Plasmids, viruses and virus-like membrane vesicles from Thermococcales

Nicolas Soler*, Marie Gaudin†, Evelyne Marguet* and Patrick Forterre*‡

†Institut de Génétique et Microbiologie, Univ Paris-Sud 11, UMR 8621, CNRS 91405, Orsay, France, and §Institut Pasteur, 25 rue du Dr Roux, 75015, Paris, France

Abstract

Several families of plasmids and viruses (PVs) have now been described in hyperthermophilic archaea of the order Thermococcales. One family of plasmids replicates by the rolling circle mechanism, whereas most other PVs probably replicate by the θ mode. PVs from Thermococcales encode novel families of DNA replication proteins that have only detectable homologues in other archaeal PVs. PVs from different families share a common gene pool and co-evolve with their hosts. Most Thermococcales also produce virus-like membrane vesicles similar to eukaryotic microparticles (ectosomes). Some membrane vesicles of Thermococcus nautilus harbour the plasmid pTN1, suggesting that vesicles can be involved in plasmid transfer between species.

Introduction

A plethora of plasmids has been described so far in Bacteria, with sizes ranging from a few to several hundred kilobases, some of them blurring the border between megaplasmids and chromosomes (chromids sensu [1]). Although much less studied, ‘extrachromosomal circle DNA’ seems to be also widespread in Eukarya, including humans [2]. In Archaea, plasmids have been found quite frequently, either in Euryarchaeota or in Crenarchaeota [3]. Interestingly, the study of archaeal plasmids has revealed extensive evolutionary relationships between plasmids and viruses (PVs). Hybrid elements combining viral and plasmid genes, pSSVx and pSSVi, have been isolated in Sulfolobus species [4,5]. Recently, Bamford and colleagues have described closely related viruses in Haloarchaea with either single- or double-stranded DNA genomes [6]. These viruses are themselves highly similar to cryptic plasmids and integrated PVs present in other haloarchaeal species. Analysis of PVs integrated into archaeal and bacterial genomes has indeed indicated that sequence spaces of plasmids and viruses greatly overlap [7].

The origin and evolution of PVs, their role in cellular evolution and their position in the tree of life are controversial issues [8–10]. The study of archaeal PVs should help to fuel the debate. Indeed, with a robust phylogeny of the archaeal domain being available [11], it should be easier to test in this domain to which extent PVs co-evolve with their hosts and how this interplay affects genome evolution. As a prerequisite, this requires the exploration of all major archaeal groups for PVs. We have undertaken such a task a few years ago for the order Thermococcales, in collaboration with the laboratory of Daniel Priere in Brest [12]. Thermococcales (Thermococcus, Pyrococcus and Palaeococcus genera) are model archaea to study life at high temperature. They are heterotrophs, anaerobes and grow at temperatures between 70 and 105 °C. Thermococcales are relatively easy to cultivate and appear first in enrichment cultures from hot spring samples performed in rich medium [13]. Examination by electron microscopy of enrichment cultures performed in conditions that favour Thermococcales also suggests the existence of a diverse virosphere associated with these archaea in deep sea hydrothermal vents [14].

Approx. 40% of Thermococcales strains that have been screened for PVs harbour at least one extrachromosomal element [12,13]. Until now, only one virus and six plasmids have been described, although many more are in the pipeline. Despite this low number, the analysis of PVs from Thermococcales has already provided valuable data, including the discovery of a new DNA polymerase family [15]. Originally, we were interested in using plasmids from Thermococcales as probe to study the intracellular DNA topology of hyperthermophiles [16]. More recently, some plasmids isolated in our groups have been used to develop genetic tools [17–19]. In the future, these plasmids should be also useful to study different molecular systems (e.g. DNA replication) at high temperature. In the present review, we summarize data already obtained on PVs from Thermococcales and their applications. We also discuss how their analyses can be interpreted in the framework of current debates on PV evolution. Recently, we have observed that most Thermococcales species produce virus-like membrane vesicles, some of them strongly associated with extracellular DNA [20]. We show that some of these vesicles contain plasmidic DNA, raising new questions on their possible role in gene transfer or in the generation of new viral forms.
Plasmid DNA topology in hyperthermophiles

