Rinda Naresh, Notting Hill and Ealing High School

Discuss the origins and evolution of eukaryotic viruses.

Humans underestimate viruses. Smug in our anthropocentric view that they are small and simple, we have forgotten viruses have a beautifully complex history. As pathogenic, parasitic biological entities that are neither inert nor alive, all viruses consist of two elements: a nucleic acid and protein-coat (capsid)(and some have a lipid envelope). This nucleic acid may be DNA, as in adenoviruses, or RNA, as in influenza. The frequently mutating eukaryoviruses are extremely diverse and lack fossil evidence; their origins and evolution are therefore unresolved.

Viruses could be classed as living; they replicate (within hosts), their nucleic acid can adapt, demonstrated by evolving influenza viruses, and on infection, they become virocells[13](virion-producing cells) which follow the criteria for 'life'. However, viruses are deemed non-living as virions lack organelles and ribosomes so cannot metabolise, move, respire or replicate independently – they are obligate intracellular parasites.

Eukaryoviruses are extensively sequenced as they economically, agriculturally and medicinally impact humans. To our knowledge, eukaryotes are mostly susceptible to RNA viruses. This may be because chitinous fungal and cellulose plant cell walls, together with eukaryotic nuclear-envelopes (which prokaryotes lack) are barriers impeding viral entry. DNA viruses typically replicate within the host nucleus, whereas RNA viruses can replicate in the cytoplasm so need not penetrate nuclear-envelopes. This selective advantage over competitor DNA viruses justifies RNA eukaryoviruses' reproductive success.[30,22]

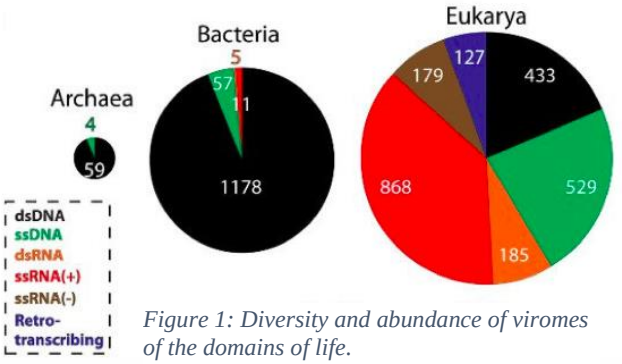

Figure 1: Diversity and abundance of viromes of the domains of life.

Why eukaryovirus origins are complex

Knowledge of viral diversity is limited as there are millions undiscovered. Furthermore, eukaryoviruses specifically are extremely diverse; eukaryotes host all viral classes.[Figure.1]Consequently, viral-origin theories are highly debated as different classes support different hypotheses.

The placement of eukaryoviruses relative to cells on the evolutionary timeline is also ambiguous. We assume all cells have a Last Universal Common Ancestor (LUCA) as their common use of ribosomes unites them genetically. The question remains whether viruses predated LUCA, or if LUCA-descendants were the first cells facing a new selection pressure – the viruses.

Evolutionary geneticist Koonin argues although viruses lack a common gene, many share viral hallmark genes(VHG). These are often highly conserved capsid-proteins – their presence in diverse viral families suggests an ancient common ancestor.[22]However, horizontal gene transfer (HGT) may explain VHG prevalence, as genetic material can pass 'horizontally' between organisms' genomes (rather than 'vertical' parent-offspring transmission). Furthermore, virologists Moreira and López-García argue convergent evolution justifies VHGs' pervasiveness.[26] The debate pursues as Koonin defends the near ubiquity of some VHGs would require multiple unlikely convergence events; he therefore believes shared VHGs imply pre-LUCA (but not monophyletic) origins.[17]

EVEs (endogenous viral elements), once thought to support the pre-LUCA theory, are fragmented proviral sequences integrated into non-viral germlines. When dated, these 'fossils' illustrate unexpected viral antiquity – the *Hepadnaviridae* family has EVEs dated at over 200 million years.[14]However, virologists believe LUCA is around 2 billion years old – viruses may be ancient but it is a non sequitur to assume they precede LUCA.

There are three leading eukaryovirus-origin theories: the virus-first, escape and reduction hypotheses.

