Chapter 43

Bacteria and Viruses: Communal Interacting Agents

GÜNTHER WITZANY*

Learning Objectives

For a long time bacteria have been assumed to be the most primitive organisms and consequently have been investigated as single-cell individuals determined by mechanistic input-output reactions. Now this picture has changed radically. Today we know that bacteria are part of a bacterial community that interacts in a highly sophisticated manner. The medium of every bacterial coordination is communication, i.e. sign-mediated interaction. A wide range of chemical molecules serve as signs through which bacterial communities exchange information and act in reaching a "quorum", which is the starting-point for decision-making: one of many different behavioral patterns will thereby be organised, such as biofilm organisation, bioluminescence, virulence or sporulation. Quorum sensing includes not only chemotaxis, but also interpretation, which mean that the incoming signs are measured against the background memory of the species-colony in their real life world. Interpretation before decision-making, coordination and organisation, such as fruiting body formation and cooperative hierarchical organisation, is context-dependent.

Bacteria, which in former times were viewed as lower life forms, have now been recognised as masters of monitoring, computing, interpretation, coordination and organisation. Bacterial communicative competences are sign-mediated interactions between the same or related species, but also between non-related species according to different situational contexts (pragmatic level of analyses) and the coherent combinatorial patterns of signs according to the molecular syntax (syntactic level of analyses), both determining the content of the messages (semantic level of analyses), the meaning of sign-molecules for a bacterial community that shares a common background memory, and a competence for culture-dependent interpretation which is an advantage for adaptational purposes.

In addition, bacteria seem to have been a suitable habitat for the natural genome editing competences of persistent viruses throughout the entire history of life. Looking at their evolutionary roots opens the perspective on communal-acting pre-cellular species which drove the evolution of cellular life.

Telos – Philosophische Praxis, Vogelsangstraβe 18c, A-5111-Buermoos, Austria
*E-mail: witzany@sbg.at

1. INTRODUCTION

Intraorganismic communication in bacteria includes generation, modifications, regulation of prokaryotic gene word order and its evolutionary roots. Interestingly, prokaryotic gene order is not as conserved as the sequences, which code for proteins. Only some higher order regulations (operons) that code for physically interacting proteins are found in almost all bacterial (and archaeal) genomes. Recent research indicates high dynamics of new gene orders as documented in the horizontal gene transfer events with their intensive intragenomic recombination. This exchange of whole genes or gene-blocks enables bacterial lifestyles to combine several bacterial competences, i.e. phenotypes. The transformation process includes the release of naked DNA, followed by the uptake and recombination, i.e. the integration, with 17 steps identified to date exemplified excellently by Thomas and Nielsen (2005). Thus we can recognize the outcomes of a diversity of mobile DNA contents, not a mass of individualized genetic texts, but a bacterial gene pool as a text repertoire which is available for each individual bacteria and the resource for bacterial genome innovation and evolution. Horizontal gene transfer is a main resource for integrating newly evolved genes into existing genomes and does not need the slow steps of chance mutations to alter the genomes but accelerated genome innovations in both bacteria and archaea. Important in this context of genomic innovation is not the sequence acquisition alone but also the contextualization, it means also their loss. It seems now that the phylogeny of microbial species is not a tree of life, but an evolutionary network or a ring of life, mediated by genetic exchange, i.e. acquisition and loss of genetic data sets, all of them depending on natural genome editing competences of viruses.

2. INTRACELLULAR COMMUNICATION

Sign-mediated transcription regulation of the DNA by diversity of signalling molecules serves for a great variety of response behaviour. One of the most interesting phenomena is the fact that in the first two billion years of life on planet earth the immense density of bacterial life has not been an event of the mass of individual organisms but their commonly shared gene pool which was in constant flux, as we now know, through investigations on horizontal gene transfer. It means that the evolution of bacteria was not a random event of chance mutations and their selection but transfer of whole genes and gene-blocks representing real phenotypes that were transferred. This leads to different combinatorial patterns of genetic encoded phenotypes and the rise of bacterial diversity. It also enables bacterial pathogens to optimize their disease-causing coordination and is therefore targeted to special kinds of drug developments for medical purposes. New empirical data seem to suggest that the phenomenon of horizontal gene transfer is driven by viral competences inherent in bacterial settlers such as phages, plasmids, retroplasmids and transposons. This means, to understand intraorganismic communication of bacteria we have to look at the roles of viral settlers in bacteria as outlined in the following chapters.

