

Review

## Viral Ancestors of Antiviral Systems

Luis P. Villarreal <sup>1,2</sup>

<sup>1</sup> Center for Virus Research, University of California, Irvine, CA 92697, USA;  
E-Mail: lpvillar@uci.edu; Tel.: +1-949-824-6074

<sup>2</sup> Department of Molecular Biology and Biochemistry, School of Biological Sciences, University of California, Irvine, CA 92697, USA

Received: 7 September 2011; in revised form: 1 October 2011 / Accepted: 10 October 2011 /

Published: 20 October 2011

---

**Abstract:** All life must survive their corresponding viruses. Thus antiviral systems are essential in all living organisms. Remnants of virus derived information are also found in all life forms but have historically been considered mostly as junk DNA. However, such virus derived information can strongly affect host susceptibility to viruses. In this review, I evaluate the role viruses have had in the origin and evolution of host antiviral systems. From Archaea through bacteria and from simple to complex eukaryotes I trace the viral components that became essential elements of antiviral immunity. I conclude with a reexamination of the ‘Big Bang’ theory for the emergence of the adaptive immune system in vertebrates by horizontal transfer and note how viruses could have and did provide crucial and coordinated features.

**Keywords:** virus evolution; prophage; evolution of immunity; restriction modification; addiction modules; toxin antitoxin; apoptosis; CRISPR; RNAi; interferon; adaptive immunity; T-cell receptor; MHC; ERV; HERV; LTR

---

### 1. Introduction

Metagenomic studies inform us that viruses are the most numerous and diverse genetic agents in essentially every habitat that has been examined [1–10]. Comparative genomics also informs us that virus derived information is abundant and lineage specific in all life forms. Although such viral information is often defective and historically considered junk, it is also an excellent record of the paleontology of genomes. As genetic parasites, viruses must be compatible with yet extend the existing genetic code of the host. This seemingly simple feature has deep implications and it has been proposed

that viruses are inherent and essential editors of host genetic code [11,12]. Wherever we find life we also find abundant genomic or extragenomic viruses. However, awareness of this virus-host ecology was not part of original Darwinian theory. The most common view is that viruses are harmful genetic parasites that can often kill their host. This well accepted and supported paradigm strongly colors our perceptions of how viruses and host interact and evolve. Viruses are thus modeled as predators to their prey host. Lytic virus destruction of host, sets up an arms race in which clonal selection of host variant (often virus receptors) results in virus resistant new host population [13]. Viruses can then adapt to new populations to ‘kill the winner’. This is classical Darwinian selection of the fittest individual type based on host variation and natural selection. Such a process should result in a slow and stepwise accumulation of genes that resist viruses derived from ancestral genes that can provide host immunity. This process also would not need to invoke any ‘big bang’ or complex horizontal transfer event to create virus immunity. However, even the earliest virus-host studies proposed virus-induced virus resistance [14], and others observed that host virus resistance could also be mediated by other virus [15], even defective ones [16]. A long standing problem with understanding the origin of host virus resistance is explaining how the host wins this battle. Since viruses replicate exponentially, can tolerate even lethal errors at high rates and can freely exchange information with host and other virus, they would seem to always have the upper hand in evolution. Indeed they have been measured to evolve up to a million times faster than the host [17]. Given this situation, how can the host build defense network or system against even yet unseen virus? But viruses are not simply run-away replicators. They are also tenacious colonizers of host able to resist and preclude competition. In this review, I outline the evidence and arguments that viruses themselves often provide the host with antiviral defenses.

Early in my introduction to virology, I became interested in cryptic virus-host states that differed from the classical lytic or pathogenic state. Persistence of virus information seemed to be broadly seen in most life forms, generally with no pathology. Persistent relationships are now known to be ubiquitous, stable and highly host specific [18]. They define a form of virus-host living together that fits within the definition of symbiosis [19]. However, symbiotic does not equate to mutualistic which leads us to ask if these viruses can provide good outcomes to their host. Indeed, examples of beneficial symbiotic viruses are now starting to be acknowledged [20]. However, one aspect of this situation that has received less attention is whether symbiotic viruses have consequences to the same or other viruses. Is there a virus-virus dynamic (such as antiviral immunity) that extends beyond virus-host dynamic? Do these cryptic and often endogenous viruses have consequence to the viral immunity of their host? In some cases (for example, see [21]), the answer is clearly yes, but the generality of this question remains unanswered. Below, I compile evidence for the general role of viruses in the origin of antiviral immunity.

## 2. Prokaryotes

Early in the study of lytic viruses of bacteria, it was observed that host virus resistance could be due to the presence of a second persisting and often genomic virus that would oppose the lytic virus. For example, various common *E. coli* lab strains were discovered to harbor P1 (an episomal temperate phage initially considered a plasmid) that would render the cells resistant to the iconic T4 and lambda

phage infections. And even defective (cryptic) prophage (resembling junk DNA) of various types (such as lambda) could affect host susceptibility to various lytic viruses such as T4. These resident viral agents were also capable of horizontal transfer (infection, colonization), thus quickly imparting other host with antiviral states. Thus, prophages can clearly provide an immune function to host. For example, the iconic *E. coli* K-12 harbors an e14 defective lambdoid cryptic prophage that encodes T4 exclusion, and cytosine restriction functions [22]. Indeed, the long studied but poorly understood iconic *rexB* gene of lambda appears to be an anti-death gene most likely associated with phage exclusion [23]. In addition, other viruses can mediated the phenomena of superinfection exclusion in which a temperate phage efficiently prevents further infection with the same or related virus, see [24]. And some viruses (e.g., P2) can more generally inhibit other viruses (e.g., T-even) [25]. Like viruses, plasmids also provide antiviral systems. Plasmid exclusion is a well known phenomena, see [26]. Historically, viruses and plasmids have been considered as distinct. However, in many cases the mosaic nature of gene exchange between them increasingly makes these distinctions vague, see [27–31]. The P1 example also blurs the distinction between virus and plasmids. However, it was with P1 that a role for T/A (toxin antitoxin) gene sets in inhibiting second infectious agents was first clarified. P1 is an episomal phage that was originally also thought to be a plasmid, see [32]. However, P1 opposes both its loss from host populations and infection with other viruses by the action of programmed cell death (via T/A disruption) [33]. Indeed, the existence of programmed cell death for clonal organisms like bacteria seemed antithetical. One of the gene pairs that P1 uses as a T/A gene set is a restriction-modification (R/M) set. Since R/Ms are major antiviral systems of prokaryotes, the viral role in R/M evolution is considered below. We know of numerous processes involving virus/plasmid by which prokaryotes become resistant to virus [34]. We also now know that most of the genetic variation in the initially sequenced *E. coli* genomes could be ascribed to the action of phage [35]. Comparative genomics makes clear that the dominant process of bacterial and Archaea evolution is through horizontal gene transfer involving gene sets [36]. Viruses are clearly major contributors to this mosaic acquisition process [37] as integration to adjacent tRNA sites is mostly phage associated [38]. Interestingly, there seem to be general differences in phage–plasmid interactions between bacteria and Archaea [39], suggesting domain wide distinctions in their virus and antiviral relationships.

### 3. The R/M System, T/As and Virus

As mentioned above, R/M genes can also be part of a maintenance and apoptosis system in which self destruction is induced with loss from the bacteria [33]. Thus R/Ms can be considered a T/A gene set (genotoxic R, protective M) that induce post segregation self killing [40]. This gene set can also be called an addiction module. Similarly, Eco R1 in *E. coli* is also involved in post segregation killing (as an addiction module). Type II R/Ms are the most common, diverse and a prokaryote-wide system of phage defense genes [41]. Since restriction enzymes recognize small palindrome sequences, host genomes are under R/M mediated phage selection and have shown clear biases regarding the frequency of these sequences in all prokaryotes [42]. R/M gene sets seem especially prevalent in the colony forming filamentous cyanobacteria [43] whose sessile and colonial life strategy may need to resist many viruses [40]. Clearly, phage can encode R/M sets as a T/A set and these have the features of an addiction module. Other phage are also known to encode T/A gene sets [44]. R/M is also found

in plasmids, but here they are often also flanked by phage derived sequences [45]. In *E. coli*, horizontal transfer or variation in R/M systems can be mediated by plasmids that are clearly P1 like [46]. Thus, in many cases, R/M horizontal transfer has occurred with phage or phage assistance. Furthermore, codon word bias confirms that R/M systems (and prophage genes) are often of ‘alien’ origin and acquired by horizontal means [47]. In some bacteria (*Haemophilus influenzae*), phase variation exists in repeat tracks expressing the R/M system along with the lipooligosaccharide phage receptor promoter to generate more diverse phage resistance [48]. This represents a more elaborate but still R/M derived defense against phage. Thus, the R/M system, the most diverse prokaryotic antiviral system known, shows strong evidence that viruses are crucial for their origin and evolution and are guarding the host genome.

#### 4. Toxins, Host Adaptation and Phage

Our focus is on the origin of antiviral systems. Yet, ironically, virus (e.g., provirus) can be crucial for bacterial adaptability, thus a fully effective antiviral system could limit host adaptability. This is particularly apparent regarding the link between bacterial prophage and toxins, including defensive R/M genes, see [49]. Toxin and R/M genes are crucial for certain habitat adaptations and such habitat adaptation is also associated with virus defense genes. Toxins are particularly abundant in marine bacteria [50]. Similar bacteria from distinct locations (*Candidatus Accumulibacter phosphatis*) will differ predominantly in regions encoding phage defense genes [51]. Prophages also contribute to habitat adaptation outside of R/M genes. In one example of deep sea hydrothermal vents, metagenomic analysis indicates that inducible temperate prophages are common and contain surprisingly numerous novel genes [52]. In other cases, it seems clear that bacterial phage are not simply promoting clonal sweeps as they appear to maintain host diversity [53]. The link between adaptability, toxin and phage has especially been studied with toxigenic strains of *E. coli* (O157) that differ in toxin carrying cryptic phage sets (especially lambdoid) [54–59]. These horizontally acquired genes are transferred via the promoting activities of defective and cooperating prophages [60], which can also provide globally coordinated gene regulation [61]. Even commensal *E. coli* strains harbor similar but strain specific cryptic phage [62]. These observations are consistent with a general role of lysogenic prophages in pathogenicity (and adaptability) of bacteria as has been previously proposed [63]. There is also a relationship between cryptic phage/plasmids and biofilm formation, another process of bacterial adaptability. Elimination of cryptic prophages from *E. coli* (including their encoded T/As) results in lost biofilm formation [64,65]. Abortive virus infection can also be toxin mediated. Diverse bacteria have T/A gene pairs, like R/M, that can preclude or abort infections with diverse phage via toxin mediated self destruction [44,66]. One prevalent version of this abortion system uses tandemly repeated Toxi RNA to counteract ToxN gene function [67]. Indeed, pathogenic island transfer between bacterial species has also been observed mediated by phage [68]. Furthermore, T/A gene sets also seem involved in persister cell formation [69], which can also affect virus production. Thus, phage derived ‘antiphage T/A defense systems’ often provide the host not only with adaptations against other viruses, but also with adaptations for better habitat colonization.