The first plasmid isolated from Thermococcales was the small plasmid pGT5 (3444 bp) from Pyrococcus abyssi GE5 [21]. pGT5 was first used as reporter for the topological analysis of intracellular DNA at high temperature [16]. The discovery of reverse gyrase in Sulfolobus, an enzyme which introduces positive superturns in circular DNA duplex in vitro, indeed suggested in the 1980s that intracellular DNA was stabilized by positive supercoiling in hyperthermophiles [22,23]. This hypothesis was supported by the positively supercoiled state of the genome of the Sulfolobus virus SSV (Sulfolobus spindle-shaped virus) 1 isolated from infected cells [24]. The topological analysis of pGT5 DNA did not confirm this view, since purified pGT5 turned out to be slightly negatively supercoiled in vitro [16]. Further analyses, using either pGT5 or various Sulfolobus plasmids, revealed that intracellular plasmid topology changes with temperature, the plasmid becoming more negatively supercoiled after a cold shock [25] and slightly positively supercoiled after heat shock [26]. An increase in the linking number (relaxation or positive supercoiling) with increasing temperature turned out to be a general property of bacterial or archaeal plasmids (for a review, see [27]). Further studies revealed that reverse gyrase apparently does not determine intracellular DNA topology. Indeed, plasmids from organisms containing both gyrase and reverse gyrase are negatively supercoiled, whereas plasmids isolated from organisms lacking gyrase (reverse gyrase being present or not) are more or less relaxed [27]. The existence of hyperthermophiles with negatively supercoiled plasmids, such as Archaeoglobus and Thermotoga, is not so surprising, since a topologically closed circular DNA, either positively or negatively supercoiled, is highly resistant to thermal denaturation [28]. The role of reverse gyrase remains a mystery (for recent reviews, see [29,30]). However, reverse gyrase is most likely to be required for some crucial adaptation to high temperature, since it is ubiquitous in hyperthermophiles [31], and a reverse-gyrase-knockout mutant of Thermococcus kodakaraensis is thermostable [32]. Novel plasmids recently isolated from Thermococcales should be helpful to resume DNA topological studies in organisms harbouring reverse gyrase.

The pGT5 RC (rolling circle) plasmid family

The plasmid pGT5 encodes two proteins of 46 and 75 kDa [33]. The larger one (Rep75) contains the three amino acid motifs (1, 2 and 3) characteristic of RC replication initiator proteins [34]. Rep proteins are site-specific endonuclease-ligases that initiate DNA replication by introducing a nick at the double-stranded replication origin (dso), producing a free 3′-hydroxy group that serves as a primer for the host DNA polymerase. This reaction is catalysed by a tyrosine residue present in motif 3. The Rep protein remains covalently linked to this 3′ extremity after cleavage, and the energy of this link is used for re-ligation. The RC mode of replication of pGT5 was confirmed by the detection of the circular single-stranded intermediate in P. abyssi cells [33] and biochemical characterization of Rep75 [35–37]. The recombinant purified Rep75 protein cleaves specifically in vitro an oligonucleotide containing the 11 bp sequence of the dso found within the rep75 gene (Figure 1). This sequence is similar to the dso of bacterial RC plasmids of the pC194 family. Rep 75 also exhibits DNA ligase and type I DNA topoisomerase activity in vitro [35,37]. In addition, Rep75 exhibits an unusual deoxynucleotidyldtransferase activity [35]. Only one dATP or ATP can be incorporated at the 3′-OH end of the cleavage site, reflecting that an adenosine is needed at the 5′-P end for effective ligation. Specific mutations of the protein have also been studied, demonstrating that motif 3 is essential for all Rep75 activities [36].

