Old Herborn University Seminar Monograph 31:
Evolutionary biology of the virome, and impacts in human health and disease.
Editors: Peter J. Heidt, Pearay L. Ogra, Mark S. Riddle and Volker Rusch.
Old Herborn University Foundation, Herborn, Germany: 5-13 (2017).

EVOLUTIONARY BIOLOGY OF THE VIROME AND IMPACTS ON HUMAN HEALTH AND DISEASE: AN HISTORICAL PERSPECTIVE

PEARAY L. OGRA[1] and MARK S. RIDDLE[2]

[1]Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA;
[2]Uniformed Services University of Health Sciences, Bethesda, MD, USA

"THE SINGLE BIGGEST THREAT TO MAN'S CONTINUED DOMINANCE ON THE PLANET IS A VIRUS"
(Joshua Lederberg)

This quote opens the Hollywood movie "Outbreak" to introduce the epidemics of Haemorrhagic Virus fever in Zaire in 1967 and again in the mid 1990's. The movie is a fascinating commentary on the contemporary perceptions of serious or fatal virus infections, and subsequent impact on human rights and the societal good versus evil. The perception of viruses as evil life forms has been part of human society for thousands of years, and the word "virus" has been used in Latin, Greek and Sanskrit languages for centuries to describe the venom of snake, a dangerous slimy liquid, a fatal poison, or a substance produced in the body as a result of disease, especially one that is capable of infecting others.

Although numerous studies have attempted to explain the evolution of viruses and their biologic ancestry, the precise origins of viruses continue to remain a mystery. Recent studies have proposed that modern viruses are possibly derived from multiple ancient (now extinct) cell types that harboured segmented RNA genomes and co-existed with the ancestors of modern living cells (*Nasir* and *Caetano-Anolles*, 2015). Despite the mystery surrounding their origins, viruses have played a key role in the exploration of modern cellular and subcellular biologic systems, especially in defining the structure and function of the genes, and in the exploration of many other aspects of molecular biology. Of particular importance has been the role of viruses in human diseases.

Viruses represent the single most successful (numerically) biologic entity to date. It is estimated that $>10^{31}$ virus particles inhabit the earth and over 10^{7} viruses exist per ml of water in the oceans. It has been suggested that viruses outnumber their hosts globally by tenfold or more (*Proctor*, 1997). Virus as an infectious entity was initially introduced to describe a process of non-bacterial pathology by Ivanovsky in 1892. In 1898, Beijerinck independently applied the term to identify the causative agent of tobacco mosaic disease (*Johnson*, 1942). Three decades later, Stanley succeeded in crystalizing the tobacco mosaic virus (*Stanley*, 1935). The crystalized preparations retained all the biologic properties, including the ability to infect live cells. As a result, Stanley concluded that the viruses could not be truly alive.

Although the existence of viruses as a distinct morphologic and biologic entity has now been established for over a century, most modern classifications of existing life forms have continued to exclude viruses from the tree of life, in part, because they did not

exhibit the ribosomal genes which are routinely employed in microbial phylogenetic analysis. Recent identification of a giant virus species with dimensions and size of the genomes as big as many other microbes, and the discovery of mimivirus, a parasite of amoeba with a 1.2 megabase-pair DNA genome, and of other giant viruses whose genomes contain over 1.9 to 2.5 million bases have challenged these perceptions of the viruses (*Raoult* et al., 2006; *LaScola* et al., 2008; *Desnues* et al., 2012).

Based on ribosomal gene analysis, the living organisms have now been classified in the following domains:

1) Subcellular organisms such as viruses.
2) Bacteria and prokaryotes, as unicellular organisms without a nucleus or any membrane bound organelle in the cell. The classical examples include bacteria such as *E. coli, Streptococcus pyogenes, and cyanobacteria*. However, *Proteobacteria* are phylogenetically and physiologically related to eukaryotes.
3) Archaea. These resemble bacteria, are prokaryotes and contain no nucleus. They make up their own domain, live in a wide range of extreme environmental conditions.
4) Eukaryotes. Include virtually all nucleated unicellular or multicellular organisms including, fungi, plants, animal (including man), and
5) Protista. These include some multicellular and single cellular organisms which don't fit in other life categories. They are characterized in two major taxons; Animal like protists include amoeba, trypanosomes, plasmodium (malarial parasites), crithidia. Plant like protists include red or green algae and possibly other still to be classified life forms (*Woese* and *Fox*, 1977).

