Current Topics in Virology

Review

Viruses in host evolution: General principles and future extrapolations

Luis Villarreal[1,*] and Frank Ryan[2,#]

[1]Center for Virus Research, Department of Molecular Biology and Biochemistry, 2208 Biological Science 3, University of California, Irvine, Irvine CA 92697, USA. [2]Department of Animal & Plant Sciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

ABSTRACT

There is growing evidence that viruses have contributed to host evolution early in the evolution of life and throughout the subsequent evolution of biodiversity. It is important, from both a biological and medical perspective, that we grasp the essential principles involved. While virologists are familiar with the contribution of mutation to the evolutionary dynamics of the virus-host interaction, they may be less familiar with the role of partnerships and the principles and dynamics of symbiosis. We believe that an understanding of viral symbiosis complements the existing understanding of virology and offers the potential for novel virological research. This is of particular importance at a time when animal genomes, and the human genome in particular, have been found to contain large amounts of virus-derived DNA, formerly dismissed as junk but now increasingly recognized as contributing to host evolution, embryogenesis and physiology. In this paper we define the concept of viral symbiosis, clarifying the general principles involved before focusing on the particular example of the ERVWE1 locus in human evolution. From this we extrapolate the general principles to future biological and medical research.

KEYWORDS: natural selection, virus-host interactions, cooperation, partnership, symbiosis, aggressive symbiosis, genetic symbiosis, symbiogenesis, holobiont, holobiontic genome

*lpvillar@uci.edu
#F.Ryan@sheffield.ac.uk

INTRODUCTION

Symbiology is a large and expanding discipline within evolutionary biology, yet, though it has long been a source of understanding and research in microbiology in general, it is not conventionally taught to undergraduates studying virology and few virologists have been involved with symbiological research. The reasons for this are complex but in part derive from historical contingency, and in part from the long-standing difficulty in agreeing on the essential nature of viruses. Virologists may be familiar with some partnerships (virus-virus complementation, virus-satellite and virus-defective interactions) but need to familiarize themselves with the definitions and mechanics of symbiogenesis if they are to grasp the potential it offers for novel virological research and understanding. They also need to grasp how symbiogenesis differs, conceptually and dynamically, from mutation-plus-selection, while acknowledging that the two dynamics are essentially complementary rather than contradictory to a broader understanding of the evolutionary situation.

In the 1930s and 1940s the synthesis of Darwinian natural selection, mutation and Mendelian genetics gave rise to the paradigm of "modern Darwinism", also known as "neo-Darwinism". This has contributed greatly to our understanding of evolutionary dynamics, and particularly so within field of virology. But increasing knowledge of other

forces for hereditary change, such as symbiogenesis, hybridogenesis, horizontal gene transfer in prokaryotes, epigenetics, and mechanisms specific to evolutionary development, have expanded our horizons to a diversity of evolutionary possibilities. Many of these dynamics apply to viruses, both in the sense of viral evolution *per se* as well as the potential of viruses to co-evolve with, and influence, the evolution of their hosts. Viruses are genetic parasites long familiar as the cause of endemic and epidemic diseases. Over the last decade or so we have become increasingly aware of a wider role for viruses, possibly in the origins of life itself and certainly in its subsequent evolution to biodiversity as we see it today. This relatively new, but rapidly growing, field has implications for our understanding of life, ecology, human and veterinary medicine, and agriculture. It has led to a rapid expansion of novel research extending across all of the biological and applied fields. Many virologists, and workers in related disciplines, are currently involved in these advances. This makes it all the more relevant that we consider the role of viruses in symbiogenesis.

Defining symbiosis in relation to virology

Mutation, as first coined by De Vries [1] and subsequently developed through most of the twentieth century [2], was defined as genetic change arising from copying errors during cell division. This definition is conceptually and mechanistically meaningful and it predicates that mutational change, arising during meiosis and mitosis, is essentially random in its origins and effects. Thus we can predict that mutation may result in insignificant genetic change, deleterious change (giving rise to disease, or death), or occasionally to significant genetic change in individuals. Such mutations, or more likely a series of mutations, are capable, under positive selection, of giving rise to the hereditable genetic change that is essential to the origin of new species. Subsequently, as additional mechanisms of hereditable genetic change came to be discovered, there was a tendency to lump them all within the umbrella concept of mutation. But this ignores the fact that mechanisms such as genetic symbiogenesis, hybridogenesis and hereditable epigenetic change, are conceptually and mechanistically different from genetic change arising from copying errors in cell division. These differences are illustrated in Table 1, where it will be apparent that they not only involve selection at different levels, but also include genetic symbiosis, a mechanism that allows rapid and reticulate virus-host (holobiont) evolution, and which has hitherto been a historically peripheral concept to virologists. The broader evolutionary implications of these different mechanisms of hereditary change

Table 1. The operative mechanisms for hereditary change.

Operative mechanism	Genetic change	Nature of change	Level of selection[#]	Pattern of phylogeny	Amenable to environmental influences
Mutation[*]	Yes	Random, cumulative	Mainly individual gene, organism	Linear, branching	No
Genetic symbiogenesis	Yes	Non-random, rapid	Mainly holobiont	Reticulate	No
Hybrido-genesis	Yes	Non-random, rapid	Mainly hybrid genome	Reticulate	No
Epigenetic	No	Non-genetic change in gene expression	Epigenetic inheritance system	Linear, branching	Yes

*Mutation in this table is defined as random change during cell division. We acknowledge that mutation affecting developmental pathways could potentially give rise to rapid change.