## 5. CRISPRS

CRISPRS are clusters of regularly interspaced palindromic repeat DNA sequences linked to CAS genes that are found in both Archaea and bacteria but more prevalent in Archaea. The locus provides RNA directed immunity against viruses and plasmids, see [70,71]. The process involves dsDNA recognition via guide CRISPR RNA (crRNA) [72]. Some CRISPR loci can also counteract plasmids via self splicing intron insertion into plasmid DNA [73]. CRISPR is an effective but mostly antiviral and anti-plasmid system that retains memory of past viruses by incorporating identity-recognition relevant information from these viruses [74]. This memory thus provides a powerful and advanced adaptive feature. Indeed, the variable spacer regions of CRISPR loci has been used to detect previously unknown viruses that have infected the host [75]. Memory is thus a distinct feature of CRISPR. In Archaea, the majority of virus infections are chronic, although some are also clearly lytic [7]. These archaeal viruses have numerous distinct morphotypes, which are thought to represent ancient lineages [76]. Some archaeal plasmids can also be packaged and spread by virus [77]. In contrast to bacteria, however, Archaea appear to lack RNA viruses. It is not yet known if CRISPR loci can account for this difference. CRISPRS (like T/As) can also be needed for group (biofilm) behavior [78]. Since the distribution of the CRISPR palindromic repeats in virus/plasmid genomes is random, the prevailing hypothesis is that these repeats are derived from these viruses and plasmids [79]. The CRISPR loci shows uneven coevolution with their host. In general, CRISPRS are species conserved and stable during host evolution. However in *E. coli*, the corresponding CAS gene is rapidly changing and incongruent with CRISPR loci and host species, thus it has a distinct and horizontal evolutionary history [80]. This observation indicates the CRISPR role is not simply to protect bacteria from phage, but likely has additional functions, such as a possible role in community structure. However, some have proposed that an effective CRISPR system may limit prophage colonization (especially in Archaea). It is thus interesting that Archaea and bacteria also differ considerably with respect to toxin production in general, given that toxins are most often prophage encoded [81].

## 6. Eukaryotes Shift Virus and Antivirus Defenses

Interestingly, R/M systems are almost entirely absent from eukaryotes, although they are well conserved in the chloroviruses (PBCV1), large DNA viruses that infect unicellular eukaryotic green algae [82]. In addition, type II restriction-like genes (R lacking M) are reported in R4 family retrotransposons found in invertebrate and vertebrate [83]. The prokaryotic wide R/M defense systems are thus clearly no longer significant components of eukaryotic antiviral defense. Nor is the CRISPRS system retained in eukaryotes as neither the repeat loci nor the CAS gene is found therein. Given the major viral role in bacterial adaptability discussed above, why might this relationship have changed so completely with the emergence of eukaryotes? Along with this global change in antivirus defense, the integration of large ds DNA viruses into host chromosomes (*i.e.*, prophage life strategy) became most uncommon in eukaryotes. Instead, we see a large scale increase in the integration activity of retroviruses and retroposons into eukaryotic chromosomes. And although the CRISPR system was not retained, we instead observe the emergence of an RNA based antivirus defense system that resembles

CRISPRs in many regards (such as dsRNA directed virus recognition). Below I consider the likely role of viruses in the emergence of eukaryotic antiviral systems.

### 7. *C. elegans*, RNAi and Antivirus Systems

The model nematode worm *C. elegans* has been productively studied regarding RNA based antiviral systems. Like all other eukaryotes, *C. elegans* lacks the CRISPRs loci, R/M systems, DNA prophage, or plasmid exclusion systems as found in prokaryotes. But *C. elegans* does have an effective RNA based antiviral response. The process operates via small or micro RNAs; small ssRNA that interfere (RNAi) with gene expression [84] or silences (siRNA) gene expression [85]. Components of this system are conserved in most, but not all, eukaryotes. The small RNA response can sometimes affect genes in groups [86]. Some of these small RNAs can be expressed from genomic DNA (corresponding to LTR regions of endogenous retroviruses and transposons, presented below). This system can provide antiviral state via virus derived siRNA [87]. The origins of this system have not been defined and no prokaryotic homologues are found. The system involves an RNA dependent RNA polymerase (RDRP) to amplify the signal RNA [88]. As dsRNA based genetics is only in the domain of viruses (not host) this clearly resembles a virus-like strategy. This gene has no counterparts in the Archaea or bacteria [89]. Nor has a cellular ancestor to this RNA pol been identified. However, the RNA pol does show some domain similarity to a phage DNA dependent DNA pol, thus some type of viral mosaic origin seems possible [89]. In addition, filamentous fungi (considered basal to metazoans) are typically chronically infected with diverse dsRNA viruses that encode related genes and these genes can be truncated and/or acquired by horizontal gene transfer [90] (see below). Another component of this system involves the cleavage of dsRNA via the action of RNase III (Dicer PAZ domain). Like RDRP, RNase III also lacks homologues in Archaea and bacteria [90] and appears to have been acquired by horizontal transfer. Adding to this evidence of virus-like features is the highly transmissible nature of the RNAi response in *C. elegans* [91–93]. Such amplifying cell to cell transmission clearly resembles a virus-like strategy, further suggesting viral origins. However, all these observations are only inferential and more direct evidence is needed.

Nematodes seem to have a very effective antiviral system. Given the large and old collections of cultured nematodes, especially with respect to their roles as plant disease vectors, and given the length of the use of *C. elegans* as a laboratory model, there is a surprising paucity of reports of viruses that infect nematodes. Recently, persisting non-pathogenic infections with noda-like [94] and rhabdo-like viruses have been reported for nematodes [95]. Clearly, nematodes have yet to show the large virus mediated population crashes we describe below for marine vertebrates and invertebrates or are well established for bacteria. Their effective antiviral system is based on dsRNA response that targets RNA via small RNAs [96]. And if incapacitated, this system renders the worm susceptible to pathogenic infection with various RNA viruses [87,97–100]. A main component is a dsRNA binding protein and an endonuclease to degrade the corresponding mRNA into siRNAs. Indeed, dsRNA recognition via small RNAs (such as RNAi, siRNA, microRNA or miRNA) is a general theme of antiviral immunity not restricted to invertebrates that we can follow to plants and vertebrates [96]. Unlike the R/M system described above, however, it does not generally appear to act as a T/A system.

In *C. elegans* the majority of small RNAs are endogenously expressed, mostly deriving from endogenous LTR transcription. Indeed, the *C. elegans* genome underwent a substantial colonization by Gypsy-like ERVs (chromovirus) during its evolution and although most are defective some are essentially intact [101]. The LTRs from these retroviruses were scattered throughout the genome and this has led to proposals that dsRNA response was initially a defense against retroviral elements [102,103]. If, however, the ERV colonization was itself important for the origin of this antiviral system, we can instead assert that this dsRNA system was derived from retro and other viruses to preclude self and other viral competitors (by providing self identity and allowing stable ERV maintenance) [104]. Thus, this could explain why it appears to be a defense against other viruses yet account for why LTRs remain responsible for the expression of endogenous small RNAs. This system is thus guarding the genome against non-self elements while allowing coexistence with ‘self’ ERVs [105,106]. siRNA was clearly conserved in other more complex invertebrates, but here the small RNA may more often be derived from exogenous viral, not LTR RNAs. Accordingly, it may be possible to use siRNA to identify previously unknown exogenous viral genomes that have infected or are colonizing their host. This idea is very much like that applied to the CRISPRs system noted above [107]. Interestingly, although this approach was effectively used in *Drosophila* (ovary) to find new viruses [108], in *C. elegans* no viral genomes were thus identified. This suggests that there are indeed few persisting viruses in *C. elegans*, consistent with a virus paucity discussed above. Curiously, in *Drosophila*, piwi interacting RNA is also thought to be an anti-transposon RNA (via flamenco rasi RNA, antisense to gypsy elements) [108], but piwi RNAs are expressed only in the ovary [109], as are gypsy ERVs in some fly strains. This is a curious overlap of ERV and anti-ERV putative function. However, all fly strains retain gypsy defective elements. Indeed, one-quarter of all cloned small RNAs from *Drosophila* are also derived from such retro elements [110]. Together, these observations can support the assertion that retroviruses are indeed involved in the origin of these antiviral systems.

## 8. Paleovirology: The Filamentous Fungi

Like other early metazoans, fungi have most of the small RNA genes (including dicer) noted above, but the specific occurrence and composition of such genes is more diverse in the early metazoans [111,112]. The viral susceptibilities of filamentous fungi is also distinct from those of both prokaryotes and simple animals (*C. elegans*). Many RNA viruses are known to infect fungi, but especially interesting is the ubiquitous persistent and nonpathogenic colonization by various dsRNA viruses [113,114]. These silent, cryptic viruses [115] mostly have no extracellular mode of transmission and depend on fusion of hypha for transmission [116,117]. Indeed, some have proposed that the existing hyphal mating type restriction is for the purpose of virus control. Furthermore, some of these viruses can provide selective and pathogenic (toxin) phenotypes to their host fungi [118], whose transmission is also mating type restricted [20,115,119,120]. A core RNA viral gene is the RDRP. Recently, reports of widespread horizontal ‘module’ transfer of dsRNA viruses (especially RDRP) into the nuclear genomes of eukaryotes have been published [121]. Such ubiquitous infection of filamentous fungi with dsRNA viruses clearly distinguishes the situation from that seen in *C. elegans* as noted above. Consistent with this inference, the conservation of the antiviral small RNA genes is also most uneven in fungi. These genes have been conserved and multiplied in some lineages

but are also truncated in various other filamentous and nonfilamentous fungi [111,122]. It seems likely that the presence of these antiviral systems will also affect virus colonization. Of great interest with respect to likely origins is that some filamentous fungi do not use cellular RDRP for inverted repeat RNA silencing [123]. Fungi also seem to have some distinct recognition systems not seen in other species. For example, the RIP system of neurospora precludes the accumulation of genomic (non-nucleoli and non-plastid) repeats by heavy mutation of LTRs and ERVS during sexual reproduction [124,125]. Yet other lineages of fungi have significant LTR content which is associated with pathogenic adaptation [126]. The fungal ERVs are especially of chromoviral origin [127]. Thus in the fungi, viral specific associations and involvement in antiviral systems seems basic to their evolution.

