Virology 468-470 (2014) 322–329

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Review

Multiple roles of genome-attached bacteriophage terminal proteins

Modesto Redrejo-Rodríguez, Margarita Salas*

Instituto de Biología Molecular "Eladio Viñuela" (CSIC), Centro de Biología Molecular "Severo Ochoa" (Consejo Superior de Investigaciones Científicas—Universidad de Madrid), Universidad Autónoma, Nicolás Cabrera, 1, Cantoblanco, 28049 Madrid, Spain

ARTICLE INFO

Article history:
Received 11 July 2014
Returned to author for revisions
31 July 2014
Accepted 4 August 2014
Available online 17 September 2014

Keywords:
Bacteriophage
Protein-primed replication
Terminal protein
Nucleoid
Nuclear localization signal
Horizontal gene transfer

ABSTRACT

Protein-primed replication constitutes a generalized mechanism to initiate DNA or RNA synthesis in linear genomes, including viruses, gram-positive bacteria, linear plasmids and mobile elements. By this mechanism a specific amino acid primes replication and becomes covalently linked to the genome ends. Despite the fact that TPs lack sequence homology, they share a similar structural arrangement, with the priming residue in the C-terminal half of the protein and an accumulation of positively charged residues at the N-terminal end. In addition, various bacteriophage TPs have been shown to have DNA-binding capacity that targets TPs and their attached genomes to the host nucleoid. Furthermore, a number of bacteriophage TPs from different viral families and with diverse hosts also contain putative nuclear localization signals and localize in the eukaryotic nucleus, which could lead to the transport of the attached DNA. This suggests a possible role of bacteriophage TPs in prokaryote-to-eukaryote horizontal gene transfer.

© 2014 Elsevier Inc. All rights reserved.

Contents

Replication of linear genomes

DNA polymerases (DNAPs) are unable to start *de novo* DNA synthesis and require a free hydroxyl group to incorporate each new deoxynucleotide. In the case of circular genomes, a nick at the replication origin can generate an available 3′OH group to initiate the new strand synthesis. Alternatively, replication often starts by the synthesis of an RNA primer oligonucleotide that must be subsequently removed and refilled. However, at linear chromosome ends, removal of the primer fragment would generate an unreplicated single-stranded DNA (ssDNA) portion, which may lead to end shortening and loss of genetic information. This so-called end-replication problem can be overcome by different strategies. As far back as 1972 it was clear that previously observed head–tail concatemer intermediates during replication of bacteriophages T4 or T7 would allow them to evade the end-replication problem (Watson, 1972). Other phage linear genomes avoid end shortening by undergoing circularization prior to DNA replication, like λ phage (Campbell, 1994). Many eukaryotic viruses have similar strategies that also involve the presence of inverted terminal repetitions (ITR) that allow the formation of concatemeric replicative intermediates, e.g., in parvovirus or poxvirus replication, or genome circularization prior to replication, as in the case of herpesvirus (Boehmer and Lehman, 1997; Moss, 2013). In eukaryotic chromosomes telomerase extends directly the 3′

* Corresponding author. Tel.: +34 91 196 4675.
E-mail address: msalas@cbm.csic.es (M. Salas).

http://dx.doi.org/10.1016/j.virol.2014.08.003
0042-6822/© 2014 Elsevier Inc. All rights reserved.

end, giving rise to a G-rich overhanging ssDNA that is protected from nucleolytic attacks and recombinational activities by various strategies, including folding into higher structures such as T-loops or G-quadruplexes (Blackburn, 1990). Alternatively, other organisms including bacteriophages, animal viruses like Adenovirus, mitochondrial and cytoplasmatic plasmids of yeast and filamentous fungi, linear chromosomes and plasmids of *Streptomyces*, as well as human Hepatitis B virus and some plant and animal RNA viruses (Salas, 1991), possess a terminal protein (TP), which is involved in priming the genome replication by different mechanisms (see below) and becomes covalently linked to the genome ends. More recently, the presence of a protein-primed replication strategy has been suggested in a wide variety of replicons, including archaeal viruses, eukaryotic transposable elements (TE) and virophages (Bamford et al., 2005; Kapitonov and Jurka, 2006; Fischer and Suttle, 2011; Peng et al., 2007). These predictions are based on the homology of the protein-primed DNA polymerases, a subgroup of B-family DNA polymerases that contain two specific subdomains, named TPR1 and TPR2, which are involved in the interaction with TP and in processivity and strand-displacement capacity, respectively (Dufour et al., 2000; Rodríguez et al., 2005).

The evolutionary origin of protein-primed DNA replication mechanisms remains uncertain, although the incidence of genes encoding protein-primed DNA polymerases, with an apparent monophyletic origin (Filee et al., 2002; Braithwaite and Ito, 1993), among taxa associated with all domains of life, has been highlighted as consistent with a replication mechanism aroused early in the evolution, the genes being acquired from the primordial gene pool (Peng et al., 2007; Koonin, 2006). In addition, terminal proteins of human Hepatitis B virus and RNA viruses, though they would have independent origins, most likely also appeared early in the evolution, most likely prior to or simultaneously with key events of eukaryogenesis (Koonin et al., 2008). Hence, the utilization of a terminal protein as a primer for genome replication seems to be an ancient solution for the end-replication problem, which might have arisen independently several times and, as we discuss below, provides a useful instrument that has often acquired several additional functions associated with genome maintenance and replication, like compartmentalization or gene transfer.

Alternative protein-primed genome replication mechanisms

Protein-primed based replication of Adenoviruses and several bacteriophages from different families like Φ29 or PRD1 can duplicate the full-length genome continuously from both origins, without requiring Okazaki fragments or other accessory proteins (Salas, 1991). The replication mechanism of *Bacillus subtilis* phage Φ29 DNA has been extensively studied, thanks to the development of an *in vitro* replication system with purified proteins and TP-DNA as template (reviewed in De Vega and Salas (2011)). Fig. 1 shows a scheme of Φ29 genome replication. Briefly, the TP-DNAP heterodimer specifically recognizes the TP-containing DNA ends, which are opened with the help of the double-stranded DNA binding protein (DBP). The DNAP then catalyzes the addition of the first dAMP to the TP present in the heterodimer complex, using the second base of the genome end as a template. The presence of repetitive sequences at the replication origins in the Φ29 genome allows maintenance of the terminal nucleotides by a backward movement of the initiation complex called sliding-back, described also in other bacteriophages, or jumping-back in the case of Adenovirus. At least in the case of Φ29-like phages, the specificity of the internal template nucleotide that directs the first nucleotide incorporated is determined by the amino acids sequence near to