Based on the extensive use of whole genome sequencing, it now appears that most living organisms may in fact be chimeras containing genes from many different sources, including earlier eukaryotic life forms, prokaryotic unicellular non-nucleated organisms; and subcellular life forms including viruses. Recently, based on a series of elegant studies, Forterre and his colleagues have suggested that all life forms on earth could be more appropriately divided into two main groups:

1) ribosome encoding organisms (modern cells) and,
2) capsid encoding organisms (viruses).

Forterre has proposed that all modern cells descend from "the last universal cellular ancestor or the last universal CenAncester (LUCA). The ribosomes of LUCA contain 33 ribosomal proteins and 3 rRNA molecules. He and his colleagues have proposed that the modern universal optimized genetic code was possibly operational in the LUCA. Therefore, it may be possible to draw a new tree of life connecting together all ribosome-encoding organisms. However, there is not a single informational molecule identified to date that is common to all viruses (*Forterre* and *Prangishvilli*, 2009). Although, these authors concluded that "understanding how modern viruses originated, thus appears to be a more complex problem from the start than understanding the evolutionary history of modern cells", they have also argued strongly in favour of the possibility that viruses are essentially parasitic organisms which infect living cells and produce virions in order to spread their genes. Viral genes possibly originated in the "virosphere" during replication of the virus genome, or were recruited from other cellular lineages which are now extinct (*Forterre*, 1992). Certain specific viral proteins are present in the existing cellular domains of life forms

Table 1: Relative distribution of different viruses in different life forms*.

Virus types	No. of virus genera in the life forms	
	Prokaryotic	Eukaryotic
RNA Viruses	Few	Many
(+) RNA	<10	150-160
(-) RNA	<2	50-60
dsRNA	<2	30-35
Retroviral	5-6	15-20
ssDNA	8-10	>100
dsDNA (All)	40-50	70-80

*Adapted from King, et al., 2011.

infected by different viruses, raising the distinct possibility that viruses are very ancient, and at least two types of virions may have originated independently before the evolution of LUCA (*Forterre*, 1992). More recently, these investigators have proposed that DNA was "invented" by viruses, helping to convert a world of RNA based organisms to one of DNA-based hereditary ancestry, and giant viruses represent the origin of the nucleus of eukaryotic life forms (*Forterre*, 2010).

ENVIRONMENT AND HUMAN VIROME

Environmental Virome

Viruses inhabit virtually the entire spectrum of unicellular or multicellular life forms described above and which exist on earth today. However, limitations in our ability to detect, isolate and classify many known and still to be discovered viruses have limited our ability to explore in detail the comprehensive evolution of the global environmental virome. It has been proposed that of the over 10^{31} viral particles present in the current global microbial population, only about 2200 genomes from double stranded DNA viruses (dsDNA) and retroviruses are well established, as opposed to over 45,000 bacterial genomes identified to date (*Koonin*, 2015).

The viromes of the major cellular life domains are strikingly different in different life forms. Several forms of dsDNA viruses are present in both bacteria and archaea; but no viruses are known to be shared by eukaryotes with other life forms. Based on the classification of viruses introduced by Baltimore nearly 40 years ago, seven classes of viruses have now been identified in the environmental virome (*Baltimore*, 1971; *Koonin* et al., 2015). These include:

1) positive stranded RNA viruses, characterized by virions containing RNA of the same polarity as messenger RNA (mRNA),
2) negative stranded RNA viruses,
3) double stranded (ds)RNA viruses,
4) reverse transcribing viruses with positive stand RNA genome,
5) reverse transcribing virus with dsDNA genome,
6) single stranded DNA (ssDNA) viruses, and
7) dsDNA viruses.

Table 2: Relative distribution of DNA or RNA viral families infecting different life forms.

