HHS Public Access

Author manuscript

Curr Opin Virol. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:

Curr Opin Virol. 2017 August ; 25: 7–15. doi:10.1016/j.coviro.2017.06.008.

Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity

Eugene V. Koonin[1,*] and **Mart Krupovic**[2,*]

[1]National Center for Biotechnology Information, National Library of Medicine. National Institutes of Health, Bethesda, MD 20894, USA

[2]Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Department of Microbiology, Institut, Pasteur, Paris, France

Abstract

Virophages are satellite DNA viruses that depend for their replication on giant viruses of the family *Mimiviridae*. An evolutionary relationship exists between the virophages and Polintons, large self-synthesizing transposons that are wide spread in the genomes of diverse eukaryotes. Most of the Polintons encode homologs of major and minor icosahedral virus capsid proteins and accordingly are predicted to form virions. Additionally, metagenome analysis has led to the discovery of an expansive family of Polinton-like viruses (PLV) that are more distantly related to *bona fide* Polintons and virophages. Another group of giant virus parasites includes small, linear, double-stranded DNA elements called transpovirons. Recent in-depth comparative genomic analysis has yielded evidence of the origin of the PLV and the transpovirons from Polintons. Integration of virophage genomes into genomes of both giant viruses and protists has been demonstrated. Furthermore, in an experimental coinfection system that consisted of a protist host, a giant virus and an associated virophage, the virophage integrated into the host genome and, after activation of its expression by a superinfecting giant virus, served as an agent of adaptive immunity. There is a striking analogy between this mechanism and the CRISPR-Cas system of prokaryotic adaptive immunity. Taken together, these findings show that Polintons, PLV, virophages and transpovirons form a dynamic network of integrating mobile genetic elements that contribute to the cellular antivirus defense and host-virus coevolution.

Introduction

Viruses and transposons are two major classes of mobile genetic elements (MGE). Many viruses can behave as transposons or at least integrate into the host genome. Also, several groups of viruses show common evolutionary origins with transposons indicating that these two classes of MGE comprise interconnected regions of the MGE space [1,2]. The classic case in point are reverse-transcribing elements, a class of MGE that includes both an

*For correspondence: koonin@ncbi.nlm.nih.gov; krupovic@pasteur.fr.

enormous variety of non-viral elements, and bona fide viruses with an essential step of integration into the host genome in their life cycle [3,4]. Two families of viruses, *Metaviridae* and *Pseudoviridae*, primarily lead the transposon life style but retain the ability to form virions and share common ancestry with bona fide reverse-transcribing viruses [5,6].

More recently, it has become clear that a class of eukaryotic self-synthesizing DNA transposons known as Polintons (after DNA POLymerase and INTegrase, the two enzymes encoded by all these elements), or Mavericks, possess similar features. Polintons are integrated into the genomes of diverse unicellular eukaryotes as well as animals, in highly variable copy numbers [7–10]. These elements have genomes of 15 to 20 kilobase (kb), the largest among the known transposons, flanked by long terminal inverted repeats (TIRs). In addition to the two universal genes coding for a protein-primed DNA polymerase (pDNAP) and a retrovirus-like (RVE) integrase, most of the Polintons also encode a DNA-packaging ATPase and a maturation protease that are homologous to the morphogenetic enzymes of Nucleo-Cytoplasmic Large DNA Viruses (NCLDV) of eukaryotes [11,12]. Because of their large size (on the transposon scale) and the presence of genes for morphogenetic enzymes, Polintons have been considered ‘virus-like’ transposons since the time they were discovered, but neither actual virions nor structural proteins have been detected [7,8]. This conundrum has been resolved when in-depth computational analysis of the Polinton protein sequences has shown that most of the Polintons encode homologs of the major and minor capsid proteins of a broad range of viruses with icosahedral virions including the NCLDV, adenoviruses and several groups of prokaryotic viruses, most notably tectiviruses [13]. The capsid protein genes are clearly ancestral in the Polintons but some elements have lost them which seems to be associated with their proliferation in the host genome, e.g. in the protist *Trichomonas vaginalis* [11]. These findings imply that Polintons lead a dual life style that involves both a transposon and a virus stage although the virions and the conditions of their formation remain to be characterized [11,12]. The discovery of the capsid protein genes prompted the proposal to rename the Polintons to polintoviruses [11,12] but pending the experimental validation of the prediction, we generally refer to these elements by their original name, using the term polintoviruses only when virions are specifically discussed.

An additional link from Polintons to viruses is their relationship with the virophages (currently classified into the family *Lavidaviridae*), satellite viruses of giant NCLDV in the family *Mimiviridae* [14–17]. The virophages are similar to the Polintons in size and share the morphogenetic gene module, and in some instances, also the DNA polymerase and integrase. Recent metagenome analysis resulted in the identification of a diverse group of polinton-like viruses (PLV) (with a single prototype virus that has been actually isolated) [18]. The PLV resemble Polintons in size and gene repertoire but lack some genes typical of the Polintons, such as that coding for the maturation protease, and show only distant similarity at the protein sequence level [18]. The Polintons, PLV and virophages jointly comprise a distinct class of eukaryotic viruses and transposons that seems to have been key players in the evolution of eukaryotic dsDNA virosphere [11,12].

In this brief review, we discuss recent advances in comparative genomics that expand the polinton-like class of MGE and clarify the direction of evolution among these elements. We further address new findings on virophage biology that implicate the virophages in the host

defense against giant NCLDV, in a striking parallel with the CRISPR-Cas adaptive immunity in prokaryotes. These discoveries suggest that some if not many of the Polintons, similar to virophages, could protect the host against superinfecting viruses.