3. THE EVOLUTION OF BACTERIAL IDENTITIES BY NATURAL GENOME EDITING COMPETENCES OF VIRUSES

To elucidate communicative competences of bacteria we have to look at the roles of viruses and their relationship to bacteria. Viruses have long been accepted only as disease causing, epidemic phenomena with lytic and therefore extremely dangerous consequences for infected organisms. However, new research has corrected this picture. Viruses are part of the living world, in most cases integrated in the cytoplasm or the nucleoplasma of cells without harming the host. Viruses are on their way to representing the best examples of symbiotic relationships, because there is no living being since the start of life that has not been colonised by them, in most often cases in the form of multiple colonisations. This means, that the identity of bacteria which plays an important role in communication processes of the same or related species as it functions in quorum sensing depends on features of exogenous or endogenous viral settlers. The longest period of these symbiotic relationships during evolutionary history share viruses, archaea and bacteria. As viruses are extremely biosphere specific, i.e. they adapt to special host tissues, the identification of various forms of, e.g. bacteria is to identify primarily the viruses that colonise them. This is also the concept of 'bacteriophages', in that bacteria are identified best by identifying the viruses that are associated with them. Host identification in this way is a special method called phage typing.

3.1 Lytic Versus Persistent Viral Life-strategies

As mentioned in recent years, the lytic consequences of viral infection are a special case if viruses are not able to develop a sessile lifestyle without harming the host. In most cases viruses living within organisms help to ward off competing parasites from the host and becoming part of its evolutionary history. Persistent, non-lytic viruses are decisive for species diversity and host genome editing. Nearly all natural genome editing competences represented in the conservation of expression, transcription, translation and recombination with all their detailed steps seem to derive from viral aptitudes. Even the DNA replication pathways, after the period of evolutionary early RNA influence, seems to be a special viral strategy for the conservation of coded phenotypes by warding off RNA parasites.

Since observations have become more evident that viruses are able to integrate genetic material into the host genome, it has become clear that some viruses have lytic infection lifestyles but others also endosymbiotic and even symbiogenetic lifestyles. They bestow phenotypic capabilities on the host, which non-infected hosts from the same species do not possess. As endosymbiotic viruses, which are dependent on the host's replication, they are part of the host history in that they are inheritable and part of the genomic identity of the host as documented in some several 10,000 infection events in the human genome by endogenous retroviruses.

The two viral lifestyles are not in strict opposition but, in most cases, are part of a symbiotic process. It starts with an infection by a virus. In the infected host it arrives at an equilibrial status where the immune system does not eliminate the virus but controls its replication without fatal consequences for the host organism. The persistent status lasts during most phases of the host's

life, but may return to the lytic lifestyle if the host-immune system is under stress. Most often the integration occurs by mutual neutralisation of toxic capabilities by an antitoxin of competing genetic settlers. The whole range of toxin/antitoxin addiction modules we can find throughout all genetic contents in living nature most likely is of viral origin. Therefore the persistence is sometimes called temperate lifestyle. A good example is the persistent virus in all *Symbiodinium* species being the essential endosymbiotic partner for coral animals. Coral bleaching as a worldwide phenomenon of coral disease is the consequence of dying of the coral endosymbiont because of global (water) warming. As we know now, death occurs because the persistent viruses of *Symbiodinium* become lytic as a reaction to the changing water temperature.

Also bacteria may be infected by viruses without being harmed. If infected bacteria meet non-infected bacteria it may be that the non-infected acquires lysis; the lysogenic strain does not lyse itself, but is lethal to the non-infected one. The colonized bacteria have a virus-derived molecular genetic identity, which has an advantage against the non-infected one through an acquired ability. This lysogenic bacteria, termed prophage, has an immunity function for the bacteria, which the non-infected bacteria lack. Prophage is a virus that is integrated into the bacterial host genome. Both the acute lytic phages and the persistent prophages are highly abundant in oceans and in the soil and seem to be the most dynamic life form on the entire planet. Some viruses are not integrated in the host genome but persist as plasmids and replicate independently from the host genome.

