Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Virology

Viral mutation and substitution: units and levels

Robert Belshaw[1], Rafael Sanjuán[2] and Oliver G Pybus[1]

Viruses evolve within a hierarchy of organisational levels, from cells to host species. We discuss how these nested population structures complicate the meaning and interpretation of two apparently simple evolutionary concepts: mutation rate and substitution rate. We discuss the units in which these fundamental processes should be measured, and explore why, even for the same virus, mutation and substitution can occur at very different tempos at different biological levels. In addition, we explore the ability of whole genome evolutionary analyses to distinguish between natural selection and other population genetic processes. A better understanding of the complexities underlying the molecular evolution of viruses in natural populations is needed before accurate predictions of viral evolution can be made.

Addresses
[1] Department of Zoology, South Parks Road, University of Oxford, Oxford OX1 3PS, United Kingdom
[2] Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Departament de Genetica, and Unidad Mixta de Investigación en Genómica y Salud (CSISP-UV), Universitat de València, C/ Catedrático Agustín Escardino 9, Paterna 46980, València, Spain

Corresponding author: Belshaw, Robert (robert.belshaw@zoo.ox.ac.uk)

Current Opinion in Virology 2011, **1**:430–435

This review comes from a themed issue on
Virus evolution
Edited by Peter Simmonds and Esteban Domingo

Available online 15th September 2011

1879-6257/$ – see front matter

DOI 10.1016/j.coviro.2011.08.004

Introduction

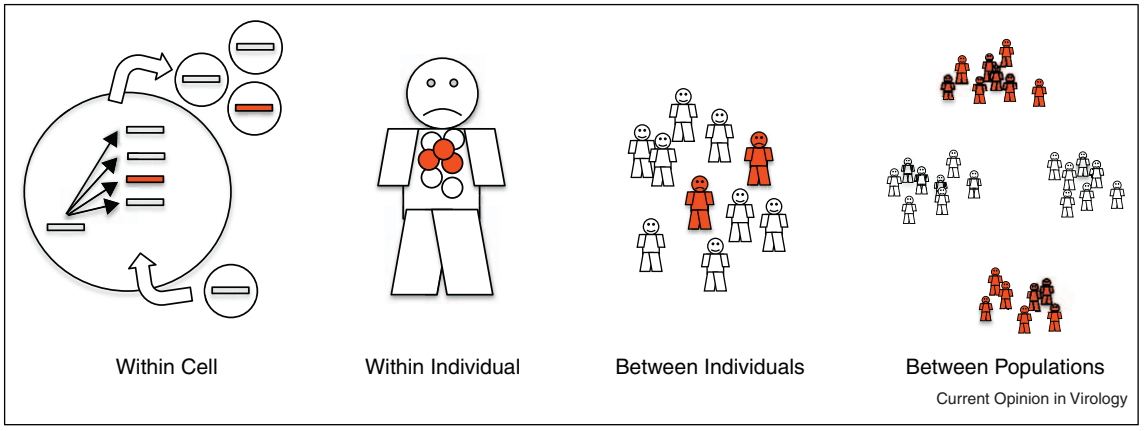
The rapid evolution and small genomes of viruses make them ideal model systems in which to study evolution [1,2••]. However, despite their apparent simplicity, viral populations in nature exhibit features that significantly complicate the study of evolutionary dynamics. Although the fundamental processes of *mutation* (the generation of genetic change) and *substitution* (the subsequent fixation of mutations in populations) are well described by standard population genetic theory, such models typically consider individual, homogeneous populations. In reality, viruses exist in a complex hierarchy of populations, from single infected cells to global pandemics, and can exhibit different evolutionary behaviors at each level (see Figure 1). Here we explore how the concepts of mutation, substitution and natural selection can be most usefully applied in the context of multi-level viral populations. We hope that a greater appreciation of these issues will lead to a more complete understanding of viral evolution.

