Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Virology

The major role of viruses in cellular evolution: facts and hypotheses

Patrick Forterre[1,2] and David Prangishvili[1]

Viral particles are much more abundant than cells and viral genes outnumber cellular ones in the biosphere. Cellular genomes also harbour many integrated viruses whereas cellular genes are rare in viral genomes. The gene flux from virus to cell is thus overwhelming if compared with the opposite event. Novel viral genes continuously arose during replication/recombination of viral genomes in the virocell. These genes can become 'cellular genes' when viral genomes integrate into cellular ones. Together with the arm race between viruses and cells, this explains why viruses have played a major role in shaping cellular gene contents. Several documented cases show that viruses have been involved in the emergence of evolutionary innovations. This gives credit to hypotheses suggesting that viruses have played an important role in the formation of modern cells.

Addresses
[1] Institut Pasteur, 25 rue du Docteur Roux, Paris 75015, France
[2] Univ Paris-Sud, CNRS UMR8621, Orsay Cedex 91405, France

Corresponding author: Forterre, Patrick (forterre@pasteur.fr)

Current Opinion in Virology 2013, **3**:xx–yy

This review comes from a themed issue on **Virus evolution**

Edited by **Valerian Dolja** and **Mart Krupovic**

1879-6257/$ – see front matter,

http://dx.doi.org/10.1016/j.coviro.2013.06.013

Introduction

Microbial ecologists have shown that viral particles are much more abundant than cellular organisms on earth and that viral genes outnumber cellular ones in metagenomes [1,2,3•,4•]. Moreover, genomics studies revealed an extraordinary abundance of virus and related mobile elements in cellular genomes [5–9,10••]. Finally, viruses turned out to be more diverse and ancient that previously suspected, suggesting that they were already present at the time of the last common ancestor and even further back in time [11–14,15••]. These observations logically suggest that viruses have played a major role in the evolution of cellular organisms [3•,12–14,16,17,18••,19]. However, this view is still disputed by some scientists who argue that viruses are mainly passive vehicles of cellular genes. For instance, Moreira and Lopez-Garcia wrote that: '*the cell to virus gene flux is overwhelming if compared with the opposite event: this suggests that viruses have played only a minor role in shaping the gene content of cells*' [20]. We will contradict this view here, assuming instead that the natural gene flux mainly occurs from viruses to cells. We will present several examples demonstrating how the integration of viral genes into cellular genomes could have dramatically affected cellular evolution. These examples illustrate the fact that viruses play a major, not a minor role, in cellular evolution, and make credible several hypotheses regarding their possible roles in the emergence of modern organisms.

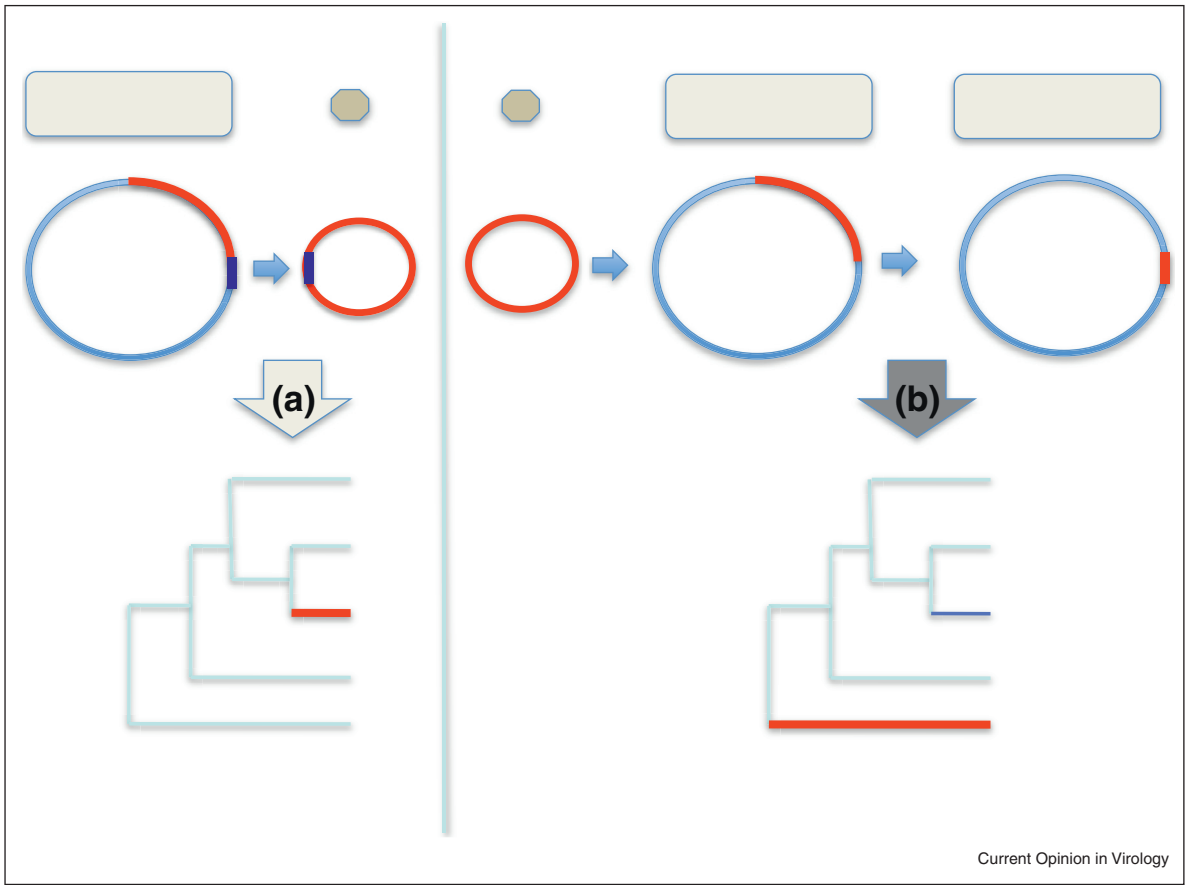
The capture of cellular genes by viruses and plasmids is a rare event

