



Synonymous nucleotide changes drive papillomavirus evolution

Kelly M. King^a, Esha Vikram Rajadhyaksha^{a,b}, Isabelle G. Tobey^c,
Koenraad Van Doorslaer^{a,c,d,*}

^a School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ, USA

^b Department of Physiology and Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

^c Cancer Biology Graduate Interdisciplinary Program, University of Arizona, Tucson, AZ, USA

^d The BIOS Institute, The Department of Immunobiology, Genetics Graduate Interdisciplinary Program, UA Cancer Center, University of Arizona Tucson, Arizona, USA

ABSTRACT

Papillomaviruses have been evolving alongside their hosts for at least 450 million years. This review will discuss some of the insights gained into the evolution of this diverse family of viruses. Papillomavirus evolution is constrained by pervasive purifying selection to maximize viral fitness. Yet these viruses need to adapt to changes in their environment, e.g., the host immune system. It has long been known that these viruses evolved a codon usage that doesn't match the infected host. Here we discuss how papillomavirus genomes evolve by acquiring synonymous changes that allow the virus to avoid detection by the host innate immune system without changing the encoded proteins and associated fitness loss. We discuss the implications of studying viral evolution, lifecycle, and cancer progression.

1. The family *Papillomaviridae*

The family *Papillomaviridae* contains small, non-enveloped viruses with double-stranded genomes ranging between 5748 bp (*Sparus aurata* papillomavirus type 1; SaPV1) to 8607 bp (canine papillomavirus type 1; CPV1). Papillomaviruses have been isolated from fish, reptiles, birds, and mammals. As the most sampled host, 441 unique papillomaviruses have been isolated from humans. The Papillomavirus Episteme (PaVE; [1,2]) contains 664 genetically distinct papillomavirus types, classified into two subfamilies, >50 genera and >130 species [1–3] (Fig. 1A). Classification of papillomaviruses is based on pairwise nucleotide sequence identity across the L1 open reading frame [3]. Distinct papillomavirus types share less than 90% sequence identity across the L1 open reading frame [3]. The original criteria distinguishing genera stated: “Most types within a PV genus show less than 60% sequence identity to types of other genera based on global multiple sequence or pairwise alignments of the L1 genes. Nevertheless, the suggested percentage identities that define PV genera have to be taken as general, but not absolute criteria as curation is necessary” [4,5]. Practically, papillomavirus genera and species are primarily delineated by visual inspection of phylogenetic trees derived from concatenated E1, E2, and L1 nucleotide sequences [3]. Efforts are underway to refine the papillomavirus classification scheme and update viral species names in line with recent changes by the International Committee for the Taxonomy of Viruses (ICTV) [6].

2. The viral lifecycle

Members of the *Papillomaviridae* family primarily infect mucosal and keratinized epithelia. Papillomavirus virions are non-enveloped with a capsid made up of 360 copies (arranged as 72 pentamers) of the major capsid protein, L1, and up to 72 molecules of the L2 minor capsid protein [7]. Each capsid packages a single copy of the viral circular dsDNA associated with core histone proteins [8,9].

Temporal expression of the viral genome is associated with tissue differentiation and is tightly regulated at the level of transcription and RNA processing [10–12]. A typical mammalian papillomavirus encodes six to nine proteins (Fig. 1B). However, only 4 proteins (E1, E2, L1, and L2) can be identified in all papillomaviruses sequenced to date [13,14]. The viral helicase, E1, is essential for the replication and amplification of the viral chromosome in the nucleus of infected cells [15]. The E2 protein regulates viral transcription, initiation of DNA replication, and partitioning of the viral genome [16]. The additional viral proteins likely play essential yet supporting roles in the viral lifecycle.