Polintons, polinton-like viruses, virophages, and transpovirons: expansion of the polinton-like class of mobile genetic elements and insights into its evolution

By definition, all Polintons share the pDNAP and RVE genes, and most also possess the virus morphogenesis module, i.e. genes for the two capsid proteins (major and minor capsid proteins, MCP and mCP, respectively), genome packaging ATPase and virion maturation protease [16] (Figure 1). Apart from the conserved replication and morphogenetic blocks, the Polintons encode diverse 'cargo' proteins which often include helicases and occasionally primases. The PLV encode the MCP, mCP (in most cases) and the packaging ATPase but lack the maturation protease [18] (Figure 1). Furthermore, most of the PLV lack the pDNAP and RVE integrase genes but instead encode a predicted tyrosine recombinase of a distinct family and often also methyltransferases as well as helicases of superfamilies 1 or 3 (SF1 and SF3, respectively) that are implicated in viral genome replication [18]. Integrated PLV genomes have been detected in the genomes of several unicellular eukaryotes, primarily green algae. However, several PLV genomes were fully assembled from metagenomic sequences, complete with TIRs at both ends of the respective contig, and thus are likely to represent free viruses. Moreover, one previously isolated but effectively uncharacterized virus, *Tetraselmis viridis* virus S1 (TVS1) [19,20], represents a phylogenetically distinct group of PLV. A distinct group of PLV, named Tlr1, has been discovered in the ciliate *Tetrahymena thermophila* prior to the description of the Polintons [21]. The Tlr1 elements encode the RVE and the morphogenetic module related to those of the Polintons [13], suggesting that they also form virions (Figure 1), butlack genes for the protease and pDNAP. Instead, these elements encode several genes shared with other PLV [18], most notably a distinct SF1 helicase and a conserved uncharacterized protein Tlr6F (Figure 1).

The virophages possess circular or linear dsDNA genomes of 17 to 34 kb and share with the Polintons the 4 genes comprising the morphogenetic module [16,22]. However, the virophage capsid proteins are highly derived forms of the double-beta-barrel fold that are only distantly related to the more typical capsid proteins of the Polintons [23]. The first discovered virophage, Sputnik, has not been initially connected to the Polintons [14]. The link has been established with the discovery of the second virophage, the Mavirus (a parasite of the giant Cafeteria roenbergensis virus, CroV, a distant relative of the mimiviruses), which shares with the Polintons not only the morphogenetic module but also the pDNAP and RVE genes [24]. Subsequent isolation of virophages and metagenome mining led to the expansion of both the Sputnik-like and Mavirus-like groups of virophages and identification of two additional groups [16,22,25,26] (Figure 1). The virophages outside the Mavirus group lack the RVE but many of them encode a distinct subfamily of tyrosine recombinases, shared with the PLV, that are likely to function as integrases [18]. Although, apart from the Mavirus group, most virophages lack the pDNAP, a group of putative hybrid virophages has been

assembled from the sheep rumen metagenome and shown to encode a polinton-like pDNAP [27].

All this gene sharing connects the Polintons, Tlr1-like elements, PLV and virophages into a dense network that comprises a compact, distinct module in the overall genomic network of dsDNA viruses as demonstrated by quantitative analysis of network modularity [28].

Another group of Mimivirus parasites includes the transpovirons, small (~7 kb), linear dsDNA plasmid-like molecules with TIRs [29]. Transpovirons accumulate in extreme high copy numbers during mimivirus reproduction, are incorporated into mimivirus particles and can also integrate into the mimivirus genome [29]. Transpovirons do not encode viral structural proteins but encompass a gene for a SF1 helicase that is related to the helicases of PLV and Tlr1 elements [16,29]. An in-depth analysis of the transpoviron-encoded protein sequences yielded notable results [30]. In the phylogenetic tree of the SF1 helicases, transpovirons, Tlr1 and PLV form a strongly supported clade. Furthermore, it has been shown, with a high statistical significance, that the transpoviron helicase branch emits from within the PLV. Besides the helicase gene, transpovirons share a gene for a Zn-ribbon protein with certain PLV and virophages [16,30]. The helicase and the Zn-ribbon genes are adjacent and divergently oriented in both types of elements (Figure 1).

The helicases of Tlr1 elements, PLV and transpovirons are large proteins of 1,000 to 1,400 amino acids, in which the actual SF1 helicase domain (~350 aa) comprises the C-terminal portion of the polypeptide. A detailed analysis of the rest of the sequence of this protein using sensitive profile-profile database searches has unexpectedly shown similarity between the central region of the Tlr1-like helicases and family B DNA polymerases (Figure 1) [30]. The region of sequence conservation corresponds to the polymerization domain of the pDNAPs (the proofreading exonuclease domain is missing) and, given the disruption of some of the catalytic motifs, appears to encompass an inactivated pDNAP. In the phylogenetic tree of the pDNAPs, the inactivated helicase-associated derivatives from PLV, Tlr1 and transpovirons form a clade that is nested within the Polinton branch (Figure 2) [30].

Identification of the remnants of pDNAPs fused to SF1 helicases in many PLV, Tlr1 elements and transpovirons defines the vector of evolution in this entire class of MGE that otherwise remained unclear [30]. Inactivation of the pDNAP in these elements implies that their ancestors encoded an active pDNAP, and accordingly, that both PLV and the Tlr1 elements are descendants of typical Polintons. This scenario is compatible with the topology of the phylogenetic tree of the pDNAPs (Figure 2). Moreover, the enigmatic transpovirons appear to be highly derived descendants of PLV, in which the morphogenetic module has been lost (Figure 3). The evolutionary position of the virophages is less clear but the presence of the pDNAP and RVE in only one virophage group and the structural distortions in the capsid proteins suggest that the virophages also are a derived, perhaps, fast-evolving offshoot of the Polintons. Virophages strictly share only the morphogenetic module, whereas the complements of other genes, including those from the replication module, are distinct and partially shared with PLV and Polintons (Figure 1). In particular, similar to Polintons, RVP and Mavirus-like virophages encode pDNAPs, whereas Sputnik-like and OLV-like virophages share with certain PLV a unique fusion protein encompassing an N-terminal

TVpol primase-polymerase domain and a C-terminal SF3 helicase domain [18]. Thus, it cannot be ruled out that the virophages are polyphyletic and evolved independently from different groups of Polintons and PLV (Figure 3).