Recently, a second RC plasmid related to pGT5, pTN1, has been isolated from a Thermococcus species (T. nautilus 30/1) [38]. pTN1 encodes a protein of 74 kDa, Rep74, which is highly similar to Rep75. Rep74 and Rep75 are much larger than typical RC proteins (30–40 kDa on average). The three RC motifs are located in the middle of these proteins, suggesting that they are formed by the fusion of a central RC initiator module with two additional modules of uncharacterized function in the N- and C-termini. PSI-BLAST analysis showed that the RC module of Rep74/75 proteins is distantly related to transposases encoded by bacterial IS (insertion sequence) elements using RC mechanism for transposition. A fourth motif downstream of motif 3 is indeed conserved with these RC transposases [38].

In addition to their Rep proteins, pGT5 and pTN1 encode unrelated proteins of 46 and 24 kDa (p46 and p24) respectively. The protein p46 of pGT5 exhibits weak similarities with transcriptional regulators and a homologue is present in the genome of the crenarchaeon Ignisphaera aggregans (Iag_1977) (P. Forterre, unpublished work). The p24 protein of pTN1 harbours a zinc-finger domain, a highly charged region and a hydrophobic segment. A truncated soluble p24 protein (lacking the hydrophobic segment) was purified and turned out to bind double- and single-stranded DNA without sequence specificity [38]. This soluble protein also compacts and aggregates DNA at a high protein/DNA ratio [38]. Deletion of the p24 gene has no effect on the stability of a shuttle vector derived from pTN1 [18]. The p24 protein that can bind both DNA and membrane could be involved in the packaging of pTN1 in membrane vesicles produced by T. nautilus (see below).

The plasmids pGT5 and pTN1 have been used to design shuttle vectors for Thermococcales. pGT5 was first engineered to produce pYS2, a plasmid that could replicate in P. abyssi using the pyrE gene of Sulfolobus acidocaldarius as a selection marker [17]. Recently, this vector was improved by Hausner’s group and used to transform Pyrococcus furiosus [19]. The pyrE gene was replaced by a gene conferring resistance to the antibiotic simvastatin. The plasmid pTN1 was engineered by Reeve’s group to produce pLC70, a plasmid that could replicate in T. kodakaraensis, using
Figure 1 | Schematic representation of characterized (pGT5/pTN1) or putative (others) replication cassettes of Thermococcales plasmids, viruses or virus-related region inserted in Thermococcales genomes

Homologous genes are drawn with the same colour. Yellow circles represent described (pGT5) or putative (others) replication origins. ORFs surrounded by a bold line have an inverted orientation regarding genes located on the other side of the putative origins, and then could be involved in replication or its regulation.

The trpE gene of T. kodakaraensis and an HMG-CoA (3-hydroxy-3-methylglutaryl-CoA) gene as selection markers [18]. The recombinant plasmids based on pGT5 and pTN1 can be used to express proteins in their respective archaeal hosts. As proof of principle, they were used to purify by affinity-chromatography RNA polymerase holoenzymes using His6-tagged subunits overexpress in P. furiosus and T. kodakaraensis cells respectively [18,19]. The pTN1-based vectors have been also used successfully to study various aspects of translation and transcription in Archaea [18,39].