[#]In certain circumstances selection may work at other levels, such as group level, and in specific circumstances it almost certainly works at more than one level, but this is the most important level at first incorporation of genomic change.

have been discussed elsewhere [3]. Here we confine ourselves to a more detailed discussion of symbiogenesis as a mechanism for hereditary change, and in particular its implications for virology.

A common misconception is to equate symbiosis with mutualism. In fact symbiosis was first defined by Anton de Bary in 1878 merely as "the living together of differently named organisms" [4]. By "differently named organisms" he meant different species. De Bary's definition of symbiosis embraced parasitism and commensalism as well as mutualism. All that is required is a living interaction. Viruses are clearly obliged to 'live with' their host during infection and their ubiquity and diversity suggest major potential for affecting host evolution. However, some biologists might object by declaring viruses are neither living nor organisms so that such partnerships are inconsequential. But, while acknowledging the difficulties posed by viruses in relation to the various definitions of life, it is evident that, from the evolutionary perspective, viruses behave like living organisms and they follow the basic rules of evolutionary biology. And many viruses form life-long partnerships with host as persistent infections. Indeed, such partnerships are much more common than many realize [5].

Virologists have traditionally extrapolated the evolutionary dynamics of mutation-plus-selection to explain the evolution of viruses *per se*, or to explain the evolutionary interactions between viruses and their hosts. It is also becoming increasingly evident that virus-host interactions have many of the features of symbiotic evolution. As long ago as 1926, Felix d'Herelle, co-discoverer with Twort of the bacteriophage, argued the case for phage as a living organism that, in certain circumstances, was symbiotic with its bacterial host [6]. The concept of viral symbiosis is now finding acceptance in symbiology [7-9] and medicine [10], and the principles of symbiogenesis merit exploration in the widest sense to the evolution of viruses per se, and to the evolutionary implications of viruses in relation to their hosts.

The simultaneous operation of natural selection, at selfish individual, or genetic, level within virus and host, and at the level of the symbiotic interaction, or "partnership", of virus and host, is the key to understanding the real evolutionary dynamic of the virus-host relationship. Indeed it may be difficult to rigidly separate parasitism from mutualism since most mutualisms arise from parasitic states. Given the nature of the evolutionary dynamic, such partnerships, particularly when they involve viral persistence, have the potential for evolutionary change at the level of the partnership itself, to bring it to any stage between the extremes of absolute parasitism or absolute mutualism. And viral persistence in itself can involve partnerships with mixtures of viral elements, including defective viruses. It is unnecessary, and probably misleading, to think of the partnership as operating exclusively in terms of either extreme, though over time a particular dynamic may come to predominate. This will become easier to understand when we deal with specific examples. We also need to grasp some differences in perspective. Where mutation-plus-selection focuses on the concept of "fitness" deriving from competition for reproductive success, symbiogenesis focuses on the dynamics of interactive partnerships between different species. The interacting partners in a symbiotic relationship are termed "symbionts" and the partnership, which can involve two or more symbionts, is termed the "holobiont". Where symbiosis gives rise to evolutionary change it is defined as symbiogenesis [11].

Virologists may be sceptical of a symbiotic perspective in what appears to be a situation of outright parasitism. But if for example we examine the evolutionary dynamics of the AIDS pandemic, we find that the rate of disease progression is strongly associated with a particular HLA-B but not HLA-A allele expression in the human host. Here Kiepiela *et al.* have reported substantially greater selection pressure imposed on HIV-1 by HLA-B alleles than by HLA-A. At the same time HLA-B gene frequencies in the population are reported to be those likely to be most influenced by HIV disease [12]. In other words even at the stage of an acute pandemic, where selection is obviously working at selfish individual, or genetic, level on both host and virus, selection is also operating at the level of the partnership, with virus influencing host gene pool frequencies and host influencing viral evolution.

Symbiotic interaction also takes place at various levels. We are familiar with the cleaner station

symbioses in the oceans, where shrimps or small fish, remove debris and parasites from the skins, or even within the mouths, of large predators, such as sharks and groupers. These symbioses involve genetically, or epigenetically, programmed mutualistic behaviour in otherwise independently programmed species. At a microbiological level, cryptic prophages help bacteria to cope with adverse environments, such as oxidative and acid stresses, or increasing growth, or influencing biofilm formation [13]. Mixtures of cryptic (defective) prophage are also crucial for the highly adaptable nature of E. coli O157:H7 [14-15]. Others assist host survival through "aggressive symbiosis", for example Gifsy-2, which provides its Salmonella host with a competitive advantage by killing competitors [16]. A similar "aggressive symbiosis" may be a feature of the plague dynamics of virus-host relationships in human disease, for example when smallpox gave a competitive advantage to the Conquistadors in their wars with the Incas [17]. A very common level of symbiotic interaction involves the sharing of metabolites between symbionts. 97% of land plants depend on metabolic symbiosis between fungi and roots, known as mycorrhizae. Many bacterial symbioses also work at a metabolic level: for example the rhizobia which fix nitrogen in the root nodules of legumes, the chemosynthetic bacteria in the trophosomes of tubeworms in the vicinity of hydrothermal vents, which help the worm to metabolize hydrogen sulphide, and the many different bacteria and protists that aid digestion in a wide variety of animals, particularly insects. Recent study suggests that viruses may also participate in metabolic symbioses, for example by providing proteins or other metabolites that are helpful to host metabolism [13], or even essential for host survival [18]. In similar manner, cyanophage, encoding both photosystem I and II genes as well as electron transport proteins, provide major metabolic functions to their infected partners [19, 20]. Large DNA viruses of protists have also contributed to the lateral transfer of complex metabolic pathways (i.e. sphingolipid biosynthesis) [21], including a viral p450 gene [22].