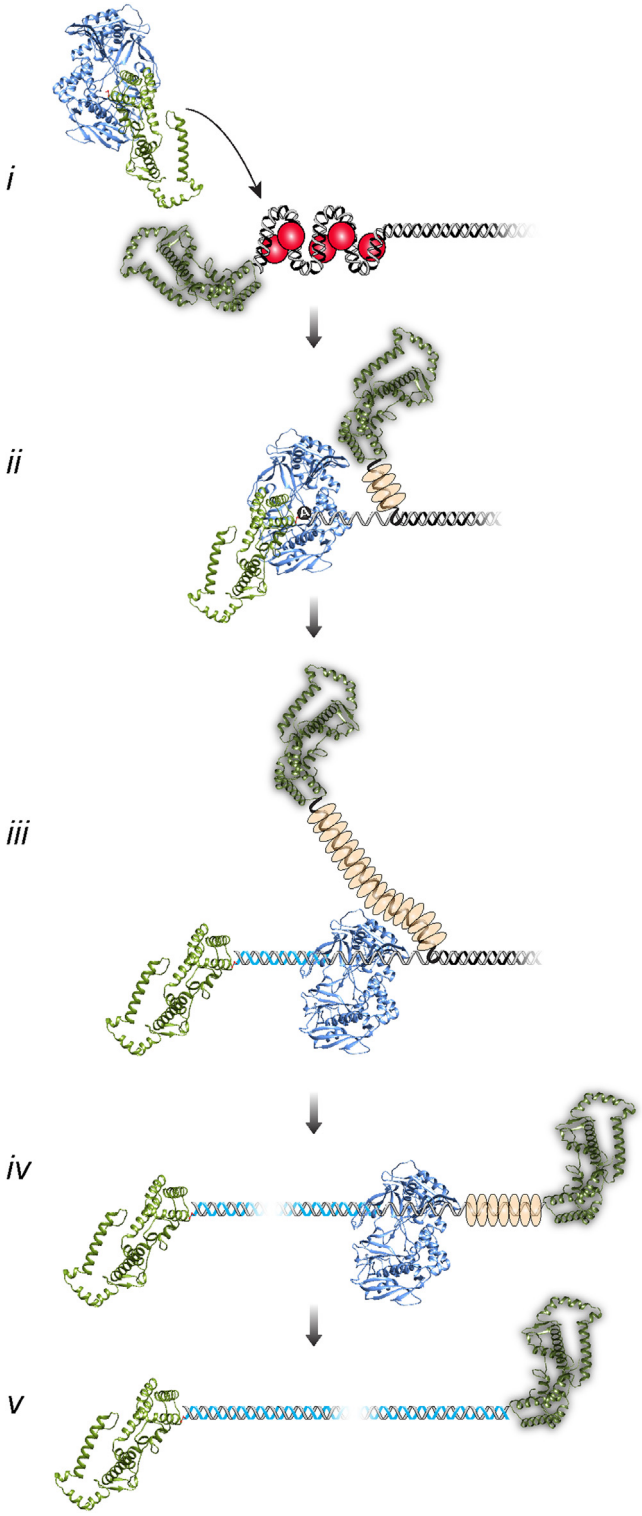

Fig. 1. Schematic representation of bacteriophage Φ29 TP-DNA replication. The DNAP (blue) and the TP (green) form an heterodimer that specifically binds to each of the genome ends (i). This binding is enhanced by the presence of the parental TP (dark green) and the end opening by the viral double-stranded DNA binding protein (DBP, red spheres). The synthesis of the new strand starts when the DNAP catalyzes the addition of the first nucleotide to its associated TP, using the second base of the genome end as a template (ii). Then, the DNAP dissociates, and the elongation progress asymmetrically from both ends (iii). The viral-encoded SSB protein (yellow ovals) binds to the displaced ssDNA and is removed by the DNAP during later stages of the replication process (iv). Eventually, continuous processive polymerization results in the generation of two fully replicated genomes (v). For simplicity only one of the TP-DNA ends is represented. See text for more details.

the priming residue of the TP, i.e. the priming loop (Longás et al., 2008). Then, after a transition step, the DNAP dissociates, and the elongation continues until replication of the nascent DNA strand is completed, thanks to the high processivity of the DNAP and its intrinsic strand displacement activity. The viral-encoded SSB protein binds to the displaced ssDNA and is removed by the DNAP during later stages of the replication process. Eventually, continuous processive polymerization results in the generation of two fully replicated genomes.

The structure of Φ29 TP has been partially solved (Kamtekar et al., 2006); it contains three structural domains: (i) the C-terminal domain (Ct), which has the serine-232 priming residue, (ii) the intermediate domain (I) that contributes to the surface of interaction with the DNAP, and (iii) the N-terminal domain (Nt) that has sequence-independent DNA-binding capacity required for recruitment of the viral genome (TP-DNA) and DNAP at the host nucleoid (Muñoz-Espín et al., 2010) and also contains a nuclear localization sequence (NLS), functional when the protein is expressed in mammalian cells (Redrejo-Rodríguez et al., 2012). In the case of Adenovirus, the TP is synthesized as pre-terminal protein (pTP), which is processed by a viral-encoded protease to yield mature TP after the viral replication cycle starts. Thus, the infective TP-DNA is the template for early transcription and the first round of replication, whereas the newly synthesized pTP-DNA is the template for subsequent rounds prior to pTP processing (Hoeben and Uil, 2013). Moreover, the pTP contains a NLS that is required for nuclear localization of both TP and DNAP (Zhao and Padmanabhan, 1988).

A number of cytoplasmic and mitochondrial linear plasmids from eukaryotic microbes also use a protein-primed replication mechanism. In this case, a single open reading frame encodes a protein with two domains, the TP and the DNAP, that starts genome replication using an unknown self-priming mechanism in the TP domain, which is subsequently proteolytically processed, releasing the DNA polymerase and the genome-linked TP (Klassen and Meinhardt, 2007). The same TP-DNAP bifunctional protein conformation has been proposed for eukaryotic mobile elements, polintons or mavericks (Kapitonov and Jurka, 2006).

In addition, as previously mentioned, filamentous bacteria from the genus *Streptomyces* and other members of the phylum Actinobacteria possess linear chromosomes and linear plasmids whose 5′-end is capped by a terminal protein (denoted as Tpg). In this case, full-length genome replication entails two steps. Replication starts from an internal origin in a bidirectional way, which eventually results in single-stranded gaps at the 3′-ends of about 300 nt. Those ssDNA patches are filled in using a mechanism that requires the fold-back pairing of distant palindromes prior to Tpg recruitment by a specific telomere-binding protein and subsequent deoxynucleotidylation of the priming residue of the Tpg, which remains covalently linked to the newly synthesized DNA strand (Hopwood, 2006). The C-terminal portion of Tpgs contains the priming residue, whereas the N-terminal half comprises different predicted functional domains, including a segment similar to the thumb domain of HIV reverse transcriptase, which overlaps a helix-turn-helix (HTH) motif that is required for successful end-patching. Moreover, several Tpgs include an amphiphilic β-strand involved in protein–protein and protein–membrane interactions and a nuclear localization signal (NLS), functional in animal and plant cells (Yang et al., 2002; Tsai et al., 2011, 2008).

Replication of human Hepatitis B virus (HBV) is a complex mechanism comprising different types of genome molecules. Infecting particles have a double-stranded DNA genome, whose synthesis requires a RNA replicative intermediate, called pregenomic RNA (pgRNA). The synthesis of pgRNA is carried out by a multifunctional viral polymerase (P protein), with reverse transcriptase and RNase H activities. Moreover, the N-terminus of P is a terminal protein, which contains the priming tyrosine-63 that is required to initiate the viral DNA minus-strand synthesis, using a specific loop in the pgRNA, called ε, as template. This first step of protein-primed DNA synthesis gives rise to a short DNA oligonucleotide, which, after a template switch to a 3′ specific region, initiates full-length minus-strand DNA synthesis (Nassal, 2008).

