Influence of viral genome properties on polymerase fidelity

Gabriel Dupré, Romain Volmer

▶ **To cite this version:**

Gabriel Dupré, Romain Volmer. Influence of viral genome properties on polymerase fidelity. Trends in Genetics, 2023, 39 (1), pp.9-14. 10.1016/j.tig.2022.10.008 . hal-04003118

HAL Id: hal-04003118
https://hal.inrae.fr/hal-04003118

Submitted on 23 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Trends in
Genetics

Opinion

Influence of viral genome properties on polymerase fidelity

Gabriel Dupré [1,*] and Romain Volmer [1,*]

The first step of viral evolution takes place during genome replication via the error-prone viral polymerase. Among the mutants that arise through this process, only a few well-adapted variants will be selected by natural selection, renewing the viral genome population. Viral polymerase-mediated errors are thought to occur stochastically. However, accumulating evidence suggests that viral polymerase-mediated mutations are heterogeneously distributed throughout the viral genome. Here, we review work that supports this concept and provides mechanistic insights into how specific features of the viral genome could modulate viral polymerase-mediated errors. A predisposition to accumulate viral polymerase-mediated errors at specific loci in the viral genome may guide evolution to specific pathways, thus opening new directions of research to better understand viral evolutionary dynamics.

Highlights

Evolution can be defined as a two-step process. The first step is the accumulation of mutations in the viral genome thanks to error-prone viral polymerases. The second step is the selection of descendants that harbour the genetic pool conferring the highest fitness. Thus, specific parts of the viral genome are under intense selection pressure and often appear as the most plastic loci because they accumulate more mutations than other parts of the genome.

However, various studies indicate that the base composition and the structure of the genome sequence can themselves modulate viral polymerase fidelity, resulting in the non-homogenous accumulation of mutation events along the viral genome.

A better understanding of the influence of the template genome properties on the viral polymerase fidelity may improve our understanding of the evolutionary trajectories taken by viruses.

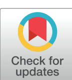
Viral polymerase errors during genome copy and the filter of selection

It is straightforward to define an evolving virus by its capacity to create numerous genome variants in its progeny. The cloud of viral genome descendants, initially described in the case of RNA viruses because of the high error rate of viral RNA polymerases, is defined as viral quasispecies [1,2]. This property has profound consequences on viral evolutionary dynamics, regulating viral fitness, immune escape, drug resistance, virulence, host adaptation, and 'social' interactions between variants, detailed in the following reviews [3–5]. Specific errors or mutations are fixed by natural selection and then passed down, after the action of multiple filters or bottlenecks that reduce the size of the mutant spectrum population [6]. Parts of the genome encoding for more plastic protein domains are more tolerant to mutations, favouring mutation accumulation in these loci. Moreover, parts of the viral particle are subjected to strong immune response-mediated bottlenecks, which tend to select immune-escaping viruses harbouring new antigenic sites. Consequently, these antigenic sites are associated with high nucleotide variability. Viral glycoproteins (GPs), which constitute major targets of the immune response, evolve more rapidly than other viral proteins [7–10]. On the contrary, some internal proteins are less tolerant to mutations and are more evolutionary stable, probably because they are hidden from neutralising antibodies and are composed of domains that are structurally and/or functionally more constrained [8,9,11–13].

However, accumulating evidence also suggest that the viral genome template itself can modulate mutation events. Host enzymes have been shown to modify genomes of several RNA and DNA viruses [14–23]. These host editing enzymes target specific genome locations enriched in particular base composition or RNA/DNA structures, revealing the influence of the viral genome properties on host-dependent editing [24–26]. Host editing enzymes will not be further described here. By contrast, this review will focus on recent findings indicating that the viral RNA genome template properties also modulate viral polymerase fidelity, causing nonrandom mutation events, which ultimately could influence viral evolution.