The pRT1/pAMT11 plasmid family

A third small plasmid, pRT1 (3373 bp) isolated from the Pyrococcus strain JT1, resembles pGT5 and pTN1 in size and genome organization [40]. This plasmid encodes two proteins of 63 kDa (probably the plasmid Rep protein) and 29 kDa. The protein p29 has an homologue encoded in the genome of PAV1 (Pyrococcus abyssi virus 1) (see below). Rep63 was first described as homologous with Rep75 from pGT5 [40]. However, Rep63 cannot be aligned convincingly with both Rep74 and Rep75 [38] and lacks clear-cut motifs 1, 2 and 3 in the proper orientation. Recently, Rep72, a homologue of Rep63 has been detected in the plasmid pAMT11 (20534 bp) from a new Thermococcus species [41]. Although pAMT11 is larger than typical RC plasmids, Gonnet and co-workers could detect a sequence identical with the dso of some RC plasmid upstream of the Rep72 gene (Figure 1). The mode of replication of pRT1/pAMT11 thus remains presently unclear. In any case, Rep63 and Rep72 appear to be the prototypes of a new family of Rep proteins and interesting models for future biochemical studies. Most pAMT11 ORFs (open reading frames) have homologues in the integrated element TKV1 (Thermococcus kodakaraensis virus 1; 23800 bp) from T. kodakaraensis. However, the Rep63 homologue present in pAMT11 is replaced in TKV1 by a gene encoding a MCM (minichromosome maintenance)-like helicase at the corresponding position in the alignment between these two elements [41].
PVs that replicate by the θ mode
PV1 from *P. abyssi* GE23 and three plasmids (pTN2 from *T. nautilus*, pP12-1 from *Pyrococcus* sp. 12-1 and pT26-2 from *Thermococcus* 26-2) probably replicate by the θ mode [15,42,43]. The genome of PV1 and these three plasmids are larger than RC plasmids (between 12 and 21.5 kb) and do not encode RC Rep protein. Furthermore, their predicted replication origins resemble those of bacterial plasmids that replicate via the θ mode. These origins have been detected by the cumulative GC skews method [44]. Their positions correspond to inflexions in the cumulative GC skew curves and they are found in the largest intergenic region of each element. These putative origins are also AT-rich, corresponding to a region of low stability, and they contain multiple direct and inverted repeated sequences, similar to iterons present in the replication origins of bacterial θ plasmids [45]. In PV1 and the three plasmids, most genes are transcribed in the same direction. Interestingly, one or two genes transcribed in the opposite direction are often located close to the replication origin (Figure 1). These genes could encode proteins involved in the control of plasmid replication.

**PAV1**
Many viruses whose virions exhibit unique morphotypes have been characterized from hyperthermophilic crenarchaea [46]. In contrast, PV1 is presently the only known virus for hyperthermophilic euryarchaea (*P. abyssi* GE23) [42]. This virus is not lytic, and infectivity has not yet been demonstrated. The host cell spontaneously releases PAV1 particles, with production increasing in stationary phase. The PAV1 genome is present as a high-copy-number plasmid in *P. abyssi* GE23. PAV1 strikingly resembles the SSVs from *Sulfolobus* spp., both in genome size (a double-stranded circular genome of 18098 bp) and in virion morphology (a spindle-shaped particle reminiscent of *Fuselloviridae*) [42]. However, none of its 25 predicted proteins exhibits similarity with SSV proteins [43]. Instead, PAV1 is evolutionarily related to plasmids of Thermococcales, with three genes having homologues in the plasmids pTN2, two in the plasmid pP12-1 and one in pRT1. In addition, PAV1 shares one gene with the provirus TKV4 of *T. kodakaraensis* [47].

Transcript analysis by Q-PCR (quantitative PCR) has shown that 16 of the 25 PAV1 predicted proteins (including the major structural protein) are co-transcribed. Ten of these proteins, which harbour hydrophobic segments, are probably involved in formation of the virion and/or interact with the host cell membrane. Two large proteins contain LamG domains, suggesting a role in cell adhesion. This operon also encodes an ATPase with homologues in plasmids pTN2 and pP12-1. This ATPase could be involved in the segregation of the viral genome between daughter cells. The remaining nine proteins (most of them small) are encoded by five transcripts from genes located on both sides of the putative replication origin. These proteins are likely to be involved in the regulation of the virus life cycle. Interestingly, one of them encodes a relatively large DNA-binding protein (Figure 1) that has distant homologues in several genomes and large plasmids of halocarchaea (P. Forterre, unpublished work). These proteins could correspond to a novel type of Rep proteins involved in the initiation of DNA replication by the θ mode.