Finally, a number of vertebrate and plant positive-sense (+)-RNA viruses from *Picornaviridae*, *Caliciviridae* and *Potiviridae* families, and some members of the genus *Sobemovirus*, have a small TP (2–24 kDa) known as VPg, which is released from a polyprotein precursor and primes RNA synthesis, being thus covalently linked to the 5′-terminal uridine of the genome. VPg proteins perform additional functions *in vivo*; thus, the VPg precursor has a chaperone activity that stimulates the RNA polymerase, and a possible role in encapsidation of calicivirus RNA has been suggested. Moreover, plant virus VPg proteins contain an NLS that directs the protein to the nucleus (Jiang and Laliberte, 2011; Goodfellow, 2011).

Bacteriophage TPs

Contrary to DNAPs, TP sequences are very poorly conserved, even among the same viral family. Among bacteriophages, only TPs from Φ29 closely related phages (the Φ29-like viral genus), like Nf or GA-1, show a clear similarity with Φ29 in Blast analysis (E-values 8×10^{-55} and 3×10^{-122}, respectively). TPs from other *Podoviridae* members, like Cp-1 or ΦCP24R, show no significant similarity (*E*-value > 0) with Φ29 TP, and the same can be found

Table 1
Summary of bacteriophage terminal proteins (TPs) from representative phages from diverse families and hosts and their main properties.

Virus	Family	Host	TP Genbank accession number	Size (aa)	pI	Priming residue	N-terminal Nucleoid/DNA-binding	Nuclear targeting
Φ29	*Podoviridae*	*B. subtilis*	P03681.1	266	8.57	Ser232	Yes	Yes$^{(E)}$
Nf	*Podoviridae*	*B. subtilis*	ACH57070.1	266	5.59	Ser232	Yes	Yes$^{(E)}$
GA-1	*Podoviridae*	*B. subtilis*	NP_073686.1	266	4.88	Ser232	Yes	No$^{(E)}$
PRD1	*Tectiviridae*	*E. coli* and other Gram-negative	P09009.1	259	10.79	Tyr190	Yes	Yes$^{(E)}$
Bam35	*Tectiviridae*	*B. thuringiensis*	NP_943750.1*	245	10.59	n.d.	n.d.	Yes$^{(E)}$
Cp-1	*Podoviridae*	*S. pneumoniae*	NP_044816.1	230	10.77	Thr189	Yes	Yes$^{(E)}$
Av-1	*Podoviridae*	*Actinomyces sp*	YP_001333658*	360	10	n.d.	n.d.	Yes$^{(P)}$
ΦCP24R	*Podoviridae*	*Clostridium perfringens*	AEW47836.1*	246	10.12	n.d.	n.d.	Yes$^{(P)}$
AsccΦ28	*Podoviridae*	*Lactococcus lactis*	ACA21480.1*	167	10.19	n.d.	n.d.	No$^{(P)}$
ΦYS61	Myoviridae	*Thermus thermophilus*	YP_006560295.1*	261	10.72	n.d.	n.d.	No$^{(P)}$

When indicated (*) the TP was annotated or suggested based on the location of an ORF adjacent to the DNA polymerase (see the text for references). Nuclear targeting is based on localization of the protein when expressed in mammalian cells (E) or in the presence of putative NLSs (P), as reported in Redrejo-Rodríguez et al. (2012).

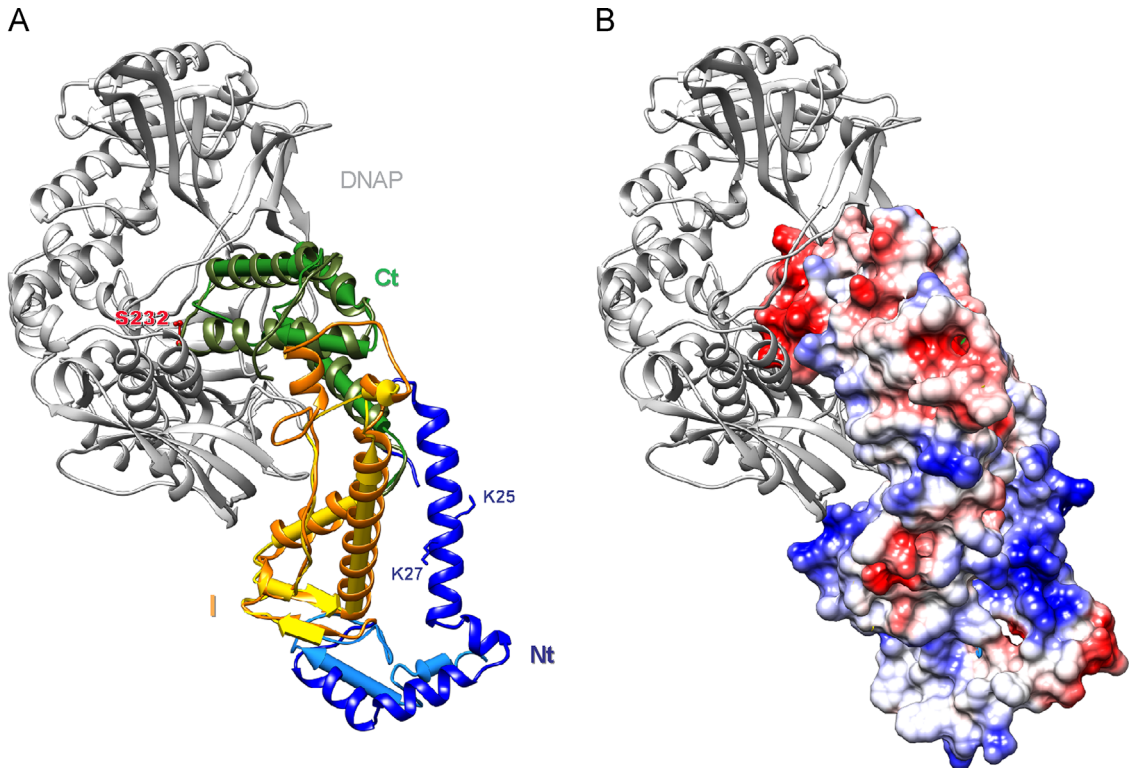

Fig. 2. Structure of Φ29 terminal protein (TP). A. Structure of the Φ29 DNA polymerase (DNAP) and TP complex. The DNAP (in gray ribbon) and the TP (cylindrical helices) are represented as in Kamtekar et al. (2006). Using this structure as a major template, a TP homology model, obtained with I-Tasser (Roy et al., 2010), is superimposed on the crystallographic structure. TP domains Nt, I and Ct are colored in blue, yellow and green, respectively; darker hues are used for the modeled structure. Lateral chain of the Serine-232 priming residue and the Lysines-25 and -27 are also shown. B. Structure of the Φ29 DNA polymerase (DNAP) and TP complex in the same orientation than in panel A, but with the electrostatic surface representation of the modeled TP. The Coulombic electrostatic surface is colored blue and red for electropositive and electronegative charges, respectively. The figure was rendered with UCSF Chimera software (Pettersen et al., 2004).

for phage TPs from other viral families, either when compared with Φ29 or between themselves. Moreover, structure–function studies from other phage TPs beyond Φ29 or Φ29-like phages are not available, therefore their structural or functional domains are not easy to recognize. However, as expected from proteins that perform the same function, they share some structural features (Table 1), like a relatively small size and a high proportion of basic residues, some of them grouped in positively charged clusters in the N-terminal half of the protein (Redrejo-Rodríguez et al., 2012), which usually results in a high isoelectric point. Moreover, although the priming residue may be a serine, threonine or tyrosine, it is always in a loop between two α-helices and followed by an acidic residue (D/E), which has been recently shown to be involved in the stabilization of the priming loop/DNAP catalytic active site interaction (del Prado et al., 2013). With the exception of PRD1, the priming loop is highly hydrophilic, due to a negatively charged surface that mimics DNA in the interaction with the DNAP (see below). Interestingly, the priming residue of *Streptomyces* Tpg is also within a hydrophilic pocket (Yang et al., 2013) and is followed by an acidic residue, despite the fact that the two proteins may have evolved independently.