[1]Ecole nationale vétérinaire de Toulouse, ENVT, INRAE, IHAP, UMR 1225, Université de Toulouse, 23 Chemin des Capelles, Toulouse 31076, France

*Correspondence:
gabriel.dupre@envt.fr (G. Dupré) and
romain.volmer@envt.fr (R. Volmer).



© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Impact of the genomic structure and sequence on viral polymerase fidelity

Beta-coronaviruses evolve through different genetic mechanisms, including recombination events or, more precisely, polymerase jumps followed by template switching. Several **recombination breakpoints** (see Glossary) and crossing overs have been described on both structural and non-structural genes, mostly on the spike encoding gene [27,28]. Recent chemical mappings revealed the presence of extensive RNA structures in the genome of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that colocalise with hotspots of template switching, characterised by insertion and deletion (indel) scars in the viral genome, suggesting that the RNA structures present in the viral genome could promote polymerase template switching [11]. In line with this observation, the brome mosaic virus, which belongs to the *Bromoviridae* family, undergoes non-homologous recombination events that seem to be dependent on both the primary sequence of the viral genome and its folding [29]. These findings support early reports suggesting that RNA structures present on influenza A viral ribonucleoproteins could trigger polymerase-jumps and promote the production of defective genome segments [30].

In addition, viral polymerase-mediated nucleotide mutations, consisting of small indels, have also been shown to be influenced by the viral genome properties. The cucumber mosaic virus of the *Bromoviridae* family is known to harbour parasitic RNA called satellite RNA (satRNA), which are small well-structured elements with specific functions depending on nucleic acid folding. Their replication and persistence are strictly dependent on the helper virus [31]. Indel events have been detected predominantly in very structured regions on the satRNA of cucumber mosaic virus [32]. Moreover, viral polymerase-mediated RNA editing, leading to the generation of additional mRNA and new open reading frames (ORFs) by the insertion of one or several nontemplate nucleotides, has been shown to depend on the RNA template properties. In the case of the Turnip mosaic virus, belonging to the *Potyviridae* family, a conserved sequence consisting of a stretch of six adenosines preceded by one guanosine induces insertion events via polymerase stuttering [33]. Similarly, Ebola virus (EBOV) of the *Filoviridae* family produces its GP after the insertion of additional adenosine at a precise location during transcription of the GP-ORF. When this genetic mechanism is absent, only truncated forms of the GP are produced (secreted and small secreted GPs: sGP and ssGP), which are involved in neutralising antibodies. Cis-acting elements corresponding to a stretch of seven/eight consecutive uridines and flanking sequences present on the genome control RNA editing [34–36]. Similar findings were observed for viruses of the *Paramyxoviridae* family involving different editing sequences [34,37,38]. Importantly, RNA editing also occurs during viral genome replication, modifying around 10% of the Zaire EBOV newly synthesised genomes [39] and so could potentially influence viral evolution. In the case of influenza A virus, variants harbouring two to 13 supplementary U/C insertions at the poly-U stretch in the 5′ noncoding region were identified, suggesting a stuttering mechanism comparable with the RNA editing in filoviruses and paramyxoviruses [40].

All the aforementioned studies point to an influence of the viral RNA genome structure or base composition on the viral polymerase fidelity. However, it is warranted to note that these observations were made in the context of selective pressure on the functions of proteins encoded by the viral genome. As a consequence, mutation events associated with a deleterious effect on protein function are counter selected. To strictly assess the effect of genomic properties on mutation events, grasp their diversity, and progress in the understanding of the mechanisms, the use of experimental systems allowing selective pressure to be bypassed is recommended.

Driving a wedge between genome properties and protein function to grasp the full diversity of mutation events

Using a shuttle vector experimental system, freeing the sequence of interest from selection pressure, the most external part of the HIV viral GP, corresponding to the V1–V5 loops, was found to be less

Glossary

Nucleotide insertions: nucleoside insertions or deletions are genetic mutations where one or several nucleotides/nucleosides are inserted/deleted from a viral DNA or an RNA genome.