The E6 and E7 proteins are critical in creating a cellular milieu that supports the viral lifecycle [17,18]. Specifically, since papillomaviruses replicate in differentiating tissues, E6 and E7 uncouple viral replication from cellular differentiation. In a subset of papillomaviruses, the E6 and E7 proteins are the primary oncogenes [19,20]. The E1/E4 gene product is typically translated from a spliced mRNA fusing approximately the first four E1 codons to the E4 ORF, present in an alternative reading frame to the E2 ORF. The role of the E4 protein is not entirely clear, but

* Corresponding author. School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ, USA.

E-mail address: vandoorslaer@arizona.edu (K. Van Doorslaer).

<https://doi.org/10.1016/j.tvr.2022.200248>

Received 29 June 2022; Received in revised form 11 October 2022; Accepted 12 October 2022

Available online 17 October 2022

2666-6790/© 2022 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

E4 serves as a biomarker of active virus infection [21]. E8^E2 is a similarly alternatively spliced gene product. The E8 exon is embedded within E1 and utilizes the same splice acceptor site as E1^E4 mRNA, generating mRNA for the E8^E2 protein. E8^E2 inhibits viral replication and gene expression [22]. These alternatively spliced genes appear to be unique to papillomaviruses that infect mammals. A subset of viral mRNA encodes a short, hydrophobic, transmembrane protein, E5 or E10. E5 proteins are typically encoded in the 3'-end of the early coding region [23,24]. The E10 proteins are located in this region without an E6 gene [14]. Most of these 'early genes' are transcribed as polycistronic mRNAs using the same promoter and identical poly-A termination signal.

Interestingly, the E8^E2 gene product is derived from a unique promoter in E1, suggesting that E8^E2 can be separately regulated throughout the viral lifecycle. As the infected cell differentiates, the viral genome is amplified, and the viral capsid proteins, L1 and L2, are expressed. The major capsid protein L1 contributes to the main structural component of the viral capsid [25]. In contrast, the minor capsid protein L2 is critical for viral assembly and subsequent infection of a new host cell [26–28].

3. Host and virus evolution are linked, but it is complicated

Most papillomavirus infections cause benign, mostly unapparent, persistent infections in their hosts. In addition, papillomaviruses are highly host-restricted and cause abortive infections in non-host species [29,30]. The observation that papillomaviruses cause highly

host-restricted and primarily benign infections led to the hypothesis of virus-host co-speciation or "host-linked evolution" [13,31].

Traditionally, co-speciation theory implies that the phylogenetic tree of the viruses and the phylogenetic tree of infected hosts should closely mirror each other (Fahrenholz's rule; [32]).

However, with the increasing number of viruses being identified in different animal hosts, it is clear that strict co-speciation between virus and host is not supported by the data [24,30,33–44]. Indeed, it seems unlikely that a single ancestral papillomavirus infected an ancestral host and both virus and host co-speciated, generating the extant diversity in papillomavirus genome sequences. Several violations of a strict-virus-host co-speciation can be seen in the phylogenetic tree in Fig. 1A. A similar but interactive and zoomable phylogenetic tree can be accessed through the PaVE website (https://pave.niaid.nih.gov/analyze/phylogenetic_tree; accessed June 2022). For example, human papillomaviruses can be found in five different genera (*Alphapapillomavirus*, *Betapapillomavirus*, *Gammapapillomavirus*, *Mupapillomavirus*, and *Nupapillomavirus*) in different locations throughout the phylogenetic tree. Under strict co-speciation, all human viruses would be expected to cluster together in a single clade.

Another striking example is two papillomaviruses isolated from snakes. BcosPV1 [43] and MsPV1 [44] were isolated from a Boa constrictor and Diamond Python, respectively. Based on co-speciation and the location of other viruses isolated from reptile hosts, we expect these viruses to cluster near the root of the phylogenetic tree (parrot silhouette in Fig. 1A). However, these snake viruses cluster among