Life forms	Percent Viral Families	
	RNA viruses	DNA viruses
Eukaryotic:		
Fungi	80	20
Plants	86	14
Animals (invertebrates and vertebrates)	60	40
Protista-like and algae	57	43
Bacteria:		
Proteobacteria	28	79
Other bacteria	0	100
Archaea	0	100

In prokaryotic organisms most viruses possess dsDNA genomes, with a significant minority of ssDNA containing viruses, and only a very small number of RNA viruses. On the other hand, RNA viruses represent the majority of virome diversity in Eukaryotic life forms, with a significant, although a smaller number of ssDNA, and dsDNA viruses (*King* et al., 2011), as shown in Table 1. It should however be noted that genome sizes of RNA viruses are much more restricted compared to the DNA viruses which exhibit significantly more genome size diversity, often over 4 orders of magnitude higher genome size ranges (*Campillo-Balderas* et al., 2015). Of all the known viruses, about 9% have 2-3 segments and about 27% exhibit more than 4 segments. RNA viruses appear to be segmented more frequently (50%), than DNA viruses. Of all the segmented DNA and RNA viruses 60% infect plants, 21% infect other phylogenetically distant plants and other animals including 18% vertebrates and, about 1% infects bacteria (*Campillo-Balderas* et al., 2015). The relative distribution of RNA or DNA viruses in other eukaryotic life forms, (including, fungi, plants, animals), bacteria, archaea and Protista-like organisms (including algae) are shown in Table 2 (*Letunic* and *Bork*, 2007).

It is clear that the environmental virome is immense and virtually all life forms are inhabited by viruses. It appears that the genome size of dsDNA viruses have the highest diversity and surpass the far more restricted genome sizes of RNA and ssDNA viruses. However the RNA viruses exhibit a broader range of hosts in eukaryotic life forms, and infect a lot fewer bacterial species when compared to the DNA viruses.

Human virome

The human virome represents a diverse spectrum of viruses found in other prokaryotic and eukaryotic life forms which inhabit various systemic sites, skin or the mucosal surfaces in humans. The viral component of the human microbiome is also referred to as the human virome.

The repertoire of viruses in human virome includes traditional human viruses, bacterial viruses (bacteriophage) associated with bacterial species which constitute the human microbiome, and viruses infecting fungi and plants associated with the human microbiome.

Table 3: Listing of the viral families in the human virome identified to date.

RNA viruses:	Viral family:
dsRNA	Reoviridae
ssRNA(+)	Coronaviridae; Astroviridae; Caliciviridae; Flaviviridae; Picornaviridae; Togaviridae; Herpesviridae
ssRNA(-)	Rhabdoviridae; Feloviridae; Paramyxoviridae; Arenaviridae; Bunyaviridae; Orthumyxoviridae; Deltavirus.
DNA viruses:	Viral family:
dsDNA	Herpesviridae; Adenoviridae; Papillomaviridae; Polyomaviridae; Poxviridae
ssDNA	Anelloviridae; Parvoviridae
Retro viruses:	Viral family:
ssRNA (RT)	Retroviridae
dsDNA (RT)	Hepadnaviridae

Human viral infections can manifest as:

a) asymptomatic and acute self-limiting, or as fulminant and progressive infections,
b) chronic symptomatic or asymptomatic infections,
c) endogenous retroviral infections (retroviruses comprise over 8% of the human genome), and
d) still to be discovered viruses related to many diseases currently considered to being of unknown aetiology.

The virome of existing bacterial cell populations is immense, and bacteriophages affect human health, because they have major influence on bacterial cell population, structure, toxin production and virulence. For example, bacteriophage communities in the respiratory tract of healthy individuals have been found to be unique in each individual, representing a random and often transient sampling of the external environment. On the other hand, the bacteriophage communities in patients with pathologic states, such as cystic fibrosis, were found to be similar to other patients with cystic fibrosis. Such colonization is presumably facilitated by the similar underlying airway pathology (*Willner* et al., 2009).

Recent technologic advances to study viruses now include metagenomics analyses, employing comparison of genetic information from next generation sequencing of clinical samples to genomes of all known viruses, and techniques designed to translate viral genes into proteins and computationally search for similar protein sequences in newly discovered agents (*Wylie* et al., 2012). These advances have made it possible to characterize rapidly the existing viral genomes and identify new viral agents, as listed in Table 3.