Nucleotide substitution: genetic mutation where a single nucleotide or nucleoside base is changed from a viral DNA or an RNA genome, respectively. This modification can have no effect (synonymous mutation) or modify the protein sequence (non-synonymous mutation).

Recombination breakpoints: location in the viral genome where the genetic material has been exchanged from the parental sequence to another sequence ('donor') during genome replication.

subject to spontaneous mutations than another part of the genome, suggesting the existence of hypo-mutative sequences or genetic elements reducing mutation rate [41]. Conversely, this same study found that the C2 region and the Rev responsive element (RRE) of the HIV-1 *env* gene were significantly more subjected to substitutions and recombination or template-switching events than other parts of the genome [41]. Interestingly, extended and complex RNA secondary structures have been described in the C2 region and RRE, suggesting that the topology of the viral genome could impact viral polymerase fidelity [42–44].

In the case of the human hepatitis C virus (HCV), deep sequencing of HCV genome fragments obtained via a replicon system, freed from selection pressure on protein function, indicates that the HCV genome exhibits non-homogenous mutation rates along its genome. Substitution biases were found to be dependent on the genome base composition and can largely differ between two adjacent genomic sites [45]. However, whether the RNA structures may play a role in HCV polymerase-mediated mutation events remains unclear [45].

Avian influenza viruses of the H5 and H7 subtypes, initially low pathogenic avian influenza viruses (LPAIV), can become highly pathogenic (HPAIV) through the acquisition of a multi-basic cleavage site (MBCS) motif on the hemagglutinin (HA), ultimately responsible for increased pathogenicity in poultry (Figure 1) [46–48]. **Nucleotide substitutions**, insertions (codon duplication), and/or polymerase jump-template switching have been shown to be responsible for MBCS acquisition [48–53]. Work using a minigenome system, again freeing the sequence of interest from selection pressure, showed that the HA nucleotide sequence had a profound effect on nucleotide insertion rates, a mechanism possibly leading to the acquisition of a MBCS [52]. To our knowledge, this study represents the first demonstration of a possible link between nucleotide composition and/or RNA topology of the viral genome and a probability to acquire a phenotypic change, here acquisition of a HA MBCS, thus linking genome properties to an evolutionary pathway [52]. Interestingly, several studies highlighted the presence of a conserved predicted RNA stem-loop encompassing the HA cleavage site coding sequence, that may play a role in LPAIV-to-HPAIV transition through polymerase slippage/stuttering and polymerase jump-template switching [48,50–57], even though the detailed mechanisms involved in these modifications remain to be identified (Figure 1).

Concluding remarks

Here, we compiled evidence that specific features of the viral genome, including its folding and the presence of specific nucleotide motifs, directly influence the rate and type of viral RNA-dependent RNA polymerase mutations. Interestingly, a recent unreviewed report also showed that the propensity of eukaryotic DNA-dependent RNA polymerase II to cause **nucleotide insertions** depended on the DNA template nucleotide composition [58], thus extending the concept from viruses to eukaryotes. To the best of our knowledge, there is currently no strong evidence that the nucleotide template properties modulate the fidelity of DNA-dependent DNA polymerases. The only studies found to refer to possible mutation hotspots in DNA viruses at precise genomic motifs were obtained under strong selective pressures [59,60]. A straightforward and probable explanation is the proofreading activity of DNA-dependent DNA polymerases that will considerably reduce/correct mutation load [1]. Interestingly, however, a recent study performed on *Arabidopsis thaliana* demonstrated that epigenetic modifications and physical genomic features are linked, resulting in mutation rate biases, which impact genome evolution [61]. Whether this phenomenon can also be translatable to viral genomes is unknown and constitutes an exciting route for future research.