### 9. Paucity of Tunicate Virus, Antivirus States and the Problem of Gradualism

It is a long evolutionary jump from worms to people if we think about the continuing emergence of immune systems. Our interest thus turns to the tunicates, as the most likely ancestor to living vertebrates [128]. Tunicates have a highly polymorphic fusion/histocompatibility locus (FuHC) associated colony fusion or destruction [129,130]. In addition, lampreys and hagfish have distinct, highly diverse and modular variable lymphocyte receptors (VLRs) that promote lymphocyte pattern recognition [131,132]. These loci are both considered to also provide pathogen immunity. Neither of these loci bear similarity to classic genes of the adaptive immune system of jawed vertebrates. Nor do they have the classic interferon alpha (dsRNA response) or interferon gamma system, although they do have complement [133] and Thy-1 like genes [134,135]. It is clear that these tunicate systems emulate various features of the adaptive immune system, even though they lack specific homology [136,137]. Thus their emergence was clearly a distinct evolutionary event. Chromoviruses (gypsy) remain abundant and active in tunicate genomes [138,139]. Yet tunicates retain the smallest of all vertebrate genomes [139,140]. Currently, and most curiously, there is no evidence that these tunicate systems function as antiviral systems. Unlike the small RNA system of *C. elegans* presented above, the antiviral consequences of mutations in the tunicates systems have not been evaluated. This may partially be due to the paucity of viruses so far found in tunicates. There are no laboratory models for tunicate infection with virus, possibly indicating that tunicates have effective antiviral systems. One recent report using metagenomic approaches did find the presence of rhabdovirus RNA (VHSV) in the European river lamprey [141]. However, this appeared to be a cryptic, nonpathogenic infection with unknown linkage to the VLR system. Thus tunicates resemble nematodes in the low occurrence of natural virus susceptibility. Although we cannot link these tunicate systems to that of jawed vertebrates, many researchers are still proposing a gradual transition to the origin of adaptive immunity via convergent evolution [137,142]. One of their main criticisms against the long held 'big bang' theory is that the acquisition of all the complex linked features of adaptive immunity is considered too 'extreme' [142,143]. Gradual evolutionary mechanisms are instead favored. However, arguments that acknowledge the evidence for or major effects of horizontal transfer do not appear to underlie these views. Rather they appear to be based solely on the belief that only point mutation and natural selection is needed to account for such genetic novelty, thus it is posited that whole genome duplication and slow adaptation of individual genes co-opted from various other uses can fully explain the origin of adaptive immunity [142–143]. That no living organism resembles the intermediate



required for this gradualist process is apparently not a major concern. More fundamentally, however, I disagree with these views because they fully ignore a much deeper issue of the regulatory complexity and network coherence that is required for originating the adaptive immune system. The adaptive immune system is an especially troubling example of a novel complex regulatory network of interacting genes that was absent from any living ancestor. Even ignoring all the needed new genes, simply solving the dispersed editing and regulatory requirements of adaptive immunity presents a major conundrum by itself. This indeed requires a ‘big bang’ like event of complex horizontal colonization by a population of linked agents with coherent regulatory competence. Indeed, a most basic problem regarding editing of the genetic code in general is that needed to create a new network. The fundamental hurdles to coordinate distributed code to the degree needed to generate such regulatory complexity are much more challenging and difficult than has previously been acknowledged [144,145]. Indeed, this problem relates directly to the nature of code ‘meaning’ itself since edited preexisting code (*i.e.*, co-opted genes and regulators) must be assigned new meaning in the context of the new network. An incremental and individual based process for this is not sufficient. Point mutation errors followed by selection of the fittest type cannot account for such network coherence. However, the colonization by a population of linked agents can promote coherence or network formation [146]. I propose here that mosaic populations of mixed viruses and their regulatory elements are naturally competent in host code usage (especially that used by immune systems) and are also natural extenders (editors) of new (virus derived) code. Let us now examine the origin of the adaptive immune system from this perspective.

## 10. Components of Adaptive Immunity of Jawed Vertebrates: Big Bang Reexamined

The early ‘big bang’ view of the origin of adaptive immunity required the horizontal transfer into emerging vertebrates of almost the entire system *en toto*. This system could then evolve into greater complexity as seen from fish to humans [147]. But where did such a complex system come from if not from an ancestral organism? In a prior publication, I presented a detailed and in depth proposal of how and why viruses and other genetic parasites would provide the likely origin of the adaptive immune system [148]. As none of the recent ‘gradualist’ reviews on the adaptive immune system have addressed or critiqued the main assertions made therein, the evidence and arguments previously presented remain pertinent. Below I outline those main issues and provide some crucial references that relate to the role viruses and genetic parasites played in the origin and evolution of adaptive immunity. Teleost fish genomes underwent a big virus colonization (ERV) and genome expansion relative to tunicates [149]. This involved a significant shift including both chromovirus expansion (Gypsy-like ERV) and colonization by many other new retrovirus families [150,151]. Along with this major genome change, we see that vertebrates acquired all of the core elements of adaptive immunity. This also included a major shift in innate immunity (including altered apoptosis and interferons—alpha and gamma). The specific elements of adaptive immunity include the antigen specific cytotoxic T-cell (CTL), antigen processing and presentation (via MHC locus), adaptive antibody, memory, and IFN alpha and core regulation by IFN gamma [152,153]. The organs and tissues needed to produce all this also emerged with the vertebrates. This system functions as a complicated and coherent network. However, the regulatory IFN system of fish also shows direct antiviral activity (not antigen

specific) [152]. Vertebrates also show a big increase in the complexity of their apoptotic response [110]. It has been argued that most of these changes appear to have originated via horizontal gene transfer.

It is commonly assumed that the adaptive immune system with its anticipatory capacity is much more advanced and capable than the ancestral immune system. We have seen, however, that effective antiviral systems based on RNA were widely present in various invertebrate ancestors in highly virus infested habitats. In contrast to tunicates, fish have lots of viral disease. This includes a diverse set of viruses that can cause large population crashes in both natural and farmed fish populations. For example, teleost fish population crashes due to rhabdoviruses, herpesviruses and infectious retroviruses that transform immune cells have all been frequently observed [154,155]. Indeed, viral infection of immune cells is rather common in the jawed vertebrates. So was the adaptive immune system better for controlling viruses? Possibly not. Nor is there evidence for large selective sweeps by viral pathogens of ancestor populations in the geological record. Thus it is not clear what selective pressure led to the emergence of adaptive immunity. One interesting idea concerns the invention of teeth and jaws used by predators in the consumption of living prey (mostly marine invertebrates). The resulting exposure to wounds, blood and ubiquitous virus from marine invertebrates could have promoted the origin of adaptive immunity [156]. Another view is that an ancient form of antiviral immunity led to the adaptive immune system [157]. Let us now follow this perspective and consider how it might explain the various extant features of adaptive immunity. These central features are conserved from fish to humans (who also have the most complex MHC locus) and are outlined below [147]. Along these lines, cytotoxic T-cells (CTLs) would have initially emerged to kill virus infected cells. The T-cell receptor is found within a 1 Mbp cluster of interacting genes (via protein-protein interaction), which formed contemporaneously with the origin of the adaptive immune system [158]. These did not originate by genome duplication, nor are they found in chordates. Of particular interest is that the MHC locus is the most dynamic and gene dense region in the chromosomes of fish and humans. However, it made dynamic via the action of ERV and other retroposons as described below.

## 11. The T-Cell Receptor (TCR and Ig Family)

It has been pointed out that a cellular virus receptor will bind its virus with good affinity, but if the receptor is also released into the media, it can potentially also bind and inactivate free virus, thus providing the basis for soluble antibody to bind virus (Ig) [159]. The TCR appears to be the most basal form of the Ig family members of the immune system. However, even more ancestral features of the Ig superfamily are found in various virus receptors, such as CTX (Coxsackie receptor), PRV (poliovirus receptor), JAM (ds RNA reovirus). Similar receptors can be found in *Ciona*. Thus it has been proposed that these virus-binding properties were recruited for origin Ig superfamily in the adaptive immunity, then secreted to produce early virus neutralizing activity [159,160]. Along these lines, the adenovirus CAR receptor is known to be shed as a soluble form [161]. However, as outlined in my paper, and noted briefly above, viruses themselves are often the source of the most variable virus receptors [148]. In addition, there are several examples of virus receptor systems that also encode a linked system for generating variation in the receptor. The colonization of the host genome by such a virus that encoded a variable receptor system would provide the foundation TCR (and Ig) variation. Along these lines, a very interesting situation is found in sharks (basal vertebrates with adaptive immunity). Sharks express

high levels of natural antibody (from unrearranged Ig genes) [162]. These antibodies have affinity to TCR but also for numerous retroviral proteins, especially envelope proteins. This natural affinity may identify the original target of the Ig immune proteins and would be consistent with a very early retroviral involvement.

## 12. The TCR ‘Receptor’ Must Link to a Recombination System

The recombination and variation of TCR and the Ig family members occurs via the site specific action of RAG1/2 recombinase involving sequence specific binding for V(D)J recombination. These two genes are not found in tunicates or most other invertebrates [163]. It has long been noted that this integrase has features of phage and viral integrases (especially that of phage Mu needed for replication). As RAG1/2 also targets small repeat sequences for integration, several early proposals suggested viral involvement in the horizontal acquisition of RAG1/2 [164,165]. Clearly the RAG1/2 recombinase along with its target sequences resembles an ancient site specific transposable element. More current views, however, propose that RAG1 and RAG2 were derived from selfish DNA transposons via horizontal transfer [166]. However, how this recombinase became regulated and linked to the TCR genes or how this might have occurred without involving viruses is mysterious. It is typically assumed that classical Darwinian selection was sufficient to provide this needed linkage and coordination. In my prior review, however, I provide a detailed scenario of how a ‘preevolved’ virus based system for receptor gene variation with inherent links to an integration process could colonize the genome and provide the foundation of this needed TCR-RAG link [148]. Numerous and highly conserved examples of homing introns found in specific viral core genes are known. Similar systems for viral receptor gene variation are also known. Thus host colonization of mixed viral receptor module with a coordinated variation system would allow the acquisition of a coherent system much more quickly, resembling a big bang like event. In this case, an intermediate organism with only some of these components would have never existed.

## 13. MHC Locus Is Dynamic via ERVs

The MHC gene loci (I, II and III) encodes the TRC and Ig genes but also numerous other genes for processing and presentation of antigen, see [167]. These loci are the most polymorphic of all human genes. MHC class III is also the most gene dense (intron poor) region of human chromosome [168]. Many of the genes found within these loci are mammal and fish specific [168]. These genes have been linked together since the evolution of fish [169]. This raises the question regarding what was the source of most of these genes. The high plasticity of the MHC loci results from the action of many densely clustered retroelements [170]. The primate MHC loci are particularly dense with ERVs as they are in 10-fold greater abundance compared to other regions of the chromosomes [170,171]. High ERV MHC density is also seen in sharks, even though shark genomes do not host the same general high ERV level seen in primates [172]. MHC I appears to be ancestral to MHC II and their evolution has been by duplication and divergence of a core alpha block of genes [173,174]. ERVs are clearly involved in this duplication process, and with humans and chimp DNA, ERV9 can be used to categorize the types of duplication that has occurred [175]. The core MHC I ‘duplicon’ thus contains an ERV (HERV 16) [176–178]. MHC retroposon variation is especially found in introns of genes,

resulting from HERVs and retrotransposon integration [179]. These retroposons are generating the diversity that also distinguish humans from chimp MHC loci. Interestingly, from a regulatory perspective, some of these MHC ERVs are retained in antisense orientation [180]. Thus there is little question that ERVs and retroposons provide the major variation in the MHC locus. Greater details regarding why viruses would be expected to have played such a role (from the perspective of identity systems) are presented in my previous review [148].