Viruses have been recovered from all human mucosal surfaces, skin and many systemic sites. In recent studies of the DNA virome in several thousand human samples of blood, faecal samples, respiratory and gastrointestinal secretions, saliva, milk, urine, and other body fluids have yielded a wealth of information about the distribution, genetic diversity and pathogenicity of human viruses identified to date. They are listed in Table 4 (*Pride* et al., 2012; *Delwart*, 2016; *Moustafa* et al., 2017). In one study, sequences for 94 different viruses were identified in different human blood samples, including for 19 human DNA viruses, proviruses and RNA viruses (*Moustafa* et al., 2017).

Table 4: Tissue distribution of different viral families in different human body surfaces or secretions.

Virus	Genetic diversity	Tissue distribution	Pathogenic or commensal
Adenoviridae	High	GI, Resp., Urine, Blood	Both
Anelloviridae	High	Blood	Commensal/Pathogen?
Astrovirus	Medium	Blood, GI*	Both
Flaviviridae	Low	Blood	Both
Herpesviridae	Low	Blood, Skin	Both
Papillomaviridae	High	Skin	Both
Parvoviridae	High	Blood, GI*	Both
Picobirnaviridae	High	GI*	Commensal/Pathogen?
Picornaviridae	High	Blood, GI*, Resp.**, Skin	Both
Polyomaviridae	Medium	Blood, Resp.** Skin	Both

GI*: Gastrointestinal tract. Resp.**: respiratory tract.
Adapted from: Delwart, 2016; Moustafa et al., 2017; and Pride et al., 2012.

Introduction of metagenomics has indeed revolutionized the study of human viruses. For example until recently, only two polyomaviruses were known as human pathogens. However, about 13 other human polyomaviruses have now been identified and some of these have been implicated in several neurologic or renal disorders and other pathological states in immunocompromised hosts. Some papillomaviruses are also found in asymptomatic healthy skin, while a few can induce anal or cervical cancers (*Foulongne* et al., 2012; *DeCaprio* and *Garcea*, 2013). Another important recent observation relates to the identification of a rarely studied ssDNA virus family, the anelloviruses. It appears that they may be the most common human viral infections identified to date. These viruses have been detected in almost 100% of blood samples from human adults. These viruses are acquired shortly after birth and multiple strains have been identified in the same individual. The anelloviruses exhibit the highest level of genetic diversity known in any viral family to date. Most anelloviruses function as commensals and the infections are generally asymptomatic. Increasing amount of anelloviruses in immunosuppressed subjects have been associated with possible immunologically mediated chronic inflammation and expression of clinical symptoms. However, their role in the pathogenesis of the disease still remains to be clearly defined (*Spandole* et al., 2015).

The spectrum and the load of different viruses in the human body are truly immense. Over 10^{10}-10^{11} viral particles/gram of faeces representing bacteriophage predators of bacteria and archaea, as well as many other viruses have been identified to date in the humans. Furthermore, as pointed out earlier, endogenous retroviruses constitute at least 8% of human DNA (*Minot* et al., 2011, 2012). In spite of their overwhelming number (possibly billions),

as residents of the human body, only few hundred human viruses have been clearly associated with serious disease in man (*Delwart*, 2016).

Available epidemiologic evidence suggests that the acquisition of the environmental virome by mankind is an evolutionary (and possibly a requisite) adaptation, and the organisms are neither consistently commensal nor pathogenic. The outcome of the microbiome-host interaction at a cellular level is determined by a complex balance between the pathogenic potential of the virus and the immunologic status of the host at the time of their initial and subsequent interactions.

Maternally derived antibodies acquired transplacentally or via breast milk induce significant protection against disease against a variety of bacterial as well as viral infections in the neonatal period and early infancy. Acquisition of such viral infections during early years of human growth and development appear to play a very important role in the maturation and functional development of the host immune system at mucosal and systemic sites. Such early infections may also protect against more pathogenic infections with other viruses later in life. For example, there is evidence to suggest that pegivirus or GBV-C virus belonging to the larger family of viruses formerly known as of hepatitis G or HPgV, dengue and zika viruses can significantly improve the outcome of HIV infection. Pegivirus-c infection appears to be quite common and asymptomatic. It is estimated that to date over 700 million humans are infected with the virus. HIV infected subjects with pegivirus-c infection tend to live longer and experience healthier lives (*Linen* et al., 1996) Similarly, experimental animal studies have demonstrated that mouse norovirus, a common human commensal, restored immunologic functions that were previously altered by an induced germ-free status or the use of antibiotics (*Kernbauer* et al., 2014). In another recent study, it has been demonstrated that early enteric infections can influence the development of the gut and murine immune system in a very similar manner as observed with the bacterial microbiome on development of the human gut and maturation of the immune system (*Caldwell*, 2015).