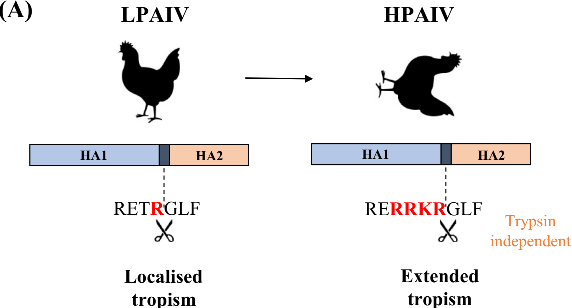
Strong evidence suggests that RNA structures in viral genomes modulate viral polymerase-induced mutation events [41,52,57]. These observations provide an impetus to further investigate

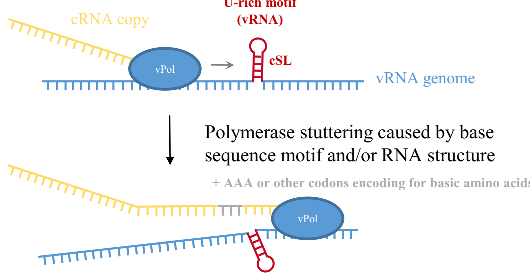
Outstanding questions

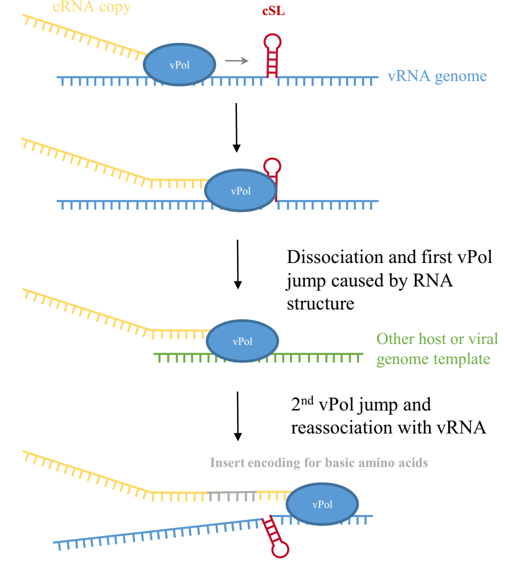
What are the genetic elements that can influence mutation rates?

What are the mechanisms causing increased viral polymerase-mediated mutation events?

What is the extent of this evolutionary force in both RNA and DNA viruses?







Trends in Genetics

Figure 1. Possible contribution of the hemagglutinin (HA) genome properties to avian influenza A virus evolution from low pathogenic avian influenza viruses (LPAIV) to highly pathogenic avian influenza viruses (HPAIV). (A) Schematic representation of the influenza A HA protein of LPAIV and HPAIV. The HA of LPAIV is proteolytically matured by tryspin-like proteases located in the respiratory and digestive tract of birds, which determine its tropism for these tissues. HPAIV originate from LPAIV that evolve through the acquisition of a multibasic HA cleavage site motif via nucleotide substitutions, insertions, and/or non-homologous recombination. This enables the HA protein to be matured by ubiquitous intracellular furin-like proteases, leading to an extended viral tropism and high pathogenicity in poultry. (B) Model of the nucleotide insertion events leading to LPAIV–HPAIV transition. In the HA1/HA2 encoding region, a specific U-rich stretch and/or the conserved stem-loop structure (cSL) could promote polymerase stuttering during genome replication, leading to the insertions of one or multiple bases in the neo-synthetised strand. (C) Model of recombination or polymerase jump/template switching leading to LPAIV-to-HPAIV transition. In this model, the viral polymerase copies the viral template until it reaches a cSL at the HA1/HA2 boundary region. The cSL, with the action of possible cofactors, induces the dissociation of the polymerase from the template. The polymerase then lands on another template, which can be a host RNA or another viral RNA, and copies a portion of this new template. Eventually, the polymerase goes back to the initial viral template and elongates the remaining part of the viral genome template.

this association. However, it is warranted to note that despite the power of predicting algorithms available [62], *in vitro*, *in cellulo*, and *in virion* chemical structural probing is required to verify the authenticity of predicted RNA structures [44,62–68]. In addition, because there is currently no method capable of assessing RNA structures in the context of the genome replication process, and more precisely during template copy in the catalytic site of viral polymerases, the interpretation of a suggested link between RNA structures and mutability has to remain prudent.