When we speak about the relationship of bacteria and viruses in most cases we speak about phage ecology. Most prokaryotic viruses are double-stranded DNA viruses with either linear or circular genome morphology and are packaged in an icosahedral capsid. Whereas acute viruses in most cases code for their own replication, recombination and repair proteins, the persistent phages lack such genes and use the host-cellular replication. This involves a totally different gene word order in acute lytic and in persistent phages. This is documented in the very different nucleotide words (di-, tri- and tetranucleotides). Nucleotide word frequencies of acute phages are very dissimilar to those of their hosts while persistent or temperate phages share nucleotide word frequencies with the host. This means the molecular syntax from acute and persistent phages is constructed totally differently according to the different strategies. Different life strategies with different behavioral pathways need a completely different semantic content in the genome expressed in a different syntactic arrangement of nucleotides.

As the bacterial cell walls differ substantially between different types of bacteria a different behaviour is necessary for viruses for recognition, attachment and penetration. Owing to these diverse barriers of the bacterial cell walls, the prokaryotic viruses do not enter the host cells physically but attach to the cell surface and inject their genomes through contractile tails or pilot proteins. Also, the progeny of the virus has to deal with this barrier.

Bacterial DNA does not have highly stable structures as do eukaryotes and can interact with the cellular replication and transcription. In most cases it is circular with a unique origin of replication system. In contrast to that viral double-stranded DNA is a linear DNA with integrated short terminal repeats which indicates persistent infection of the dsDNA virus by a retrovirus. Since bacterial viruses do not use a transport technique as they need in eukaryotes to be transported out of the nucleus, bacterial viruses differ a great deal from eukaryotic viruses.

All bacteria have a restriction/modification system, which is a connected form of two viral competences. Only the descendants of mitochondria lack this system, which causes them not to be exposed to viral selection. It may be that they have transposed their ability to the eukaryotic nucleus, which cares in a more efficient way for cell immunity.

3.2 Bacteria as Biotic Matrix for Natural Genome Editing Competences of Viruses

Horizontal gene transfer between bacteria as being responsible for genetic plasticity in prokaryotes may be a capability, which is acquired by viral infections. Then, viral genetic inventions are transferred to bacteria via persistent lifestyles of viruses and are not an exchange phenomenon performed by bacteria.

As new research indicates the agents of horizontal gene transfer are plasmids, retroplasmids, bacteriophages and transposons. They effect DNA movements and act in all prokaryotes. DNA movement is achieved through transformation, conjugation and transduction. Transformation is the transfer of DNA between related bacteria mediated by encoded proteins. Conjugation is performed by conjugative plasmids, which are independently replicating genetic elements. These elements code for proteins, which facilitate their own transfer. Transduction is a DNA transfer mediated by phages, which can package host DNA in their capsid and inject it into a new host followed by integration into the host genome. Phages, plasmids, retroplasmids and transposons therefore played a crucial role in bacteria evolution. Bacteria are the most genetically adaptable organisms with enormous capabilities to react appropriately to extreme changes of their ecological habitats. This does not stem from their high reproductive rates but from their great ability to acquire DNA segments by plasmids, bacteriophages and transposons which transport complete and complex sets of genes from external sources.

When we consider the age of the ocean and the dense abundance of bacterial and viral life in it, then we can say that the possibility of genetic arrangements, rearrangements and exchange does not need long time periods to create the basics of the complexity of life, because the exchange rate is of astronomical order. If we imagine that 1ml of seawater contains one million bacteria and ten times more viral sequences it can be determined that 10^{31} bacteriophages infect 10^{24} bacteria per second. Since the beginning of life this behavioral pattern has been an ongoing process. The enormous viral genetic diversity in the ocean seems to have established pathways for the integration of complete and complex genetic data sets into host genomes, e.g. acquisition of complex new phenotypes via a prophage can include the acquisition of more than 100 new genes in a single genome editing event.

Owing to the virus-induced genomic plasticity of bacteria they are an ideal global biotic matrix to evolve and develop varieties in genome editing, i.e. competent content arrangement of bacterial gene word order coherent with its regulation network. Bacteria are the smallest living organisms with relatively simple genomic structures where the competitive situation between an abundance of viral infective elements leads to the adaptation of lytic viruses to temperate viruses integrated as plasmids in cytoplasma and even persistent viruses integrated in the host genome. The viral competences can develop in this dense global bacterial habitat as the bacterial species due to their immense genetic flux between viral colonization events and immunity reactions such as restriction/modification.