Integrating virophages function as an adaptive immune system

Polintons have been discovered as transposons integrated into the host genomes, and only subsequent detailed genome analysis has led to the prediction that they could be true viruses [13], which remains to be tested experimentally. For the virophages, the reverse had been the case: these MGE have been originally identified as viruses [14,22] but more recent findings indicate that at least some of them can integrate into the host cell (and virus) genome, and the biological effects of such integration have been revealed. The first observation of virophage integration came from the analysis of a mimivirus genome, in which an integrated copy of the Sputnik 2 virophage genome (provirophage) has been identified [29]. This remained an isolated finding until a large scale survey of eukaryotic genome assemblies has revealed multiple remnants of integrated virophages of the Sputnik group as well as giant viruses in the genome of the alga (Chlorarachniophyte) *Bigelowiella natans* [31,32] The majority of the virophage-derived sequences in the *B. natans* genome were fragments of deteriorated virophage genomes but 6 potentially active copies with TIRs have been identified [31]. Conserved virophage inserts have been detected in multiple strains of *B. natans* showing that the virophages can persist in the host genome for substantial time intervals. Furthermore, the integrated virophage genes have been shown to be expressed, some at a high level. Apart from the virophage sequences, it has been reported that the *B. natans* genome contained several integrated transpoviron sequences. However, subsequent analysis has shown that these sequences actually represent a distinct variety of PLV that have been taken for transpovirons due to the presence of the homologous SF1 helicase (see above) [30]. In addition to the virophage and PLV sequences, the *B. natans* genome contains remnants of NCLDV genomes of somewhat uncertain provenance but generally affiliated with phycodnaviruses [31].

The genomic survey by Blanc et al [31] included more than 1,000 eukaryotic genomes but *B. natans* was the only one, in which multiple inserts of virophage sequences have been detected. Thus, virophage integration might not be a particularly wide spread event, although such a result could primarily reflect the poor representation of diverse unicellular eukaryotes in the current genomic database. Regardless of how (un)common this phenomenon might be, the discovery of integrated virophages [31,33] boosted the hypothesis on their protective role in eukaryotes infected by giant viruses that has been originally proposed by Fischer and Suttle in connection with the discovery of the Mavirus [24]. Under this scenario, the virophage would enter the host cell either jointly with a giant virus during an abortive infection, or independently, and after integration, would be activated by new giant virus infection, resulting in a protective effect.

The hypothesis on the protective role of integrated virophages is compatible with the experimental data showing that both Sputnik and Mavirus substantially lower the yield of the respective helper giant viruses and increase the survival rate of the host [14,24,34,35]. A similar protective effect towards the host cells has been previously demonstrated in co-

infection experiments with satellite and helper viruses of plants and animals [36–38], illuminating the generality of a phenomenon whereby inter-viral competition benefits the host. Theoretical models of the dynamic of tripartite host-parasite systems motivated by the discovery of virophages also suggest a mitigating effect of virophages on the giant virus growth and stabilization of the tripartite host-parasite system compared to the simple giant virus-host case [39–41]. More recently, this hypothesis has received direct, resounding support in experiments of Fischer and Hackl showing that, when marine flagellate *C. roenbergensis* is coinfected with CroV and Mavirus, in about 30% of the infected cells, multiple Mavirus genome copies integrate into the host genome [42]. When surviving cells with integrated Mavirus are infected with CroV, expression and replication of the virophage are induced, apparently through transcription activation of the integrated Mavirus copies by a CroV-encoded transcription factor. This activation of Mavirus expression becomes possible because, as previously noticed, the Mavirus promoters are nearly identical in sequence to the CroV promoters, an apparent adaptation of the Mavirus to the satellite life style [24]. Expression of the integrated Mavirus genes then leads to the formation of virophage particles. Perhaps unexpectedly, the reproduction of the endogenous Mavirus does not seem to exert a measurable effect on the reproduction of CroV in the same *Cafeteria* cell [42]. The infected host cell dies, which seems to be the inevitable outcome once infected with CroV, and both CroV and Mavirus are released into the environment. The released Mavirus then titers out the CroV and protects neighboring cells from subsequent CroV infection, preventing its spread in the protist population. Thus, the virophage is employed by the cellular host as the key component of an altruistic defense mechanism (Figure 4).

Altruistic defense in unicellular organisms might appear controversial, but evidence is accumulating that this form of defense, in particular programmed cell death (PCD) and dormancy induction, is often activated by virus infection in both prokaryotes and unicellular eukaryotes [43–45]. Whether or not the Mavirus case qualifies as PCD, depends on which virus actually kills the virophage-carrying host cells: CroV, although its reproduction is aborted, or the Mavirus. Regardless, the Mavirus protection of *C. roenbergensis* against CroV can be interpreted as a mechanism of adaptive (acquired) immunity that involves immunological memory of past infections imprinted onto the host genome [46]. There is a striking analogy between the virophage-mediated immunity against giant viruses in protists and the CRISPR-Cas immunity in bacteria and archaea [46], which also involves formation and maintenance of immune memory [47,48] (Figure 4). The CRISPR-Cas systems memorize the infectious agent directly, through incorporation of its genome fragments (protospacers) into the genome of the host, specifically, as spacers into a CRISPR array [47,49]. In the virophage case, it is not genetic material of the principal infectious agent itself (the giant virus) but that of its satellite (the virophage) that integrates into the host genome and serves as means of defense once activated by a new giant virus infection. As with CRISPR-Cas, it remains to be fully clarified how a host cell survives to store immunological memory if virus (bacteriophage or giant virus) generally kills the infected cell. In the case of CRISPR-Cas, adaptation apparently occurs in cells infected with defective phage particles [50], and the same might apply in the CroV-virophage case. Additionally or alternatively, endogenization and integration of the virophage (“Adaptation” in Figure 4) could occur in the absence of co-infection given that Mavirus has been shown to

enter the host cell via clathrin-mediated endocytosis independent of the helper virus [24]. The similarity between virophage-mediated immunity and CRISPR-Cas extends to the outcomes of the virus-host interaction. At least one variety of CRISPR-Cas systems, the RNA-targeting type VI, seems to function primarily in the altruistic mode, via PCD or dormancy induction [51,52], and other types of CRISPR-Cas might switch to this suicidal strategy when their immune function fails [53,54]. To complete the analogy, it should be pointed out that both these defense systems originate, at least in part, from MGE, and specifically, from self-synthesizing transposons. Indeed, a prokaryotic self-synthesizing transposon, named Casposon, appears to have been the ancestor of the CRISPR adaptation module [12,55,56], whereas Polintons, from which Mavirus apparently has evolved, belong to a eukaryotic branch of self-synthesizing transposons [11,12]. These far-reaching parallels between two unrelated mechanisms of acquired immunity [46] fit the general concept of ‘guns for hire’, under which, during evolution, enzymes shuttle between MGE and defense systems, being employed, alternately, for offense and defense [57].