The structure of the Φ29 TP was partially solved in complex with the DNAP (Kamtekar et al., 2006). As mentioned above, in this complex (Fig. 2A), the TP forms an elongated three domain structure. The C-terminal priming domain (Ct, residues 173–266) is comprised of a four-helix bundle, buried inside the DNAP, mimicking DNA thanks to a highly electronegative surface. This domain is connected through a hinge to the intermediate domain (I, residues 74–122), which spans two long α-helices and a short β-turn-β structure and interacts with the DNAP mainly through the TPR1 subdomain. Several residues of the Ct and I domains have been identified as polymerase ligands (del Prado et al., 2012). Finally, the Nt domain (residues 1–73) was strongly disordered within the crystal lattice and thus remains unsolved. Bioinformatic predictions indicated a high proportion of α-helix, and two poly-alanine α-helices were modeled in the structure (Kamtekar et al., 2006). Recently, Far-UV circular dichroism (CD) spectroscopic analysis has further confirmed the high content of α-helices in the Nt domain (Holguera et al., 2014). We have used the I-Tasser protein modeling server (Roy et al., 2010) to construct a new model of the full-length TP. It shows an Nt domain conformed by two large α-helices, spanning residues 13–35 and 47–67, linked by a loop that also contains a small α-helix portion (residues 38–42). This model also allows us to locate the position of the priming serine-232, which lies in a disordered loop in the crystallographic structure, within an electropositive exposed region of the TP (Fig. 2B).

Phage TPs can localize in the bacteria nucleoid and also in the eukaryotic nucleus

Co-evolving along with their hosts, viruses have developed strategies to compartmentalize their replication cycle by directing the subcellular localization of viral proteins within infected cells. Thus, many proteins from eukaryotic viruses contain NLSs (Cohen et al., 2011), including Adenovirus (Zhao and Padmanabhan, 1988) and plant positive-strand RNA viruses (Jiang and Laliberte, 2011), which direct or contribute to localization of viral replication and transcription in the host nucleus. Analogously, Φ29 TP targets the host nucleoid by means of a sequence-independent DNA-binding ability, which provides the correct localization for virus early replication and transcription, and facilitates recruitment of the DNAP (Muñoz-Espín et al., 2010). This DNA binding capacity

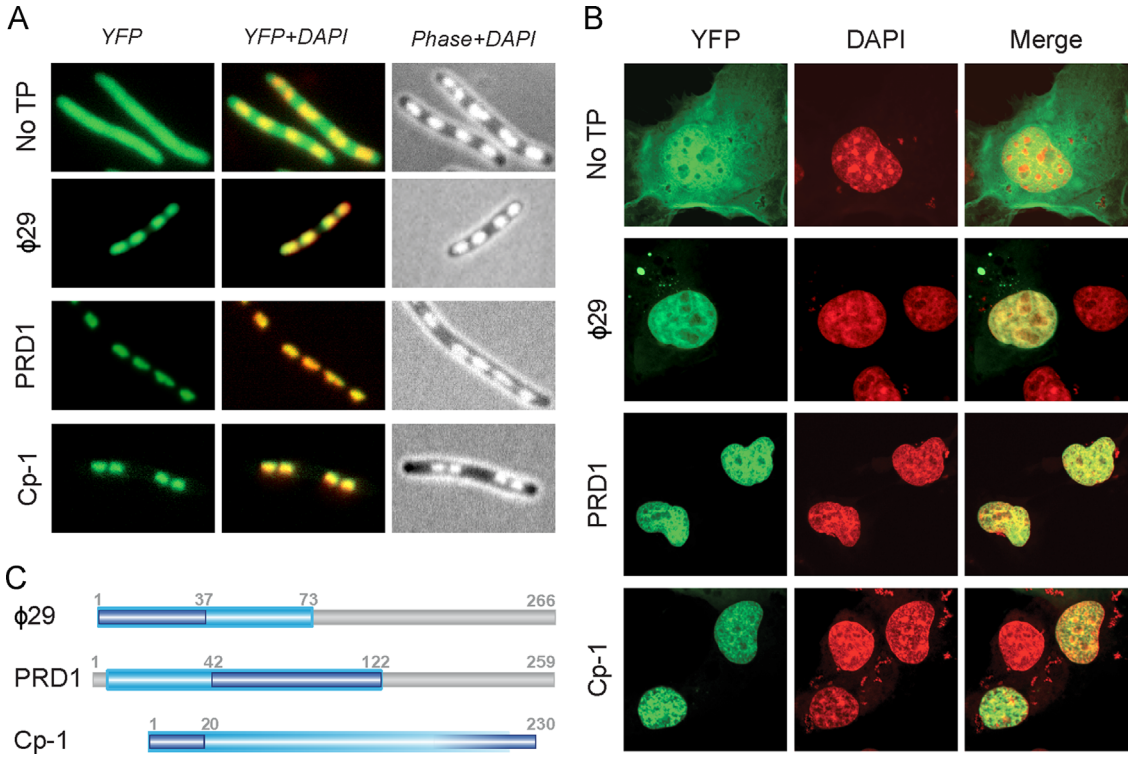

Fig. 3. Nucleoid and nuclear localization of representative phage TPs. Subcellular localization of Φ29, PRD1 and Cp-1 TPs in *E. coli*, expressed as YFP-TP fusions (A) and mammalian cells, as double YFP fusions (B). Reproduced with modifications from Redrejo-Rodríguez et al. (2012, 2013). Panel C shows a schematic representation of Φ29, PRD1 and Cp-1 TPs with the fragments required for nucleoid localization and NLs function are highlighted blue and navy blue, respectively.

resides in the Nt domain of the TP, thanks to the high density of basic amino acids. Nucleoid targeting requires almost the full Nt domain (residues 1–70). Furthermore, the double mutation K25A/K27A produces a strong decrease of DNA-binding capacity in vitro and in vivo and thus host nucleoid localization and DNAP recruitment (Holguera et al., 2014). Consequently, that mutation gives rise to an impaired genome replication in vivo, as shown by complementation assays of a TP-mutant phage (Holguera et al., 2014). These two residues are surface-exposed in the TP model (see Fig. 2), which could explain their major contributions to the DNA-binding capacity of the TP.