The highly conserved genome edited functions such as replication, transcription, translation, recombination and all the substeps evolved primarily in the competitive situation between viral competences to colonize a host and to ward off competing parasites. This includes that biotic self and non-self recognition functions, as we know it from diverse immunity systems are also of viral origin, i.e. the integration and all genetic/genomic modification steps that what we call natural genome editing are of viral origin. Therefore the immense importance of horizontal gene transfer for bacterial species evolution, diversity and competences is derived from viral genome editing competences and is, in most cases, infection induced by persistent non-lytic viruses. As phylogenetic analyses demonstrate, the main protein enzymes for natural genome editing are viral inventions and not of cellular origin. Also, the origin of eukaryotic nucleus was thought to be an ancient prokaryote but phylogenetic analyses show that its ancestor seems to be a large DNA virus. Interestingly, the early genetic invention of capsid proteins detected in viruses infecting archaea seems also to be of viral origin and of common ancestry to eukaryotic and bacterial viruses.

3.3 From Prebiotic Assemblies into Functional Agents

For a long time bacteria have been considered to be the forerunners of the eukaryotic superkingdom. Although the evolution of eukaryotes did not occur by random mutations of bacterial genomes but by integration and natural genetic engineering of former free-living prokaryotes, the key features of the eukaryotic nucleus have less in common with prokaryotic competences than with some double-stranded (ds) DNA viruses (Witzany, 2006). The textbook conviction of the early 21st century on the evolutionary history of eukaryotes was that an ancient prokaryotic cell was colonised by a large dsDNA virus and afterwards by mitochondria-like and chloroplast-like bacteria, which together built the first eukaryotic cell. This scenario makes sense from a cytological perspective, because prokaryotes are much simpler than eukaryotes. From the perspective of an early RNA world, however, this view changes.

A "virus-first"-scenario from biocommunicative perspective would look like this: At the beginning there were single-stranded unencapsulated RNA molecules with an aptitude to replicate themselves, which through both their coding and catalytic capabilities, built structures with multiple functions to form dsRNA genomes in a pre-DNA world. If we term these pre-cellular RNA replicators as viruses then ssRNA viruses evolved into dsRNA viruses. Via a reverse transcriptase function present in a RNA-dependent RNA-polymerase these dsRNA viruses evolved later on into dsDNA viruses. Now the stable DNA of dsDNA viruses was advantageous for colonising the unstable nucleotide word order in the genomic contents of RNA viruses. In parallel, DNA of dsDNA viruses served as an appropriate habitat for infection events by retroid agents. By holding these colonisation interrelations in a non-lytic but persistent inheritable status, infection forced the colonised RNA viruses to establish a bi-layered cell membrane and to encapsulate the genome in a porous nuclear envelope. Currently this could be a coherent explanation for the three remaining cellular DNA replication competences, i.e., the beginning of cellular life would have been entangled completely with three population-like genetic lineages, similar to those suggested by Forterre and co-authors.

These steps from ssRNA to dsRNA and from dsRNA to DNA are hallmarks in the evolution of life from prebiotic assemblies of ribonucleotides into functional agents with simple nucleotide grammar-editing abilities. However, these agents also had to include a self/non-self differentiation capability that means identity, being able to ward off competing agents through the first immune function similar to RNAi. In parallel this would have been an advantage for colonising RNA replicators that lacked this capability.

3.4 Communal Evolution

For a long time the first living cell has been imagined as a single agent, known as the last universal common ancestor (LUCA), which was the forerunner of all later evolutionary steps into the three domains of life. According to our bio-communicative scenario the pre-cellular agents that were competent in RNA-, DNA- and Retro-editing were population-like nucleic acid sequence-editing species with commonly shared group behaviour (self/non-self identification), group identity, and group interpretation.

This would be coherent with linguistic research, which states that (i) every biotic agent that is competent to use and interpret linguistic-like signs needs a community (either as a gene-pool and/or interactional) with which it shares these capabilities, i.e. linguistic competences cannot emerge in isolated individuals and (ii) if linguistic-like competence for editing of nucleic acid languages evolved, the capability to generate not only simple but also complex new sequences would grow exponentially and not arithmetically.