Mutation: the problem of units

To begin, we consider a simple question with a surprisingly complex answer: how should rates of viral mutation be measured? The genome replication mechanisms of viruses are more varied than are those of higher organisms, leading to different possible definitions of their mutation rate. Viruses have two basic replication modes (Figure 2). In *binary* (geometric) replication, every progeny strand serves as a template for further replication and there are therefore repeated rounds of copying per cell infection cycle; whereas in *linear* (stamping machine) replication, the parental template is repeatedly copied but the progeny strands do not become templates until they invade new cells, hence there is only one copying round per cell infection cycle. The mutation rate of a virus can therefore be expressed as 'per round of strand copying' or 'per round of cell infection' [3•]. These may differ substantially if replication is predominantly binary and the number of progeny virions (burst size) is large.

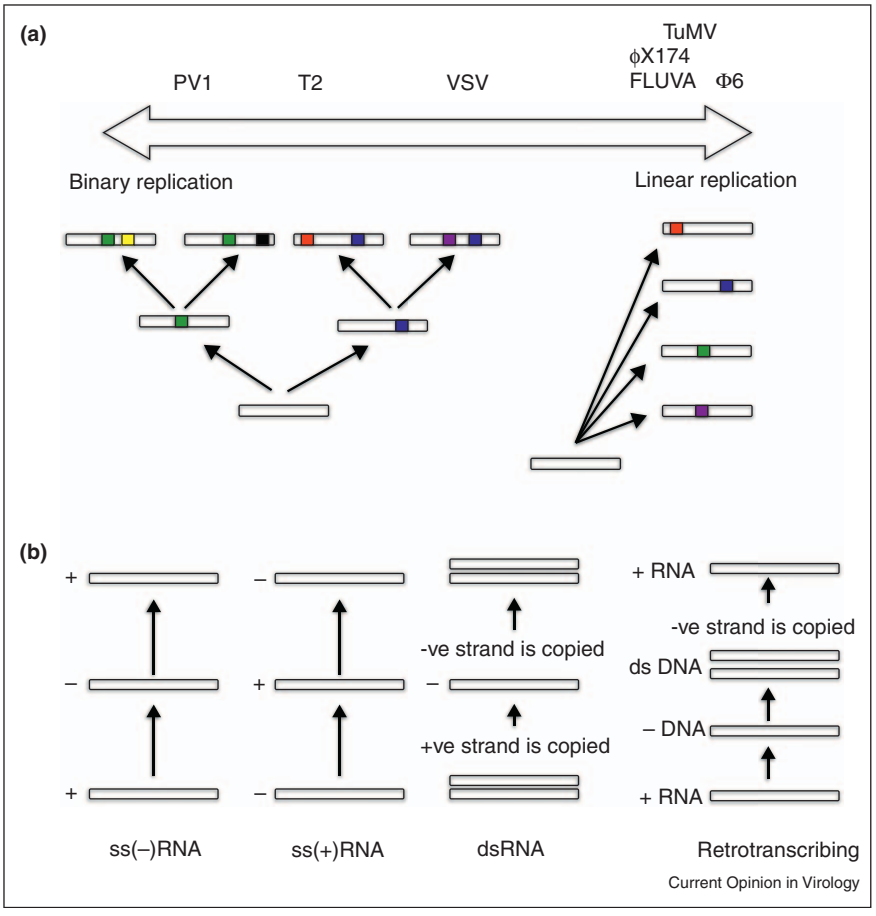
For most purposes we suggest using 'per round of cell infection'. This cycle (defined as the time between consecutive cell infection events) most closely corresponds to a viral generation within its host, enabling viral mutation rates to be compared with those of other organisms. As well as mutations caused by misincorporations during strand copying, this rate will include other sources of mutation, such as host-mediated editing by APOBECs [4] and ADARs [5], spontaneous damage of viral nucleic acid, and host replication of viruses inserted into the host genome. Using this definition also allows us to distinguish conceptually the mutation rate experienced by the virus from the fidelity of its replicase. For example, a retroviral genome can be copied by its own reverse transcriptase or the host's RNA and DNA polymerases, each of which exhibits different fidelities. Furthermore, the mutation rate experienced by the virus *in vivo* may differ substantially from the fidelity of its polymerase as measured outside the cell [6], most probably because the cellular environment determines fidelity. For instance, retrovirus dUTPases reduce mutation rates by preventing uracil incorporation into DNA [7]. Copying fidelity is not necessarily the same as the mutation rate per round of strand copying: although proofreading can occur concurrently with strand copying (3′ exonuclease activity) it might also occur afterwards (cellular mismatch DNA repair of DNA and retrotranscribing viruses).