Since the bacterial nucleoid lacks an envelope that isolates the chromosome from the rest of the cell cytoplasm, DNA-binding proteins should have access to the chromosome and eventually bind to the nucleoid. In agreement with this, the TPs nucleoid localization is conserved beyond the natural host since Φ29 TP, as well as TPs from the related phages Nf and GA-1, also target the *Escherichia coli* nucleoid (Table 1 and Fig. 3A) (Muñoz-Espín et al., 2010; Redrejo-Rodríguez et al., 2013). Also, as in the case of the host nucleoid, the Φ29 TP Nt domain is responsible for nucleoid localization in *E. coli* (Redrejo-Rodríguez et al., 2013). Furthermore, the TPs of the Φ29-unrelated phages PRD1 and Cp-1, which infect a variety of Gram-negative bacteria and the Gram-positive pathogen *Streptococcus pneumonia*, respectively, also target the nucleoid (Redrejo-Rodríguez et al., 2013) by means of N-terminal positively charged motifs (Fig. 3A). PRD1 TP nucleoid targeting is disrupted by deletion of the first 41 amino acids, and a fragment spanning the 1–122 N-terminal residues is required for nucleoid localization. In the case of Cp-1 TP, removal of only 11 residues at the N-termini impaired nucleoid accumulation of the protein, and elimination of a longer portion of 22 amino acids fully disrupted the nucleoid targeting (Redrejo-Rodríguez et al., 2013).

As mentioned above, pTP of Adenovirus, *Streptomyces* Tpg and VPg of plant viruses contain an NLS that can transport the protein and its attached nucleic acid to the eukaryotic nucleus. Other prokaryotic proteins that are tightly bound to DNA molecules, like *Agrobacterium* VirD2, also have intrinsic NLSs required for DNA transfer to the eukaryotic nucleus (Rossi et al., 1993; Pelczar et al., 2004). Interestingly, we found that Φ29 TP also localizes in the eukaryotic nucleus when expressed in mammalian cells and, moreover, the presence of Φ29 TP at both ends of a linear DNA enhances gene delivery (Redrejo-Rodríguez et al., 2012). This nuclear targeting relies on the N-terminal residues 1–37 and, like cellular NLSs, its function requires energy, thereby suggesting that the phage TP contains an intrinsic *bona fide* NLS. Moreover, TPs from other phages from different families and with diverse host range, like PRD1 or Cp-1 (Fig. 3B), among others (Table 1), also localize or are predicted to localize in the eukaryotic nucleus by means of putative NLSs (Redrejo-Rodríguez et al., 2012).

Although analogous in terms of targeting the host genome, nucleoid and nucleus targeting features have very different structural requirements. Since the bacterial nucleoid lacks an envelope that isolates the chromosome from the cytosol, the genome is available to all proteins and those with DNA-binding capacity would eventually bind to the nucleoid. On the other hand, eukaryotic nucleus is enclosed by a double-membrane with pores that only allow the diffusion of small proteins or nucleic acids. Larger proteins, either from the cell or from a foreign origin, contain NLSs that comprise an small amino acid stretch, which is recognized by the importin/karyopherin transport machinery that actively import them into the nucleus (Gorlich and Mattaj, 1996). The interaction with nuclear transport mediators depends on sequence recognition of the NLS motifs, which usually are enriched in lysine and arginine residues (Chook and Süel, 2011). On the other hand, DNA-binding domains are usually composed of several α-helix segments enriched in basic amino acids (Rohs et al., 2010), which in the case of TPs are commonly enriched near the N-terminus of the protein (Fig. 3C). The presence of positively charged amino acids in both NLS and DNA-binding domains

have led some authors to propose an evolutionary origin of NLSs in previously existing DNA-binding motifs as a mechanism for compartmentalization of proteins inside the nucleus of primitive eukaryotic cells (Cokol et al., 2000; Dingwall and Laskey, 1991). In line with this, we have shown that NLSs from three different phages, Φ29, PRD1 and Cp-1, are located within the region required for nucleoid targeting in bacteria (Fig. 3C) (Redrejo-Rodríguez et al., 2013). However, we could design Φ29 TP mutants with disrupted nuclear localization that still target the nucleoid and, moreover, we have found two independent NLSs in the Cp-1 TP, only one of them being required for nucleoid localization. Thus, nuclear and nucleoid localizations (i.e. NLS and DNA-binding motifs) are partially overlapping in phage TPs, but they can be uncoupled, suggesting that they have been independently conserved along the evolution (Redrejo-Rodríguez et al., 2013).

TPs provide a new possible mechanism of bacteriophage-mediated inter-kingdom horizontal gene transfer

Ancient viruses evolved along with their hosts from the beginning of life and might have contributed to eukaryogenesis (Bell, 2009; Claverie, 2006; Forterre, 2011). Therefore, NLSs of viral and cellular proteins may have arisen simultaneously, which would explain the origin of those motifs in current eukaryotic viruses. However, although the presence of NLSs in bacteriophages might be a remnant of ancient evolutionary processes, they must have a biological function in order to be maintained throughout evolution.

Virus–host interactions are highly complex and often include gene exchange processes (Canchaya et al., 2003). Bacteriophages are the most abundant biological entity on Earth (Clokie et al., 2011) and they coevolved along with their bacterial hosts but also with eukaryotic organisms, from soil or marine unicellular phagotrophs to the internal environment of mammals. Given the capacity of Φ29 TP to enhance gene delivery, we suggested that TPs may ferry new traits from bacteria to eukaryotic cells, thereby mediating horizontal (or lateral) gene transfer (HGT) (Redrejo-Rodríguez et al., 2012).

The HGT phenomenon has been extensively reported in prokaryotes, fungi, plants, and also humans (Abby et al., 2012; Marcet-Houben and Gabaldon, 2010; Won and Renner, 2003; Lander et al., 2001; Huang, 2013; Koonin, 2014; Bruto et al., 2014), and it is an accepted driving force of evolution, although with some controversy about its relative contribution (Salzberg et al., 2001; Kurland et al., 2003). HGT has been extensively studied in prokaryotes to decipher the transfer mechanisms, as well as to understand its biological importance for acquisition of new traits. In bacteria, natural cell transformation, transduction, and conjugation account for most of gene transfer (Thomas and Nielsen, 2005). Bacteriophages can promiscuously package foreign DNA fragments into viral particles and also collect host DNA fragments and transfer them to a new bacterial host (Canchaya et al., 2003).

For a long time it was believed that HGT was limited to the bacterial and archaeal domains, because of the eukaryotic cell barriers that make it difficult to access the nuclear genome, and the fact that in most multicellular eukaryotes the transmission of genes to offspring is made more difficult because of the separation of somatic and germ cells (Keeling and Palmer, 2008). An increasing number of prokaryote-to-eukaryote HGT events have been described in the last years (Dunning Hotopp et al., 2007; Bruto et al., 2013) and, moreover, actual inter-kingdom gene transfers from bacteria to human somatic cells have been detected (Riley et al., 2013). The molecular mechanisms of those transfers remain unclear. For instance, in the case of phagotrophic protists, in which HGT has been shown to be more common (Andersson, 2005), DNA transfer from bacterial food was postulated as a mechanism for the acquisition of new genes (Doolittle, 1998). However, a number of HGT candidates observed in photoautotrophic algae can hardly be explained by this mechanism (Bruto et al., 2014; Schonknecht et al., 2013; Bhattacharya et al., 2013), which suggests the requirement for an active transfer vector, like a virus. Moreover, once inside the cell, the molecular mechanism by which foreign DNA can overcome the nuclear envelope and enter the eukaryotic nucleus is unknown (Dunning Hotopp, 2011; Schonknecht et al., 2014). A well-known example is the HGT between intracellular bacteria *Wolbachia sp.* and its eukaryotic host cells, which has been shown to be facilitated by *Wolbachia* prophages (Metcalf and Bordenstein, 2012; Kent et al., 2011).