This could explain both the great diversity of single-celled life soon after pre-cellular life-processes started and the common feature of evolutionary processes, i.e. the invention of new genetic data as whole sequences, genes or gene-blocks. When the pre-cellular consortium of three different viral - or even subviral - lineages developed a common genetic code, which further on served as a stable DNA storage medium for the evolution of the three cellular domains, I suggest we should be talking about the last universal common ancestors (LUCAs), because one single ancestor couldn't evolve both (i) sequence editing competences (ii) a competence for sign-mediated interaction necessary for coordination of common behaviour (group identity) and self/non-self identification.

3.5 Current Competences from an Ancient World

Interestingly, even today we can look at relics of pre-cellular evolution in both RNA viruses and viroids. Viroids and their monophyletic sister group, satellite RNAs, are short circular ssRNAs, viroids being unencapsulated whereas satellite RNAs are encapsulated. We know that viroids have extreme plasticity in their nucleotide sequences, being the most rapidly evolving biological agents. Important features of small RNAs such as RNA silencing seem to derive from viroid competences. The most conserved competences of RNA viruses and viroids are RNA stem-loop structures, which play important roles in priming and replication, with an inherent self/non-self differentiation in that they determine RNA replication to viral and not to host RNA molecules.

We now can imagine eukarya-like dsDNA viruses with the ribozymatic function of endonucleases competent in RNA-splicing, excision of introns out of tRNAs, integration of

retroid DNA and its key features, a double membrane, linear chromosomes with telomere ends, intronic elements with regulatory functions, segregation of transcription and translation and the subviral competences which we find in the ribonucleoprotein structures of pre-mRNA, pre-tRNA and pre-rRNA, all processed by small nucleolar (sno)RNAs and small nuclear (sn)RNAs. As we know today, the precursor RNAs are a highly sophisticated network of regulatory agents, each of them with a separate RNA processing pathway. Although in prokaryotes we do not find linear chromosomes with telomere repeats, the ancient nuclear pore complex or the highly mobile genetic settlers inherent in introns that are competent in RNA-splicing, we do find them in eubacterial and archaeal phages.

In addition, prokaryotes share a circular genome with nearly intron-free genetic syntax, whereas the seemingly evolutionarily later eukaryotes have linear chromosomes with telomere repeats to protect their ends against genetic invaders and genomes that are highly colonised by virus-derived agents such as transposons, retroposons and related genetic settlers.

Although the "error-prone" coding-fidelity of the RNA world at the beginning was an advantage for fast adaptation, the evolutionary target evolved into both the relatively stable DNA configuration (via the reverse transcriptase competence - the only encoded function common to all retroelements) and the resistant protein world necessary in the high temperature environments of archaeal populations. Prokaryotes lack the key features of the early RNA world and therefore they would appear to be specialised fast-adapting single-celled organisms that used the advantages of the stable DNA storage medium to code for highly temperature resistant protein structures to protect this storage medium.

Although accelerated ssRNA processing of mRNA, tRNA and rRNAs in linear RNA genomes built core competences for natural genome editing in the early RNA world, those ssRNAs without cellular habitats are extremely thermolabile and could not survive in high temperature environments. The lack of RNA correction and repair and the high rate of replication combined with innovation allowed a rate of recombination events $1–10 \times 10^6$ times faster than in DNA genomes. RNA-based life forms could evolve millions of times faster than DNA-based systems. This was an advantage for the exploration and invention of new sequence space, i.e. new genomic content with phenotypic competences and functions. In contrast, circular genomes with few higher order regulatory elements (represented by a diversity of genetic parasites present in intron-like genomic habitats) had more advantages in a high temperature environment and could adapt faster because of their ability to exchange selected phenotypes within and between protein-coding data-sets, as happens in horizontal gene transfer. So RNA cultures with eukaryote-like RNA-processing seem to predate the evolution of prokaryotes, which adapted to fast-changing environmental conditions by reducing their genomic content to a DNA with nearly analog (intron-free) protein-coding data-sets. This could be the evolutionary pathway from ribozymes of the early RNA world to ribonucleoproteins via low complexity RNA-chaperones to a DNA-protein-based life. That eukarya-like genomes predated prokaryotic genomes is consistent with the existence of telomeres and telomere-like functions in ancient dsDNA viruses that seem to be the ancestors of the eukaryotic nucleus and is not part of prokaryotic genomes, although some are found in persistent bacteriophages.