Finally, the retroviral components of the human genome have also been adapted to enhance functional development as well as the survival of the mammalian host. It has been proposed that proteins expressed by such endogenous retroviruses can effectively bind or block receptors for other exogenous retroviruses (*Griffiths*, 2001; *Malfavon-Borja* and *Feschotte*, 2015). In an elegant study, endogenous retroviral envelope proteins were found to be responsible for the fusion of foetal trophoblast cells with the mammalian placental cells during the onset of pregnancy. Such fusion is essential for nutritional exchange between maternal and foetal system during gestation (*Mi* et al., 2000).

Based on the information summarized above, it is clear that we are just beginning to explore and define the complexity of human virome relative to the immense load of environmental virome. Our understanding of viruses has evolved as a result of specific infections associated with many eukaryotic life forms, beginning with the discovery of tobacco mosaic disease virus. However, the human virome, as we understand it today, is immensely larger than the few hundred disease producing viruses. While some viruses are prone to produce acute, chronic, fulminant or fatal infections in man (*Hulo* et al., 2011), the human virome predominantly represents a

large reservoir of symbiotic interactions between the viruses and the human and other life forms on earth. In fact, the human virome may provide significantly far more beneficial effects on human health, immunologic homeostasis, and maternal-neonatal interactions, than the development of serious or fatal disease.

LITERATURE

Baltimore, D.: Expression of animal virus genomes. Bacteriol. Rev. 35, 235-241 (1971).

Caldwell, K.: Expanding the role of the virome: Commensalism in the gut. J. Virol. 89, 1951-53 (2015).

Campillo-Balderas, J.A., Lazcano, A., and Becerra, A.: Viral genome size distribution does not correlate with the antiquity of the host lineages. Front. Ecol. Evol. 3, doi: 10.3389/fevo.2015.00143 (2015).

DeCaprio, J.A. and Garcea, R.L.: A cornucopia of human polyomaviruses. Nat. Rev. Microbiol. 11, 264-276 (2013).

Delwart, E.: The good viruses. The Scientist 115, 41-46 (2016)

Desnues, C., La Scola, B., Yutin, N., Fourhous, G., Robert, C., Azza, S., Jardot, P., Monteil, S., Campocasso, A., Koonin, E.V., and Raoult, D.: Provirophage and transpovirions as the diverse mobilome of giant viruses. Proc. Nat. Acad. Sci. USA 109, 18078-18083 (2012).

Forterre, P.: New hypotheses about the origins of viruses, prokaryotes and eukaryotes. In.: Frontiers of life (Eds.: Trân Thanh Vân, J. & K., Mounolou, J.C., Shneider, J., and McKay, C.). Edition Frontieres, Gif-sur-Yvette, France, 221-234 (1992).

Forterre, P.: Giant viruses: conflicts in revisiting the virus concept. Intervirology 53, 362-378 (2010).

Forterre, P. and Prangishvili, D.: The origin of viruses. Res. Microbiol. 160, 466-472 (2009).

Foulongne, V., Sauvage, V., Dereure, H.C., and Cheval, O.: Human skin microbiota: High diversity of DNA viruses identified on the human skin by high throughput sequencing. PLoS One 7, e38499 (2012).

Griffiths, D.J.: Endogenous retroviruses in the human genome sequence. Genome Biol. 2, reviews1017.1-reviews1017.5 (2001).

Hulo, C., deCastro, E., Masson, P., Bougueleret, L., Bairoch, A., Xenarios, I., and Le Mercier, P.: Viral zone: a knowledge resource to understand virus diversity. Nucleic Acids Res. D576-582 (2011).

Johnson. J. (Ed.): Phytopathologic classics, Vol. 7. American Phytopathologic Society, St. Paul, MN, USA, 27-30 (1942).