Studies using replicon-based systems, which free the genome from selection pressure on the protein functions, provided evidence in favour of a possible link between mutation events and RNA structure constraints and/or nucleotide motifs [41,45,52]. Replicon-based experiments are powerful in multiple ways, including their relative simplicity of setup, their flexibility to assess various combinations of factors (different viral polymerases, nucleotide sequences), and the absence of infectious-related risks. However, replicon systems do not allow accurate recapitulation of the exponential production of new genome variants by multiple infection cycles, due to their 'non-infectious' nature. Thus, the design of viable infectious viruses, in which part of their genome can be freed from selective pressures, could help to more accurately investigate how the genetic environment could impact mutation rates. A better understanding of the mechanisms

modulating viral polymerase-mediated mutations and their potential consequences on viral evolution represents an interesting challenge in viral evolutionary biology (see Outstanding questions).

Declaration of interests

The authors declare no competing interests.

References

1. Sanjuán, R. and Domingo-Calap, P. (2016) Mechanisms of viral mutation. *Cell. Mol. Life Sci.* 73, 4433–4448
2. Sanjuán, R. *et al.* (2010) Viral mutation rates. *J. Virol.* 84, 9733–9748
3. Andino, R. and Domingo, E. (2015) Viral quasispecies. *Virology* 479–480, 46–51
4. Domingo, E. *et al.* (2021) Historical perspective on the discovery of the quasispecies concept. *Annu. Rev. Virol.* 8, 51–72
5. Sanjuán, R. (2021) The social life of viruses. *Annu. Rev. Virol.* 8, 183–199
6. McCrone, J.T. and Lauring, A.S. (2018) Genetic bottlenecks in intraspecies virus transmission. *Curr. Opin. Virol.* 28, 20
7. Lee, J.M. *et al.* (2018) Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc. Natl. Acad. Sci. U. S. A.* 115, E8276–E8285
8. Thyagarajan, B. and Bloom, J.D. (2014) The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife* 2014, e03300
9. Lyons, D.M. and Lauring, A.S. (2018) Mutation and epistasis in influenza virus evolution. *Viruses* 10, 407
10. Heaton, N.S. *et al.* (2013) Genome-wide mutagenesis of influenza virus reveals unique plasticity of the hemagglutinin and NS1 proteins. *Proc. Natl. Acad. Sci. U. S. A.* 110, 20248–20253
11. Chrisman, B.S. *et al.* (2021) Indels in SARS-CoV-2 occur at template-switching hotspots. *BioData Min.* 14, 1–16
12. Munis, A.M. *et al.* (2021) Genomic diversity of SARS-CoV-2 in Oxford during United Kingdom's first national lockdown. *Sci. Rep.* 11, 21484
13. Harvey, W.T. *et al.* (2021) SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* 19, 409–424
14. Harris, R. and Dudley, J. (2015) APOBECs and virus restriction. *Virology* 479–480, 131–145
15. Hirose, Y. *et al.* (2018) Within-host variations of human papillomavirus reveal APOBEC signature mutagenesis in the viral genome. *J. Virol.* 92, e00017-18
16. Suspène, R. *et al.* (2005) Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8321–8326
17. Vartanian, J. *et al.* (2008) Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* 320, 230–233
18. Sadler, H.A. *et al.* (2010) APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *J. Virol.* 84, 7396–7404
19. Suspène, R. *et al.* (2011) Double-stranded RNA adenosine deaminase ADAR-1-induced hypermutated genomes among inactivated seasonal influenza and live attenuated measles virus vaccines. *J. Virol.* 85, 2458
20. Martínez, I. and Melero, J. (2002) A model for the generation of multiple A to G transitions in the human respiratory syncytial virus genome: predicted RNA secondary structures as substrates for adenosine deaminases that act on RNA. *J. Gen. Virol.* 83, 1445–1455