Admittedly, the analogy between the Mavirus-CroV-*Cafeteria* and CRISPR-Cas systems is not complete. Despite considerable horizontal mobility and the apparent major contributions of MGE including the Casposons and other transposons to the evolution of CRISPR-Cas [58–60], the CRISPR-Cas loci as such do not show features of MGE let alone viruses. In other words, the MGE that gave rise to different components of the CRISPR-Cas systems have been fully domesticated by bacteria and archaea – and substantially changed in the process. This is not the case of Mavirus whose relationships with Cafeteria can be viewed as symbiotic whereby the virophage benefits from vertical propagation while integrated in the cellular host genome and concomitantly protects the host against the giant virus at the population level. The two mechanisms represent different stages in the evolution of defense systems. Further study of integrated virophages should show whether some of them have gone a longer way along the path to domestication than Mavirus.

Concluding remarks

Recent comparative genomic studies clarify an evolutionary scenario, in which Polinotns (polintoviruses) that evolved from prokaryotic tectiviruses during eukaryogenesis gave rise to a plethora of diverse MGE including PLV, virophages, transpovirons, and through more complex chains of evolutionary events, adenoviruses, bidnaviruses, cytoplasmic plasmids and NCLDV (Figure 3).

The discovery of the role of virophages in anti-giant virus immunity raises the general question on the nature of the biological relationships between polinton-like MGE and autonomous viruses, in particular, NCLDV. As suggested by Katzourakis and Aswad [61], could it be the case that the Polintons actually are integrated and domesticated virophages? The evolutionary scenario outlined in Figure 3 implies that the Polintons, descending from tectiviruses, started as autonomous viruses (polintoviruses) that subsequently acquired the integrase and adopted the dual, virus-transposon life style. The NCLDV evolved at a later stage of evolution (even if the time gap could have been relatively short), so that the virophages must have been an even later addition to the eukaryotic host-virus ecosystems [11]. Thus, in the grand scheme of things, it appears that virophages evolved from Polintons

not the other way around as originally proposed [24]. However, the extant Polintons are highly diverse, and it is impossible to tell, without further, detailed investigation, which of them are autonomous and which depend on other viruses. The conservation of the morphogenetic gene module in most Polintons implies that they retain the capacity to form virions, and continued NCLDV infections might be the selective factor maintaining that ability. Should that be the case, many more NCLDV, including giant viruses, remain to be discovered, considering the wide spread of Polintons in eukaryotic genomes. Further, unbiased genomic and metagenomic sequencing, hopefully, combined with experimental discovery and characterization of polintoviruses, should shed light on this important and fascinating area of eukaryotic virology.

Acknowledgments

EVK is supported by intramural funds of the US Department of Health and Human Services (to the National Library of Medicine)

References

1. Koonin EV, Dolja VV. Virus world as an evolutionary network of viruses and capsidless selfish elements. Microbiol Mol Biol Rev. 2014; 78(2):278–303. [PubMed: 24847023]
2. Koonin EV, Dolja VV, Krupovic M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. Virology. 2015; 479–480:2–25.
3. Kazazian HH Jr. Mobile elements: drivers of genome evolution. Science. 2004; 303(5664):1626–1632. [PubMed: 15016989]
4. Goodier JL, Kazazian HH Jr. Retrotransposons revisited: the restraint and rehabilitation of parasites. Cell. 2008; 135(1):23–35. [PubMed: 18854152]
5. Krupovic M, Koonin EV. Homologous capsid proteins testify to the common ancestry of retroviruses, caulimoviruses, pseudoviruses and metaviruses. J Virol. 2017; 91(12) pii: e00210–17.
6. Gladyshev EA, Arkhipova IR. A widespread class of reverse transcriptase-related cellular genes. Proc Natl Acad Sci U S A. 2011; 108(51):20311–20316. [PubMed: 21876125]
7. Kapitonov VV, Jurka J. Self-synthesizing DNA transposons in eukaryotes. Proc Natl Acad Sci U S A. 2006; 103(12):4540–4545. [PubMed: 16537396]
8. Pritham EJ, Putliwala T, Feschotte C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. Gene. 2007; 390(1–2):3–17. [PubMed: 17034960]
9. Haapa-Paananen S, Wahlberg N, Savilahti H. Phylogenetic analysis of Maverick/Polinton giant transposons across organisms. Mol Phylogenet Evol. 2014; 78:271–274. [PubMed: 24882428]
10. Jurka J, Kapitonov VV, Kohany O, Jurka MV. Repetitive sequences in complex genomes: structure and evolution. Annu Rev Genomics Hum Genet. 2007; 8:241–259. [PubMed: 17506661]
11**. Krupovic M, Koonin EV. Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. Nat Rev Microbiol. 2015; 13(2):105–115. Comparative genomic analysis reported in this paper implies that the polintons are direct descendants of tectiviruses and played a key role in the evolution of diverse groups of eukaryotic dsDNA viruses and plasmids. [PubMed: 25534808]
12. Krupovic M, Koonin EV. Self-synthesizing transposons: unexpected key players in the evolution of viruses and defense systems. Curr Opin Microbiol. 2016; 31:25–33. [PubMed: 26836982]
13. Krupovic M, Bamford DH, Koonin EV. Conservation of major and minor jelly-roll capsid proteins in Polinton (Maverick) transposons suggests that they are bona fide viruses. Biol Direct. 2014; 9:6. [PubMed: 24773695]
14. La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, Raoult D. The virophage as a unique parasite of the giant mimivirus. Nature. 2008; 455(7209):100–104. [PubMed: 18690211]