#### **14. The Sudden Emergence of IFN Alpha and Gamma System**

Along with the emergence of jawed vertebrates, we also see the emergence of both the interferon alpha and the interferon gamma system, crucial control systems that are both directly antiviral and initiate and maintain the adaptive immune response [181,182]. Comparative analysis of these two systems indicates that the interferon alpha system is more basal [152,183]. Although marine invertebrates have a virus associated dsRNA response with clear similarities to vertebrate interferon [184], this system differs notably from that of vertebrates. In jawed vertebrates, dsRNA longer than 30 bp can induce interferon which can also trigger many undesirable toxic side effects, such as apoptosis. Since dsRNA response systems were present before the emergence of vertebrates, it is simplest to propose that these systems were simply co-opted for use by the adaptive immune system. Yet the many tightly coordinated components of this response (such as NF-kappaB, [182]) still presents an evolutionary problem for attaining regulatory coherence. Invertebrates do not have the IFN alpha or gamma network, but various viruses clearly have or regulate essentially all of the genes involved.

#### **15. T Cells Must Clonally Transform and Differentiate**

As noted above, some major and coordinated changes were also needed in cell reprogramming (code editing) in order to create the tissue and cell types associated with the adaptive immune system. No single event would seem able to accomplish such a complex process. Yet, host cell transformation along with reprogramming differentiation is precisely the type of change viruses are evolved to perform. Indeed, HHV-6 for example can integrate [185], replicate nonlytically in dendritic cells and be required for efficient transmission and productive infection when cocultured with stimulated CD4 T cells (needing cell-cell contact) [186]. Similarly, CMV induces CD4+ and CD8+ cells to differentiate and over proliferate [187], resulting in T cell memory inflation [188]. EBV integration is distributed in chromosomes, providing potential regulatory links [189]. All these viral abilities would also have been important for the normal function of adaptive immunity. As genetic parasites viruses must be compatible with existing host code, but also extend it with new viral derived code to ensure virus maintenance or reproduction (and preclude competition). Indeed, both retroviruses and large DNA viruses of vertebrates and invertebrates are known to infect, transform and alter differentiation of immune cells or their stem cell precursors. Plus, gene sets within viruses would necessarily be under pre-coordinated viral regulatory control to operate with some coherence. Thus virus colonization provides a natural solution for reprogramming host cell fates. This reasoning was also previously developed to much greater detail [148].

## 16. Conclusions: Ongoing Viral Solutions to Adaptive and Other Immunity

We can find no extant organism that displays the gradual acquisition of the various components of the adaptive immune system. The complex coordinated regulation, along with complex coordinated dense gene sets involved do not exist outside of the jawed vertebrates. Yet, examples of most of these coordinated processes and genes can be found in viruses of various prokaryotes and eukaryotes. For example, viruses using various types of phase variation for receptor genes and addiction modules for self destruction are well known, see [148]. Thus a viral origin of such a linked process seems more plausible than a stepwise acquisition requiring reassignment of function. Viruses are known to encode and manipulate almost all aspects of the interferon network, see [190]. This is especially true for large persisting DNA viruses of the herpes family. The large DNA viruses of vertebrates are master modulators of MHC mediated immunity and persistence [190,191]. These herpesviruses show synchronous virus host evolution [192]. Some members, such as CMV, encode G coupled proteins and chemokines [193]. However such MCMV genes are usually assigned to evolve from host genes by gene piracy [194]. Thus, the similarities that viruses show to host interferon (or MHC) genes are broadly accepted to be due to viral piracy of host genes. In my judgment, such views are based on an incorrect assumption that any similarity between virus and host genes will necessarily support the idea of viral piracy of host. Yet many of these same viral genes are clearly unlike and often much simpler than any host version and are highly conserved in viral lineages [195]. Along these lines, I have argued vigorously that the converse situation is more likely true; that viruses originate these coordinated and dense gene sets associated with antiviral responses and have provided them to the host via colonization (a big bang like horizontal transfer). For example, CMV can completely suppress CTL mediated lysis [196]. This is a complex function that would also be essential for the origin of adaptive immunity itself in order to control the potential self destructive capacity of the system. CMV is a member of the herpesvirus family with ancient evolutionary links to phage. But unlike phage, the herpesviruses generally evolve via action of retrovirus and are intron poor. This contrasts with other DNA viruses including phycodnaviruses of algae (distant relative of herpes). Indeed some herpesvirus members can integrate fully functional retrovirus which results in an enhanced cooperating phenotype [197,198]. Clearly such a mixed virus must have attained regulatory coherence, yet it also has acquired the capacity for wide scale and distributed host integration with related regulatory elements. Thus virus evolution indeed can be mixed and complex. No single lineage of virus appears capable of providing all the complex changes needed to originate adaptive immunity. Complicated and cooperating mixtures of virus, however, do seem do able to provide these functions. Others have also previously considered whether viruses might have been involved in the origin of the adaptive immune system [199]. However, this was a much more narrow idea limited to the source of the RAG recombination system.

The acquisition of regulatory complexity, especially involving networks, appears to identify other key problems for recent human evolution. The great leaps forward or major evolutionary transitions associated with the emergence of humans (such as a large social brain) all seem to require coordinated regulatory changes, not simply new genes [200–202]. Thus the recent changes in the human genome are mostly found in non-coding DNA [203]. Here, HERVs and their LTRs appear to have had major effects on promoter or intron function in a process that appears to involve RNA interference and other regulatory mechanisms [204,205]. Some suggest a direct effect of LTRs on human promoters is

small [206] although others claim human LTR promoter use is significant [207,208]. Few HERVs that are actively expressed in humans have open reading frames however [209]. Thus HERV/LTR effects are mostly RNA regulatory with an inherent potential to coordinate, even if sometimes the effects are deleterious [210]. With respect to immune system, here too it looks like HERV mediated changes predominate. Thus in spite of close overall conservation of human and chimpanzee ORFs, the dispersed indel differences in the MHC locus are considerable which results in species specific difference in retro and other viral sensitivities [211]. For example, the major MHC II polymorphism between human and chimpanzee via intronic sense HERVK31 and antisense for other ERVs is associated with lentivirus susceptibility [179]. Interestingly, EBV mediated introduction of HERV-K into primate cell lines has been reported suggesting virus cooperation [212]. Thus viruses are modifying antiviral systems in recent human evolution.

My objective for this review was to bring to the attention of readers the general involvement of viruses in antiviral systems. We have examined restriction-modification, CRISPERS, RNAi, and the adaptive immune system all from this perspective. In all cases, complex coordinated changes mediated by viruses can be observed. Essentially all antiviral systems appear to have some virus derived components. Perhaps because viruses can provide complex solutions, oppose virus competition and are able to inherently edit and extend host code they may be well suited for this general role.

## References and Notes

1. Breitbart, M.; Hewson, I.; Felts, B.; Mahaffy, J.M.; Nulton, J.; Salamon, P.; Rohwer, F. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* **2003**, *185*, 6220–6223.
2. Angly, F.E.; Felts, B.; Breitbart, M.; Salamon, P.; Edwards, R.A.; Carlson, C.; Chan, A.M.; Haynes, M.; Kelley, S.; Liu, H.; *et al.* The marine viromes of four oceanic regions. *PLoS Biol.* **2006**, *4*, e368.
3. Santos, F.; Meyerdierks, A.; Pena, A.; Rossello-Mora, R.; Amann, R.; Anton, J. Metagenomic approach to the study of halophages: The environmental halophage 1. *Environ. Microbiol.* **2007**, *9*, 1711–1723.
4. Breitbart, M.; Haynes, M.; Kelley, S.; Angly, F.; Edwards, R.A.; Felts, B.; Mahaffy, J.M.; Mueller, J.; Nulton, J.; Rayhawk, S.; *et al.* Viral diversity and dynamics in an infant gut. *Res. Microbiol.* **2008**, *159*, 367–373.
5. Desnues, C.; Rodriguez-Brito, B.; Rayhawk, S.; Kelley, S.; Tran, T.; Haynes, M.; Liu, H.; Furlan, M.; Wegley, L.; Chau, B.; *et al.* Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature* **2008**, *452*, 340–343.
6. Dinsdale, E.A.; Edwards, R.A.; Hall, D.; Angly, F.; Breitbart, M.; Brulc, J.M.; Furlan, M.; Desnues, C.; Haynes, M.; Li, L.; *et al.* Functional metagenomic profiling of nine biomes. *Nature* **2008**, *452*, 629–632.
7. Schoenfeld, T.; Patterson, M.; Richardson, P.M.; Wommack, K.E.; Young, M.; Mead, D. Assembly of viral metagenomes from yellowstone hot springs. *Appl. Environ. Microbiol.* **2008**, *74*, 4164–4174.

8. Williamson, S.J.; Rusch, D.B.; Yooseph, S.; Halpern, A.L.; Heidelberg, K.B.; Glass, J.I.; Andrews-Pfannkoch, C.; Fadrosch, D.; Miller, C.S.; Sutton, G.; *et al.* The sorcerer ii global ocean sampling expedition: Metagenomic characterization of viruses within aquatic microbial samples. *PLoS One* **2008**, *3*, e1456.
9. Santos, F.; Yarza, P.; Parro, V.; Briones, C.; Anton, J. The metavirome of a hypersaline environment. *Environ. Microbiol.* **2010**, *12*, 2965–2976.
10. Roossinck, M.J.; Saha, P.; Wiley, G.B.; Quan, J.; White, J.D.; Lai, H.; Chavarria, F.; Shen, G.; Roe, B.A. Ecogenomics: Using massively parallel pyrosequencing to understand virus ecology. *Mol. Ecol.* **2010**, *19*, 81–88.
11. Witzany, G. Natural genome-editing competences of viruses. *Acta Biotheor.* **2006**, *54*, 235–253.
12. Villarreal, L.P.; Witzany, G. Viruses are essential agents within the roots and stem of the tree of life. *J. Theor. Biol.* **2010**, *262*, 698–710.
13. Luria, S.E.; Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **1943**, *28*, 491–511.
14. d'Herelle, F. *The Bacteriophage and Its Behavior*; Williams and Wilkins: Baltimore, MD, USA, 1926.
15. Delbrück, M.; Luria, S.E. Interference between bacterial viruses. *Arch. Biochem.* **1942**, *1*, 111–141.
16. Redfield, R.J.; Campbell, A.M. Origin of cryptic lambda prophages in escherichia coli k-12. *Cold Spring Harb. Symp. Quant. Biol.* **1984**, *49*, 199–206.
17. Wain-Hobson, S. Retrovirus evolution. In *Origin and Evolution of Viruses*, 2nd ed.; Domingo, E., Parrish, C.R., Holland, J.J., Eds.; Academic Press: San Diego, CA, USA, 2008.
18. Villarreal, L.P.; Defilippis, V.R.; Gottlieb, K.A. Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology* **2000**, *272*, 1–6.
19. Ryan, F.P. Viruses as symbionts. *Symbiosis* **2007**, *43*, 11–21.
20. Roossinck, M.J. The good viruses: Viral mutualistic symbioses. *Nat. Rev. Microbiol.* **2011**, *9*, 99–108.
21. Yan, Y.; Buckler-White, A.; Wollenberg, K.; Kozak, C.A. Origin, antiviral function and evidence for positive selection of the gammaretrovirus restriction gene fv1 in the genus mus. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3259–3263.
22. Mehta, P.; Casjens, S.; Krishnaswamy, S. Analysis of the lambdoid prophage element e14 in the e. Coli k-12 genome. *BMC Microbiol.* **2004**, *4*, 4.
23. Engelberg-Kulka, H.; Reches, M.; Narasimhan, S.; Schoulaker-Schwarz, R.; Klemes, Y.; Aizenman, E.; Glaser, G. Rexb of bacteriophage lambda is an anti-cell death gene. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 15481–15486.
24. Sun, X.M.; Gohler, A.; Heller, K.J.; Neve, H. The ltp gene of temperate streptococcus thermophilus phage tp-j34 confers superinfection exclusion to streptococcus thermophilus and lactococcus lactis. *Virology* **2006**, *350*, 146–157.
25. Mosig, G.; Yu, S.; Myung, H.; Haggard-Ljungquist, E.; Davenport, L.; Carlson, K.; Calendar, R. A novel mechanism of virus-virus interactions: Bacteriophage p2 tin protein inhibits phage t4 DNA synthesis by poisoning the t4 single-stranded DNA binding protein, gp32. *Virology* **1997**, *230*, 72–81.
26. Hartskeerl, R.; Zuidweg, E.; van Geffen, M.; Hoekstra, W. The inci plasmids r144, r64 and colib belong to one exclusion group. *J. Gen. Microbiol.* **1985**, *131*, 1305–1311.