Figure 1



A proposed hierarchy of viral populations illustrating virus genetic diversity at different biological levels. The units of replication at each level are, from left to right, (i) viral genomic templates, (ii) infected cells, (iii) infected hosts, (iv) infected host populations. At each level, reproductive units are coloured white or red to indicate viral genetic variation.

Figure 2



An illustration of viral replication mechanisms. **(a)** Binary replication (left) can result in a more rapid accumulation of mutations compared to linear replication (right). Each coloured box represents a mutation, occurring here at one per round of strand copying. Viruses often use a combination of these two mechanisms and, where this combination is known, we have placed viruses on a continuum between fully binary and fully linear [3•,49]. This illustration shows a double-stranded (ds) DNA virus, which uses semi-conservative replication. Single-stranded (ss) DNA viruses undergo an initial single round of strand copying to become double-stranded. **(b)** The replication of RNA viruses is more complicated: ssRNA viruses (−ve and +ve sense) can in theory employ either binary or linear replication. Whether dsRNA viruses employ more than two rounds of strand copying is not known.

Furthermore, if we can measure viral mutation rates in both units (per strand copying and per round of cell infection) we may be able to infer the dominant mode of replication – linear or binary – from the difference between the two values [3•]. This might help to clarify some viral replication mechanisms, for example whether or not the so-called late transcription of reoviruses and $\phi 6$ results in more viral genomes being formed [8].

Substitution: the problem of levels

Substitution rates are most commonly reported in units of 'substitutions per nucleotide site per year'. Although statistical methods for estimating such rates are well established [9], the values we obtain can differ according to the biological and temporal level at which they are studied.

The substitution rate of viruses that cause chronic infections may differ when measured at the 'within-host' (i.e. individual infection) and 'between-host' (i.e. epidemiological) levels (Figure 1). For example, substitution rates of HIV-1 tend to be both higher and more variable when estimated within-host than between-host [10]. This has also been noted for the envelope genes of the hepatitis C virus (HCV) [11]. This phenomenon, possible explanations of which are considered below, highlights the need to clarify which biological level is being investigated in a study. Substitution rates typically vary substantially among genome regions, so the genes from which substitution rates are derived should also be specified.

Estimated substitution rates are also time-dependent: rates estimated from sequences whose ancestry spans a few months or years can be hugely discrepant from those derived from sequences that have diverged over thousands or millions of years [12]. Time-dependence is most striking in the case of retroviruses, for which there is a contrast between fast short-term and apparently slow long-term substitution rates. For example, if the evolutionary history of SIVs (simian immunodeficiency viruses) is calibrated using the 10,000 year old isolation of one host population then we obtain a substitution rate that is >100 times slower than that estimated by following HIV or SIV evolution over 10 or 20 years [13]. Perhaps the most striking example comes from the observation of spumavirus elements in the genomes of sloths. Combined biogeographic and genomic data suggest that these elements are at least 40 million years old, yet they share recognisable amino acid homology with rapidly evolving contemporary spumaviruses [14]. Indeed, the time-dependency of substitution rates appears to be a general phenomenon observed in many organisms, and rates estimated over short timescales always appear to be higher [15]. We have previously proposed that selective constraint tightly limits the region of sequence space that can be explored by a viable, fast-changing virus, which therefore behaves like a 'restless beast pacing a small cage' [16].