**The pTN2 plasmid family**
The plasmids pTN2 from *T. nautilus* and pP12-1 from *Pyrococcus* sp12-1 share six homologous genes, two of them encoding a SF (superfamily) I helicase and a new type of DNA polymerase respectively [15]. These two genes and a *copG* like gene are clustered close to the putative replication origin (Figure 1). SF1 helicases of pTN2 and pP12-1 are closely related to helicases encoded by mobile elements present in the genomes of *Thermococcus gammatolerans* [TGV2 (*Thermococcus gammatolerans* virus-related region 2), see below] and of *Thermococcus onnurineus* [15]. The DNA polymerases of pTN2 (whose activity has been demonstrated experimentally) and pP12-1 are distantly related to Rep proteins encoded by some *Sulfolobus* plasmids and by PVs integrated in the genomes of Methanococcales. Alignment of these proteins allows defining three motifs in their N-terminal regions [15]. The motifs 1 and 3 include the three aspartate residues present in the active sites of enzymes from the AEP (archaeoeukaryotic primase) and related palm domain protein superfamily [48]. Indeed, the pTN2 polymerase also exhibits primase and nucleotidyltransferase activities in vitro (G. Sezanov, C. Desnoue and P. Forterre, unpublished work). The DNA polymerase/primase of the *Sulfolobus* pRN plasmids also belong to the AEP superfamily. However, the pRN and pTN2/pP12-1 DNA polymerases cannot be recovered by reciprocal BLAST searches. The pTN2/pP12-1 type polymerases thus form a new family of DNA polymerases/primases within the AEP superfamily.

**The pT26-2 plasmid family**
The plasmid pT26-2 (21566 bp) encodes 32 proteins, including a SSV1-like integrase, two ATPases and a resolvase [15]. The gene encoding one of the two ATPases is located just upstream of the putative replication origin (Figure 1). This ATPase is probably a replicative helicase, since homologous ATPases replace either SFIII helicases or MCM helicases in bimodular Rep proteins of several *Sulfolobus* pRN plasmids (where they are fused to DNA polymerase/primase). This pT26-2 ATPase is also replaced by a MCM-like helicase in a set of contiguous genes whose syntheny is conserved between pT26-2 and TKV1 [15]. The number of genes common to pT26-2 and TKV1 is limited, and these genes have no homologues in pAMT11. In contrast, 19 genes of pT26-2 have homologues in two other integrated PVs of *T. kodakaraensis*, TKV2 (27301 bp) and TKV3 (28020 bp) [49]. Closely related PVs are also present in the genomes of *T. gammatolerans* (TGV1, 22802 bp) [50] and...
Pyrococcus horikoshi (PHV1, 21901 bp) (Figure 2). All of these PVs are inserted in tRNA genes and framed by the N-terminal and C-terminal moieties of pTN2-like integrases, a mode of insertion typical of archaeal type I integrases and first described in the case of Sulfolobus SSV [51]. All other PVs detected in Thermococcales genomes, such as TKV1, TKV4 (see below) and TGV2 (see below), use similar integrases. Interestingly, seven genes of pT26-2 have homologues not only in integrated PVs of Thermococcales, but also in integrated PVs of Methanococcales [15] (Figure 2). Homologues of these genes are also present in a free plasmid recently detected in Methanocaldococcus fervens (pMEFER01, direct GenBank® submission) (Figure 3). Most of these genes encode proteins with hydrophobic segments, being reminiscent of the similar cluster observed in the PAV1 genome. Finally, several genomes of Methanococcales encode only the two larger proteins common to all PVs of the pT26-2 family, suggesting that these two proteins have been recruited for cellular functions. The structure of one of them (from pT26-2) has been solved, revealing three new folds in a single polypeptide [52].