21. Cuevas, J. *et al.* (2016) Human norovirus hyper-mutation revealed by ultra-deep sequencing. *Infect. Genet. Evol.* 41, 233–239
22. Atkinson, N.J. *et al.* (2014) The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids Res.* 42, 4527–4545
23. Cattaneo, R. *et al.* (1988) Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell* 55, 255–265
24. Holtz, C. *et al.* (2013) APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. *Nucleic Acids Res.* 41, 6139–6148
25. Sharma, S. and Baysal, B.E. (2017) Stem-loop structure preference for site-specific RNA editing by APOBEC3A and APOBEC3G. *PeerJ* 5, e4136
26. Kuttan, A. and Bass, B.L. (2012) Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci. U. S. A.* 109, E3295–E3304
27. Rowe, C. *et al.* (1997) Generation of coronavirus spike deletion variants by high-frequency recombination at regions of predicted RNA secondary structure. *J. Virol.* 71, 6183–6190
28. Forni, D. *et al.* (2020) Recombination and positive selection differentially shaped the diversity of betacoronavirus subgenera. *Viruses* 12, 1313
29. Figlerowicz, M. (2000) Role of RNA structure in non-homologous recombination between genomic molecules of brome mosaic virus. *Nucleic Acids Res.* 28, 1714
30. Jennings, P.A. *et al.* (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* 34, 619–627
31. Palukaitis, P. (2016) Satellite RNAs and satellite viruses. *Mol. Plant-Microbe Interact.* 29, 181–186
32. Pita, J.S. *et al.* (2007) Environment determines fidelity for an RNA virus replicase. *J. Virol.* 81, 9072–9077
33. Olspert, A. *et al.* (2015) Transcriptional slippage in the positive-sense RNA virus family *Potyviridae*. *EMBO Rep.* 16, 995–1004
34. Le Mercier, P. and Kolakofsky, D. (2019) Bipartite promoters and RNA editing of paramyxoviruses and filoviruses. *RNA* 25, 279–285
35. Volchkova, V.A. *et al.* (2011) Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. *J. Infect. Dis.* 204, S941–S946
36. Mehedi, M. *et al.* (2013) Ebola virus RNA editing depends on the primary editing site sequence and an upstream secondary structure. *PLoS Pathog.* 9, e1003677
37. Welch, B.D. *et al.* (2012) Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16672–16677
38. Kolakofsky, D. (2016) Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: a review. *Virology* 498, 94–98
39. Shabman, R.S. *et al.* (2014) Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. *MBio* 5, e02011
40. Zheng, H. *et al.* (1999) Influenza A virus RNA polymerase has the ability to stutter at the polyadenylation site of a viral RNA template during RNA replication. *J. Virol.* 73, 5240
41. Geller, R. *et al.* (2015) The external domains of the HIV-1 envelope are a mutational cold spot. *Nat. Commun.* 6, 1–9
42. Simon-Loriere, E. *et al.* (2009) Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus. *PLoS Pathog.* 5, e1000418
43. Simon-Loriere, E. *et al.* (2010) RNA structures facilitate recombination-mediated gene swapping in HIV-1. *J. Virol.* 84, 12675–12682
44. Tomezsko, P.J. *et al.* (2020) Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* 582, 438–442
45. Geller, R. *et al.* (2016) Highly heterogeneous mutation rates in the hepatitis C virus genome. *Nat. Microbiol.* 1, 1–7
46. Swayne, D.E. and Pantin-Jackwood, M. (2006) Pathogenicity of avian influenza viruses in poultry. *Dev. Biol. (Basel)* 124, 61–67
47. Lee, D.-H. *et al.* (2020) Pathobiological origins and evolutionary history of highly pathogenic avian influenza viruses. *Cold Spring Harb. Perspect. Med.* 11, a038679
48. Abdelwhab, E.-S.M. *et al.* (2013) Genetic changes that accompanied shifts of low pathogenic avian influenza viruses toward higher pathogenicity in poultry. *Virulence* 4, 441–452

