW (2001) Presence of prokaryotic and eukaryotic species in all subgroups of the PPi-dependent group II phosphofructokinase protein family. J Bacteriol 183:6714–6716
122. Venkatesh R, Ganesh N, Guhan N, Reddy MS, Chandrasekhar T & Muniyappa K (2002) RecX protein abrogates ATP hydrolysis and strand exchange promoted by RecA: insights into negative regulation of homologous recombination. Proc Natl Acad Sci USA 99: 12091–12096
123. Stohl EA, Brockman JP, Burkle KL, Morimatsu K, Kowalczykowski SC & Seifert NS (2003) *Escherichia coli* RecX inhibits RecA recombinase and coprotease activities in vitro and in vivo. J Biol Chem 278:2278–2285
124. Lin J, Chen Z-Z, Tian B & Hua Y-J (2007) Evolutionary pathways of an ancient gene recX. Gene 387, 15–20 doi: 10.1016/j.gene.2006.07.031
125. Schlieper D, Oliva MA, Andreu JM & Löwe J (2005) Structure of bacterial tubulin BtubA/B: evidence for horizontal gene transfer. Proc Natl Acad Sci USA 102:9170–9175

126. Fuxelius H-H, Darby A, Min C-K, Cho, N-H & Andersson Siv GE (2007) The genomic and metabolic diversity of *Rickettsia*. Res Microbiol 158:745–753
127. Eppinger M, Baar C, Linz B, Raddatz G, Lanz C, Keller H, Morelli G, Gressmann H, Achtman M & Schuster SC (2006) Who ate whom? Adaptive *Helicobacter* genomic changes that accompanied a host jump from early humans to large felines. PLoS Genet 2, e120
128. Backert S & Meyer TF (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol 9:207–217
129. Linz B & Schuster SC (2007) Genomic diversity in Helicobacter and related organisms. Res Microbiol 158:737–744
130. Gophna U, Ron EZ & Graur D (2003) Bacterial type III secretion systems are ancient and evolved by multiple horizontal transfer events. Gene 312:151–163
131. Rahme LG, Ausubel FM, Cao H, Drenkard E, Goumnerov BC, Lau, GW, Mahajan-Miklos S, Plotnikova J, Tan M-W, Tsongalis J, Walendziewicz CL & Tompkins RG (2000) Plants and animals share functionally common bacterial virulence factors. Proc Natl Acad Sci USA 97(16): 8815–8821
132. Qiu X, Gurkar AU & Lory S (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. Proc Natl Acad Sci USA 103:19830–19835
133. Hughes AL & Friedman R (2005) Poxvirus genome evolution by gene gain and loss. Mol Phyl Evol 35:186–195
134. Pisurek O & Okada, N (2007) Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. Proc Natl Acad Sci USA 104:12046–12051
135. Gophna U, Charlebois RL & Doolittle WF (2004) Have archaeal genes contributed to bacterial virulence? Trends Microbiol 12: 213–219
136. Kechris KJ, Lin JC, Bickel PJ & Glazer AN (2006) Quantitative exploration of the occurrence of lateral gene transfer by using nitrogen fixation genes as a case study. Proc Natl Acad Sci USA 103:9584–9589
137. Raizada N, Sonakya V, Anand V & Kalia VC (2002) Waste management and production of future fuels. J Sci Ind Res 61:184–207
138. Kalia VC & Purohit HJ (2008) Microbial diversity and genomics in aid of bioenergy. J Ind Microbiol Biotechnol 35:403–419
139. Kalia VC, Jain SR, Kumar A & Joshi AP (1994) Fermentation of biowaste to hydrogen by *Bacillus licheniformis*. World J Microbiol Biotechnol 10:224–227
140. Kalia VC & Joshi AP (1995) Conversion of waste biomass (pea–shell) into hydrogen and methane through anaerobic digestion. Bioresour Technol 53:165–168
141. Kalia VC, Anand V, Kumar A & Joshi AP (1997) Efficient biomethanation of plant materials by immobilized bacteria. In: Proceedings of R'97 Congress (Recovery, Recycling, Re-integration) Geneva, Switzerland Vol I: 200–205
142. Sonakya V, Raizada N & Kalia VC (2001) Microbial and enzymatic improvement of anaerobic digestion of waste biomass. Biotechnol Lett 23:1463–1466
143. Porwal S, Kumar T, Lal S, Rani A, Kumar S, Cheema S, Purohit HJ, Sharma R, Patel SKS & Kalia VC (2007) Hydrogen and polyhydroxybutyrate producing abilities of microbes from diverse habitats by dark fermentative process. Bioresour Technol 99:5444–5451
144. Kalia VC, Chauhan A, Bhattacharyya G & Rashmi XX (2003) Genomic databases yield novel bioplastic producers. Nat Biotechnol 21:845–846
145. Kalia VC, Lal S, Ghai R, Mandal M & Chauhan A (2003) Mining genomic databases to identify novel hydrogen producers. Trends Biotechnol 21:152–156
146. Kalia VC, Rani A, Lal S, Cheema S & Raut CP (2007) Combing databases reveals potential antibiotic producers. Expert Opin Drug Discov 2:211–224
147. Kalia VC, Lal S & Cheema S (2007) Insight in to the phylogeny of polyhydroxyalkanoate biosynthesis: Horizontal gene transfer. Gene 389: 19–26
148. Richardson AO & Jeffrey DP (2007) Horizontal gene transfer in plants. J Expt Bot 58:1–9
149. Hall C, Brachat S & Dietrich FS (2005) Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. Eukaryot Cell 4:1102–1115
150. Kondo N, Nikoh N, Ijichi N, Shimada M & Fukatsu T (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci USA 99:14280–14285
151. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, 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 doi:10.1126/science.1142490
152. Davis CC & Wurdack KJ (2004) Host-to-parasite gene transfer in flowering plants: Phylogenetic evidence from Malpighiales. Science 305:676–678
153. Nickrent DL, Blarer A, Qiu Y-L, Vidal-Russell R & Anderson FE (2004) Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol Biol 4(40). doi:10.1186/1471-2148-4-40
154. Woloszynska M, Bocer T, Mackiewicz P & Janska H (2004) A fragment of chloroplast DNA was transferred horizontally, probably from non-eudicots, to mitochondrial genome of Phaseolus. Plant Mol Biol 56:811-820
155. Doolittle FW (2000) Uprooting the tree of life. Scientific American: 72–77
156. Kunin V, Goldovsky L, Darzentas N & Ouzounis CA (2005) The net of life: Reconstructing the microbial phylogenetic netwrok. Genome Res 15: 954–959
157. Prakash O, Verma M, Sharma P, Kumar M, Gupta SK, Khanna M & Lal R (2007) Polyphasic approach of bacterial classification – An overview of recent advances. Ind J Microbiol 47:98–108
158. Oh SJ, Joung JG, Chang JH & Zhang BT (2006) Construction of phylogenetic trees by kernel-based comparative analysis of metabolic networks. BMC Bioinformatics 7:284
159. Wenzl P, Wong L, Kwang-won K & Jefferson RA (2005) A functional screen identifies lateral transfer of beta-glucuronidase (gus) from bacteria to fungi. Mol Biol Evol 22:308–316
160. Baumann P & Jackson SP (1996) An archaebacterial homologue of the essential eubacterial cell division protein FtsZ. Proc Natl Acad Sci USA 93:6726–6730

161. Wang X & Lutkenhaus J (1996) FtsZ ring: the eubacterial division apparatus conserved in archaebacteria. Mol Microbiol 21:13–319
162. Ricard G, McEwan NR, Dutilh BE, Jouany JP, Macheboeuf D, Mitsumori M, McIntosh FM, Michalowski T, Nagamine T, Nelson N, Newbold CJ, Nsabimana E, Takenaka A, Thomas NA, Ushida K, Hackstein JH & Huynen MA (2006) Horizontal gene transfer from bacteria to rumen ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. BMC Genomics 7:22
163. Gomes JP, Bruno WJ, Borrego MJ & Dean D (2004) Recombination in the genome of *Chlamydia trachomatis* involving the polymorphic membrane protein C gene relative to *ompA* and evidence for horizontal gene transfer. J Bacteriol 186:4295–4306
164. Omelchenko MV, Makarova KS, Wolf YI, Rogozin IB & Koonin EV (2003) Evolution of mosaic operons by horizontal gene transfer and gene displacement in situ. Genome Biol 4:R55
165. Viale AM & Arakaki AK (1994) The chaperone connection to the origins of eukaryotic organelles. FEBS Lett 341: 146–151
166. Gupta RS (1995) Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. Mol Microbiol 15:1–11
167. Jackson CR & Dugas SL (2003) Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin arsenate reductase. BMC Evol Biol 3:18
168. Sheveleva EV & Hallick RB (2004) Recent horizontal intron transfer to a chloroplast genome. Nucleic Acids Res 32:803–810
169. Klenk H-P, Palm P & Zillig W (1993) DNA-dependent RNA polymerases as phylogenetic marker molecules. Syst Appl Microbiol 16:138–147
170. Pühler G, Leffers H, Gropp F, Palm P, Klenk H-P, Lottspeich F, Garrett RA & Zillig W (1989) Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc Natl Acad Sci USA 86: 4569–4573
171. Ponting CP, Aravind L, Schultz J, Bork, P & Koonin EV (1999) Eukaryotic signalling domain homologues in archaea and bacteria, Ancient ancestry and horizontal gene transfer. J Mol Biol 1289:729–745
172. Kennelly PJ (2002) Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. FEMS Microbiol Lett 206:1–8
173. Zhang W & Shi L (2004) Evolution of the PPM-family protein phosphatases in Streptomyces: duplication of catalytic domain and lateral recruitment of additional sensory domains. Microbiology 150:4189–4197
174. Won H & Renner SS (2003) Horizontal gene transfer from flowering plants to *Gnetum*. Proc Natl Acad Sci USA 100: 10824–10829
175. Pruss B, Meyer HE & Holldorf AW (1993) Characterization of the glyceraldehyde 3-phosphate dehydrogenase from the extremely halophilic archaebacterium *Haloarcula vallismortis.* Arch Microbiol 160:5–11
176. Nolling J & Vos WM (1992) Characterization of the archaeal, plasmidencoded type II restriction-modification system MthTI from *Methanobacterium thermoformicicum* THF: homology to the bacterial NgoPII system from *Neisseria gonorrhoeae*. J Bacteriol 174: 5719–5726
177. Brinkman FS, Blanchard JL, Cherkasov A, Av-Gay Y, Brunham RC, Fernandez RC, Finlay BB, Otto SP, Ouellette BF, Keeling PJ, Rose AM, Hancock RE, Jones SJ & Greberg H (2002) Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. Genome Res 12:1159–116796.
178. Lawson FS, Charlebois RL & Dillon JA (1996) Phylogenetic analysis of carbamoylphosphate synthetase genes: evolution involving multiple gene duplications, gene fusions, and insertions and deletions of surrounding sequences. Mol Biol Evol 13:970–977