A whole genome perspective

If we suppose that mutation rates are constant across viral genomes, then variation in substitution rates among genome regions should signify differences in selective pressures. For example by comparing the ratio of non-synonymous to synonymous changes (d_N/d_S), or by comparing the relative frequency of these classes among unfixed and fixed sites [17], many studies have shown that viral attachment genes (whose proteins are often targeted by antibodies) are positively selected, for example [18–20,21•]. Two factors complicate the whole genome comparison of substitution rates. First, mutation rates may not be uniform across the genome: HIV-1 replication errors *in vitro* are not randomly distributed [22] and it is argued that the genomic distribution of substitutions can be explained to some extent by the frequency of mutation [23]. DNA polymerases can generate mutational hotspots if replication fidelity falls during homopolymeric runs owing to template-primer misalignment [24]. Second, synonymous changes are often not neutral in RNA viruses, whose secondary structures may be functionally important [25]. For instance, conservation of base pairing in HIV stem-loops affects both nonsynonymous and synonymous sites [26]. Selective factors other than RNA structure also act on synonymous variation, such as codon bias [27]. This does not preclude the detection of selection using methods such as d_N/d_S, but does complicate their interpretation: they should be considered as measures of the relative selective pressures acting on synonymous and non-synonymous sites, not as absolute measures of selection on the latter.

If mutation rates are constant, uniform changes in substitution rate across viral genomes should, in most cases, signify either shifts in viral population size (such as the genetic bottleneck at transmission), or changes in viral generation time (defined here as the time between cell infections). For viruses, the most extreme changes in population size occur during transmission, when typically only one or a few virions are successfully transmitted [28–30]. Such bottlenecks can accelerate substitution rates at all sites, because more nearly neutral substitutions reach fixation when effective population sizes are small. Concerning changes in viral generation time, many viruses persist in host cells by ceasing or slowing replication. For example, HTLV-II replication by viral replicases ceases shortly after transmission, hence its substitution rate is many times slower when measured in endemically infected Amerindians, among whom transmission is vertical, than in injecting drug users, among whom transmission via needle sharing occurs frequently [31]. Some HIV lineages can remain as proviruses in quiescent lymphocytes for years without replication [32], and this may explain the high variance in substitution rate among

HIV lineages within individual infected hosts [33•]. Other viruses, such as herpes simplex virus, persist for decades in non-permissive nerve cells. At the between-host level, inactive viruses may persist in the environment. The duration of environmental persistence may be long (faecal/oral or waterborne transmitted viruses) or short (those transmitted by direct contact). All forms of inactive persistence or latency should manifest themselves as a reduced substitution rate across all genome regions, and are therefore distinguishable from rate variation owing to natural selection.

Viral natural selection at different biological levels

Not only can selection act separately on different genome regions, as discussed above, but it can also act differentially at different levels of biological organisation (Figure 1). Each level is characterised by a different reproductive unit (nucleic acid strand, infected cell, infected host or infected population). If there is viral genetic variation in reproductive success (i.e. transmission) among these units then the viral population at that level will be selected. However, selection may act variably at different levels, necessitating that each is studied separately.

It is estimated that 20–40% of random single-nucleotide mutations fully inactivate viral replication and that many of the remainder significantly reduce viral fitness [34]. Therefore, the substitution rate (per cell infection) will always be lower than the mutation rate. However, selection at the within-cell level can be offset by complementation, whereby a mutation that abrogates viral replication can be propagated because replicative ability is rescued by functional viral proteins from non-mutated genomes (co-infecting virions, or sibling/parental genomes). Several studies have demonstrated complementation in cell culture and in laboratory plant infections [35] but the phenomenon can also occur at the between-host level: complementation is thought to explain the spread and persistence among hosts of a premature stop codon in the dengue virus envelope protein [36]. Complementation requires co-infection or superinfection of the same cell with genetically different viruses, and therefore is expected to be less frequent in viruses with low within-host variation (those that result in acute infection or low viraemia). In addition, specific aspects of the cell infection cycle may preclude complementation: in poliovirus, for instance, the replicase molecule that copies a positive-sense genome will have been translated from that same strand (cis-acting), thus mutations knocking-out this function cannot be transmitted to other cells [37]. Further studies of co-infection and superinfection are required to better understand the significance of complementation in natural populations, but synergistic interactions certainly seem important in some plant viruses [38,39].

Selection among viruses within a host has been widely reported, especially for chronic infections like HIV, whose virions differ in their ability to evade host immunity. Within-host selection may also lead to variation in substitution rates between lineages, and there is evidence that distinct HIV lineages in different host tissues exhibit different substitution rates [40,41]. By contrast, there is much less evidence for selection at the between-host level, as indicated by the much greater concurrent circulation of HIV lineages at this scale [42•] (given the chance, all viable HIV can infect all non-CCR5Δ32 hosts).