In practice symbioses frequently operate at more than one level, for example pollination mutualisms involve both behavioural and metabolic aspects. But if there is a core concept that is important to grasp, it is the fact that in all symbioses selection, while still operating at selfish individual, or genetic level, will also operate, to a significant degree, at partnership, or holobiontic, level. For example, in the cleaner station symbioses, we can see how the interacting partners display significant behavioural changes, or adaptations, in relation to how they interact with each other, and in hummingbird-flower symbioses we can see how the elongated curved bill of the violet sabrewing hummingbird and the depth and shape of the columnia flower have accommodated to the partnership [23].

From an evolutionary perspective, the most important level of symbiotic interaction is at the genetic level. This is of major importance to viral symbiosis and will be discussed in more detail.

Genetic symbiosis

In symbioses at genetic level, symbionts contribute whole pre-evolved genes or regulatory sequences to their partners, or at its most innovative level, symbionts unite whole pre-evolved genomes to form a new holobiontic genome. Striking examples of genetic symbiosis include mitochondria and plastids, which derived from holobiontic unions of the ancestors of animals, plants and fungi with photosynthetic and oxygen-breathing prokaryotes and protists [24]. Viruses, notably retroviruses and bacteriophages, have an extraordinary potential to invade, and subsequently unite with, host genomes. Virologists are familiar with the process of endogenization, in which infectious retroviruses insert themselves into the chromosomes of the host germ line, to become endogenous retroviruses, or ERVs. This process is currently being observed in the koala retroviral epidemic in Australia. All vertebrate lineages, and many non-vertebrate lineages, have been extensively colonized in this way by a variety of retroviral lineages. Meanwhile the related genetic entities of LINEs and SINEs have spread through widespread retroposition throughout the chromosomes of animal genomes, from marine invertebrates to humans.

From the symbiological perspective, the endogenization of retroviruses into vertebrate lineages might be seen as the viral equivalent of the symbiogenetic union of genomes with the

ancestral microbes of mitochondria and plastids. However, we also need to consider that symbiotic viruses will make contributions that are fundamentally different to bacteria, in ways that depend on the quintessential nature and evolutionary dynamics of viruses. To understand what this implies we need to consider the basic genomic structure, and functional dynamics, of a retrovirus.

Virologists are familiar with the basic retroviral genome, which comprises three genetic domains, conventionally referred to as the genes, *env*, *gag* and *pol*, which code for a variety of proteins (see Figure 1). For example *gag* codes for the proteins necessary for viral assembly, including matrix and core shell proteins, *pol* codes for the enzymes necessary for viral replication, such as reverse transcriptase, protease, ribonuclease and integrase, while *env* codes for the surface and trans-membrane glycoproteins. The flanking LTRs contain the viral regulatory regions. Thus, when a retrovirus invades the host chromosomes, it introduces a pre-evolved genome that is not only novel to the host genome, but is also a holistically functional unit whose genes, and LTRs, are pre-evolved to interact with and regulate key aspects of host genetics, immunity, and physiology. Indeed, retroviral integration is not random and shows a very strong bias towards 5', 3' and intron regulatory sequences [25, 26]. The invasion of the host germ line is potentially hazardous, and inbuilt defence mechanisms will attempt to switch off the invading viral genome, initially through epigenetic mechanisms such as methylation, and more long term through stop mutations, insertions and deletions. However, at the same time, the invading viral genome will also offer the potential of novel holobiontic genomic evolution - much as the invasion of the oxygen-breathing bacterium offered the potential of oxygen breathing to the protist host long ago. Indeed, as retroviral invasions typically involve numerous scattered elements (LTRs), they also provide the potential for introducing new complex regulatory networks. There are very few mechanisms by which such coordinated multipoint genetic change might otherwise occur by classical point mutation selection. This symbiogenetic potential will include a number of possibilities, such as the contribution of viral genes novel to the host and the contribution of the viral regulatory regions, the LTRs, to holobiontic genomic function.

Invading retroviruses will often endogenize the germ lines of their hosts again and again, with up to a thousand or so different sites of integration randomly distributed throughout the chromosomes, each integration site offering the possibility of future symbiogenetic potential. In human terms, although many such insertions have been "switched off" by the policing action of selection, such is the massive volume of viral insertions that a significant number have positively contributed to the virus-host genetic union, adding new complexity to the evolving holobiontic genome. Indeed reproductive tissue in particular will often retain the capacity to express many retroviral elements as presented below.

Such genetic symbioses afford considerable potential for evolutionary change both in the short and long term, and this will inevitably include novel potential deriving from the quintessential nature of viruses and their pre-evolved potentials to control and manipulate key aspects of the host physiology and genome. We must also consider that with the embedding of whole viral genomes within the host germ line - with pre-evolved abilities to interact with, and manipulate the host genome - there now exists a new potential for symbiogenetic interaction at genetic level. As we have seen with symbiosis at all levels, where selection will often begin in large degree at selfish individual, or genetic, level, in such symbiogenetic unions of viruses and their hosts, there will be a new, and powerful selection pressure, that will operate at partnership - or holobiontic - level.