We proposed (Redrejo-Rodríguez et al., 2012) a possible general model in which infected bacteria containing large amounts of viral particles and/or genomes might encounter eukaryotic cells, and might be internalized by phagocytosis or other processes. Normally, viral DNA would be degraded inside the cell, but if the phage DNA contains 5′-linked TPs, it may enter endocytic pathways and subsequently be driven across the nuclear envelope by the NLS of the TP. In agreement with our hypothesis, the attachment of a single NLS has been shown to increase both nuclear transport and subsequent expression of large DNA fragments (Ludtke et al., 1999; Zanta et al., 1999). Furthermore, phages have been shown to be excellent gene delivery vectors (Tao et al., 2013; Lankes et al., 2007; Poul and Marks, 1999), although going through the nuclear envelope would be expected to be a limiting step for subsequent gene expression and integration of foreign DNA into the cell chromosomes. Whatever the mechanism, once inside the cell nucleus, the DNA may be expressed and provide a new trait to the eukaryotic cell. In line with this idea, a recent report has shown that prokaryotic promoters may drive gene expression in eukaryotic cells (Bentancor et al., 2013). Finally, the foreign DNA might eventually be incorporated into the cell genome and transferred to progeny cells.

Concluding remarks and perspectives

Different terminal proteins have evolved in very diverse replicons to perform a common function of priming genome (DNA or RNA) synthesis in spite of the lack of clear homology, even amongst the same kind of terminal proteins like phage TPs. However, as has been already pointed out, despite the fact of poor sequence conservation, they share a modular architecture with the priming function in a C-terminal domain and a positively charged N-terminal portion that contributes to DNA- or RNA-binding activity. Furthermore, the N-terminus often includes other motifs that provide multiple alternative functions in replication/transcription compartmentalization or pathogenic determinants, like the presence of nuclear localization signals (Muñoz-Espín et al., 2010; Redrejo-Rodríguez et al., 2012, 2013; Zhao and Padmanabhan, 1988; Jiang and Laliberte, 2011; Goodfellow, 2011). New structure-function studies in different systems and resolution of more TPs structures, which may reveal structural similarities beyond the amino acids sequence, would help to understand the origin and evolution of TPs and protein-primed genome replication.

The recurrent presence of NLSs in TPs from different origins such as Adenovirus, plant RNA viruses, as well as prokaryotic TPs from *Streptomyces* chromosomes and plasmids, bacteriophages and other genome-associated proteins like *Agrobacterium* VirD2 is remarkable and suggests an unknown evolutionary role. Particularly, the existence of NLSs in bacteriophage proteins was unforeseen. Since karyopherin and importin transport mediators

are highly conserved and the diversity of nucleocytoplasmatic transport pathways arose prior to the radiation of current eukaryotes (O'Reilly et al., 2011), investigation of cellular partners that mediate the nuclear internalization of TPs may shed some light on the origin and biological role of those NLSs. Furthermore, it is tempting to speculate that other bacteriophage proteins, including DNA packaging terminases, recombinases or structural proteins, could contain NLSs that target them to the cell nucleus and might facilitate interactions of phages and eukaryotic cells.

In the last years, more and more evidences have underlined the role of inter-domain HGT as major driving force of eukaryotic evolution, although the molecular mechanisms of transfer are unclear (Huang, 2013; Bruto et al., 2013; Dunning Hotopp, 2011). Moreover, it has been predicted that new algorithms providing a more comprehensive analysis of potential transfers would reveal the biological importance of HGT in vertebrates (Huang, 2013; Koonin, 2014; Schonknecht et al., 2014). Those new methods should implement novel strategies in order to allow the inclusion of sequences from plasmids or bacteriophages that are currently ruled out in gene transfer analysis on the basis of possible contamination. However, some of those transfers could be genuine (Riley et al., 2013).

The gene pool of non-eukaryotic organisms (bacteria and archaea) and their viruses may have served as a huge reservoir of potential new traits during eukaryotes evolution. The widespread presence of NLSs in phage TPs (Redrejo-Rodríguez et al., 2012) and the fact that TPs are covalently linked to viral genomes provide a possible molecular mechanism for common genetic interactions between phages and eukaryotic cells, previously disregarded, that may have contributed to shape eukaryotic pangenome along evolution. The actual impact of that process in the evolution remains to be confirmed.

Finally, the Φ29 *in vitro* replication system is a useful tool for amplification of heterologous DNA with a covalently linked TP (Mencía et al., 2011) and straightforward delivery of genes efficiently into the eukaryotic nucleus with increased efficiency in cultured cells (Redrejo-Rodríguez et al., 2012). Further development of this tool may include the use of fusion TPs with several functional motifs in order to enhance gene delivery by, for instance, reducing the requirement of transfection reagents, increasing stability *in vivo* or modulating target cell specificity (Redrejo-Rodríguez et al., 2013).

Acknowledgments

Research in MS laboratory is supported by Grants BFU2011-23645 from the Spanish Ministry of Economy and Competitiveness and by an institutional Grant from Fundación Ramón Areces to the Centro de Biología Molecular 'Severo Ochoa'. We are grateful to the members of the laboratory that have made contributions to the topic of this paper.