Each new host represents a novel environment for the virus. This is exemplified by the reversion of immune escape mutations in new hosts, when such mutations carry a fitness cost in the absence of the original immune response [43]. This could explain why within-host substitution rates are sometimes higher than corresponding between-host rates: the fast within-host substitution rate of the HCV envelope genes possibly reflects repeated reversions of mutations selected for in previous hosts [11]. An alternate explanation for the observed differences in within-host and between-host substitution rates is that lineages within a host with a high substitution rate may have a lower probability of being transmitted – lineages that have undergone fewer rounds of replication (perhaps owing to latency) will have accumulated fewer host-specific adaptations and may therefore be more transmissible to (or more fit within) new hosts, as has been hypothesised for HIV-1 [33•].

In the same way that different host individuals represent variable environments for a virus, so can different host populations (or species). Host populations that carry epidemics that are particularly long-lived, or that have high rates of dispersal, are more likely to be the progenitors of new epidemics elsewhere. However, as noted above at other levels, selection is blind and can adapt viruses to their local environment at the cost of reduced onward transmission. It is notable, therefore, that HIV-1 has adapted differentially to the local HLA background of different human populations [44••]. Viruses that are vector-borne also exhibit host-specific evolution. The substitution rate of dengue virus in cell culture differs between human and mosquito cells [45]. Vector-borne RNA viruses tend to have reduced substitution rates [46] compared to non-vector-borne RNA viruses, which possibly reflects stronger negative selection on the former.

Conclusion

We hope that our discussion sheds some light on the difficulties inherent when applying fundamental evolutionary concepts to natural viral populations. For example, there is little evidence for a correlation between known viral mutation and substitution rates [47•]. Partly this will reflect the paucity of available data, but variation in viral life history might play a significant role: viruses

with shorter generation times, defined here as the time between transmissions, tend to have higher substitution rates [47•,48•]. We have also already mentioned the differences between within-host and between-host substitution rates observed in some viruses. In both these cases, we suggest that the relative roles of the underlying causal factors, for examples latency, transmission bottlenecks or differences in selection, can begin to be separated once large sets of whole viral genomes become available.

Acknowledgements

RB is partly funded by the Wellcome Trust. RS is funded by the Spanish MICINN. We also thank Nicola Chirico and John Drake for earlier discussions.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S: **Rapid evolution of RNA genomes**. *Science* 1982, **215**:1577-1585.

2. •• Holmes EC: *The Evolution and Emergence of RNA Viruses*. Oxford University Press; 2009.
Recent book that covers many of the outstanding questions in viral evolution.

3. • Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R: **Viral mutation rates**. *J Virol* 2010, **84**:9733-9748.
Our compilation and reanalysis of all published viral mutation rates that we could find. Data also available and updated at http://www.uv.es/rsanjuan/virmut

4. Harris RS, Liddament MT: **Retroviral restriction by APOBEC proteins**. *Nat Rev Immunol* 2004, **4**:868-877.

5. Samuel CE: **Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral**. *Virology* 2011, **411**:180-193.

6. Varela-Echavarria A, Garvey N, Preston BD, Dougherty JP: **Comparison of Moloney murine leukemia virus mutation rate with the fidelity of its reverse transcriptase *in vitro***. *J Biol Chem* 1992, **267**:24681-24688.

7. Lerner DL, Wagaman PC, Phillips TR, Prosperogarcia O, Henriksen SJ, Fox HS, Bloom FE, Elder JH: **Increased mutation frequency of feline immunodeficiency virus lacking functional deoxyuridine-triphosphatase**. *Proc Natl Acad Sci USA* 1995, **92**:7480-7484.

8. Frilander M, Gottlieb P, Strassman J, Bamford DH, Mindich L: **Dependence of minus-strand synthesis on complete genomic packaging in the double-stranded RNA bacteriophage ϕ6**. *J Virol* 1992, **66**:5013-5017.

9. Drummond A, Pybus OG, Rambaut A: **Inference of viral evolutionary rates from molecular sequences**. *Adv Parasitol* 2003, **54**:331-358.