Here we see how a symbiotic perspective adds to our understanding of the potential evolutionary outcome. It implies is that there will be selection for

Figure 1. Schematic of the basic retrovirus genome.

former host or viral genes, or regulatory sequences, or whole functioning genetic units comprising genes plus relevant regulatory sequences, that enhance survival of the new holobiontic organism; meanwhile there will be selection against former host or viral genes, or regulatory sequences, or whole functioning genetic units, that impair survival of the new holobiontic organism. This level of understanding is helpful in constructing a logical experimental methodology, capable of confirming holobiontic evolution of virus and host. This is illustrated, in a series of logical steps, in Table 2, which outlines the type of evidence that supports virus-host holobiontic selection.

Some 8% of the human genome is derived from human endogenous retroviruses (HERVs) and, if we extend this to HERV fragments and virus-dependent entities, the retroviral legacy amounts to almost half of the human DNA. Where this was previously dismissed as "junk DNA", there is growing evidence that HERVs, and related retroviral sequences, have made a major contribution to human evolution, and are playing important roles in human embryology, reproduction and day-to-day physiology. HERVs are conventionally divided into 30-50 different HERV families, subdivided into some 200 different subgroups, which originated from exogenous retroviral invasion and resulted in multiple integrations. The families and subgroups are commonly characterised on the basis of the single-letter code for the amino acid complementary to the t-RNA primer binding site that initiates transcription. Thus, for example, a HERV-K initiates transcription with lysine and HERV-W with tryptophan. Each family and subgroup is thought to represent an independent evolutionary lineage, which would imply that the ancestral human genome has been subject to a large series of independent retroviral colonisations. Although many of these unions took place during our mammalian ancestry, at least eight full length HERV-K viruses are unique to humans having entered the human germ line after humans diverged from chimpanzees [27, 28]. The symbiogenetic potential of such massive viral invasion, and subsequent holobiontic interaction, is likely to be major.

It is instructive to examine one important locus of symbiogenetic viral interaction as the exemplar. Known as the ERVWE1 locus, this is a genetic region that humans share with the other great apes, and which plays a vital role in reproductive biology, consistent with early and recent proposals for such a role [29, 30].

The ERVWE1 Genetic Locus

Syncytin-1 is a viral protein that plays an essential role in the differentiation of the syncytiotrophoblast cells of the human placenta [31, 32]. It's genetic coding, and regulation, is through the ERVWE1 locus, based around a HERV-W insertion located on human chromosome 7 (7q21.2), as shown in

Table 2. Step-wise evidence for viral symbiosis at genetic level with selection operating at holobiontic level.

1.	Strict preservation of viral genetic sequences in host genome over long periods by comparing sequences between widely separated populations, or races, within species.
2.	Strict preservation of viral gene sequences in host genomes between related species within genera or more widely disparate evolutionary taxa, for example between great apes including humans.
3.	Strict preservation of viral gene plus its viral promoter or other regulatory sequences as in 1 and 2.
4.	Evidence for functional interaction between viral and host genetic sequences in the embryogenesis or physiology of the host.
5.	Discovery of conserved viral regulatory sequence having replaced earlier host regulatory sequence, with viral sequence now playing a functional role in host embryology or physiology.
6.	Discovery of conserved viral genetic sequence within intron or open reading frame of functional host gene or regulatory sequence.
7.	Discovery of expressed protein product of viral gene, with viral protein playing a significant role in host embryology or physiology.

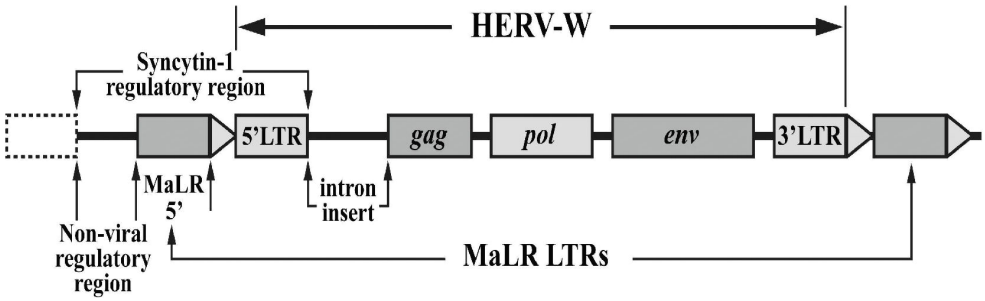

Figure 2. Schematic illustration of the ERVWE1 locus.

Figure 2. This has been adapted from the genetic study published by Gimenez and Mallet [33]. It is clear that the locus contains a complete HERV-W, with *gag*, *pol* and *env* genes flanked by the regulatory LTRs. The viral genome is further flanked by the isolated LTRs of a second endogenous retrovirus of the MaLR family. The HERV-W is thought to derive from germ-line invasion by an exogenous retrovirus between 20 and 30 million years ago, and while the *gag* and *pol* genes have been inactivated by deletions and stop codons, the *env* gene has been strictly conserved in humans, chimpanzees, gorillas and orang-utans. This *env* gene thus fulfils steps 1, 2 and 7 of Table 2. It also fulfils step 3 in that its promoter is within the selectively conserved 5'LTR of the HERV-W, and it fulfils step 4 in the finding that additional upstream regulation is provided by host-derived elements within the non-viral regulatory region.