The virus-first hypothesis

Conceived by d'Hérelle,[11]the virus-first hypothesis suggests RNA virus-like particles predated cells, originating from a primordial gene-pool as self-replicating entities. Through increasing complexity, these replicons could have facilitated abiogenesis. Alternatively, they may have emerged simultaneously and

coevolved with cells. However, many repudiate this theory, as it paradoxically rejects the definition of viruses as *intracellular* parasites. How could viruses – reliant on cellular replication-mechanisms – proliferate before cells themselves?
Most biologists accept the 'RNA-world' hypothesis where pre-LUCA life-forms adopted ribozymes as their genome and for catalysis. RNA's fragility meant cells gradually evolved to use DNA and proteins.[39]Therefore, Koonin argues RNA's instability could not have supported complex bacterial and archaeal sequences, but smaller, more error-prone genomes, as in viruses.[18]However, Koonin's argument of cellular absence for most of the RNA-world contradicts itself; without cells there could not be *bona fide* viruses. Therefore, modern RNA eukaryoviruses cannot come directly from the RNA world.(J.Iranzo, personal communication February 18th, 2018).

Conversely, modern viroids appear remnants of an RNA-world; their small genome is only 250-400 nucleotides, they do not encode proteins (the RNA-world lacked ribosomes) and some use ribozymes.[15]Perhaps virions are legacies of other RNA-world replicons, some of which evolved into viruses that parasitized cells after abiogenesis. This may explain the abundance of negative-sense RNA viruses whose genomes hosts cannot directly translate, but must first transcribe into a complementary 'sense' strand.[2]
Furthermore, Koonin uses RNA eukaryovirus *diversity* as evidence of an RNA-world origin. However, the validity of this depends on the supposed eukaryogenesis hypothesis.

As parasites, eukaryovirus-origins are tethered to eukaryote origins. Two models are the autogenous and symbiogenetic scenarios.[23]The former proposes a proto-eukaryote – the basic amitochondrial eukaryote – engulfed the ancestral mitochondrion. The latter advocates an archaea-bacteria fusion resulting in symbiosis and a hybrid (eukaryotic) cell.[32] The symbiogenetic model challenges a pre-LUCA origin; an archaea-bacteria fusion implies RNA eukaryoviruses would have evolved from RNA bacteriophages or archaeoviruses. However, RNA bacteriophage diversity is limited and there are no known RNA archaeoviruses, rendering subsequent RNA eukaryovirus diversity unlikely. Conversely, the autogenous model somewhat allows pre-LUCA origins, as proto-eukaryotes would likely be genetically dissimilar to prokaryotes, so perhaps allowed diversified viral evolution *within* their lineage.[22]
Despite its paradoxes, the virus-first hypothesis is currently most popular for RNA eukaryoviruses due to the supposed primordial RNA-world. Retroviruses are believed to have originated after them,[22]perhaps via the escape hypothesis.

<u>The escape hypothesis</u>
Retroviruses are unique to the eukaryotic virome,[30]suggesting they originated directly from eukaryotes. The escape hypothesis submits the emancipation of mobile genetic elements (retro-elements) such as plasmids or transposons from post-LUCA eukaryotic genomes. These components, lacking essential replication systems, may have become parasitic to proliferate genetically, culminating in retroviruses.[13]For instance, the CPEB3 ribozyme in a human intron is biochemically similar to human hepatitis delta viral (HDV) ribozyme sequences. As HDV solely infects humans and needs hepatitis B virus to replicate, HDV may have 'escaped' from the human transcriptome.[13]
Retroviral and eukaryotic phenotypes of a similar ilk support this theory. For instance, retrovirus replication is comparable to eukaryotic retrotransposon activity. In replication, retroviruses release viral reverse-transcriptase into hosts. This enzyme uses the virus' ssRNA to synthesise viral dsDNA which integrates into host DNA. The host' ribosomes translate the provirus and synthesise viral ssRNA and capsid-proteins. After assembly, progeny virions are exocytosed.[1]
Retrotransposons have a remarkably homologous lifecycle. Transposons are mobile DNA-segments, colloquially called 'jumping genes'. Eukaryotes transcribe retrotransposons into mRNA encoding the (eukaryotic)enzyme reverse-transcriptase which reverse-transcripts the mRNA into dsDNA retrotransposon-copies.[47]This integrates into another location in the eukaryotes' genome (homologous to viral dsDNA integration into host DNA). With random mutations, retro-elements such as retrotransposons could acquire a protein-coat and abandon cells as the first retrovirus – just as virocells exocytose virions.