It is often assumed that viruses evolve by capture and accretion of cellular genes (the virus pickpocket paradigm) and that a major role of viruses in cellular evolution is to facilitate the lateral gene transfers (LGT) of cellular genes between cellular lineages [20]. However, in contradiction with this view, the percentage of genes encoding a protein with cellular homologues is rather low in viral genomes [9,21–23]. This percentage increases with genome size, but can be very low (or null) for genomes in the range of 5–50 kb (for instance, none of the genes encoded by the archaeal viruses from the families *Globuloviridae* [24] and *Clavaviridae* [25] has a homologue in the databases). In large viral DNA genomes, this percentage can reach 10% or more, but the majority of genes remain ORFans or viral specific [22,23]. Furthermore, the origin of viral proteins with cellular homologues is itself controversial. In agreement with the pickpocket paradigm, it is usually assumed that all of them derived from cellular proteins captured by a viral ancestor [26]. However, these proteins can also derived from viral proteins whose genes have been captured by cellular genomes (Figure 1) [23]. For instance, Hamacher and co-workers recently suggested that red and green algae recruited their potassium ion channel proteins from a phycodnavirus [27••].

It is sometimes possible to determine the sense of the transfer by phylogenetic analysis: the nesting of viral genes within cellular sequences argues in favour of a cell to virus transfer, whereas the location of a 'cellular' gene on a provirus suggests a virus to cell transfer (Figure 1a). However, it is often not possible to conclude about the direction of transfer because viral and cellular genes are very divergent and/or because all traces of a possible ancestral provirus have disappeared (Figure 1b). Unfortunately, very few studies have been performed to quantify the flux of cellular genes to viruses and vice versa based on comprehensive phylogenetic analyses and

Please cite this article in press as: Forterre P, Prangishvili D. The major role of viruses in cellular evolution: facts and hypotheses, Curr Opin Virol (2013), http://dx.doi.org/10.1016/j.coviro.2013.06.013

Figure 1



Lateral gene transfers (LGT) from cells to viruses (left panel) and from viruses to cells (right panel); cellular genomes and genes are coloured in blue, viral genomes and genes in red. LGT from cells to viruses usually transfers into viral genome only a small piece of DNA adjacent to the integration site; **(a)** in phylogenetic analysis, viral proteins of cellular origin are usually nested within cellular organisms related to their hosts (see an example in Figure 2). LGT from viruses to cells can transfer massive amount of DNA into cellular genomes, viral genes can become cellular by fixation; **(b)** in phylogenetic analyses, cellular genes of viral origin appear often as outgroup of cellular organisms related to their hosts (e.g. see Ref. [27••]). The viral origin can only be rigorously demonstrated for recent transfer when the gene is still embedded into an integrated viral genome.

identification of viral genes integrated into cellular genomes. However, preliminary data confirm the paucity of cellular genes in viral genomes. For example, an exhaustive analysis of 136 archaeal and bacterial genomes identified only 1.2% of cellular genes (defined as core genes conserved in all genomes of a given lineage) in viral database, and only 3.2% in a database of integrated viruses and plasmids (see Figure 4 in Ref. [9]). The rarity of cellular genes in viral genomes is logical considering the mechanism used by viruses to disseminate their genomes: the production of virions. The virions size necessarily limits the amount of foreign DNA that can be integrated into viral genomes, except if this DNA replaces a portion of the viral genome, something difficult without compromising virus viability. These constraints are less pronounced with increasing virion size, explaining why the proportion of cellular genes is higher in large DNA viruses. In particular, some members of the order *Caudavirales* and proposed order *Megavirales* exhibit a high proportion of bacterial genes at the extremities of their chromosomes, due to a specific mechanism of gene capture linked to the replication of their linear chromosomes [22].

The paucity of cellular genes in viral genomes, together with the strict virus/host specificity, suggests that viruses play a lesser role than usually assumed in LGT between cells. For instance, phylogenetic analyses of archaeal the replicative helicases MCMs have shown that firstly, MCM proteins encoded by viruses and plasmids systematically branch with MCM proteins of their hosts and secondly, the MCM tree is congruent with the archaeal phylogeny based on ribosomal proteins (Figure 2) [28••]. These results confirm that viruses and plasmids coevolved with their hosts and were not involved in LGT between distantly related archaea. Viruses mediated

Please cite this article in press as: Forterre P, Prangishvili D. The major role of viruses in cellular evolution: facts and hypotheses, Curr Opin Virol (2013), http://dx.doi.org/10.1016/j.coviro.2013.06.013