ERVWE1 is also enlightening in illustrating the potential for cooperation between disparate viral elements in host evolution. The regulation of syncytin is further dependent on conserved elements within the 5'LTR of the MaLR virus. We also know that placentation in humans is dependent on the cooperative interaction of the HERV-W that codes for syncytin with three additional HERVs, HERV-FRD [34], ERV-3 [35] and HERV-K [36], as well as a number of additional HERVs [37-42] that are also thought to play various, if as yet poorly defined, roles in human reproduction (Table 3). Many retroviral *env* glycoproteins possess an immunosuppressive domain (ISU), so that the cooperative interaction of the ISUs of several placentally-expressed HERVs may also play an important role in immune regulation at the maternal-foetal interface. Much remains to be clarified, but the complexity of such cooperative interactions between the various viral entities and non-viral components of the human genome in human placentation is a remarkable feature of the holobiontic genome [43]. In effect, selection at holobiontic level must involve selection for cooperative interactions between different viral elements as well as between viral and non-viral elements. It is likely that similar cooperative interactions between viral and non-viral elements play a part in reproduction in other animals. Two endogenous retroviruses are also known to be involved in murine placentation [44], and viral elements are critical to placentation in sheep [45] and possibly cattle [46], suggesting that viral symbiosis has been central to the evolution of placentation throughout the mammals. The complexities of such retrovirus-host symbioses has been illustrated in particular by the ongoing work of Massimo Palmarini and others in the example of the sheep Jaagsietke virus [47].

Virus-host symbioses will fundamentally alter the interactive relationship that results from the union. One potential outcome might be the capacity of viral release from the holobiontic genome in some later circumstance, as a provirus or prophage that might contribute to host competitiveness or survival. This is seen in phage-bacterial interactions as described above. Another potential outcome might involve the induction of resistance in the host for the same exogenous virus, or a related strain. Virus-host symbioses will fundamentally alter the interactive relationship that results from the union. One potential outcome might be the capacity viral release, as a provirus or prophage, in some later circumstance. Another potential outcome could involve the induction of resistance in the host for

Table 3. HERVs that play various roles in human reproduction, notably placentation. Many of these roles are not yet fully determined but the cooperative interplay between the different viral entities and non-viral genes and regulatory regions is remarkable [31-42].

HERV	Role in reproduction
ERVWE1 (HERV-W)	Syncytiotrophoblast differentiation and cell fusion to form syncytium.
HERV-FRD	Syncytiotrophoblast fusion. Also thought to play a role in placental defence against maternal immune rejection of foetus.
ERV-3	Syncytiotrophoblast fusion (May not be essential since a small minority lack this role but can still create a functional placenta).
HERV-K	Expressed in cytotrophoblasts and probably contributes to normal placentation. Possible contribution to immune protection of the foetus.
HERV-E.PTN	HERV insert in growth factor pleiotrophin gene (PTN) led to a phylogenetically new promoter driving the expression of functional HERV-PTN transcripts which are expressed in the proliferative and invasive trophoblasts of gestational tissue as well as the malignant trophoblasts of choriocarcinomas.
LTR10A of HERV-1	Placenta-exclusive expression of nitric acid synthase gene (*NOS3*).
HERV-H7/F(XA34)	Trophoblast-specific transcription, with reduced expression in placentas from pregnant women suffering from pregnancy-induced hypertension.
HERV-Fb1	Trophoblast-specific transcription, with reduced expression in placentas from pregnant women suffering from pregnancy-induced hypertension.
HERV-HML6-c14	Trophoblast-specific transcription. In contrast to syncytin, HERV-FB 1 and HERV-H7/F(XA34) transcripts were localized to the nucleus and expression was raised in placentas from women suffering from pregnancy-induced hypertension.
Unknown HERV LTR	Human-placenta-specific expression of Human Insulin family gene, *INSL4*, during differentiation of cytotrophoblast into syncytiotrophoblast cells is regulated to a significant degree by a HERV-3LTR.
HERV-E LTR	Contributes to the expression of endothelin B receptor (EDNRB), one of two receptors that mediate the vasoconstrictor effects of endothelins.
HERV-E LTR	Alternative tissue-specific transcriptional regulator of the human Opitz gene, *Midl*, which encodes a microtubule-associated protein.

the same exogenous virus, or a related strain. Indeed, small regulatory RNA transcribed from the integrated viral elements should have significant affect on the virus-host selection. The sheep Jaagsietke virus exists in two closely homologous forms, one of which has been endogenized into the sheep genome and the other remains exogenously active as a pandemic virus infection in sheep. The endogenous symbiont within the sheep genome has enabled the host to present a dual physiological blocking of entry, through the genital tract, meanwhile the exogenous virus has evolved ingress through the respiratory route [48].

There is some evidence, albeit preliminary, that different HERV families may be contributing as yet unknown roles in the normal human brain [49-53]. Although not yet published, Larsson *et al.* have demonstrated widespread and dense expression of the *env*-coded proteins syncytin-1 and syncytin-2 in the cortex of the developing human brain, indicating what may prove to be structural or physiological function, or both [54]. Given the complex regulatory functions that must be used during human brain evolution, symbiosis of HERV and other elements would appear an attractive way of generating such complexity. A symbiotic perspective may also be relevant to the growing evidence that HERVs may contribute to diseases involving the central nervous system [55]. There are two additional considerations of holobiontic

viral integration that need to be grasped. Where a specific HERV has inserted many times into different chromosomes, and where one or more of its insertions have become conserved for an essential genetic function, selection is likely to exert a "policing" control on the unwanted expression of homologous genetic sequences from other insertion sites, which might otherwise give rise to dysregulation of the conserved role. This is further complicated by the fact that viral sequences, even if truncated or silenced by stop codons, may be capable of reactivation and expression through recombination between related HERVs, *in vitro* and possible even *in vivo* [56, 57]. They may also be capable of inhibiting retroviral and other gene expression via small regulatory RNA expression from ERV elements [58].