However, the generally low homology of eukaryovirus gene-sequences to host sequences is peculiar.† If retroviruses 'escaped' from cells, VHGs (typically absent from cells) must be ancestrally cellular and once advantageous.[17]It then seems untenable viral emergence instigated a coordinated cellular VHG loss. A pre-LUCA escape theory is easier to defend; retrovirus-particles would have 'escaped' from a pre-LUCA entity unlikely to resemble modern cells, therefore may be *expected* to be genetically dissimilar to cells.[13]Consequently, by fusing the escape and virus-first hypotheses,[28]retroviruses may have emerged from proto-cells and coevolved with LUCA descendants. (A.Nasir, personal communication February 28th, 2018)

The reduction hypothesis

DNA viruses supposedly originated after RNA viruses and retroviruses on the evolutionary timeline, perhaps via the reduction hypothesis. This states unicellular organisms gradually lost their cellular elements and adopted parasitism. Mitochondria have their own circular DNA, suggesting they originated from an ingested bacterium which simplified in evolution, eventually becoming an endosymbiont.[46]Eukaryotic nucleocytoplasmic large dsDNA viruses (NCLDVs) may have evolved in a similar trajectory.

NCLDVs include the giant amoebal Mimivirus whose size catalysed a revolution in redefining the virosphere. Mimivirus is short for '**mi**micking **mi**crobe' – its genomic (1,200,000 base-pairs) and physical size (750 nanometre(nm) total diameter) seems microbial. Most virus diameters are less than 200nm; Porcine circovirus is 17nm with just 1,760 nucleotides.[42]NCLDVs encode proteins facilitating genome-replication and reduce reliance on host transcription-mechanisms.[19]Yet Mimivirus extraordinarily encodes proteins integral in *translation*, including vital aminoacyl-tRNA synthetases (aaRS) and six tRNAs.[34,35]Nonetheless, Mimivirus, like all viruses, cannot translate without host ribosomes and enzymes.

Furthermore, Tupanvirus (*Mimiviridae* family)was recently revealed (February 27th, 2018) to have the virosphere's most comprehensive translational machinery. With an unrivalled seventy tRNAs and the complete set of twenty aaRS among other translation factors, Tupanvirus only lacks ribosomes.[7]Therefore, it is unclear why some NCLDVs carry putative superfluous translation genes when they could exploit host apparatus.

A proposed history is autonomous cells symbiotically associated with each other, resulting in over-dependence of one organism by losing essential genetic material.[44]Thus a parasitic relationship transpired – the genetically smaller, acellular entity becomes a virus. This could elucidate Mimivirus' and Tupanvirus' extant tRNA-apparatus – they may be relics of an entire ancestral cellular translation-system. NCLDVs even echo cells in their vulnerability to attack from 'virophages' (virus-infecting subviral entities). In an ingenious reversal of evolution, it seems viruses streamlined their genome, abandoning all but essential genes to ruthlessly exploit increasing cellular complexity.

However, NCLDVs also possess VHGs, which challenges the reduction hypothesis. How could these eukaryoviruses, assumed to originate from *cells*, have pervasive viral genes mostly lacking cellular homologs? Moreira and López-García therefore reject this hypothesis and suggest HGT from hosts or parasitic bacteria attacking the same amoeba[25]explains NCLDVs' genomic complexity.