15. Desnues C, Boyer M, Raoult D. Sputnik, a virophage infecting the viral domain of life. Adv Virus Res. 2012; 82:63–89. [PubMed: 22420851]

16. Yutin N, Raoult D, Koonin EV. Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. Virol J. 2013; 10:158. [PubMed: 23701946]

17. Krupovic M, Kuhn JH, Fischer MG. A classification system for virophages and satellite viruses. Arch Virol. 2016; 161(1):233–247. [PubMed: 26446887]

18**. Yutin N, Shevchenko S, Kapitonov V, Krupovic M, Koonin EV. A novel group of diverse Polinton-like viruses discovered by metagenome analysis. BMC Biol. 2015; 13:95. This paper reports an extensive analysis of metagenomic sequence data resulting in the identification of a large group of predicted viruses that resemble Polintons in genome size and architecture, and share with them homologous capsid proteins, packaging ATPases, and in some cases, pDNAP. [PubMed: 26560305]

19. Stepanova OA, Boyko AL, Gordienko AI, Sherban SA, Shevchenko TP, Polischuck VP. Characteristics of virus of Tetraselmis viridis norris (Chorophyta, Prasinophycea). Dokl Akad Nauk Ukr. 2005; 1:158–162.

20. Pagarete A, Grebert T, Stepanova O, Sandaa RA, Bratbak G. Tsv-N1: A Novel DNA Algal Virus that Infects Tetraselmis striata. Viruses. 2015; 7(7):3937–3953. [PubMed: 26193304]

21. Wuitschick JD, Gershan JA, Lochowicz AJ, Li S, Karrer KM. A novel family of mobile genetic elements is limited to the germline genome in Tetrahymena thermophila. Nucleic Acids Res. 2002; 30(11):2524–2537. [PubMed: 12034842]

22. Bekliz M, Colson P, La Scola B. The Expanding Family of Virophages. Viruses. 2016; 8(11) pii: E317.

23. Zhang X, Sun S, Xiang Y, Wong J, Klose T, Raoult D, Rossmann MG. Structure of Sputnik, a virophage, at 3.5-A resolution. Proc Natl Acad Sci U S A. 2012; 109(45):18431–18436. [PubMed: 23091035]

24. Fischer MG, Suttle CA. A virophage at the origin of large DNA transposons. Science. 2011; 332(6026):231–234. [PubMed: 21385722]

25. Zhou J, Sun D, Childers A, McDermott TR, Wang Y, Liles MR. Three novel virophage genomes discovered from Yellowstone Lake metagenomes. J Virol. 2015; 89(2):1278–1285. [PubMed: 25392206]

26. Zhou J, Zhang W, Yan S, Xiao J, Zhang Y, Li B, Pan Y, Wang Y. Diversity of virophages in metagenomic data sets. J Virol. 2013; 87(8):4225–4236. [PubMed: 23408616]

27*. Yutin N, Kapitonov VV, Koonin EV. A new family of hybrid virophages from an animal gut metagenome. Biol Direct. 2015; 10:19. This paper reports the metagenomic discovery of a family of hybrid virophages that combine a morphogenetic module typical of virophages with a Polinton-like pDNAP. [PubMed: 25909276]

28**. Iranzo J, Krupovic M, Koonin EV. The Double-Stranded DNA Virosphere as a Modular Hierarchical Network of Gene Sharing. MBio. 2016; 7(4) pii: e00978–16 This article reports a quantiative dissection of the bipartite network of gene sharing for all dsDNA viruses. In this analysis, tectiviruses, Polintons, virophages adenoviruses, and mitochondrail as well as cytoplasmic dsDNA plasmids for a robust model that, at the next clustering iteration, merges with the NCLDV.

29. Desnues C, La Scola B, Yutin N, Fournous G, Robert C, Azza S, Jardot P, Monteil S, Campocasso A, Koonin EV, Raoult D. Provirophages and transpovirons as the diverse mobilome of giant viruses. Proc Natl Acad Sci U S A. 2012; 109(44):18078–18083. [PubMed: 23071316]

30**. Krupovic M, Yutin N, Koonin EV. Fusion of a superfamily 1 helicase and an inactivated DNA polymerase is a signature of common evolutionary history of Polintons, polinton-like viruses, Tlr1 transposons and transpovirons. Virus Evolution. 2016; 2(1):vew019. This article demonstrates the presence of an inactivated pDNAP domain at the N-termini of the SF1 helicases encoded by some of the PLV, Tlr1 elements and transpovirons, thus establishing the directionality of evolution whereby all these elements are derived from Polintons. [PubMed: 28694999]

31**. Blanc G, Gallot-Lavallee L, Maumus F. Provirophages in the Bigelowiella genome bear testimony to past encounters with giant viruses. Proc Natl Acad Sci U S A. 2015;

112(38):E5318–5326. This article reports a survey of eukaryotic genomes for the presence of integrated virophages genomes resulting in the discovery of multiple virophage copies, some highliy expressed, in the genome of the alga *Bigelowiella natans*. A hypothesis is proposed on protective effect of integrated virophage on giant virus infection. [PubMed: 26305943]

32. Fischer MG. Virophages go nuclear in the marine alga Bigelowiella natans. Proc Natl Acad Sci U S A. 2015; 112(38):11750–11751. [PubMed: 26330604]

33*. Villain A, Gallot-Lavallee L, Blanc G, Maumus F. Giant viruses at the core of microscopic wars with global impacts. Curr Opin Virol. 2016; 17:130–137. This article expands the hypothesis on the protective role of virophages in giant virus infections. [PubMed: 27088734]

34. Gaia M, Pagnier I, Campocasso A, Fournous G, Raoult D, La Scola B. Broad spectrum of mimiviridae virophage allows its isolation using a mimivirus reporter. PLoS One. 2013; 8(4):e61912. [PubMed: 23596530]

35. Campos RK, Boratto PV, Assis FL, Aguiar ER, Silva LC, Albarnaz JD, Dornas FP, Trindade GS, Ferreira PP, Marques JT, Robert C, et al. Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. Virol J. 2014; 11:95. [PubMed: 24886672]