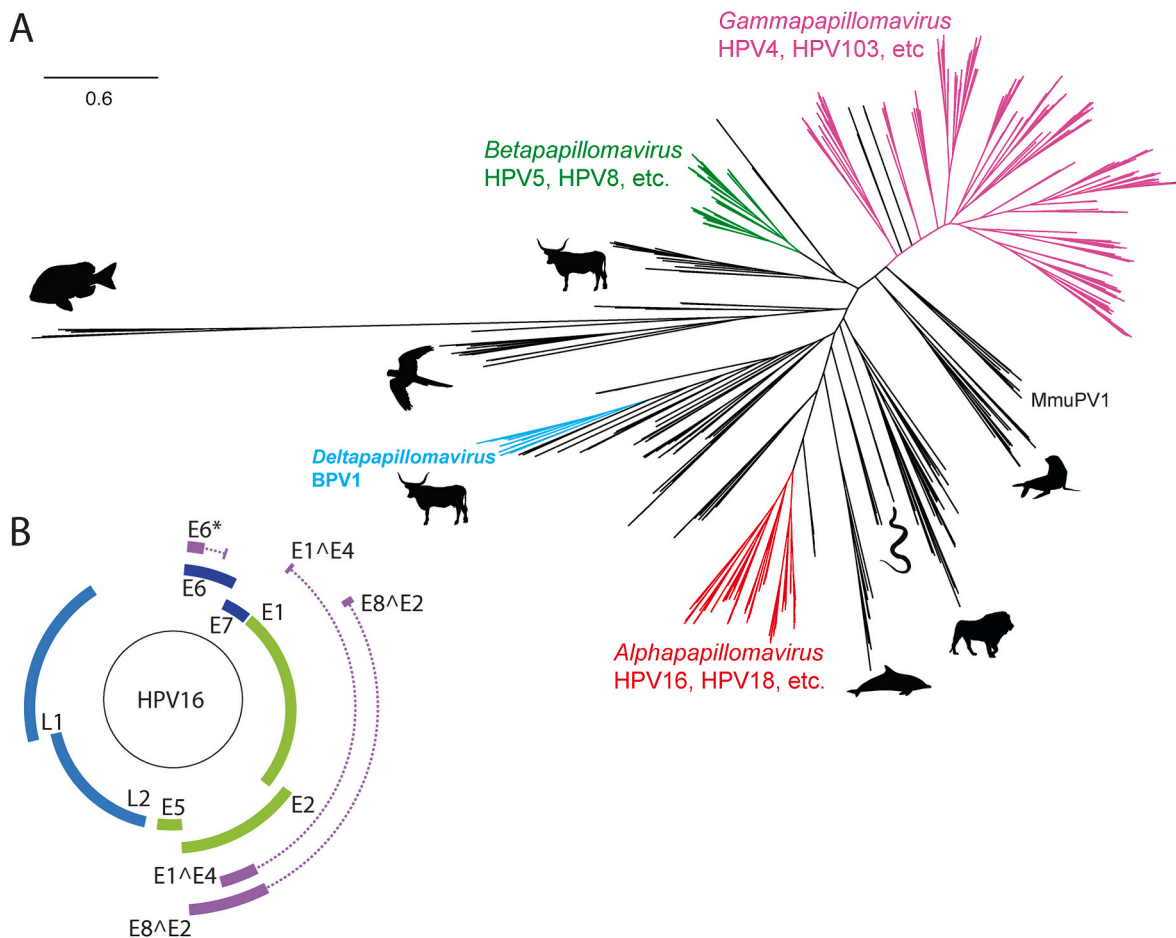


Fig. 1. (A) Papillomavirus phylogenetic tree. The protein sequence coding for E1, E2, and L1 for all 667 papillomaviruses currently on PaVE ([1,2]; June 2022) were downloaded and aligned. A concatenated phylogeny was constructed using default parameters for FastTree. Members of the genera *Alphapapillomavirus*, *Betapapillomavirus*, *Gammapapillomavirus*, and *Deltapapillomavirus* are indicated [3–5]. The different animal silhouettes provide a broad overview of animal tropism. (B) Diagram of a typical human papillomavirus genome belonging to the genus *Alphapapillomavirus*. The outer boxes indicate the protein-coding open reading frames. Dotted lines represent intron sequences. Diagram was downloaded from PaVE and edited in Adobe Illustrator.

mammalian viruses (snake silhouette).