27. Briani, F.; Deho, G.; Forti, F.; Ghisotti, D. The plasmid status of satellite bacteriophage p4. *Plasmid* **2001**, *45*, 1–17.
28. Oshima, K.; Kakizawa, S.; Nishigawa, H.; Kuboyama, T.; Miyata, S.; Ugaki, M.; Namba, S. A plasmid of phytoplasma encodes a unique replication protein having both plasmid- and virus-like domains: Clue to viral ancestry or result of virus/plasmid recombination? *Virology* **2001**, *285*, 270–277.
29. Boyd, D.; Peters, G.A.; Cloeckert, A.; Boumedine, K.S.; Chaslus-Dancla, E.; Imberechts, H.; Mulvey, M.R. Complete nucleotide sequence of a 43-kilobase genomic island associated with the multidrug resistance region of salmonella enterica serovar typhimurium dt104 and its identification in phage type dt120 and serovar agona. *J. Bacteriol.* **2001**, *183*, 5725–5732.
30. Boltner, D.; MacMahon, C.; Pembroke, J.T.; Strike, P.; Osborn, A.M. R391: A conjugative integrating mosaic comprised of phage, plasmid, and transposon elements. *J. Bacteriol.* **2002**, *184*, 5158–5169.
31. Jalasvuori, M.; Jaatinen, S.T.; Laurinavicius, S.; Ahola-Iivarinen, E.; Kalkkinen, N.; Bamford, D.H.; Bamford, J.K.H. The closest relatives of icosahedral viruses of thermophilic bacteria are among viruses and plasmids of the halophilic archaea. *J. Virol.* **2009**, *83*, 9388–9397.
32. Lobočka, M.B.; Rose, D.J.; Plunkett, G., 3rd; Rusin, M.; Samojedny, A.; Lehnherr, H.; Yarmolinsky, M.B.; Blattner, F.R. Genome of bacteriophage p1. *J. Bacteriol.* **2004**, *186*, 7032–7068.
33. Yarmolinsky, M.B. Programmed cell death in bacterial populations. *Science* **1995**, *267*, 836–837.
34. Klassen, R.; Tontsidou, L.; Larsen, M.; Meinhardt, F. Genome organization of the linear cytoplasmic element ppe1b from pichia etchellsii. *Yeast* **2001**, *18*, 953–961.
35. Binnewies, T.T.; Motro, Y.; Hallin, P.F.; Lund, O.; Dunn, D.; La, T.; Hampson, D.J.; Bellgard, M.; Wassenaar, T.M.; Ussery, D.W. Ten years of bacterial genome sequencing: Comparative-genomics-based discoveries. *Funct. Integr. Genom.* **2006**, *6*, 165–185.
36. Koonin, E.V.; Wolf, Y.I. Genomics of bacteria and archaea: The emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* **2008**, *36*, 6688–6719.
37. Osborn, A.M.; Boltner, D. When phage, plasmids, and transposons collide: Genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid* **2002**, *48*, 202–212.
38. Lehnherr, H.; Maguin, E.; Jafri, S.; Yarmolinsky, M.B. Plasmid addiction genes of bacteriophage p1: Doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J. Mol. Biol.* **1993**, *233*, 414–428.
39. Basta, T.; Smyth, J.; Forterre, P.; Prangishvili, D.; Peng, X. Novel archaeal plasmid pah1 and its interactions with the lipothrixvirus afv1. *Mol. Microbiol.* **2009**, *71*, 23–34.
40. Ichige, A.; Kobayashi, I. Stability of ecori restriction-modification enzymes *in vivo* differentiates the ecori restriction-modification system from other postsegregational cell killing systems. *J. Bacteriol.* **2005**, *187*, 6612–6621.
41. Pagie, L.; Hogeweg, P. Individual- and population-based diversity in restriction-modification systems. *Bull. Math. Biol.* **2000**, *62*, 759–774.
42. Gelfand, M.S.; Koonin, E.V. Avoidance of palindromic words in bacterial and archaeal genomes: A close connection with restriction enzymes. *Nucleic Acids Res.* **1997**, *25*, 2430–2439.



43. Zhao, F.Q.; Zhang, X.W.; Liang, C.W.; Wu, J.Y.; Bao, Q.Y.; Qin, S. Genome-wide analysis of restriction-modification system in unicellular and filamentous cyanobacteria. *Physiol. Genom.* **2006**, *24*, 181–190.
44. Dziewit, L.; Jazurek, M.; Drewniak, L.; Baj, J.; Bartosik, D. The sxt conjugative element and linear prophage n15 encode toxin-antitoxin-stabilizing systems homologous to the tad-ata module of the paracoccus aminophilus plasmid pami2. *J. Bacteriol.* **2007**, *189*, 1983–1997.
45. Naderer, M.; Brust, J.R.; Knowle, D.; Blumenthal, R.M. Mobility of a restriction-modification system revealed by its genetic contexts in three hosts. *J. Bacteriol.* **2002**, *184*, 2411–2419.
46. Tyndall, C.; Lehnerr, H.; Sandmeier, U.; Kulik, E.; Bickle, T.A. The type ic hsd loci of the enterobacteria are flanked by DNA with high homology to the phage p1 genome: Implications for the evolution and spread of DNA restriction systems. *Mol. Microbiol.* **1997**, *23*, 729–736.
47. Mrazek, J.; Karlin, S. Detecting alien genes in bacterial genomes. *Ann. N. Y. Acad. Sci.* **1999**, *870*, 314–329.
48. Zaleski, P.; Wojciechowski, M.; Piekarowicz, A. The role of dam methylation in phase variation of haemophilus influenzae genes involved in defence against phage infection. *Microbiology* **2005**, *151*, 3361–3369.
49. Pullinger, G.D.; Bevir, T.; Lax, A.J. The pasteurella multocida toxin is encoded within a lysogenic bacteriophage. *Mol. Microbiol.* **2004**, *51*, 255–269.
50. Persson, O.P.; Pinhassi, J.; Riemann, L.; Marklund, B.I.; Rhen, M.; Normark, S.; Gonzalez, J.M.; Hagstrom, A. High abundance of virulence gene homologues in marine bacteria. *Environ. Microbiol.* **2009**, *11*, 1348–1357.
51. Kunin, V.; He, S.; Warnecke, F.; Peterson, S.B.; Garcia Martin, H.; Haynes, M.; Ivanova, N.; Blackall, L.L.; Breitbart, M.; Rohwer, F.; *et al.* A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res.* **2008**, *18*, 293–297.
52. Williamson, S.J.; Cary, S.C.; Williamson, K.E.; Helton, R.R.; Bench, S.R.; Winget, D.; Wommack, K.E. Lysogenic virus-host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *ISME J.* **2008**, *2*, 1112–1121.
53. Rodriguez-Valera, F.; Martin-Cuadrado, A.B.; Rodriguez-Brito, B.; Pasic, L.; Thingstad, T.F.; Rohwer, F.; Mira, A. Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.* **2009**, *7*, 828–836.
54. Cheetham, B.F.; Katz, M.E. A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. *Mol. Microbiol.* **1995**, *18*, 201–208.
55. Ogura, Y.; Ooka, T.; Asadulghani; Terajima, J.; Nougayrede, J.P.; Kurokawa, K.; Tashiro, K.; Tobe, T.; Nakayama, K.; Kuhara, S.; *et al.* Extensive genomic diversity and selective conservation of virulence-determinants in enterohemorrhagic escherichia coli strains of o157 and non-o157 serotypes. *Genome Biol.* **2007**, *8*, R138.
56. Ooka, T.; Ogura, Y.; Asadulghani, M.; Ohnishi, M.; Nakayama, K.; Terajima, J.; Watanabe, H.; Hayashi, T. Inference of the impact of insertion sequence (is) elements on bacterial genome diversification through analysis of small-size structural polymorphisms in escherichia coli o157 genomes. *Genome Res.* **2009**, *19*, 1809–1816.
57. Zhou, Z.; Li, X.; Liu, B.; Beutin, L.; Xu, J.; Ren, Y.; Feng, L.; Lan, R.; Reeves, P.R.; Wang, L. Derivation of escherichia coli o157:H7 from its o55:H7 precursor. *PLoS One* **2010**, *5*, e8700.