36. Krupovic M, Cvirkaite-Krupovic V. Virophages or satellite viruses? Nat Rev Microbiol. 2011; 9(11):762–763. [PubMed: 22016897]

37. Palukaitis P. Satellite RNAs and Satellite Viruses. Mol Plant Microbe Interact. 2016; 29(3):181–186. [PubMed: 26551994]

38. Ziebell H, Carr JP. Cross-protection: a century of mystery. Adv Virus Res. 2010; 76:211–264. [PubMed: 20965075]

39. Wodarz D. Evolutionary dynamics of giant viruses and their virophages. Ecol Evol. 2013; 3(7): 2103–2115. [PubMed: 23919155]

40. Taylor BP, Cortez MH, Weitz JS. The virus of my virus is my friend: ecological effects of virophage with alternative modes of coinfection. J Theor Biol. 2014; 354:124–136. [PubMed: 24662503]

41. Nee S. The evolutionary ecology of molecular replicators. R Soc Open Sci. 2016; 3(8):160235. [PubMed: 27853598]

42**. Fischer MG, Hackl T. Host genome integration and giant virus-induced reactivation 1 of the virophage mavirus. Nature. 2016 This article reports integration of Mavirus into the *Cafeteria roenbergensis* genome, activation of Mavirus gene expression upon coinfection with CroV and protection of neighbor cells from new CroV infection by the released Mavirus.

43. Durand PM, Sym S, Michod RE. Programmed Cell Death and Complexity in Microbial Systems. Curr Biol. 2016; 26(13):R587–593. [PubMed: 27404254]

44. Bidle KD. Programmed Cell Death in Unicellular Phytoplankton. Curr Biol. 2016; 26(13):R594–607. [PubMed: 27404255]

45. Durand PM, Choudhury R, Rashidi A, Michod RE. Programmed death in a unicellular organism has species-specific fitness effects. Biol Lett. 2014; 10(2):20131088. [PubMed: 24573154]

46*. Koonin EV, Krupovic M. Virology: A parasite's parasite saves host's neighbours. Nature. 2016; 540(7632):204–205. This paper draws the analogy between the Mavirus-mediated immunity against CroV and the CRISPR-Cas systems in prokaryotes as adaptive immune mechanisms with genomic memory of past infections. [PubMed: 27929010]

47. Amitai G, Sorek R. CRISPR-Cas adaptation: insights into the mechanism of action. Nat Rev Microbiol. 2016; 14(2):67–76. [PubMed: 26751509]

48. Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science. 2016; 353(6299):aad5147. [PubMed: 27493190]

49. Sorek R, Lawrence CM, Wiedenheft B. CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu Rev Biochem. 2013; 82:237–266. [PubMed: 23495939]

50. Hynes AP, Villion M, Moineau S. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. Nat Commun. 2014; 5:4399. [PubMed: 25056268]

51. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, et al. C2c2 is a single-component

programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016; 353(6299):aaf5573. [PubMed: 27256883]

52. Smargon AA, Cox DB, Pyzocha NK, Zheng K, Slaymaker IM, Gootenberg JS, Abudayyeh OA, Essletzbichler P, Shmakov S, Makarova KS, Koonin EV, et al. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Mol Cell. 2017; 65(4):618–630. [PubMed: 28065598]
53. Koonin EV, Zhang F. Coupling immunity and programmed cell suicide in prokaryotes: Life-or-death choices. Bioessays. 2017; 39(1):1–9.
54. Makarova KS, Anantharaman V, Aravind L, Koonin EV. Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes. Biol Direct. 2012; 7(40)
55. Krupovic M, Makarova KS, Forterre P, Prangishvili D, Koonin EV. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. BMC Biol. 2014; 12:36. [PubMed: 24884953]
56. Krupovic M, Béguin P, Koonin EV. Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery. Curr Opin Microbiol. 2017 in press.
57. Koonin EV, Krupovic M. A Movable Defense. The Scientist. 2015 Jan 1.
58. Koonin EV, Krupovic M. Evolution of adaptive immunity from transposable elements combined with innate immune systems. Nat Rev Genet. 2015; 16(3):184–192. [PubMed: 25488578]
59. Krupovic M, Beguin P, Koonin EV. Casposons: the mobile elements that gave rise to the adaptation module of CRISPR-Cas systems. Curr Opin Microbiol. 2017; 38:36–43. [PubMed: 28472712]
60. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, Abudayyeh OO, Gootenberg JS, Makarova KS, Wolf YI, Severinov K, et al. Diversity and evolution of class 2 CRISPR-Cas systems. Nat Rev Microbiol. 2017; 15(3):169–182. [PubMed: 28111461]
61. Katzourakis A, Aswad A. The origins of giant viruses, virophages and their relatives in host genomes. BMC Biol. 2014; 12:51. [PubMed: 25184667]

- Polintons, virophages and transpovirons form a class of mobile elements
- Polinto-like elements include viruses, transposons and plasmids
- Virophages integrate into host genome and protect host against giant viruses
- There is deep analogy between virophage-mediated immunity and CRISPR-Cas