NCLDVs straddle the cellular world and virosphere; they can transcript but not completely translate, are larger than some microbes, but smaller than others, have VHGs but also hints of cellular translation-systems. To further obfuscate the situation, 86% of Mimivirus' genome has no detectable homologs in databases.[6]Nasir et al. therefore postulate NCLDVs represent a 'fourth supergroup'[27]alongside Bacteria, Archaea and Eukarya. Moreover, they argue rapid viral evolution means they may be known sequences mutated beyond recognition.[29]

Koonin, however, posits NCLDV origins links to Polintons, a group of eukaryotic transposons. Genomic sequencing has found Polintons encode two functional viral capsid-proteins; he therefore proposes they can be both transposons and *bona fide* viruses ('polintoviruses' are the proposed nomenclature).[22]The shared genes between Polintons and dsDNA eukaryoviruses imply a common ancestor – Koonin

† In a study, 90% of ORFs(open reading frames) in seven viral genomes has no matches in databases.[33]

suggests polintoviruses. Thus NCLDVs are not a fourth domain, nor evolved through genomic reduction, but originate from proposed polintoviruses which acquired cellular genes via HGT from eukaryotic hosts, as Moreira and López-García suggested.[21]

Each hypothesis faces a panoply of arguments and counterarguments due to eukaryovirus diversity – different viral classes support different hypotheses. Eukaryovirus origins is therefore a complex field, as is its subsequent evolutionary history.

Eukaryovirus evolution

The Red Queen Hypothesis comes from Lewis Carroll's Red Queen,[45]who explains to Alice, "Now, *here*, you see, it takes all the running you can do, just to keep in the same place." For billions of years, eukaryovirus and host have been relentlessly coevolving – simply for survival.

Viruses 'evolve' via recombination of viral genomes parasitising the same cell (genetic shift) or stochastic gene mutations (genetic drift);[40,41]their evolution is almost Darwinian. Viruses' short generation times mean they evolve rapidly, many at rates of 10^{-3}nucleotide-substitutions per site per year(s/n/y). Eukaryotic cells evolve a million times slower, at around 10^{-9}s/n/y.[8]However, these high mutation rates increase viral genetic load, therefore viruses compress their genomes often by gene-overlap.[3]Their frequent inimical mutations within limited phenotypes means they have metaphorically been described as "restless beats pacing a small cage."[10](Belshaw et al. 2008)

RNA viruses mutate particularly frequently as they lack proofreading systems in replication and RNA is less stable than DNA (RNA's hydroxyl groups make it more hydrolysis-prone).[38]This low replication fidelity together with large populations is propitious – a mutation allowing descendants to elude immune systems or antiviral drugs becomes more probable. Combined with natural selection and huge progenies, viruses are evolutionary professionals.

Eukaryoviruses must, however, find an optimum mutation rate. Too frequent mutations mean viruses cross the 'error threshold' and fitter genotypes are not maintained. Furthermore, rapidly-mutating viruses would be selected-against for their copious deleterious mutations. They could even suffer an 'error catastrophe' where excess mutations render the virion non-infectious. The antiviral drug Ribavirin therefore employs lethal mutagenesis, using mutagens to amplify mutation rates and annihilate eukaryovirus populations by inducing an error catastrophe. Viruses favouring lower mutation rates would therefore curtail their decimation of essential genes, yet rates too low reduce adaptive capabilities and the virus would fall behind host immune responses in the Red Queen's race.[3]

Evolution of eukaryovirus virulence is illustrated by the 'trade-off' hypothesis. It argues extreme parasitic virulence hinders long-term evolution as host death prohibits viral transmission.[9]For instance, moribund humans will transmit viruses less effectively than healthier humans continuing everyday interactions. Therefore, Nasir et al. propose viruses evolved latency to prolong transmission and mitigate negative corollaries of host lysis for propagation.[48]

Confucius once said, "a common man marvels at uncommon things; a wise man marvels at the commonplace."[36]Some everyday eukaryovirus transmission-adaptations are remarkable, albeit unnerving. Adenovirus prevents infected cells signalling viral attacks, whilst tobacco mosaic virus hijacks host DNA and enlarges plant plasmodesmata for faster transmission. The particularly macabre rabies virus accumulates in aggression-control regions of the host's brain. This induces agitation and biting – consequently infecting others. The victim's swallowing muscles are also paralysed, causing virus-riddled saliva to amass in the mouth – all to increase transmission.[4]

Therefore, studying eukaryovirus evolution is integral in epidemiology; HIV's rapid evolution quintessentially exemplifies its relevance. This retrovirus causes acquired immune deficiency syndrome(AIDS). The antiviral Zidovudine (AZT) inhibits reverse-transcriptase and prevents HIV replication. However, HIV's prolonged latency and frequent mutations means it can produce *every* possible point mutation everyday within a host, including the code for AZT-resistance.[12]Therefore, the drug can often only *delay,* rather than prevent, the onset of AIDS.