The need for a HERV transcriptome project

HERVs and their products appear to play an equally pervasive role in normal adult structure and physiology. The phylogenetically conserved *env* genes presented above are also observed to be conserved in screens of human transcripts, especially in the placenta [59]. The LTR of ERV-L controls most of the gene transcripts of the human gene *β3GAL-T5* in the human colon [60], and two other HERV LTRs help regulate the human genes *SLC4A8* (sodium bicarbonate co-transporter) and *IFT172* (intraflagellar transport protein 172) [61], and more than 50% of the human-specific HERV-K LTRs are active promoters for non-viral DNA transcription [62]. An example is the placenta specific NOS3 expression via HERV LTR10A [63]. In another illustrative example of selection working at holobiontic level, the LTR of ERV-9 plays a key role in transcriptional control of the β-globin gene cluster of humans [64, 65], a role that has been conserved throughout at least 15 million years of primate evolution, and in which the viral promoter appears to have displaced several non-viral promoters within the locus control region. Many virus-related or dependent elements, including SINEs (*Alus)*, LINEs, and LTRs, are found in a large number of human protein-coding genes, where most are inserted into introns, where they appear to influence some 533 genes [66]. LTRs also act as alternative promoters, or splice receptors, for example in the control of the endothelin B receptor and apolipoprotein C-I genes [67], and in the control of the human leptin receptor [68]. Symbiotic retroviruses have also contributed to the evolution, and tissue specific expression, of the enzyme amylase in humans [69, 70], and HERV sequences have contributed functional genes, or parts of genes, to the human genome, including integrase [71] and transaldolase [72].

In fact the viral roles in human genomic function are so widespread, yet still underestimated, that Flockerzi and colleagues have suggested the need for a specialised human endogenous retrovirus transcriptome project [73].

Extrapolating viral symbiosis to biology in general

A similar complex, multi-faceted contribution of viral symbiosis is increasingly found throughout all the biological disciplines, working at many levels of virus-host interaction, and including both exogenous and endogenous viruses [74]. Virus-host symbioses are critical to the highly successful evolution of the parasitic ichneumonid and braconid wasps, which involve approximately 25,000 species of wasps and approximately 20,000 species of polydnaviruses. These include genetic and non-genetic virus-host symbioses that have been conserved by selection at holobiontic level for more than 70 million years [75-77]. Viral genes are expressed in the Lepidopteran caterpillar prey, where they block the cellular (haemocyte) immune rejection of the eggs, disrupt the caterpillar endocrine system, suppressing the thoracic gland and thus the metamorphic development, while inducing the production of sugars to feed the wasp larvae.

By their very nature, viruses cooperate with host elements in their replication and expression systems. They must both use existing host genetic code and also extend it with new viral information. But viruses can also cooperate with other viruses. This would suggest interesting possibilities for future research into the investigation of the action of selection at holobiontic level on the interaction of exogenous viruses and host in a multitude of situations throughout biology. For example, humans appear to be especially prone to harbour various human specific herpes family viruses. And, interestingly, herpes viruses often appear to evolve via the action of retroviruses [78, 79].

Thus a human-herpesvirus-retrovirus holobiont could well have affected human evolution. Such symbiosis may also have ancient and ongoing consequence. For example, virus-host symbioses may have helped usher in the transition from an RNA to DNA world by providing DNA replication as seen in bacteria as well as the multi-start site replication seen in eukaryotes and some Archaea [80, 81]. At molecular genetics level, it is becoming increasingly established that viruses can have a significant impact on the content and regulation of the chromosomes of host cells. It is possible that viral symbiosis may have contributed to the origins of nuclei, to key enzymes involved in DNA and even whole chromosomal replication. Indeed the symbiotic role of viruses in host evolution is seen to be both major and universal. But a great deal more needs to be done.

The potential for future research into viral symbiosis would appear to be considerable. Since virus-host symbioses are universal, such potential will extrapolate to theoretical aspects of evolutionary biology as well as to the practical applications, including medicine, veterinary medicine, agriculture, and ecology.