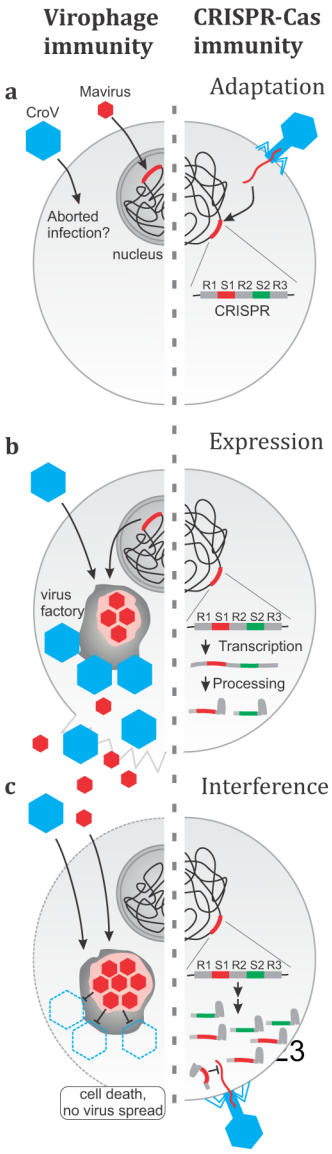

Figure 1. Representative genome architectures of polinton-like MGE

The genome schematics are drawn roughly to scale (shown at the bottom of the figure). The genes are shown by block arrows indicating the direction of transcription and identified by color code. The blank arrows show poorly conserved genes, many of them encoding uncharacterized proteins. Abbreviations: AEP, aerchaeo-eukaryotic primase; Bnat, *Bigellowiella natans*; P1-DY, Polinton 1 of *Drosophila yakuba*; P1-TC, Polinton 1 of *Tribolium castaneum*; P1-TV, Polinton 1 of *Trichomonas vaginalis*; RED, Red Sea; SAF, South Africa; DJR MCP, double jelly-roll major capsid protein; mCP, minor capsid protein; OLV, Organic Lake virophage; pDNAP, protein-primed DNA polymerase (hatched shading shows inactivated pDNAP derivatives); RVE, retrovirus-like integrase; RVP, rumen virophage [27]; SF1 and SF3, superfamily 1 and 3, respectively; TVpol, primase-polymerase homologous to the bacterial DNA polymerase I; Y-integrase, integrase of the tyrosine recombinase superfamily; ZnR, zinc ribbon; GIY-YIG, nuclease of the GIY-YIG family

(denoted after the conserved catalytic motifs). Tlr6f is an uncharacterized protein that is widespread among the polinton-like MGE [18].

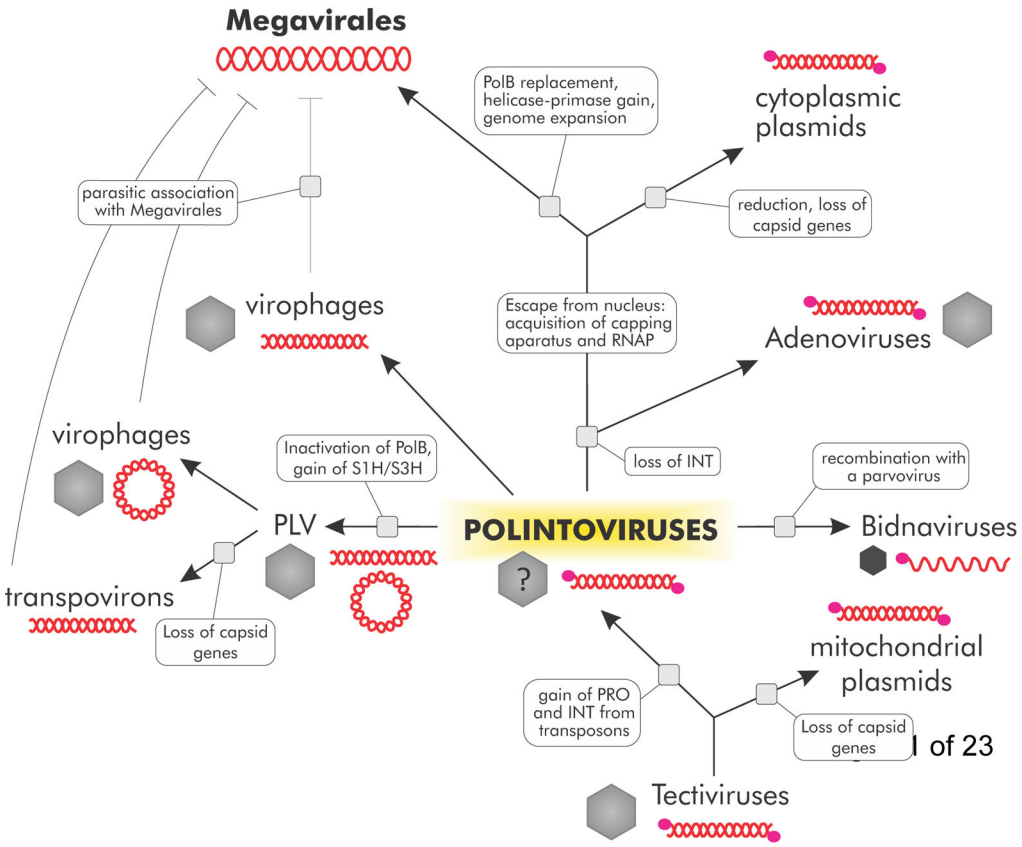

Figure 2. Schematic phylogeny of protein-primed family B DNA polymerases

The schematic shows the topology of the main, strongly supported branches in the tree [30].