References

Abby, S.S., Tannier, E., Gouy, M., Daubin, V., 2012. Lateral gene transfer as a support for the tree of life. Proc. Natl. Acad. Sci. USA 109, 4962–4967.
Andersson, J.O., 2005. Lateral gene transfer in eukaryotes. Cell. Mol. Life Sci. 62, 1182–1197.
Bamford, D.H., Ravantti, J.J., Ronnholm, G., Laurinavicius, S., Kukkaro, P., Dyall-Smith, M., Somerharju, P., Kalkkinen, N., Bamford, J.K., 2005. Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. J. Virol. 79, 9097–9107.
Bell, P.J., 2009. The viral eukaryogenesis hypothesis: a key role for viruses in the emergence of eukaryotes from a prokaryotic world environment. Ann. NY Acad. Sci. 1178, 91–105.
Bentancor, L.V., Bilen, M.F., Mejias, M.P., Fernandez-Brando, R.J., Panek, C.A., Ramos, M.V., Fernandez, G.C., Isturiz, M., Ghiringhelli, P.D., Palermo, M.S., 2013. Functional capacity of Shiga-toxin promoter sequences in eukaryotic cells. PLoS One 8, e57128.
Bhattacharya, D., Price, D.C., Chan, C.X., Qiu, H., Rose, N., Ball, S., Weber, A.P., Arias, M.C., Henrissat, B., Coutinho, P.M., Krishnan, A., Zauner, S., Morath, S., Hilliou, F., Egizi, A., Perrineau, M.M., Yoon, H.S., 2013. Genome of the red alga *Porphyridium purpureum*. Nat. Commun. 4, 1941.
Blackburn, E.H., 1990. Telomeres: structure and synthesis. J. Biol. Chem. 265, 5919–5921.
Boehmer, P.E., Lehman, I.R., 1997. Herpes simplex virus DNA replication. Annu. Rev. Biochem. 66, 347–384.
Braithwaite, D.K., Ito, J., 1993. Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res. 21, 787–802.
Bruto, M., Prigent-Combaret, C., Luis, P., Hoff, G., Moënne-Loccoz, Y., Muller, D., 2013. Horizontal acquisition of prokaryotic genes for eukaryote functioning and niche adaptation. In: Pontarotti, P. (Ed.), Evolutionary Biology: Exobiology and Evolutionary Mechanisms. Springer, Berlin, Heidelberg, pp. 165–179.
Bruto, M., Prigent-Combaret, C., Luis, P., Moenne-Loccoz, Y., Muller, D., 2014. Frequent, independent transfers of a catabolic gene from bacteria to contrasted filamentous eukaryotes. Proc. R. Soc.: Biol. Sci. 281, 20140848.
Campbell, A., 1994. Comparative molecular biology of lambdoid phages. Annu. Rev. Microbiol. 48, 193–222.
Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M.L., Brussow, H., 2003. Phage as agents of lateral gene transfer. Curr. Opin. Microbiol. 6, 417–424.
Chook, Y.M., Süel, K.E., 2011. Nuclear import by karyopherin-βs: recognition and inhibition. BBA—Mol. Cell Res. 1813, 1593–1606.
Claverie, J.M., 2006. Viruses take center stage in cellular evolution. Genome Biol. 7, 110.
Clokie, M.R., Millard, A.D., Letarov, A.V., Heaphy, S., 2011. Phages in nature. Bacteriophage 1, 31–45.
Cohen, S., Au, S., Panté, N., 2011. How viruses access the nucleus. BBA—Mol. Cell Res. 1813, 1634–1645.
Cokol, M., Nair, R., Rost, B., 2000. Finding nuclear localization signals. EMBO Rep. 1, 411–415.
De Vega, M., Salas, M., 2011. Protein-primed replication of bacteriophage ø29 DNA. In: Kusic-Tisma, J. (Ed.), DNA Replication and Related Cellular Processes. InTech, Croatia.
Dingwall, C., Laskey, R.A., 1991. Nuclear targeting sequences—a consensus? Trends Biochem. Sci. 16, 478–481.
Doolittle, W.F., 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet. 14, 307–311.
Dufour, E., Méndez, J., Lázaro, J.M., de Vega, M., Blanco, L., Salas, M., 2000. An aspartic acid residue in TPR-1, a specific region of protein-priming DNA polymerases, is required for the functional interaction with primer terminal protein. J. Mol. Biol. 304, 289–300.
del Prado, A., Villar, L., de Vega, M., Salas, M., 2012. Involvement of residues of the Φ29 terminal protein intermediate and priming domains in the formation of a stable and functional heterodimer with the replicative DNA polymerase. Nucleic Acids Res. 40, 3886–3897.
del Prado, A., Lázaro, J.M., Villar, L., Salas, M., de Vega, M., 2013. Dual role of Φ29 DNA polymerase Lys529 in stabilisation of the DNA priming-terminus and the terminal protein-priming residue at the polymerisation site. PLoS One 8, e72765.
Dunning Hotopp, J.C., 2011. Horizontal gene transfer between bacteria and animals. Trends Genet. 27, 157–163.
Dunning Hotopp, J.C., Clark, M.E., Oliveira, D.C., Foster, J.M., Fischer, P., Munoz Torres, M.C., Giebel, J.D., Kumar, N., Ishmael, N., Wang, S., Ingram, J., Nene, R.V., Shepard, J., Tomkins, J., Richards, S., Spiro, D.J., Ghedin, E., Slatko, B.E., Tettelin, H., Werren, J.H., 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317, 1753–1756.
Filee, J., Forterre, P., Sen-Lin, T., Laurent, J., 2002. Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. J. Mol. Evol. 54, 763–773.
Fischer, M.G., Suttle, C.A., 2011. A virophage at the origin of large DNA transposons. Science 332, 231–234.
Forterre, P., 2011. A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. Res. Microbiol. 162, 77–91.
Goodfellow, I., 2011. The genome-linked protein VPg of vertebrate viruses—a multifaceted protein. Curr. Opin. Virol. 1, 355–362.
Gorlich, D., Mattaj, I.W., 1996. Nucleocytoplasmic transport. Science 271, 1513–1518.
Hoeben, R.C., Uil, T.G., 2013. Adenovirus DNA replication. Cold Spring Harb. Perspect. Biol. 5, a013003.
Holguera, I., Redrejo-Rodríguez, M., Salas, M., Muñoz-Espín, D., 2014. New insights in the Φ29 terminal protein DNA-binding and host nucleoid localization functions. Mol. Microbiol. 91, 232–241.
Hopwood, D.A., 2006. Soil to genomics: the *Streptomyces* chromosome. Annu. Rev. Genet. 40, 1–23.
Huang, J., 2013. Horizontal gene transfer in eukaryotes: the weak-link model. BioEssays 35, 868–875.
Jiang, J., Laliberte, J.F., 2011. The genome-linked protein VPg of plant viruses-a protein with many partners. Curr. Opin. Virol. 1, 347–354.
Kamtekar, S., Berman, A.J., Wang, J., Lázaro, J.M., de Vega, M., Blanco, L., Salas, M., Steitz, T.A., 2006. The phi29 DNA polymerase: protein-primer structure suggests a model for the initiation to elongation transition. EMBO J. 25, 1335–1343.
Kapitonov, V.V., Jurka, J., 2006. Self-synthesizing DNA transposons in eukaryotes. Proc. Natl. Acad. Sci. USA 103, 4540–4545.
Keeling, P.J., Palmer, J.D., 2008. Horizontal gene transfer in eukaryotic evolution. Nat. Rev. Genet. 9, 605–618.