REFERENCES

1. De Vries, H. 1906, Species and Varieties, Their Origin by Mutation, Open Court, Chicago.
2. Hartl, D. L. and Jones, E. W. 2003, Genetics: Analysis of Genes and Genomes, 5th Edition, Jones and Bartlett, Boston.
3. Ryan, F. P. 2006, Biol. J. Linnean Soc., 88, 655.
4. Sapp, J. 1994, Evolution by Association: A History of Symbiosis, Oxford University Press.
5. Villarreal, L. P., Defilippis, V. R., and Gottlieb, K. A. 2000, Virology, 272, 1.
6. D'Herelle, F. 1926, Bacteriophage and its Behaviour, Ballière, Tindall and Cox, London.
7. Margulis, L., Hall, J., and McFall-Ngai, M. 2007, Symbiosis, 44, ii.
8. Villarreal, L. P. 2007, Symbiosis, 44, 1.
9. Ryan, F. P. 2007, Symbiosis, 44, 11.
10. Ryan, F. P. 2009, J. Roy. Soc. Med., 102, 324.
11. Sapp, J. 2002, History and Philosophy of Life Sciences, 24, 413.
12. Kiepiela, P., Leslie, A. J., Honeyborne, I., Ramduth, D., Thobakgale, C., Chetty, S., Rathnavalu, P., Moore, C., Pfafferott, K. J., Hilton, L., Zimbwa, P., Moore, S., Allen, T., Brander, C., Addo, M. M., Altfeld, M., James, I., Mallal, S., Bunce, M., Barber, L. D., Szinger, J., Day, C., Klenerman, P., Mullins, J., Korber, B., Coovadia, H. M., Walker, B. D., and Goulder, P. J. R. 2004, Nature, 432, 769.
13. Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturin, J. M., and Wood, T. K. 2010, Nature Communications, doi: 10.1038/ncomms1146.
14. Yang, Z., Kim, J., Zhang, C., Zhang, M., Nietfeldt, J., Southward, C. M., Surette, M. G., Kachman, S. D., and Benson, A. K. 2009, J. Bacteriol., 191, 3553.
15. Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M., Oshima, K., Kodama, T., Abe, H., Nakayama, K., Kurokawa, K., Tobe, T., Hattori, M., and Hayashi, T. 2009, PNAS, 106, 17939.
16. Bossi, L., Fuentes, J. A., Mora, G., and Figueroa-Bossi, N. 2003, J. Bacteriol., 185, 6467.
17. McNeill, W. H. 1977, Peoples and Plagues, Basil Blackwell, Oxford, England.
18. Márquez, L. M., Redman, R. S., Rodriguez, R. J., and Roossinck, M. J. 2007, Science, 315, 513.
19. Lindell, D., Sullivan, M. B., Johnson, Z., Tolonen, A. C., Rohwer, F., and Chisholm, S. W. 2004, PNAS, 101, 11013.
20. Alperovitch-Lavy, A., Sharon, I., Rohwer, F., Aro, E. M., Glaser, F., Milo, R., Nelson, N., and Beja, O. 2011, Environ. Microbiol., 13, 24.
21. Monier, A., Pagarete, A., de Vargas, C., Allen, M. J., Read, B., Claverie, J. M., and Ogata, H. 2009, Genome Research, 19, 1441.
22. Lamb, D. C., Lei, L., Warrilow, A. G., Lepesheva, G. I., Mullins, J. G., Waterman, M. R., and Kelly, S. L. 2009, J. Virol., 83, 8266.
23. Ryan, F. 2002, Darwin's Blind Spot, Houghton Mifflin, Boston/ New York.
24. Margulis, L. and Sagan, D. 1986, Microcosmos: Four Billion Years of Microbial Evolution, University of California Press.

25. Mitchell, R. S., Beitzel, B. F., Schroder, A. R., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D. 2004, PLoS Biol., 2, E234.
26. Desfarges, S. and Ciuffi, A. 2010, Viruses-Basel, 2, 111.
27. Barbalescu, M., Turner, G., Seaman, M. I., Deinard, A. S., Kidd, K. K., and Lenz, J. 1999, Current Biology, 9, 861.
28. Medstrand, P. and Mager, D. L. 1998, Journal of Virology, 72, 9782.
29. Villarreal, L. P. and Villareal, L. P. 1997, J. Virol., 71, 859.
30. Varela, M., Spencer, T. E., Palmarini, M., and Arnaud, F. 2009, Ann. NY Acad. Sci., 1178, 157.
31. Mi, S., Lee, X., Li, X-P., Veldman, G. M., Finnerty, H., Racie, L., LaVallie, E., Tang, X-Y., Edouard, P., Howes, S., Keith, J. C. Jr., and McCoy, J. M. 2000, Nature, 403, 785.
32. Blond, J-L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F-L. 2000, J. Virol., 74, 3321.
33. Gimenez, J. and Mallet, F. 2007 (September), Atlas Genet. Cytogenet. Oncol. Haematol., See also http://AtlasGeneticsOncology.org/Genes/ERVWE1ID40497ch7q21.html.
34. Blaise, S., de Parseval, N., Bénit, L., and Heidmann, T. 2003**,** PNAS, 100, 13013.
35. Rote, N. S, Chakrabarti, S., and Stetzer, B. P. 2004, Placenta, 25, 673.
36. Kämmerer, U., Germeyer, A., Stengel, S., Kapp, M., and Denner, J. 2011, J. Reprod. Immunol., doi: 10.1016/j.jri.2011.06.102.
37. Schulte, A. M., Shoupeng, L., Kurtz, A., Czubayka, F., Riegel, A. T., and Wellstein A. 1996, PNAS, 93, 14759.
38. Huh, J-W., Ha, H-S., Kim, D-S., and Kim, H-S. 2008, Placenta, 29, 602.
39. Koduka, W., Oda, T., Jinno, Y., Yoshimi, N., and Aoki, Y. 2008, Placenta, 29, 282.
40. Bièche, I., Laurent, A., Laurendeau, I., Duret, L., Giovangrandi, Y., Frendo, J-L., Olivi, M., Fausser, J-L., Danie`le Evain-Brion, D.le, and Michel Vidaud, M. 2003, Biology of Reproduction, 68, 1422.
41. Landry, J. R. and Mager, D. L. 2003, Virology, 77, 7459.
42. Landry, J. R., Rouhi, A., Medstrand, P., and Mager, D. L. 2002, Mol. Biol. Evol., 19, 1934.
43. Sugimoto, J. and Schust, D. J. 2009, Reproductive Sciences, 16, 1023.
44. Dupressoir, A., Marceau, G., Vernochet, C., Bénit, L., Kanellopoulos, C., Sapin, V., and Heidmann, T. 2005, PNAS, 102, 725.
45. Dunlap, K. A., Palmarini, M., Varela, M., Burghardt, R. C., Hayashi, K., Farmer, J. L., and Spencer, T. E. 2006, PNAS, 103, 14390.
46. Baba, K., Nakaya, Y., Shojima, T., Muroi, Y., Kizaki, K., Hashizume, K., Imakawa, K., and Miyazawa, T. 2011, J. Virol., 85, 1237.
47. Arnaud, F., Caporale, M., Varela, M., Biek, R., Chessa, B., Alberti, A., Golder, M., Mura, M., Zhang, Y-p., Yu, L., Pereira, F., DeMartini, J. C., Leymaster, K., Spencer, T. E., and Palmarini, M. 2007, PLOS Pathogens, 3, 1716.
48. Palmarini, M., Mura, M., and Spencer, T. E. 2004, J. Gen. Virol., 85, 1.
49. Perron, H., Lazarini, F., Ruprecht, K., Péchoux-Longin, C., Seilhean, D., Sazdovitch, V., Créange, A., Battail-Poirot, N., Sibaï, G., Santoro, L., Jolivet, M., Darlix, J. L., Rieckmann, P., Arzberger, T., Hauw, J. J., and Lassmann, H. 2005, J. Neurovirol., 11, 23.
50. Frank, O., Giehl, M., Zheng, C., Hehlmann, R., Leib-Mösch, C., and Seifarth, W. 2005, J. Virol., 79, 10890.
51. Kim, H-S., Ahn, K., and Kim, D-S. 2008, Arch. Virol., 153, 1587.
52. Weiss, S., Llenos, I. C., Sabunciyan, S., Dulay, J. R., Isler, L., Yolken, R., and Perron, H. 2007, J. Neurl. Transm., 114, 645.
53. Kwon, D. N., Nguyen, S., Chew, A., Hsu, K., Greenhalgh, D., and Cho, K. 2008, Virus Genes, 36, 439.
54. Professor Larsson presented this work at a lecture at the Linnean Society of London and the work is currently in preparation for publication.
55. Ryan, F. P. 2011, Current Neuropharmacology, 9, 360.
56. Ryan, F. P. 2009, J. Roy. Soc. Med., 102, 415.
57. Weiss, R. A. 2006, Retrovirol., 3, doi:10.1186/1742-4690-3-67.
58. Best, S., Le Tissier, P., Towers, G., and Stoye, J. P. 1996, Nature, 382, 826.