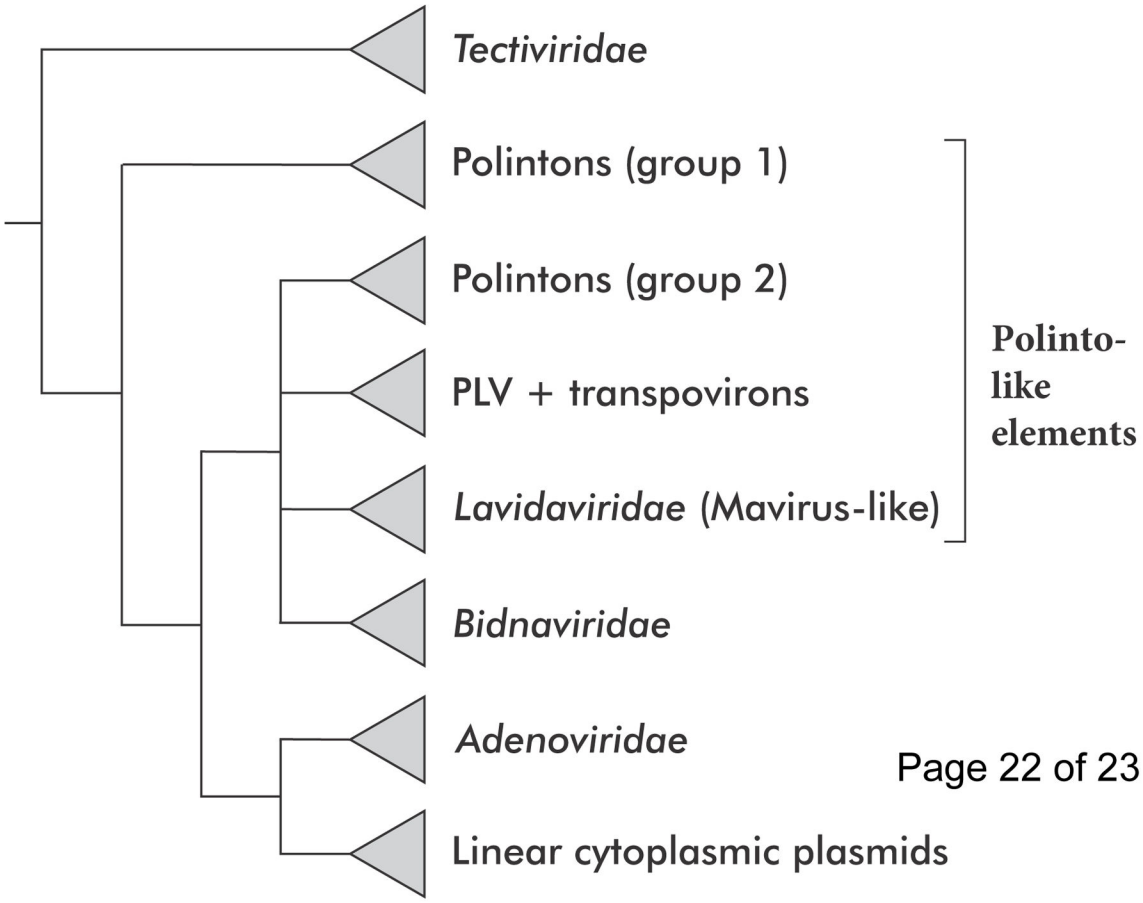

Figure 3. Evolutionary scenario for the Polinton-like MGE and their derivatives

The scenario is based on the previously published analyses [11,18,30]. Abbreviations: INT, integrase; PRO, protease involved in virion maturation; S1H, superfamily1 helicases; S3H, superfamily 3 helicases.

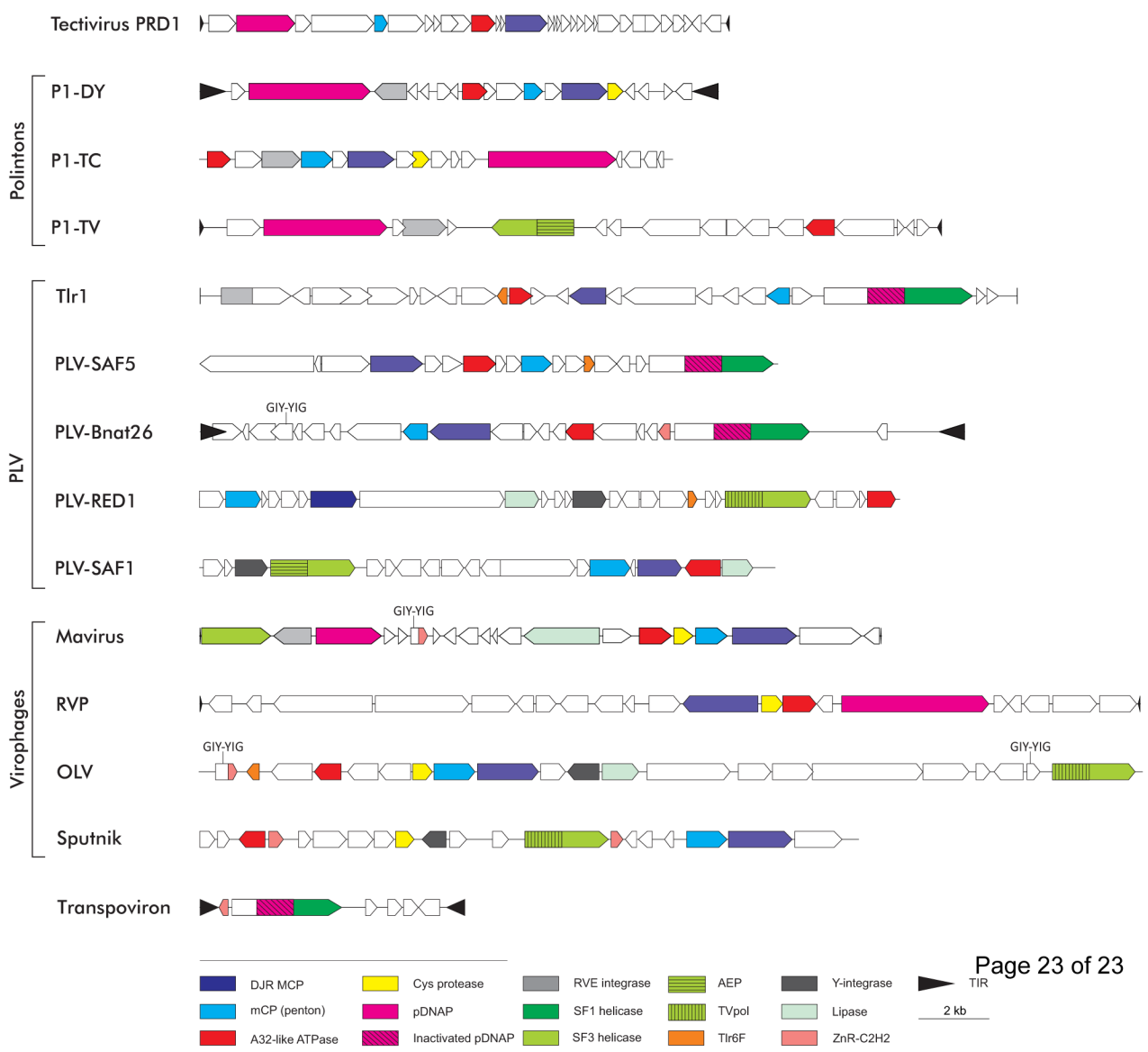

Figure 4. Virophage and CRISPR-Cas: two unrelated but analogous forms of acquired immunity with genomic memory of past infections

R1, R2, R3: CRISPR repeats; S1, S2, CRISPR spacers.