Multiple small regulatory RNAs also play important roles in the bacterial expression of target genes at the post-transcriptional level. They are immediately available after being transcribed from the non-protein-coding sections of bacterial genomes, unlike protein enzymes, which must be translated as well. From the perspective of evolutionary history, bacteria seemed to reduce the predated RNA-based metabolism of early eukarya-like genetic content arrangements to become specialised in highly-selective environmental conditions such as high temperature and/or fast-changing nutrient availability, dependent on nearly intron-free DNA-protein metabolism, and containing circular genomes with only one starting-point for replication. As intron-rich linear chromosomes are the preferred habitat for persistent retroviral infections, and because of their important role in host-genetic content (re) arrangements, the invention of bacterial circular genomes must have had an effective immune function against retroviral infections. The result was the evolution of organisms that successfully escaped the high selective pressures of the early RNA world.

FURTHER READING

Bordenstein, S.R., Reznikoff, W.S. **2005**, Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol*. 3, 688-699

Brown, J.R. **2003**, Ancient horizontal gene transfer. *Nat. Rev. Genet*. 4, 122-132

Chen, I., Christie, P.J., Dubnau, D. **2005**, The ins and outs of DNA transfer in bacteria. *Science* 310, 1456-1460

Daros, J.A., Elena, S.F., Flores, R. **2006**, Viroids, an Ariadne's thread into the RNA labyrinth. *EMBO Rep*. 7, 593-598

Flores, R., Delgado, S., Gas, M.E., Carbonell, A., Molina, D., Gago, S., De la Pena, M. **2004**, Viroids, the minimal non-coding RNAs with autonomous replication. *FEBS Lett*. 567, 42-48

Forterre, P. **2002**, The origin of DNA genomes and DNA replication proteins. *Curr. Opin. Microbiol*. 5, 525-532

Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A. **2005**, Mobile genetic elements, the agents of open source evolution. *Nat. Rev. Microbiol*. 3,722-732

Gogarten, J.P., Townsend, J.P. **2005**, Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol*. 3, 679-687

Hambly, E., Suttle, C.A. **2005**, The viriosphere, diversity, and genetic exchange within phage communities. *Curr. Opin. Microbiol*. 8, 444-450

Hammer, B.K. Bassler, B.L. **2007**, Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic Vibrio cholerae. *Proc. Nat. Acad. Sci*. USA 104, 11145-11149

Imaizumi-Anraku, H. *et al*. **2005**, Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. *Nature* 433, 527-531

Jain, R., Rivera, M.C., Lake, J.A. **1999**, Horizontal gene transfer among genomes, The complexity hypothesis. *Proc. Nat. Acad. Sci*. USA 96, 3801-3806

Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M., Johnson, J.E. **2005**, Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Nat. Acad. Sci*. USA 102, 18944-18949

Majdalani, N., Vanderpool, C.K., Gottesman, S. **2005**, Bacterial small RNA regulators. *Crit. Rev. Biochem. Mol. Biol*. 40, 93-113

Masse, E., Gottesman, S. **2002**, A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Nat. Acad. Sci*. 99, 4620-4625

Pandey, D.P., Gerdes, K. **2005**, Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucl. Acids Res*. 33, 966–976

Penny, D., Poole, A. **1999**, The nature of the last universal common ancestor. *Curr. Opin. Genet. Dev*. 9, 672-677

Rivera, M.C., Lake, J.A. **2004**, The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152-155

Tettelin, H. *et al.*, **2005**, Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*, implications for the microbial "pan-genome". *Proc. Nat. Acad. Sci.* 102, 13950-13955

Thomas, C.M., Nielsen, K.M. **2005**, Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711-721

Vetsigian, K., Woese, C., Goldenfeld, N. **2006**, Collective evolution and the genetic code. *Proc. Nat. Acad. Sci.* USA 103, 10696-10701

Villarreal, L. **2008** (ed), *Origin of group identity, viruses, addiction and cooperation.* Springer

Villarreal, L.P. **1999**, DNA Virus Contribution to Host Evolution. In, Domingo, E., Webster, R.G., Holland, J.F. (eds), *Origin and evolution of viruses.* London, Academic Press, pp 301-420

Wang, M.B., Bian, X.Y., Wu, L.M., Liu, L.X., Smith, N.A., Isenegger, D., Wu, R.M., Masuta, C., Vance, V.B., Watson, J.M., Rezaian, A., Dennis, E.S., Waterhouse, P.M. **2004**, On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites. *Proc. Nat. Acad. Sci.* USA 101, 3275-3280