Kernbauer, E., Ding, Y., and Caldwell, K.: An enteric virus can replace the beneficial function of commensal bacteria. Nature 516, 94-98 (2014).

King, A.M.Q., Lefkowitz, E., Adams, M.J., and Carstens, B. (Eds.): Virus taxonomy; Ninth report of the international Committee on Taxonomy of Viruses. Elsevier Academic Press, Amsterdam (2011).

Koonin, E.V., Dolja, V.V., and Krupovic, M.: Origins and evolution of viruses of eukaryotes: the modularity. Virology 479-480, 2-25 (2015).

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P, Koonin, E., and Raoult, D.: The virophage as a unique parasite of the giant mimivirus. Nature 455, 100-104 (2008).

Letunic, L. and Bork, P.: Interactive tree of life (ITOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23, 127-128 (2007).

Linnen, J., Wages, J. Jr., Zhang-Keck, Z.Y., Fry, K.E., Krawczynski, K.Z., Alter, H., Koonin, E., Gallagher, M., Alter, M., Hadziyannis, S., Karayiannis, P., Fung, K., Nakatsuji, Y., Shih, J.W., Young, L., Piatak, M. Jr., Hoover, C., Fernandez, J.,

Chen, S., Zou, J.C., Morris, T., Hyams, K.C., Ismay, S., Lifson, J.D., Hess, G., Foung, S.K., Thomas, H., Bradley, D., Margolis, H., and Kim J.P.: Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. Science 271,505-508 (1996).

Malfavon-Borja, R. and Feschotte, C.: Fighting fire with fire: endogenous retrovirus envelopes as restriction factors. J. Virol. 89, 4047-4050 (2015).

Mi S, Lee X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.-Y., Edouard, P., Howes, S., Keith, J.C. Jr., and McCoy, J.M.: Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403, 785-789 (2000).

Minot, S., Grunberg, S., Wu, G., Lewis, J., and Bushman, F.D.: Hypervariable loci in the human gut virome. Proc. Natl. Acad. Sci. USA 109, 3962-3966 (2012).

Minot, S., Sinha, R., Chen, J, Li, H., Keilbaugh, S.A., Wu, G., Lewis, J., and Bushman, F.D.: Dynamic response of the human virome to dietary intervention. Genome Res. 21, 1616-1625 (2011).

Moustafa, A., Xie, C., Kirkness, E., Biggs, W., Wong, E., Turpas, Y., Bloom, K., Delwart, E., Nelson, K.E., Venter, J.C., and Telenti, A.: The blood DNA virome in 8,000 humans. PLOS Pathogens 13, e1006292 (2017).

Nasir, A. and Caetano-Anolles, G.: A phylogenomic data-driven exploration of viral origins and evolution. Sci. Adv. 1, e1500527 (2015).

Pride, D.T., Salzman, J., Haynes, M., Rohwer, F., Davis-Long, C., White, R.A. III, Loomer, P., Armitage, G.C., and Relman, D.A.: Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. ISME J. 6, 915-926 (2012)

Proctor, L.M.: Advances in the study of marine viruses. Microsc. Res. Tech. 37, 136-161(1997).

Raoult, D., Renesto, P., and Brouqui, P.: Laboratory infection of a technician by mimivirus. Antern. Med. 144, 702-703 (2006).

Spandole, S., Cimponeriu, D., Berca, L.M., and Mihăescu, G.: Human anelloviruses: an update of molecular, epidemiological and clinical aspects. Arch. Virol. 160, 893-908 (2105).

Stanley, W.M.: Isolation of a crystalline protein processing the properties of tobacco mosaic virus. Science 81, 644-655 (1935).

Willner, D., Furlan, M., Haynes, M., Schmieder, R., Angly, F.E., Silva, J., Tammadoni, S., Nosrat, B., Conrad, D., and Rohwer, F.: metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. PLoS One 4, e7370 (2009).

Woese, C. and Fox, G.: Phylogenetic structure of the bacteria domain in the primary kingdom. Proc. Natl. Acad. Sci. USA 74, 5088-5090 (1977).

Wylie, K.M., Weinstock, G.M., and Storch, G.A.: Emerging view of the human virome. Transl. Res. 160, 283-290 (2012).