The provirus TKV4
A fourth *T. kodakaraensis* integrated PV, TKV4 (18818 bp), encoding 33 proteins, has no sequence similarity with TKV1, 2 or 3, except for a MCM-like helicase (Figure 1) shared with TKV1. However, phylogenetic analysis has shown that the MCM helicases of TKV1 and TKV4 have been recruited independently from ancient *Thermococcus* species [53]. One of the TKV4 proteins has a homologue in PAV1. Importantly, Krupovic and Bamford [47] have shown that TKV4 is a provirus member of the viral lineage defined by the bacteriophage PRD1. This lineage also includes human adenovirus, eukaryotic NCLDVs (nucleocytoplasmic large DNA viruses) and the crenarchaeal virus STIV (*Sulfolobus* turreted icosahedral virus). All these viruses have in common an MCP (major capsid protein) characterized by the double-jelly-roll fold [54] and a homologous packaging ATPase. Krupovic and Bamford [47] identified the MCP and packaging ATPases of TKV4 using the STIV MCP as query sequence and genome context analysis, extending the presence of the PRD1 viral lineage to Euryarchaea.

The integrated mobile element TGV2
Besides TGV1 (of the pT26-2 family), the genome of *T. gammatolerans* harbours another element TGV2 (20418 bp) unrelated to other PVs from Thermococcales [50]. TGV2 encodes two ATPases, a resolvase, a nuclease and a methylase, a putative Rep protein, and a SFI helicase related to those of pTN2 and pP12-1 (Figure 1). The nuclease and the methylase form a type III restriction/modification complex that has been probably acquired recently from bacteria by HGT (horizontal gene transfer). The best hits with both genes in BLAST searches are with proteins of *Thermocrinis albus*, a hyperthermophilic bacterium, and the two genes are located next to the resolvase gene which has itself a huge number of close bacterial relatives (P. Forterre, unpublished work). Interestingly, the three motifs characteristic of the pTN2
DNA polymerase family can be detected in the N-terminal region of the putative Rep protein of TGV1, suggesting that TGV2 also encodes its own DNA polymerase (P. Forterre, unpublished work). However, whereas the SFI helicases of pTN2 and TGV2 are closely related, their DNA polymerases are very divergent (P. Forterre, unpublished work).

**Virus-like vesicles**

Screening for viral particles in culture supernatants by electron microscopy led to the observation that most strains of Thermococcales produce membrane vesicles, including both globular and filamentous structures [20]. Some globular vesicles strikingly resemble the virion of recently described globular pleomorphic viruses in Haloarchaea [6]. The production of membrane vesicles is a feature shared by cells from the three domains of life. In proteobacteria, vesicles are produced by budding from the outer membrane, whereas, in eukaryotes, they are produced either by release of exosomes from multivesicular bodies or outward budding of ectosomes (also called microparticles) from the cytoplasmic membrane. In Thermococcales, vesicles appear to be produced by the latter mechanism (M. Gaudin, P. Forterre and E. Marguet, unpublished work). In bacteria, various publications reported the presence of DNA (including plasmidic DNA) associated with vesicles and suggest that vesicles could be vectors of HGT [55]. Some vesicles produced by Thermococcales are also associated with genomic DNA [20]. This DNA is more resistant to DNase treatment and to thermal denaturation than naked DNA, suggesting that vesicles could stabilize extracellular DNA in hot environments, indeed facilitating HGT [20]. It has been difficult to determine previously whether DNA associated with vesicles is located inside these vesicles or strongly bound to their surface. Microscopic examination of vesicles from *T. nautilus* disrupted by NaCl treatment shows some partially disrupted vesicles with extruding DNA, strongly suggesting that DNA is located within vesicles (Figure 4). Analysis of plasmid DNA isolated from vesicles by electron microscopy shows plasmids with size corresponding to pTN1 (Figure 4) and larger linear molecules (not shown). We could not find molecules corresponding to pTN2. The pTN1 DNA-binding protein p24 that contains a hydrophobic segment could be involved in linking specifically pTN1 DNA to virus-like vesicles [38].