Daniel Racey and professor Stuart West consequently advocate teaching eukaryovirus evolution in medical schools.[43] By studying antiviral resistance, we could develop targeted multi-drug or reduced treatments. Mapping phylogenetic trees could allow anticipation of epidemics and resistance-emergence,

whilst examining eukaryovirus adaptations to immunisation could allow analysis of vaccine efficacy – particularly relevant in light of the noxious 2017-2018 flu season.

Humans have also influenced eukaryovirus evolution, exemplified by AIDS. The strain infamous for the 1980s epidemic is a mutation (HIV-1 subgroup-M)of a chimpanzee immunodeficiency virus. This strain affected millions globally, yet previous mutations infected fewer.‡ Viral microevolution directly affects macroevolution, therefore the rapid transmittance is due to the mutation itself, but also in the 1980s sex-trade explosion, escalating intravenous drug-users and international travel.[4]Through these methods, humans transmitted HIV so efficiently the virus could afford to evolve increased virulence without hindering its own survival, even as parasites dependent on living hosts.

We have not proved if eukaryoviruses evolved by the virus-first, escape or reduction hypothesis. Their morphological and reproductive diversity and the lack of an identified common viral gene suggests viruses are polyphyletic,[26]therefore, microbiologists can likely conflate all three hypotheses for different eukaryovirus classes. Nevertheless, increased molecular phylogeny and bioinformatics may elucidate the truth behind eukaryovirus origins and evolution. Along the way, discoveries of resistance and human-influence in the mercurial field of virology can drive medical advances in frontline antivirals.

(Word count: 2500)

Bibliography

Books

1. Fullick A., 2016. Edexcel AS/A Level Biology B Student Book 1. 1st ed. Pearson Education Limited; pp100-101.
2. Gupta, Y. (1998). *Selected Readings in Microbiology and Genetics*. Sarup & Sons. p204
3. Holmes EC. (2009). The Evolution and Emergence of RNA Viruses. Oxford University Press; 2009: p.26
4. Nesse, R. and Williams, G. (1996) *Why We Get Sick: The New Science of Darwinian Medicine*. 1st ed. United States: Vintage Books: p46.

Magazines

5. Hamilton, G. (2016). Infect and direct: How giant viruses could rewrite the story of life on Earth. New Scientist, [online] (3057). Available at: https://www.newscientist.com/article/2073733-how-giant-viruses-could-rewrite-the-story-of-life-on-earth/
6. Van Etten, J., 2011. Giant Viruses. American Scientist, 99: pp.304-311.