49. Perdue, M.L. and Suarez, D.L. (2000) Structural features of the avian influenza virus hemagglutinin that influence virulence. *Vet. Microbiol.* 74, 77–86
50. Perdue, M.L. *et al.* (1997) Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. *Virus Res.* 49, 173–186
51. García, M. *et al.* (1996) Heterogeneity in the haemagglutinin gene and emergence of the highly pathogenic phenotype among recent H5N2 avian influenza viruses from Mexico. *J. Gen. Virol.* 77, 1493–1504
52. Nao, N. *et al.* (2017) Genetic predisposition to acquire a polybasic cleavage site for highly pathogenic avian influenza virus hemagglutinin. *MBio* 8, 1–15
53. Gultyaev, A.P. *et al.* (2021) Insertions of codons encoding basic amino acids in H7 hemagglutinins of influenza A viruses occur by recombination with RNA at hotspots near snoRNA binding sites. *RNA* 27, 123–132
54. Dietze, K. *et al.* (2018) From low to high pathogenicity—characterization of H7N7 avian influenza viruses in two epidemiologically linked outbreaks. *Transbound. Emerg. Dis.* 65, 1576–1587
55. Beerens, N. *et al.* (2020) Emergence and selection of a highly pathogenic avian influenza H7N3 virus. *J. Virol.* 94, e01818-19
56. Gultyaev, A.P. *et al.* (2019) Conserved structural RNA domains in regions coding for cleavage site motifs in hemagglutinin genes of influenza viruses. *Virus Evol.* 5, 1–11
57. Dupré, G. *et al.* (2021) Phylodynamic study of the conserved RNA structure encompassing the hemagglutinin cleavage site encoding region of H5 and H7 low pathogenic avian influenza viruses. *Virus Evol.* 7, 1–16
58. Chung, C. *et al.* (2022) The fidelity of transcription in human cells. *bioRxiv* Published online May 10, 2022. https://doi.org/10.1101/2022.05.10.491385
59. Wilson, J.B. *et al.* (1986) A frameshift at a mutational hotspot in the polyoma virus early region generates two new proteins that define T-antigen functional domains. *Cell* 44, 477–487
60. Sasadeusz, J.J. *et al.* (1997) Homopolymer mutational hot spots mediate herpes simplex virus resistance to acyclovir. *J. Virol.* 71, 3872–3878
61. Monroe, J.G. *et al.* (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 602, 101–105
62. Zhao, Q. *et al.* (2021) Review of machine learning methods for RNA secondary structure prediction. *PLoS Comput. Biol.* 17, e1009291
63. Smyth, R.P. *et al.* (2018) RNA structure-a neglected puppet master for the evolution of virus and host immunity. *Front. Immunol.* 9, 2097
64. Gilmer, O. *et al.* (2021) Chemical and enzymatic probing of viral RNAs: from infancy to maturity and beyond. *Viruses* 13, 1894
65. Lim, C.S. and Brown, C.M. (2018) Know your enemy: successful bioinformatic approaches to predict functional RNA structures in viral RNAs. *Front. Microbiol.* 8, 2582
66. Merino, E.J. *et al.* (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* 127, 4223–4231
67. Rausch, J.W. *et al.* (2018) Probing the structures of viral RNA regulatory elements with SHAPE and related methodologies. *Front. Microbiol.* 8, 2634
68. Dadonaite, B. *et al.* (2019) The structure of the influenza A virus genome. *Nat. Microbiol.* 4, 1781–1789