**Evolution**

The origin and nature of PVs, their role in evolution and their position in the tree of life are controversial topics [8–10]. It is traditionally assumed that PVs essentially evolved by recruiting cellular genes. However, homologues of cellular genes are rarely encoded in PV genomes, arguing against this hypothesis. Indeed, in the case of PVs from Thermococcales, a single protein of cellular origin, those encoding MCM helicases, can be detected out of hundreds of proteins encoded by all PVs and integrated PVs combined. In TKV1, MCM has replaced a PV-specific ATPase present in pT26-2, TKV2 and TKV3, and a PV-specific Rep protein present in pAMT11. The discovery of novel types of DNA replication proteins (Rep proteins, DNA polymerases, helicases) in archaeal plasmids confirm that PVs can act as ‘inventors’ of new genes involved in DNA replication (for a discussion on the origin of PV-specific proteins, see the virocell concept in [56]). The diversity of DNA replication proteins in the PV world can be taken as an argument supporting the idea that DNA and DNA-manipulating enzymes originated first in an ancient virosphere [31].
The presence of a high proportion of ORFans (orphan ORFs) is another general feature of PV genomes [7]. Indeed, approx. 50% of all proteins encoded by PVs from Thermococcales have no detectable homologues in databases, whereas the other 50% have homologues only in other archaeal PVs, most often in Methanococcales, secondly in Haloarchaea and rarely in Crenarchaea. The close relationships between some PVs of Thermococcales and Methanococcales is probably not due to preferential gene transfer between these two orders, since Methanococcales harbouring these PVs are mesophiles. This relationship probably reflects the proximity of these two orders in archaeal phylogeny [12]. These data strongly suggest that archaeal PVs co-evolved with their hosts. This has been confirmed in the case of the pT26-2 family. Phylogenetic analysis of their core genes has shown that pT26-2 and related PVs have co-evolved within the group formed by Thermococcales and Methanococcales [15]. The same conclusion was obtained for the whole archaeal domain in reconstructing the network of evolutionary relationships between the pT26-2 family and integrated PVs identified in silico in archaeal genomes [15]. The number of genes common to different PVs in the network indeed decreases with increasing phylogenetic distances between their hosts [11,57]. Finally, the phylogeny of MCM proteins in the archaeal domain also agrees with the co-evolution scenario. Indeed, MCM proteins encoded by archaeal integrated or free PVs systematically group with cellular MCM proteins of corresponding hosts, indicating that PVs never (or rarely) jump from one archaeal group to the other [53]. Finally, recombination between different PV lineages also occurs essentially between PVs of the same group (i.e. a hybrid Thermococcus/Sulfolobus plasmid, such as a pRN/pTN plasmid infecting either a Thermococcus or a Sulfolobus has never been observed). All these data argue against the frequent assumption that PVs are used extensively as vehicles for gene transfer between distant cellular lineages.

### Perspectives

Although the number of PVs characterized from Thermococcales is still limited, several families can be already identified on the basis of replication cassettes (Figure 1) or gene content (Figure 2), and more should follow (D. Prieur, personal communication). This variety should help to design more versatile vectors and other genetic tools for Thermococcales. Conversely, these genetic tools should help to analyse the biological functions of proteins encoded by PVs. Hopefully, PVs from Thermococcales should be also useful to set up in vitro replication systems at very high temperature. Preliminary data on PV evolution based on their study in Thermococcales are encouraging. Considering that we have in hand a robust phylogeny of Archaea [11,57], a major goal of future research will be to identify all lineages of replicons (plasmids, viruses) and major capsid proteins present in Archaea to learn more about the mode of evolution of PVs and their impact on genome evolution. Finally, the study of virus-like membrane vesicles appears to be another promising line of research. Production of membrane-derived vesicles, being universal, is probably a very ancient feature of cellular life. The existence of vesicles carrying plasmids, reported in the present paper for Thermococcales, raises fascinating questions on the role that such a phenomenon has played in the origin of viruses in ancient RNA and/or DNA cells.

### Acknowledgments

We thank Eric LeCam and Fatima Boukhalfa-Heniche for the examination of virus-like vesicles and plasmids by electron microscopy.

### Funding

This work was funded by a JST (Japan Science and Technology Agency)/CNRS (Centre National de la Research Scientifique) Marine Genome and Marine Biotechnology grant.

### References


Received 30 October 2010
doi:10.1042/BST0390036

© The Authors Journal compilation © 2011 Biochemical Society