Journal Articles

7. Abrahão, J., Silva, L., Silva, L., Khalil, J., Rodrigues, R., & Arantes, T. et al. (2018). Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nature Communications*, *9*(1). [online] Available at: https://www.nature.com/articles/s41467-018-03168-1
8. Aiewsakun, P. and Katzourakis, A. (2015). Endogenous viruses: Connecting recent and ancient viral evolution. *Virology*, 479-480, pp.26-37. [online] Available at: https://www.sciencedirect.com/science/article/pii/S0042682215000549#bib73
9. Alizon, S., Hurford, A., Mideo, N. And Van Baalen, M. (2009). Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. Journal of Evolutionary Biology, 22: 245– 259. [online] Available at: https://pdfs.semanticscholar.org/e265/cb821aa08bc96f70e8c52ab7b95aff919949.pdf
10. Belshaw, R., Gardner, A., Rambaut, A., & Pybus, O. G. (2008). Pacing a small cage: mutation and RNA viruses. *Trends in Ecology & Evolution, 23*(4), 188-193. [pdf] Available at: tree.bio.ed.ac.uk/downloadpaper.php?id=379
11. Claverie J-M. (2006). Viruses take center stage in cellular evolution. *Genome Biology*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1779534/#B9
12. Cuevas, J. M., Geller, R., Garijo, R., López-Aldeguer, J., & Sanjuán, R. (2015). Extremely High Mutation Rate of HIV-1 In Vivo. *PLOS Biology, 13*(9). [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002251#pbio.1002251.ref001
13. Durzyńska J, Goździcka-Józefiak A. (2015). Viruses and cells intertwined since the dawn of evolution. *Virology Journal*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4609113/#CR45
14. Feschotte C, Gilbert C. (2012). Endogenous viruses: insights into viral evolution and impact on host biology. *Nature Reviews Genetics*. 2012;13(4):283-296. [online] Available at: https://www.sciencedirect.com/science/article/pii/S0042682215000549
15. Flores, R., Gago-Zachert, S., Serra, P., Sanjuán, R., & Elena, S. F. (2014). Viroids: Survivors from the RNA World? *Annual Review of Microbiology, 68*(1), 395-414. [online] Available at: https://www.annualreviews.org/doi/abs/10.1146/annurev-micro-091313-103416
16. Holmes, E. C. (2011). What Does Virus Evolution Tell Us about Virus Origins? *Journal of Virology*, *85*(11), 5247– 5251. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3094976/
17. Koonin EV, Senkevich TG, Dolja VV. (2006). The ancient Virus World and evolution of cells. Biology Direct. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1594570/.
18. Koonin EV, Senkevich TG, Dolja VV. (2009) Compelling reasons why viruses are relevant for the origin of cells. Nat Rev Microbiol. [online] Available at: https://www.nature.com/articles/nrmicro2108-c5
19. Koonin EV, Yutin N. (2010). Origin and Evolution of Eukaryotic Large Nucleo-Cytoplasmic DNA Viruses. *Intervirology*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2895762/
20. Koonin EV. (2009). On the Origin of Cells and Viruses: Primordial Virus World Scenario. *Annals of the New York Academy of*

‡ Other mutations include HIV-1 subgroup-O, which comparatively affected tens of thousands in Cameroon.[37]

Sciences. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3380365/
21. Koonin, E. V., Krupovic, M. and Yutin, N. (2015), Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. Ann. N.Y. Acad. Sci., 1341: 10– 24. [online] Available at: http://onlinelibrary.wiley.com/doi/10.1111/nyas.12728/full
22. Koonin, E., Dolja, V., & Krupovic, M. (2015). Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*, *479-480*, 2-25. [online] Available at: https://www.sciencedirect.com/science/article/pii/S0042682215000859?via%3Dihub#bib12
23. López-García, P., & Moreira, D. (2015). Open questions on the origin of eukaryotes. *Trends in Ecology & Evolution*, *30*(11), 697–708. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4640172/
24. Malik, S., Azem-e-Zahra, S., Kim, K., Caetano-Anollés, G., & Nasir, A. (2017). Do Viruses Exchange Genes across Superkingdoms of Life?. *Frontiers In Microbiology*, *8*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5671483/
25. Moreira, D. (2005). Comment on "The 1.2-Megabase Genome Sequence of Mimivirus." Science. [online] Available at: http://science.sciencemag.org/content/308/5725/1114.1
26. Moreira, D. & López-García, P. (2009). Ten reasons to exclude viruses from the tree of life. *Nature Rev. Microbiol.* Available at: https://edisciplinas.usp.br/pluginfile.php/2871445/mod_resource/content/2/2009_Ten%20reasons%20to%20exclude%20viruses%20from%20the%20tree%20of%20life.pdf
27. Nasir A, Kim KM, Caetano-Anolles G. (2012). Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evolutionary Biology*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3570343/
28. Nasir A, Kim KM, Caetano-Anollés G. (2012). Viral evolution: Primordial cellular origins and late adaptation to parasitism. *Mobile Genetic Elements*. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575434/
29. Nasir, A., & Caetano-Anollés, G. (2015). A phylogenomic data-driven exploration of viral origins and evolution. *Science Advances*, *1*(8), e1500527. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643759/
30. Nasir, A., Forterre, P., Kim, K., & Caetano-Anollés, G. (2014). The distribution and impact of viral lineages in domains of life. *Frontiers In Microbiology*, *5*. [online] Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2014.00194/full
31. Nasir, A., Kim, K., & Caetano-Anollés, G. (2017). Long-term evolution of viruses: A Janus-faced balance. *Bioessays*, *39*(8), 1700026. Acquired through email correspondence with G. Caetano-Anollés.
32. Poole, A. M., & Penny, D. (2006). Evaluating hypotheses for the origin of eukaryotes. *BioEssays, 29*(1), 74-84. [pdf] Available at: http://www.d.umn.edu/~pschoff/documents/PooleEndosymbiosis.pdf
33. Prangishvili D, Garrett R. (2004). Exceptionally diverse morphotypes and genomes of crenarchaeal hyperthermophilic viruses. *Biochem Soc Trans*. 32(2):204-208. Abstract only [online] Available at: https://www.ncbi.nlm.nih.gov/pubmed/15046572
34. Raoult D, Forterre P. (2008). Redefining viruses: lessons from Mimivirus. *Nature Reviews Microbiology*. Abstract only. [online] Available at: https://www.ncbi.nlm.nih.gov/pubmed/18311164
35. Raoult, D. (2004). The 1.2-Megabase Genome Sequence of Mimivirus. *Science*. [online] Available at: http://science.sciencemag.org/content/306/5700/1344.full