The observed discrepancy between the host and virus trees has been explained by evolutionary events such as horizontal gene transfer, recombination, and virus duplication (e.g., following ecological niche adaptation) [30,45–50]. While recombination likely played a role at specific moments throughout viral evolution, its overall importance during viral evolution is likely limited (reviewed in Ref. [13]). Since papillomaviruses are highly species-specific, horizontal gene transfer likely did not play any role in the evolution of the *Papillomaviridae* [48]. It is important to point out that more complete and unbiased sequencing of animal papillomavirus diversity will likely help to clarify the evolutionary history of the *Papillomaviridae*.

An updated version of the co-evolution theory suggests that viruses follow the evolution of host resources [51]. Under this model, a generalist ancestral virus was able to thrive in multiple tissue niches. During the late Devonian, some infected ancestral animal adapted to life on land. Specific events in the evolution of these ancestral hosts (e.g., the presence/absence of fur, the evolution of sweat glands, etc.) created new ecological niches for papillomaviruses to adapt [31]. Resource-driven evolution would have driven niche adaptation allowing for specialization and virus duplication. This process would explain the diversification of papillomaviruses into four or five groups of increasingly specialized viruses (reflected in the 4–5 major clades of the phylogenetic tree; Fig. 1A) [33]. Following these niche adaptation events, virus-host co-speciation was the main driver of virus diversification. Throughout the co-speciation process, the availability of new niches would, in turn, drive viral radiation, followed by further co-speciation. A similar evolutionary process was hypothesized for the related *Polyomaviridae* [52].

Therefore, strict co-speciation between papillomaviruses and their hosts does not explain the observed diversity. However, repeated events of niche sorting followed by virus-host-linked speciation were likely a vital determinant of the papillomavirus evolutionary history [48,53].

4. Minimal evidence of traditional Darwinian selection

Evolutionary selection is traditionally quantified by calculating the ratio of the number of non-synonymous substitutions per non-synonymous site to the number of synonymous substitutions per synonymous site (dN/dS). The dN/dS ratio is a useful measure of the strength and mode of natural selection acting on protein-coding genes. A dN/dS ratio close to 1 indicates neutral evolution. An excess of non-synonymous changes vs. synonymous changes (i.e., dN/dS > 1) argues that the codon (and the encoding amino acid) is under Darwinian or diversifying (positive) selection. Diversifying selection is often seen as the result of the genetic conflict. For example, in host-parasite interactions, the host and the parasite constantly adapt to each other. This "back-and-forth" evolution can result in rapid changes in both host and viral proteins, which can be detected as positive selection [54]. Negative selection (i.e., dN/dS < 1) is indicative of purifying selection. Under purifying selection, mutation of the codon and amino acid are believed to have a detrimental effect on the fitness of the organism [55].

Currently, there is limited evidence that papillomavirus genes or specific residues are evolving under positive selection [13,41]. An important caveat is that the use of highly divergent sequences saturates the evolutionary changes, thereby complicating the detection of selection signatures. However, even when comparing closely related viral genomes (i.e., variants sharing more than 90% nucleotide identity), only a few sites are under diversifying selection (reviewed in Ref. [13]). For example, analyses of HPV16 found only seven codons to be under diversifying selection [46,56,57]. Therefore, most papillomavirus genes are under strong purifying selection, thus restricting changes in the encoded proteins. Interestingly, most residues under diversifying selection are located within the E5, E6, and E7 proteins (reviewed in Ref. [13]). While the core – E1, E2, L2, and L1 – proteins may not tolerate mutations, the adaptive proteins – E5, E6, and E7 – appear more

malleable [13,58,59]. This was demonstrated elegantly in a recent paper tracing the co-evolution between E6 and p53 [60]. The authors demonstrated a species barrier to p53 degradation by E6; human papillomavirus E6 proteins could degrade human p53 but not non-human primate p53. Vice-versa, macaque (MFPV10) E6 does not target human p53. This species specificity was mapped to a single amino acid in E6 and p53, demonstrating that E6 and p53 co-evolved in these viruses and their hosts [60].