58. Ogura, Y.; Ooka, T.; Iguchi, A.; Toh, H.; Asadulghani, M.; Oshima, K.; Kodama, T.; Abe, H.; Nakayama, K.; Kurokawa, K.; *et al.* Comparative genomics reveal the mechanism of the parallel evolution of o157 and non-o157 enterohemorrhagic escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 17939–17944.
59. Iguchi, A.; Thomson, N.R.; Ogura, Y.; Saunders, D.; Ooka, T.; Henderson, I.R.; Harris, D.; Asadulghani, M.; Kurokawa, K.; Dean, P.; *et al.* Complete genome sequence and comparative genome analysis of enteropathogenic escherichia coli o127:H6 strain e2348/69. *J. Bacteriol.* **2009**, *191*, 347–354.
60. Asadulghani, M.; Ogura, Y.; Ooka, T.; Itoh, T.; Sawaguchi, A.; Iguchi, A.; Nakayama, K.; Hayashi, T. The defective prophage pool of escherichia coli o157: Prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog.* **2009**, *5*, e1000408.
61. Abe, H.; Miyahara, A.; Oshima, T.; Tashiro, K.; Ogura, Y.; Kuhara, S.; Ogasawara, N.; Hayashi, T.; Tobe, T. Global regulation by horizontally transferred regulators establishes the pathogenicity of escherichia coli. *DNA Res.* **2008**, *15*, 25–38.
62. Oshima, K.; Toh, H.; Ogura, Y.; Sasamoto, H.; Morita, H.; Park, S.H.; Ooka, T.; Iyoda, S.; Taylor, T.D.; Hayashi, T.; *et al.* Complete genome sequence and comparative analysis of the wild-type commensal escherichia coli strain se11 isolated from a healthy adult. *DNA Res.* **2008**, *15*, 375–386.
63. Brussow, H. Bacteria between protists and phages: From antipredation strategies to the evolution of pathogenicity. *Mol. Microbiol.* **2007**, *65*, 583–589.
64. Kim, Y.; Wang, X.; Ma, Q.; Zhang, X.S.; Wood, T.K. Toxin-antitoxin systems in escherichia coli influence biofilm formation through yjgk (taba) and fimbriae. *J. Bacteriol.* **2009**, *191*, 1258–1267.
65. Harrison, J.J.; Wade, W.D.; Akierman, S.; Vacchi-Suzzi, C.; Stremick, C.A.; Turner, R.J.; Ceri, H. The chromosomal toxin gene yafq is a determinant of multidrug tolerance for escherichia coli growing in a biofilm. *Antimicrob. Agents Chemother.* **2009**, *53*, 2253–2258.
66. Blower, T.R.; Fineran, P.C.; Johnson, M.J.; Toth, I.K.; Humphreys, D.P.; Salmond, G.P. Mutagenesis and functional characterization of the rna and protein components of the toxin abortive infection and toxin-antitoxin locus of erwinia. *J. Bacteriol.* **2009**, *191*, 6029–6039.
67. Fineran, P.C.; Blower, T.R.; Foulds, I.J.; Humphreys, D.P.; Lilley, K.S.; Salmond, G.P. The phage abortive infection system, toxin, functions as a protein-rna toxin-antitoxin pair. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 894–899.
68. Chen, J.; Novick, R.P. Phage-mediated intergeneric transfer of toxin genes. *Science* **2009**, *323*, 139–141.
69. Spoering, A.L.; Vulic, M.; Lewis, K. Gldp and plsB participate in persister cell formation in escherichia coli. *J. Bacteriol.* **2006**, *188*, 5136–5144.
70. Horvath, P.; Barrangou, R. Crispr/cas, the immune system of bacteria and archaea. *Science* **2010**, *327*, 167–170.
71. Marraffini, L.A.; Sontheimer, E.J. Self versus non-self discrimination during crispr rna-directed immunity. *Nature* **2010**, *463*, 568–571.

72. Jore, M.M.; Lundgren, M.; van Duijn, E.; Bultema, J.B.; Westra, E.R.; Waghmare, S.P.; Wiedenheft, B.; Pul, U.; Wurm, R.; Wagner, R.; *et al.* Structural basis for crispr rna-guided DNA recognition by cascade. *Nat. Struct. Mol. Biol.* **2011**, *18*, 529–536.
73. Marraffini, L.A.; Sontheimer, E.J. Crispr interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **2008**, *322*, 1843–1845.
74. Horvath, P.; Romero, D.A.; Coute-Monvoisin, A.C.; Richards, M.; Deveau, H.; Moineau, S.; Boyaval, P.; Fremaux, C.; Barrangou, R. Diversity, activity, and evolution of crispr loci in streptococcus thermophilus. *J. Bacteriol.* **2008**, *190*, 1401–1412.
75. Snyder, J.C.; Bateson, M.M.; Lavin, M.; Young, M.J. Use of cellular crispr (clusters of regularly interspaced short palindromic repeats) spacer-based microarrays for detection of viruses in environmental samples. *Appl. Environ. Microbiol.* **2010**, *76*, 7251–7258.
76. Bize, A.; Forterre, P.; Prangishvili, D. Archeovirus. *Virologie* **2010**, *14*, 101–117.
77. Wang, Y.; Duan, Z.H.; Zhu, H.J.; Guo, X.; Wang, Z.Y.; Zhou, J.; She, Q.X.; Huang, L. A novel sulfolobus non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus. *Virology* **2007**, *363*, 124–133.
78. Zegans, M.E.; Wagner, J.C.; Cady, K.C.; Murphy, D.M.; Hammond, J.H.; O'Toole, G.A. Interaction between bacteriophage dms3 and host crispr region inhibits group behaviors of pseudomonas aeruginosa. *J. Bacteriol.* **2009**, *191*, 210–219.
79. Shah, S.A.; Hansen, N.R.; Garrett, R.A. Distribution of crispr spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem. Soc. Trans.* **2009**, *37*, 23–28.
80. Touchon, M.; Rocha, E.P.C. The small, slow and specialized crispr and anti-crispr of escherichia and salmonella. *PLoS One* **2010**, *5*, e11126.
81. Gill, E.E.; Brinkman, F.S.L. The proportional lack of archaeal pathogens: Do viruses/phages hold the key? *Bioessays* **2011**, *33*, 248–254.
82. Agarkova, I.V.; Dunigan, D.D.; Van Etten, J.L. Virion-associated restriction endonucleases of chloroviruses. *J. Virol.* **2006**, *80*, 8114–8123.
83. Volff, J.N.; Korting, C.; Froschauer, A.; Sweeney, K.; Scharl, M. Non-ltr retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. *J. Mol. Evol.* **2001**, *52*, 351–360.
84. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature* **1998**, *391*, 806–811.
85. Tomari, Y.; Zamore, P.D. Perspective: Machines for RNAi. *Gene. Dev.* **2005**, *19*, 517–529.
86. Anglicheau, D.; Muthukumar, T.; Suthanthiran, M. Micrnas: Small rnas with big effects. *Transplantation* **2010**, *90*, 105–112.
87. Lu, R.; Yigit, E.; Li, W.X.; Ding, S.W. An rig-i-like rna helicase mediates antiviral rnai downstream of viral sirna biogenesis in caenorhabditis elegans. *PLoS Pathog.* **2009**, *5*, e1000286.
88. Sijen, T.; Fleenor, J.; Simmer, F.; Thijssen, K.L.; Parrish, S.; Timmons, L.; Plasterk, R.H.; Fire, A. On the role of rna amplification in dsrna-triggered gene silencing. *Cell* **2001**, *107*, 465–476.
89. Iyer, L.M.; Koonin, E.V.; Aravind, L. Evolutionary connection between the catalytic subunits of DNA-dependent rna polymerases and eukaryotic rna-dependent rna polymerases and the origin of rna polymerases. *BMC Struct. Biol.* **2003**, *3*, 1.

90. Jin, L.; Kryukov, K.; Suzuki, Y.; Imanishi, T.; Ikeo, K.; Gojobori, T. The evolutionary study of small rna-directed gene silencing pathways by investigating rnase iii enzymes. *Gene* **2009**, *435*, 1–8.
91. Barstead, R. Genome-wide rna. *Curr. Opin. Chem. Biol.* **2001**, *5*, 63–66.
92. van Roessel, P.; Brand, A.H. Spreading silence with sid. *Genome Biol.* **2004**, *5*, 208.
93. May, R.C.; Plasterk, R.H. Rna interference spreading in c. *Elegans*. *Meth. Enzymol.* **2005**, *392*, 308–315.
94. Felix, M.A.; Ashe, A.; Piffaretti, J.; Wu, G.; Nuez, I.; Belicard, T.; Jiang, Y.; Zhao, G.; Franz, C.J.; Goldstein, L.D.; *et al.* Natural and experimental infection of caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol.* **2011**, *9*, e1000586.
95. Bekal, S.; Domier, L.; Niblack, T.; Lambert, K. Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *J. Gen. Virol.* **2011**, *92*, 1870–1879.
96. Bagasra, O.; Prilliman, K.R. Rna interference: The molecular immune system. *J. Mol. Histol.* **2004**, *35*, 545–553.
97. Lu, R.; Maduro, M.; Li, F.; Li, H.W.; Broitman-Maduro, G.; Li, W.X.; Ding, S.W. Animal virus replication and rna-mediated antiviral silencing in caenorhabditis elegans. *Nature* **2005**, *436*, 1040–1043.
98. Schott, D.H.; Cureton, D.K.; Whelan, S.P.; Hunter, C.P. An antiviral role for the rna interference machinery in caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 18420–18424.
99. Wilkins, C.; Dishongh, R.; Moore, S.C.; Whitt, M.A.; Chow, M.; Machaca, K. Rna interference is an antiviral defence mechanism in caenorhabditis elegans. *Nature* **2005**, *436*, 1044–1047.
100. Fritz, J.H.; Girardin, S.E.; Philpott, D.J. Innate immune defense through RNA interference. *Sci. STKE* **2006**, *2006*, pe27.
101. Bowen, N.J.; McDonald, J.F. Genomic analysis of caenorhabditis elegans reveals ancient families of retroviral-like elements. *Genome Res.* **1999**, *9*, 924–935.
102. Vastenhouw, N.L.; Plasterk, R.H. Rnai protects the caenorhabditis elegans germline against transposition. *Trends Genet* **2004**, *20*, 314–319.
103. Robert, V.J.; Sijen, T.; van Wolfswinkel, J.; Plasterk, R.H. Chromatin and rnai factors protect the c. *Elegans* germline against repetitive sequences. *Gene. Dev.* **2005**, *19*, 782–787.
104. Villarreal, L.P. *Origin of Group Identity: Viruses, Addiction, and Cooperation*; Springer: New York, NY, USA, 2009; pp. xxiv, 614, 619, plates.
105. Voinnet, O. Induction and suppression of rna silencing: Insights from viral infections. *Nat. Rev.* **2005**, *6*, 206–220.
106. Malone, C.D.; Hannon, G.J. Small rnas as guardians of the genome. *Cell* **2009**, *136*, 656–668.
107. Wu, Q.; Luo, Y.; Lu, R.; Lau, N.; Lai, E.C.; Li, W.X.; Ding, S.W. Virus discovery by deep sequencing and assembly of virus-derived small silencing rnas. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1606–1611.
108. Pelisson, A.; Sarot, E.; Payen-Groschene, G.; Bucheton, A. A novel repeat-associated small interfering rna-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the drosophila ovary. *J. Virol.* **2007**, *81*, 1951–1960.