Websites

36. A quote by Confucius. *Goodreads.com*. [online] Available at: https://www.goodreads.com/quotes/507069-a-common-man-marvels-at-uncommon-things-a-wise-man. [Accessed 15 Feb. 2018].
37. BBC News, 2014. *Aids: Origin Of Pandemic 'Was 1920S Kinshasa'*. [online] Available at: http://www.bbc.co.uk/news/health-29442642. [Accessed Feb 15, 2018]
38. EMBL-EBIRNA. RNA. [online] Available at: https://www.ebi.ac.uk/training/online/course/biomacromolecular-structures-introduction-ebi-reso/rna [Accessed 15 Feb. 2018].
39. Exploring Life's Origins: Ribozymes & the RNA World, Exploringorigins.org, http://exploringorigins.org/ribozymes.html, [Accessed 12 Feb. 2018].
40. Historyofvaccines.org., (2018) *Viruses and Evolution | History of Vaccines*. [online] Available at: https://www.historyofvaccines.org/content/articles/viruses-and-evolution [Accessed 12 Feb. 2018].
41. Khan Academy. *Evolution of viruses*. [online] Available at: https://www.khanacademy.org/science/biology/biology-of-viruses/virus-biology/a/evolution-of-viruses [Accessed 15 Feb. 2018].
42. Philips R. How big are viruses?. *Bookbionumbersorg*. [online] Available at: http://book.bionumbers.org/how-big-are-viruses/. [Accessed 25 Jan. 2018].
43. Racey D. and West S. (2008) Evolution and the curriculum. Student BMJ. [online] Available at: http://student.bmj.com/student/view-article.html?id=sbmj0804148.
44. Wessner, D. R. (2010). The Origins of Viruses. Nature Education [online] Available at: https://www.nature.com/scitable/topicpage/the-origins-of-viruses-14398218 [Accessed 15 Feb. 2018].

Videos and Documentaries

45. Stearns S. Principles of Evolution, Ecology and Behavior. [Video Lecture] Available at: https://oyc.yale.edu/ecology-and-evolutionary-biology/eeb-122/lecture-20
46. *The Hidden Life Of The Cell*. 2012. BBC.
47. Useful Genetics (2015). *10M – Retrotransposons, pseudogenes and junk DNA*. [video] Available at: https://www.youtube.com/watch?v=ZIh1vQg1BO4 [Accessed 13 Feb. 2018].
48. WBLifeSciences, Nasir, A., Kim, K., & Caetano-Anollés, G. (2017). *Bioessays: Long-term evolution of viruses: A Janus-faced balance*. [video] Available at: https://www.youtube.com/watch?v=E1TOU1JDXo4

Images

49. Figure 1: Distribution of viral classes between the three domains of life. Edited from: Nasir, A., Forterre, P., Kim, K., & Caetano-Anollés, G. (2014). The distribution and impact of viral lineages in domains of life. *Frontiers In Microbiology*, *5*. [online] Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2014.00194/full