Kent, B.N., Salichos, L., Gibbons, J.G., Rokas, A., Newton, I.L., Clark, M.E., Bordenstein, S.R., 2011. Complete bacteriophage transfer in a bacterial endosymbiont (Wolbachia) determined by targeted genome capture. Genome Biol. Evol. 3, 209–218.
Klassen, R., Meinhardt, F., 2007. Linear protein-primed replicating plasmids in eukaryotic microbes, Microbial Linear Plasmids. Springer, Germany, pp. 187–226.
Koonin, E.V., 2006. Temporal order of evolution of DNA replication systems inferred by comparison of cellular and viral DNA polymerases. Biol. Direct 1, 39.
Koonin, E.V., 2014. Bacterial genes in eukaryotes: relatively rare but real and important (comment on DOI 10.1002/bies.201300095). BioEssays 36, 8.
Koonin, E.V., Wolf, Y.I., Nagasaki, K., Dolja, V.V., 2008. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat. Rev. Microbiol. 6, 925–939.
Kurland, C.G., Canback, B., Berg, O.G., 2003. Horizontal gene transfer: a critical view. Proc. Natl. Acad. Sci. USA 100, 9658–9662.
Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921.
Lankes, H.A., Zanghi, C.N., Santos, K., Capella, C., Duke, C.M., Dewhurst, S., 2007. In vivo gene delivery and expression by bacteriophage lambda vectors. J. Appl. Microbiol. 102, 1337–1349.
Longás, E., Villar, L., Lázaro, J.M., de Vega, M., Salas, M., 2008. Phage Φ29 and Nf terminal protein-priming domain specifies the internal template nucleotide to initiate DNA replication. Proc. Natl. Acad. Sci. USA 105, 18290–18295.
Ludtke, J.J., Zhang, G., Sebestyen, M.G., Wolff, J.A., 1999. A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA. J. Cell Sci. 112, 2033–2041.
Marcet-Houben, M., Gabaldon, T., 2010. Acquisition of prokaryotic genes by fungal genomes. Trends Genet. 26, 5–8.
Mencía, M., Gella, P., Camacho, A., de Vega, M., Salas, M., 2011. Terminal protein-primed amplification of heterologous DNA with a minimal replication system based on phage Φ29. Proc. Natl. Acad. Sci. USA 108, 18655–18660.
Metcalf, J.A., Bordenstein, S.R., 2012. The complexity of virus systems: the case of endosymbionts. Curr. Opin. Microbiol. 15, 546–552.
Moss, B., 2013. Poxvirus DNA replication. Cold Spring Harb. Perspect. Biol. 5.
Muñoz-Espín, D., Holguera, I., Ballesteros-Plaza, D., Carballido-López, R., Salas, M., 2010. Viral terminal protein directs early organization of phage DNA replication at the bacterial nucleoid. Proc. Natl. Acad. Sci. USA 107, 16548–16553.
Nassal, M., 2008. Hepatitis B viruses: reverse transcription a different way. Virus Res. 134, 235–249.
O'Reilly, A.J., Dacks, J.B., Field, M.C., 2011. Evolution of the karyopherin-beta family of nucleocytoplasmic transport factors; ancient origins and continued specialization. PLoS One 6, e19308.
Pelczar, P., Kalck, V., Gomez, D., Hohn, B., 2004. *Agrobacterium* proteins VirD2 and VirE2 mediate precise integration of synthetic T-DNA complexes in mammalian cells. EMBO Rep. 5, 632–637.
Peng, X., Basta, T., Haring, M., Garrett, R.A., Prangishvili, D., 2007. Genome of the Acidianus bottle-shaped virus and insights into the replication and packaging mechanisms. Virology 364, 237–243.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Poul, M.A., Marks, J.D., 1999. Targeted gene delivery to mammalian cells by filamentous bacteriophage. J. Mol. Biol. 288, 203–211.
Redrejo-Rodríguez, M., Muñoz-Espín, D., Holguera, I., Mencía, M., Salas, M., 2012. Functional eukaryotic nuclear localization signals are widespread in terminal proteins of bacteriophages. Proc. Natl. Acad. Sci. USA 109, 18482–18487.
Redrejo-Rodríguez, M., Muñoz-Espín, D., Holguera, I., Mencía, M., Salas, M., 2013. Nuclear localization signals in phage terminal proteins provide a novel gene delivery tool in mammalian cells. Commun. Integr. Biol. 6, e22829.
Redrejo-Rodríguez, M., Muñoz-Espín, D., Holguera, I., Mencía, M., Salas, M., 2013. Nuclear and nucleoid localization are independently conserved functions in bacteriophage terminal proteins. Mol. Microbiol. 90, 858–868.
Riley, D.R., Sieber, K.B., Robinson, K.M., White, J.R., Ganesan, A., Nourbakhsh, S., Dunning Hotopp, J.C., 2013. Bacteria-human somatic cell lateral gene transfer is enriched in cancer samples. PLoS Comput. Biol. 9, e1003107.
Rodríguez, I., Lázaro, J.M., Blanco, L., Kamtekar, S., Berman, A.J., Wang, J., Steitz, T.A., Salas, M., de Vega, M., 2005. A specific subdomain in Φ29 DNA polymerase confers both processivity and strand-displacement capacity. Proc. Natl. Acad. Sci. USA 102, 6407–6412.
Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., Mann, R.S., 2010. Origins of specificity in protein-DNA recognition. Annu. Rev. Biochem. 79, 233–269.
Rossi, L., Hohn, B., Tinland, B., 1993. The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plant. Mol. Gen. Genet. 239, 345–353.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.
Salas, M., 1991. Protein-priming of DNA replication. Annu. Rev. Biochem. 60, 39–71.
Salzberg, S.L., White, O., Peterson, J., Eisen, J.A., 2001. Microbial genes in the human genome: lateral transfer or gene loss? Science 292, 1903–1906.
Schonknecht, G., Chen, W.H., Ternes, C.M., Barbier, G.G., Shrestha, R.P., Stanke, M., Brautigam, A., Baker, B.J., Banfield, J.F., Garavito, R.M., Carr, K., Wilkerson, C., Rensing, S.A., Gagneul, D., Dickenson, N.E., Oesterhelt, C., Lercher, M.J., Weber, A.P., 2013. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339, 1207–1210.
Schonknecht, G., Weber, A.P., Lercher, M.J., 2014. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. BioEssays 36, 9–20.
Tao, P., Mahalingam, M., Marasa, B.S., Zhang, Z., Chopra, A.K., Rao, V.B., 2013. In vitro and in vivo delivery of genes and proteins using the bacteriophage T4 DNA packaging machine. Proc. Natl. Acad. Sci. USA 110, 5846–5851.
Thomas, C.M., Nielsen, K.M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 3, 711–721.
Tsai, H.H., Huang, C.H., Lin, A.M., Chen, C.W., 2008. Terminal proteins of *Streptomyces* chromosome can target DNA into eukaryotic nuclei. Nucleic Acids Res. 36, e62.
Tsai, H.H., Huang, C.H., Tessmer, I., Erie, D.A., Chen, C.W., 2011. Linear *Streptomyces* plasmids form superhelical circles through interactions between their terminal proteins. Nucleic Acids Res. 39, 2165–2174.
Watson, J.D., 1972. Origin of concatemeric T7 DNA. Nat. New Biol. 239, 197–201.
Won, H., Renner, S.S., 2003. Horizontal gene transfer from flowering plants to *Gnetum*. Proc. Natl. Acad. Sci. USA 100, 10824–10829.
Yang, C.C., Huang, C.H., Li, C.Y., Tsay, Y.G., Lee, S.C., Chen, C.W., 2002. The terminal proteins of linear *Streptomyces* chromosomes and plasmids: a novel class of replication priming proteins. Mol. Microbiol. 43, 297–305.
Yang, C.C., Sun, W.C., Wang, W.Y., Huang, C.H., Lu, F.S., Tseng, S.M., Chen, C.W., 2013. Mutational analysis of the terminal protein Tpg of *Streptomyces* chromosomes: identification of the deoxynucleotidylation site. PLoS One 8, e56322.
Zanta, M.A., Belguise-Valladier, P., Behr, J.P., 1999. Gene delivery: a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. Proc. Natl. Acad. Sci. USA 96, 91–96.
Zhao, L.J., Padmanabhan, R., 1988. Nuclear transport of adenovirus DNA polymerase is facilitated by interaction with preterminal protein. Cell 55, 1005–1015.