109. Czech, B.; Malone, C.D.; Zhou, R.; Stark, A.; Schlingeheyde, C.; Dus, M.; Perrimon, N.; Kellis, M.; Wohlschlegel, J.A.; Sachidanandam, R.; *et al.* An endogenous small interfering rna pathway in drosophila. *Nature* **2008**, *453*, U798–U797.
110. Aravind, L.; Dixit, V.M.; Koonin, E.V. Apoptotic molecular machinery: Vastly increased complexity in vertebrates revealed by genome comparisons. *Science* **2001**, *291*, 1279–1284.
111. Nakayashiki, H.; Kadotani, N.; Mayama, S. Evolution and diversification of rna silencing proteins in fungi. *J. Mol. Evol.* **2006**, *63*, 127–135.
112. de Jong, D.; Eitel, M.; Jakob, W.; Osigus, H.J.; Hadrys, H.; Desalle, R.; Schierwater, B. Multiple dicer genes in the early-diverging metazoa. *Mol. Biol. Evol.* **2009**, *26*, 1333–1340.
113. Ghabrial, S.A. New developments in fungal virology. *Adv. Virus Res.* **1994**, *43*, 303–388.
114. Hillman, B.I.; Supyani, S.; Kondo, H.; Suzuki, N. A reovirus of the fungus *cryphonectria parasitica* that is infectious as particles and related to the coltivirus genus of animal pathogens. *J. Virol.* **2004**, *78*, 892–898.
115. Pearson, M.N.; Beever, R.E.; Boine, B.; Arthur, K. Mycoviruses of filamentous fungi and their relevance to plant pathology. *Mol. Plant Pathol.* **2009**, *10*, 115–128.
116. Ghabrial, S.A. Origin, adaptation and evolutionary pathways of fungal viruses. *Virus Genes* **1998**, *16*, 119–131.
117. Biella, S.; Smith, M.L.; Aist, J.R.; Cortesi, P.; Milgroom, M.G. Programmed cell death correlates with virus transmission in a filamentous fungus. *Proc. Roy. Soc. Lond. B Biol. Sci.* **2002**, *269*, 2269–2276.
118. Allen, T.D.; Nuss, D.L. Linkage between mitochondrial hypovirulence and viral hypovirulence in the chestnut blight fungus revealed by cDNA microarray analysis. *Eukaryotic. Cell* **2004**, *3*, 1227–1232.
119. Cortesi, P.; McCulloch, C.E.; Song, H.Y.; Lin, H.Q.; Milgroom, M.G. Genetic control of horizontal virus transmission in the chestnut blight fungus, *cryphonectria parasitica*. *Genetics* **2001**, *159*, 107–118.
120. Liu, Y.C.; Linder-Basso, D.; Hillman, B.I.; Kaneko, S.; Milgroom, M.G. Evidence for interspecies transmission of viruses in natural populations of filamentous fungi in the genus *cryphonectria*. *Mol. Ecol.* **2003**, *12*, 1619–1628.
121. Liu, H.; Fu, Y.; Jiang, D.; Li, G.; Xie, J.; Cheng, J.; Peng, Y.; Ghabrial, S.A.; Yi, X. Widespread horizontal gene transfer from double-stranded rna viruses to eukaryotic nuclear genomes. *J. Virol.* **2010**, *84*, 11876–11887.
122. Hammond, T.M.; Bok, J.W.; Andrews, M.D.; Reyes-Dominguez, Y.; Scazzocchio, C.; Keller, N.P. Rna silencing gene truncation in the filamentous fungus *aspergillus nidulans*. *Eukaryot. Cell* **2008**, *7*, 339–349.
123. Hammond, T.M.; Keller, N.P. Rna silencing in *aspergillus nidulans* is independent of RNA-dependent RNA polymerases. *Genetics* **2005**, *169*, 607–617.
124. Cogoni, C. Homology-dependent gene silencing mechanisms in fungi. *Annu. Rev. Microbiol.* **2001**, *55*, 381–406.
125. Daboussi, M.J.; Capy, P. Transposable elements in filamentous fungi. *Annu. Rev. Microbiol.* **2003**, *57*, 275–299.

126. Ebbole, D.J. Magnaporthe as a model for understanding host-pathogen interactions. *Annu. Rev. Phytopathology* **2007**, *45*, 437–456.
127. Gorinsek, B.; Gubensek, F.; Kordis, D. Phylogenomic analysis of chromoviruses. *Cytogenet. Genome Res.* **2005**, *110*, 543–552.
128. Delsuc, F.; Brinkmann, H.; Chourrout, D.; Philippe, H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **2006**, *439*, 965–968.
129. Klein, J. The grapes of incompatibility. *Dev. Cell* **2006**, *10*, 2–4.
130. Oren, M.; Douek, J.; Fishelson, Z.; Rinkevich, B. Identification of immune-relevant genes in histoincompatible rejecting colonies of the tunicate botryllus schlosseri. *Dev. Comp. Immunol.* **2007**, *31*, 889–902.
131. Pancer, Z.; Saha, N.R.; Kasamatsu, J.; Suzuki, T.; Amemiya, C.T.; Kasahara, M.; Cooper, M.D. Variable lymphocyte receptors in hagfish. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9224–9229.
132. Rogozin, I.B.; Iyer, L.M.; Liang, L.; Glazko, G.V.; Liston, V.G.; Pavlov, Y.I.; Aravind, L.; Pancer, Z. Evolution and diversification of lamprey antigen receptors: Evidence for involvement of an aid-apobec family cytosine deaminase. *Nat. Immunol.* **2007**, *8*, 647–656.
133. Nonaka, M.; Yoshizaki, F. Primitive complement system of invertebrates. *Immunol. Rev.* **2004**, *198*, 203–215.
134. Marino, R.; Kimura, Y.; De Santis, R.; Lambris, J.D.; Pinto, M.R. Complement in urochordates: Cloning and characterization of two c3-like genes in the ascidian ciona intestinalis. *Immunogenetics* **2002**, *53*, 1055–1064.
135. Mansour, M.H.; Cooper, E.L. Tunicate thy-1—An invertebrate member of the ig superfamily. *Progr. Clin. Biol. Res.* **1987**, *233*, 33–42.
136. Herrin, B.R.; Cooper, M.D. Alternative adaptive immunity in jawless vertebrates. *J. Immunol.* **2010**, *185*, 1367–1374.
137. Hsu, E. The invention of lymphocytes. *Curr. Opin. Immunol.* **2010**, *23*, 156–162.
138. Britten, R.J.; McCormack, T.J.; Mears, T.L.; Davidson, E.H. Gypsy/ty3-class retrotransposons integrated in the DNA of herring, tunicate, and echinoderms. *J. Mol. Evol.* **1995**, *40*, 13–24.
139. Volff, J.N.; Lehrach, H.; Reinhardt, R.; Chourrout, D. Retroelement dynamics and a novel type of chordate retrovirus-like element in the miniature genome of the tunicate oikopleura dioica. *Mol. Biol. Evol.* **2004**, *21*, 2022–2033.
140. Dehal, P.; Satou, Y.; Campbell, R.K.; Chapman, J.; Degnan, B.; De Tomaso, A.; Davidson, B.; Di Gregorio, A.; Gelpke, M.; Goodstein, D.M.; *et al.* The draft genome of ciona intestinalis: Insights into chordate and vertebrate origins. *Science* **2002**, *298*, 2157–2167.
141. Gadd, T.; Jakava-Viljanen, M.; Tapiovaara, H.; Koski, P.; Sihvonen, L. Epidemiological aspects of viral haemorrhagic septicaemia virus genotype ii isolated from baltic herring, clupea harengus membras l. *J. Fish Dis.* **2011**, *34*, 517–529.
142. Liongue, C.; John, L.B.; Ward, A. Origins of adaptive immunity. *Crit. Rev. Immunol.* **2011**, *31*, 61–71.
143. Litman, G.W.; Rast, J.P.; Fugmann, S.D. The origins of vertebrate adaptive immunity. *Nat. Rev. Immunol.* **2010**, *10*, 543–553.
144. Witzany, G. Natural genome editing from a biocommunicative perspective. *Biosemitotics* **2011**, doi:10.1007/s12304-011-9111-7.

145. Shapiro, J.A. *Evolution : A View from the 21st Century*; FT Press Science: Upper Saddle River, NJ, USA, 2011; pp. xi, 253.
146. Witzany, G. The agents of natural genome editing. *J. Mol. Cell Biol.* **2011**, *3*, 181–189.
147. Cooper, M.D.; Alder, M.N. The evolution of adaptive immune systems. *Cell* **2006**, *124*, 815–822.
148. Villarreal, L.P. The source of self: Genetic parasites and the origin of adaptive immunity. *Ann. N. Y. Acad. Sci.* **2009**, *1178*, 194–232.
149. Jaillon, O.; Aury, J.M.; Brunet, F.; Petit, J.L.; Stange-Thomann, N.; Mauceli, E.; Bouneau, L.; Fischer, C.; Ozouf-Costaz, C.; Bernot, A.; *et al.* Genome duplication in the teleost fish tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **2004**, *431*, 946–957.
150. Poulter, R.; Butler, M. A retrotransposon family from the pufferfish (fugu) *fugu rubripes*. *Gene* **1998**, *215*, 241–249.
151. Gorinsek, B.; Gubensek, F.; Kordis, D. Evolutionary genomics of chromoviruses in eukaryotes. *Mol. Biol. Evol.* **2004**, *21*, 781–798.
152. Mulero, V.; Lopez-Munoz, A.; Roca, F.J.; Meseguer, J. New insights into the evolution of ifns: Zebrafish group ii ifns induce a rapid and transient expression of ifn-dependent genes and display powerful antiviral activities. *J. Immunol.* **2009**, *182*, 3440–3449.
153. Workenhe, S.T.; Rise, M.L.; Kibenge, M.J.; Kibenge, F.S. The fight between the teleost fish immune response and aquatic viruses. *Mol. Immunol.* **2010**, *47*, 2525–2536.
154. Muroga, K. Viral and bacterial diseases of marine fish and shellfish in japanese hatcheries. *Aquaculture* **2001**, *202*, 23–44.
155. Magnadottir, B. Immunological control of fish diseases. *Mar. Biotechnol.* **2010**, *12*, 361–379.
156. Matsunaga, T.; Rahman, A. What brought the adaptive immune system to vertebrates?—The jaw hypothesis and the seahorse. *Immunol. Rev.* **1998**, *166*, 177–186.
157. Du Pasquier, L. Speculations on the origin of the vertebrate immune system. *Immunol. Lett.* **2004**, *92*, 3–9.
158. Makino, T.; McLysaght, A. Interacting gene clusters and the evolution of the vertebrate immune system. *Mol. Biol. Evol.* **2008**, *25*, 1855–1862.
159. Du Pasquier, L. Innate immunity in early chordates and the appearance of adaptive immunity. *Comptes Rendus Biologies* **2004**, *327*, 591–601.
160. Dermody, T.S.; Kirchner, E.; Guglielmi, K.M.; Stehle, T. Immunoglobulin superfamily virus receptors and the evolution of adaptive immunity. *PLoS Pathog.* **2009**, *5*, e1000481.
161. Carson, S.D. Cocksackievirus and adenovirus receptor (car) is modified and shed in membrane vesicles. *Biochemistry* **2004**, *43*, 8136–8142.
162. Adelman, M.K.; Schluter, S.F.; Marchalonis, J.J. The natural antibody repertoire of sharks and humans recognizes the potential universe of antigens. *Protein J.* **2004**, *23*, 103–118.
163. Camargo, M.M.; Nahum, L.A. Adapting to a changing world: Rag genomics and evolution. *Hum. Genom.* **2005**, *2*, 132–137.
164. Agrawal, A.; Eastman, Q.M.; Schatz, D.G. Transposition mediated by rag1 and rag2 and its implications for the evolution of the immune system. *Nature* **1998**, *394*, 744–751.
165. Hiom, K.; Melek, M.; Gellert, M. DNA transposition by the rag1 and rag2 proteins: A possible source of oncogenic translocations. *Cell* **1998**, *94*, 463–470.