10. Lemey P, Rambaut A, Pybus OG: **HIV evolutionary dynamics within and among hosts**. *AIDS Rev* 2006, **8**:125-140.

11. Gray RR, Parker J, Lemey P, Salemi M, Katzourakis A, Pybus OG: **The mode and tempo of hepatitis C virus evolution within and among hosts**. *BMC Evol Biol* 2011, **11**:131.

12. Holmes EC: **Molecular clocks and the puzzle of RNA virus origins**. *J Virol* 2003, **77**:3893-3897.

13. Worobey M, Telfer P, Souquiere S, Hunter M, Coleman CA, Metzger MJ, Reed P, Makuwa M, Hearn G, Honarvar S *et al.*: **Island Biogeography reveals the deep history of SIV**. *Science* 2010, **329**:1487.

14. Katzourakis A, Gifford RJ, Tristem M, Gilbert MTP, Pybus OG: **Macroevolution of complex retroviruses**. *Science* 2009, **325**:1512.

15. Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A: **Time-dependent rates of molecular evolution**. *Mol Ecol* 2011, **20**:3087-3101.

16. Belshaw R, Gardner A, Rambaut A, Pybus OG: **Pacing a small cage: mutation and RNA viruses**. *Trends Ecol Evol* 2008, **23**:188-193.

17. McDonald JH, Kreitman M: **Adaptive protein evolution at the ADH locus in Drosophila**. *Nature* 1991, **351**:652-654.

18. Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM: **Predicting the evolution of human influenza A**. *Science* 1999, **286**:1921-1925.

19. Sheridan I, Pybus OG, Holmes EC, Klenerman P: **High-resolution phylogenetic analysis of hepatitis C virus adaptation and its relationship to disease progression**. *J Virol* 2004, **78**:3447-3454.

20. Tee KK, Lam TT-Y, Chan YF, Bible JM, Kamarulzaman A, Tong CYW, Takebe Y, Pybus OG: **Evolutionary genetics of human enterovirus 71: origin, population dynamics, natural selection, and seasonal periodicity of the VP1 gene**. *J Virol* 2010, **84**:3339-3350.

21. • Bhatt S, Holmes EC, Pybus OG: **The genomic rate of molecular adaptation of the human influenza A virus**. *Mol Biol Evol* 2011, **28**:2443-2451.
Our recent example of using whole genome sequence data to infer selective pressures.

22. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH: **Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication**. *J Virol* 2010, **84**:9864-9878.

23. Ji JP, Loeb LA: **Fidelity of HIV-1 reverse transcriptase copying a hypervariable region of the HIV-1 env gene**. *Virology* 1994, **199**:323-330.

24. Kunkel TA: **The mutational specificity of DNA Polymerase-β during *in vitro* DNA synthesis: production of frameshift, base substitution, and deletion mutations**. *J Biol Chem* 1985, **260**:5787-5796.

25. Simmonds P, Tuplin A, Evans DJ: **Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: implications for virus evolution and host persistence**. *RNA* 2004, **10**:1337-1351.

26. Sanjuán R, Borderia AV: **Interplay between RNA structure and protein evolution in HIV-1**. *Mol Biol Evol* 2011, **28**:1333-1338.

27. Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S: **Virus attenuation by genome-scale changes in codon pair bias**. *Science* 2008, **320**:1784-1787.

28. Betancourt M, Fereres A, Fraile A, Garcia-Arenal F: **Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus**. *J Virol* 2008, **82**:12416-12421.

29. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun CX, Grayson T, Wang SY, Li H *et al.*: **Identification and characterisation of transmitted and early founder virus envelopes in primary HIV-1 infection**. *Proc Natl Acad Sci USA* 2008, **105**:7552-7557.

30. Sacristan S, Malpica JM, Fraile A, Garcia-Arenal F: **Estimation of population bottlenecks during systemic movement of Tobacco mosaic virus in tobacco plants**. *J Virol* 2003, **77**:9906-9911.

31. Salemi M, Lewis M, Egan JF, Hall WW, Desmyter L, Vandamme AM: **Different population dynamics of human T cell lymphotropic virus type II in intravenous drug users compared with endemically infected tribes**. *Proc Natl Acad Sci USA* 1999, **96**:13253-13258.

32. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C *et al.*: **Latent infection of CD4$^+$ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy**. *Nat Med* 1999, **5**:512-517.

33. • Pybus OG, Rambaut A: **Evolutionary analysis of the dynamics of viral infectious disease**. *Nat Rev Genet* 2009, **10**:540-550.
Includes a discussion of possible reasons why within-host and between-host substitution rates might differ.

34. Sanjuán R: **Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies**. *Philos Trans R Soc Lond [Biol]* 2010, **365**:1975-1982.

35. Holmes EC: **The evolutionary genetics of emerging viruses**. *Annu Rev Ecol Evol Syst* 2009, **40**:353-372.

36. Aaskov J, Buzacott K, Thu HM, Lowry K, Holmes EC: **Long-term transmission of defective RNA viruses in humans and *Aedes* mosquitoes**. *Science* 2006, **311**:236-238.

37. Novak JE, Kirkegaard K: **Coupling between genome translation and replication in an RNA virus**. *Genes Dev* 1994, **8**:1726-1737.

38. Latham JR, Wilson AK: **Transcomplementation and synergism in plants: implications for viral transgenes?** *Mol Plant Pathol* 2008, **9**:85-103.

39. Froissart R, Michalakis Y, Blanc S: **Helper component-transcomplementation in the vector transmission of plant viruses**. *Phytopathology* 2002, **92**:576-579.

40. Itescu S, Simonelli PF, Winchester RJ, Ginsberg HS: **Human immunodeficiiency virus type 1 strains in the lungs of infected individuals evolve independently from those in peripheral blood and are highly conserved in the C-terminal region of the envelope V3 loop**. *Proc Natl Acad Sci USA* 1994, **91**:11378-11382.

41. Salemi M, Lamers SL, Yu S, de Oliveira T, Fitch WM, McGrath MS: **Phylodynamic analysis of human immunodeficiency virus type 1 in distinct brain compartments provides a model for the neuropathogenesis of AIDS**. *J Virol* 2005, **79**:11343-11352.

42. • Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC: **Unifying the epidemiological and evolutionary dynamics of pathogens**. *Science* 2004, **303**:327-332.
Discusses selection in viral populations. Also introduces the term 'phylodynamics', combining epidemiological and phylogenetic approaches.

43. Friedrich TC, Dodds EJ, Yant LJ, Vojnov L, Rudersdorf R, Cullen C, Evans DT, Desrosiers RC, Mothe BR, Sidney J *et al.*: **Reversion of CTL escape-variant immunodeficiency viruses *in vivo***. *Nat Med* 2004, **10**:275-281.

44. •• Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M, Gatanaga H, Fujiwara M, Hachiya A, Koizumi H *et al.*: **Adaptation of HIV-1 to human leukocyte antigen class I**. *Nature* 2009, **458**:641-645.
Major study showing adaptation of a virus to its host.

45. Vasilakis N, Deardorff ER, Kenney JL, Rossi SL, Hanley KA, Weaver SC: **Mosquitoes put the brake on arbovirus evolution: experimental evolution reveals slower mutation accumulation in mosquito than vertebrate cells**. *PLoS Pathogens* 2009, **5**: doi:10.1371/journal.ppat.1000467 e1000467.

46. Jenkins GM, Rambaut A, Pybus OG, Holmes EC: **Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis**. *J Mol Evol* 2002, **54**:156-165.

47. • Hanada K, Suzuki Y, Gojobori T: **A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes**. *Mol Biol Evol* 2004, **21**:1074-1080.
Comparative study that attempts to explain the variation in viral substitution rates. Includes a comparison with the corresponding mutation rates.

48. • Firth C, Kitchen A, Shapiro B, Suchard MA, Holmes EC, Rambaut A: **Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses**. *Mol Biol Evol* 2010, **27**:2038-2051.
Detailed calculation of the substitution rate in double-stranded DNA viruses, suggesting they are higher than previously thought.

49. Martínez F, Sardanyés J, Elena SF, Daròs JA: **Dynamics of a plant RNA virus intracellular accumulation: stamping machine vs geometric replicatiion**. *Genetics* 2011, **188**:637-646.