59. de Parseval, N., Lazar, V., Casella, J. F., Benit, L., and Heidmann, T. 2003, J. Virol., 77, 10414.
60. Dunn, C. A., Medstrand, P., Mager, D. L. 2003, PNAS, 100, 12841.
61. Gogvadze, E., Stukacheva, E., Buzdin, A., and Sverdlov, E. 2009, J. Virol., doi:10.1128/JVI.00123-09.
62. Buzdin, A., Kovalskaya-Alexandrova, E., Gogvadze, E., and Sverdlov, E. 2006, J. Virol., 80, 10752.
63. Huh, J. W., Ha, H. S., Kim, D. S., and Kim, H. S. 2008, Placenta, 29, 602.
64. Plant, K. E., Routledge, S. J., and Proudfoot, N. J. 2001, Mol. Cell Biol., 21, 6507.
65. Routledge, S. J. and Proudfoot, N. J. 2002, J. Mol. Biol., 323, 601.
66. Tenko, A. and Li, W-H. 2001, Trend. Genet., 17, 619.
67. Medtrand, P., Landry, J. R., and Mager, D. L. 2001, J. Biol. Chem., 276, 1896.
68. Kapitonov, V. V. and Jurka, J. 1999, J. Mol. Evol., 48, 248.
69. Samuelson, L. C., Wiebauer, K., Snow, C. M., Meisler, M. H. 1990, Mol. Cell Biol., 10, 2513.
70. Ting, C-N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C., Meisler, M. H. 1992, Genes. Dev., 6, 1457.
71. Kitamura, Y., Ayukawa, T., Ishikawa, T., Kanda, T., Yoshiike, K. 1996, J. Virol., 70, 3302-6.
72. Banki, K., Halladay, D., and Perl, A. 1994, J. Biol. Chem., 269, 2847.
73. Flockerzi, A., Ruggieri, A., Frank, O., Sauter, M., Maldener, E., Kopper, B., Wullich, B., Seifarth, W., Müller-Lantzsch, N., Leib-Mösch, C., Meese, E., and Mayer, J. 2008, BMC Genomics, 9, 354.
74. Villarreal, L. P. 2005, Viruses and the Evolution of Life, ASM Press, Washington.
75. Provost, B., Varricchio, P., Arana, E., Espagne, E., Falabella, P.,Huguet, E., La Scaleia, R.,Cattolico, L., Poirié, M., Malva, C., Olszewski, J. A., Pennacchio, F., and Drezen, J. M. 2004, J. Virol., 78, 13090.
76. Whitfield, J. B. 2002, PNAS, 99, 7508.
77. Belle, E., Beckage, N. E., Rousselet, J., Poirié, M., Lemeunier, F., and Drezen, J. M. 2002, J. Virol., 76, 5793.
78. Isfort, R., Jones, D., Kost, R., Witter, R., and Kung, H. J. 1992, PNAS, 89, 991.
79. Brunovskis, P. and Kung, H. J. 1995, Virus Genes, 11, 259.
80. Forterre, P. 2005, Biochimie, 87, 793.
81. Forterre, P. 2006, Virus Res., 117, 5.

.