166. Fugmann, S.D. The origins of the rag genes—from transposition to v(d)j recombination. *Semin. Immunol.* **2010**, *22*, 10–16.
167. Hashimoto, K.; Okamura, K.; Yamaguchi, H.; Ototake, M.; Nakanishi, T.; Kurosawa, Y. Conservation and diversification of mhc class i and its related molecules in vertebrates. *Immunol. Rev.* **1999**, *167*, 81–100.
168. Xie, T.; Rowen, L.; Aguado, B.; Ahearn, M.E.; Madan, A.; Qin, S.; Campbell, R.D.; Hood, L. Analysis of the gene-dense major histocompatibility complex class iii region and its comparison to mouse. *Genome Res.* **2003**, *13*, 2621–2636.
169. Deakin, J.E.; Papenfuss, A.T.; Belov, K.; Cross, J.G.; Coggill, P.; Palmer, S.; Sims, S.; Speed, T.P.; Beck, S.; Graves, J.A. Evolution and comparative analysis of the mhc class iii inflammatory region. *BMC Genom.* **2006**, *7*, 281.
170. Andersson, G.; Svensson, A.C.; Setterblad, N.; Rask, L. Retroelements in the human mhc class ii region. *Trends Genet.* **1998**, *14*, 109–114.
171. Kulski, J.K.; Gaudieri, S.; Inoko, H.; Dawkins, R.L. Comparison between two human endogenous retrovirus (herv)-rich regions within the major histocompatibility complex. *J. Mol. Evol.* **1999**, *48*, 675–683.
172. Matsuo, M.Y.; Nonaka, M. Repetitive elements in the major histocompatibility complex (mhc) class i region of a teleost, medaka: Identification of novel transposable elements. *Mech. Dev.* **2004**, *121*, 771–777.
173. Gaudieri, S.; Kulski, J.K.; Balmer, L.; Giles, K.M.; Inoko, H.; Dawkins, R.L. Retroelements and segmental duplications in the generation of diversity within the mhc. *DNA Seq.* **1997**, *8*, 137–141.
174. Kulski, J.K.; Shiina, T.; Anzai, T.; Kohara, S.; Inoko, H. Comparative genomic analysis of the mhc: The evolution of class i duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* **2002**, *190*, 95–122.
175. Kulski, J.K.; Anzai, T.; Inoko, H. Ervk9, transposons and the evolution of mhc class i duplicons within the alpha-block of the human and chimpanzee. *Cytogenet Genome Res.* **2005**, *110*, 181–192.
176. Kulski, J.K.; Gaudieri, S.; Martin, A.; Dawkins, R.L. Coevolution of perb11 (mic) and hla class i genes with herv-16 and retroelements by extended genomic duplication. *J. Mol. Evol.* **1999**, *49*, 84–97.
177. Kulski, J.K.; Gaudieri, S.; Dawkins, R.L. Using alu j elements as molecular clocks to trace the evolutionary relationships between duplicated hla class i genomic segments. *J. Mol. Evol.* **2000**, *50*, 510–519.
178. Kulski, J.K.; Anzai, T.; Shiina, T.; Inoko, H. Rhesus macaque class i duplicon structures, organization, and evolution within the alpha block of the major histocompatibility complex. *Mol. Biol. Evol.* **2004**, *21*, 2079–2091.
179. Doxiadis, G.G.; de Groot, N.; Bontrop, R.E. Impact of endogenous intronic retroviruses on major histocompatibility complex class II diversity and stability. *J. Virol.* **2008**, *82*, 6667–6677.
180. Kulski, J.K.; Dawkins, R.L. The p5 multicopy gene family in the mhc is related in sequence to human endogenous retroviruses HERV-L and HERV-16. *Immunogenetics* **1999**, *49*, 404–412.
181. Magor, B.G.; Magor, K.E. Evolution of effectors and receptors of innate immunity. *Dev. Comp. Immunol.* **2001**, *25*, 651–682.



182. Nehyba, J.; Hrdlickova, R.; Bose, H.R. Dynamic evolution of immune system regulators: The history of the interferon regulatory factor family. *Mol. Biol. Evol.* **2009**, *26*, 2539–2550.
183. Fox, B.A.; Sheppard, P.O.; O'Hara, P.J. The role of genomic data in the discovery, annotation and evolutionary interpretation of the interferon-lambda family. *PLoS One* **2009**, *4*, e4933.
184. Robalino, J.; Bartlett, T.C.; Chapman, R.W.; Gross, P.S.; Browdy, C.L.; Warr, G.W. Double-stranded rna and antiviral immunity in marine shrimp: Inducible host mechanisms and evidence for the evolution of viral counter-responses. *Dev. Comp. Immunol.* **2007**, *31*, 539–547.
185. Tanaka-Taya, K.; Sashihara, J.; Kurahashi, H.; Amo, K.; Miyagawa, H.; Kondo, K.; Okada, S.; Yamanishi, K. Human herpesvirus 6 (hhv-6) is transmitted from parent to child in an integrated form and characterization of cases with chromosomally integrated hhv-6 DNA. *J. Med. Virol.* **2004**, *73*, 465–473.
186. Takemoto, M.; Imasawa, T.; Yamanishi, K.; Mori, Y. Role of dendritic cells infected with human herpesvirus 6 in virus transmission to cd4(+) t cells. *Virology* **2009**, *385*, 294–302.
187. Hansen, S.G.; Powers, C.J.; Richards, R.; Ventura, A.B.; Ford, J.C.; Siess, D.; Axthelm, M.K.; Nelson, J.A.; Jarvis, M.A.; Picker, L.J.; *et al.* Evasion of cd8+ t cells is critical for superinfection by cytomegalovirus. *Science* **2010**, *328*, 102–106.
188. Klenerman, P.; Dunbar, P.R. Cmv and the art of memory maintenance. *Immunity* **2008**, *29*, 520–522.
189. Lestou, V.S.; De Braekeleer, M.; Strehl, S.; Ott, G.; Gadner, H.; Ambros, P.F. Non-random integration of epstein-barr virus in lymphoblastoid cell lines. *Gene. Chromosome. Canc.* **1993**, *8*, 38–48.
190. Hengel, H.; Koszinowski, U.H.; Conzelmann, K.K. Viruses know it all: New insights into ifn networks. *Trends Immunol.* **2005**, *26*, 396–401.
191. Vossen, M.T.; Westerhout, E.M.; Soderberg-Naucler, C.; Wiertz, E.J. Viral immune evasion: A masterpiece of evolution. *Immunogenetics* **2002**, *54*, 527–542.
192. McGeoch, D.J.; Rixon, F.J.; Davison, A.J. Topics in herpesvirus genomics and evolution. *Virus Res.* **2006**, *117*, 90–104.
193. van Cleef, K.W.; Smit, M.J.; Bruggeman, C.A.; Vink, C. Cytomegalovirus-encoded homologs of g protein-coupled receptors and chemokines. *J. Clin. Virol.* **2006**, *35*, 343–348.
194. Davis-Poynter, N.J.; Degli-Esposti, M.; Farrell, H.E. Murine cytomegalovirus homologues of cellular immunomodulatory genes. *Intervirolology* **1999**, *42*, 331–341.
195. Miller-Kittrell, M.; Sparer, T.E. Feeling manipulated: Cytomegalovirus immune manipulation. *Virol. J.* **2009**, *6*, 4.
196. Pinto, A.K.; Munks, M.W.; Koszinowski, U.H.; Hill, A.B. Coordinated function of murine cytomegalovirus genes completely inhibits ctl lysis. *J. Immunol.* **2006**, *177*, 3225–3234.
197. Brunovskis, P.; Kung, H.J. Retrotransposition and herpesvirus evolution. *Virus Genes* **1995**, *11*, 259–270.
198. Becker, Y.; Asher, Y.; Bujanover, S.; Darai, G. The dynamic herpesvirus DNA genome: The case of mdv-1 and hsv-1. *Acta Virologica* **1999**, *43*, 81–89.
199. Klein, J. Did viruses play a part in the origin of the adaptive immune system? *Folia Biologica* **2004**, *50*, 87–92.

200. Hunter, P. The great leap forward. Major evolutionary jumps might be caused by changes in gene regulation rather than the emergence of new genes. *EMBO Rep.* **2008**, *9*, 608–611.
201. Zuckerkandl, E.; Cavalli, G. Combinatorial epigenetics, "junk DNA", and the evolution of complex organisms. *Gene* **2007**, *390*, 232–242.
202. Khodosevich, K.; Lebedev, Y.; Sverdlov, E. Endogenous retroviruses and human evolution. *Comp. Funct. Genom.* **2002**, *3*, 494–498.
203. Haygood, R.; Babbitt, C.C.; Fedrigo, O.; Wray, G.A. Contrasts between adaptive coding and noncoding changes during human evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 7853–7857.
204. Buzdin, A.A. Functional analysis of retroviral endogenous inserts in the human genome evolution. *Bioorganicheskaiia Khimiia* **2010**, *36*, 38–46.
205. Gogvadze, E.; Stukacheva, E.; Buzdin, A.; Sverdlov, E. Human-specific modulation of transcriptional activity provided by endogenous retroviral insertions. *J. Virol.* **2009**, *83*, 6098–6105.
206. Cohen, C.J.; Lock, W.M.; Mager, D.L. Endogenous retroviral ltrs as promoters for human genes: A critical assessment. *Gene* **2009**, *448*, 105–114.
207. Buzdin, A. Human-specific endogenous retroviruses. *Sci. World J.* **2007**, *7*, 1848–1868.
208. Jordan, I.K.; Rogozin, I.B.; Glazko, G.V.; Koonin, E.V. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* **2003**, *19*, 68–72.
209. Oja, M.; Peltonen, J.; Blomberg, J.; Kaski, S. Methods for estimating human endogenous retrovirus activities from est databases. *BMC Bioinformatics* **2007**, *8*, S11.
210. Keightley, P.D.; Lercher, M.J.; Eyre-Walker, A. Evidence for widespread degradation of gene control regions in hominid genomes. *PLoS Biol.* **2005**, *3*, e42.
211. Anzai, T.; Shiina, T.; Kimura, N.; Yanagiya, K.; Kohara, S.; Shigenari, A.; Yamagata, T.; Kulski, J.K.; Naruse, T.K.; Fujimori, Y.; *et al.* Comparative sequencing of human and chimpanzee mhc class i regions unveils insertions/deletions as the major path to genomic divergence. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7708–7713.
212. Steinhuber, S.; Brack, M.; Hunsmann, G.; Schwelberger, H.; Dierich, M.P.; Vogetseder, W. Distribution of human endogenous retrovirus herv-k genomes in humans and different primates. *Hum. Genet.* **1995**, *96*, 188–192.