5. Papillomaviruses use rare codons

The apparent lack of positive selection in papillomavirus genomes could suggest that the host is not exerting evolutionary selection on these viruses despite evolving alongside their hosts. However, evolutionary selection may act at the nucleotide level without changing the viral coding potential. Interestingly, it has been long known that papillomavirus' codon usage is not matched to their hosts [61]. While the usage of these rare codons may seem sub-optimal from a host perspective, it seems likely that viruses benefit from this unmatched codon usage. We propose that the use of these rare codons limits the rate of synonymous and non-synonymous mutations (i.e., dN/dS) and thus, in part, restricts the ability to detect positive selection in papillomavirus genomes. Nonetheless, we will discuss data suggesting that much of the viral adaptation appears to happen at the nucleotide level without extensively changing the protein-coding potential.

5.1. Overlapping open reading frames and cis-acting binding sites restrict codon usage

Like many small DNA viruses, 85% of the papillomavirus genome encodes for viral proteins. In addition, the genome has many overlapping open reading frames. A codon has three nucleotide positions, allowing for three reading frames for each stretch of nucleotides. Therefore, one nucleotide change in an overlapping sequence could be synonymous in one ORF but is likely to be non-synonymous in an overlapping frame. This will significantly reduce the evolutionary space available to both ORFs. Regulatory motifs (e.g., transcription factor binding sites, cis-acting elements regulating splicing, etc.) are located within viral ORFs. The requirements imposed by these DNA elements reduce the available codon usage. Indeed, overlapping ORFs and cis-acting binding sites further contribute to unusual codon usage patterns [62].

5.2. Risky codons restrict protein evolution

"Robust codons" code for amino acids with similar properties, while SNPs within "risky codons" are more likely to result in nonsense mutations in coding sequences. For example, codons beginning with TpA are risky due to the chance of being mutated to a stop codon [63]. Interestingly, unlike human genes that almost exclusively use robust codons [64], papillomavirus genes are biased towards "risky codons" [65]. The preponderance of risky codons implies that many mutations will dramatically change the encoded amino acid and be detrimental to viral fitness and will therefore not become fixed in the viral population. This may partly explain the overwhelming purifying selection observed in papillomavirus genomes and maintain the balance between virus and host obtained through millions of years of linked evolution. Significantly, a rare non-detrimental mutation would dramatically alter the encoded amino acid and may expand the evolutionary space available to the thus mutated protein and allow for rapid expansion in a new niche.

6. Host-pathogen interactions shape the DNA sequence of papillomavirus genomes

As pointed out, few papillomavirus amino acids are under diversifying selection. This is somewhat surprising for viruses interacting with

the host immune system [66]. However, typical methods (i.e., dN/dS-based approaches) designed to quantify evolutionary selection require amino acid changes. They do not directly quantify evolutionary selection acting at the level of the underlying nucleotide sequences.

It has been long known that papillomavirus' codons are different from their hosts. Indeed, by matching the codon usage of BPV L1 and L2 to codons more commonly used in humans, Zhou and colleagues significantly improved the expression of these proteins in human cells [61]. These data suggest that evolution may select nucleotide sequences and not only amino acid residues. By analyzing dinucleotide patterns, we and others identified that specific dinucleotide pairs are less common

than what would be expected by chance [67–70]. Specifically, there is a dramatic reduction in CpG and TpC dinucleotides (Fig. 2). The following paragraphs will discuss the potential implications of these reductions in genome content.

6.1. APOBEC3

APOBEC3 genes are a family of interferon-stimulated genes (ISGs). Thus, the activation of innate immune responses could lead to increased APOBEC3 activity [71–75]. This might especially be important for combating viruses acquired at mucosal sites of infection, where sentinel

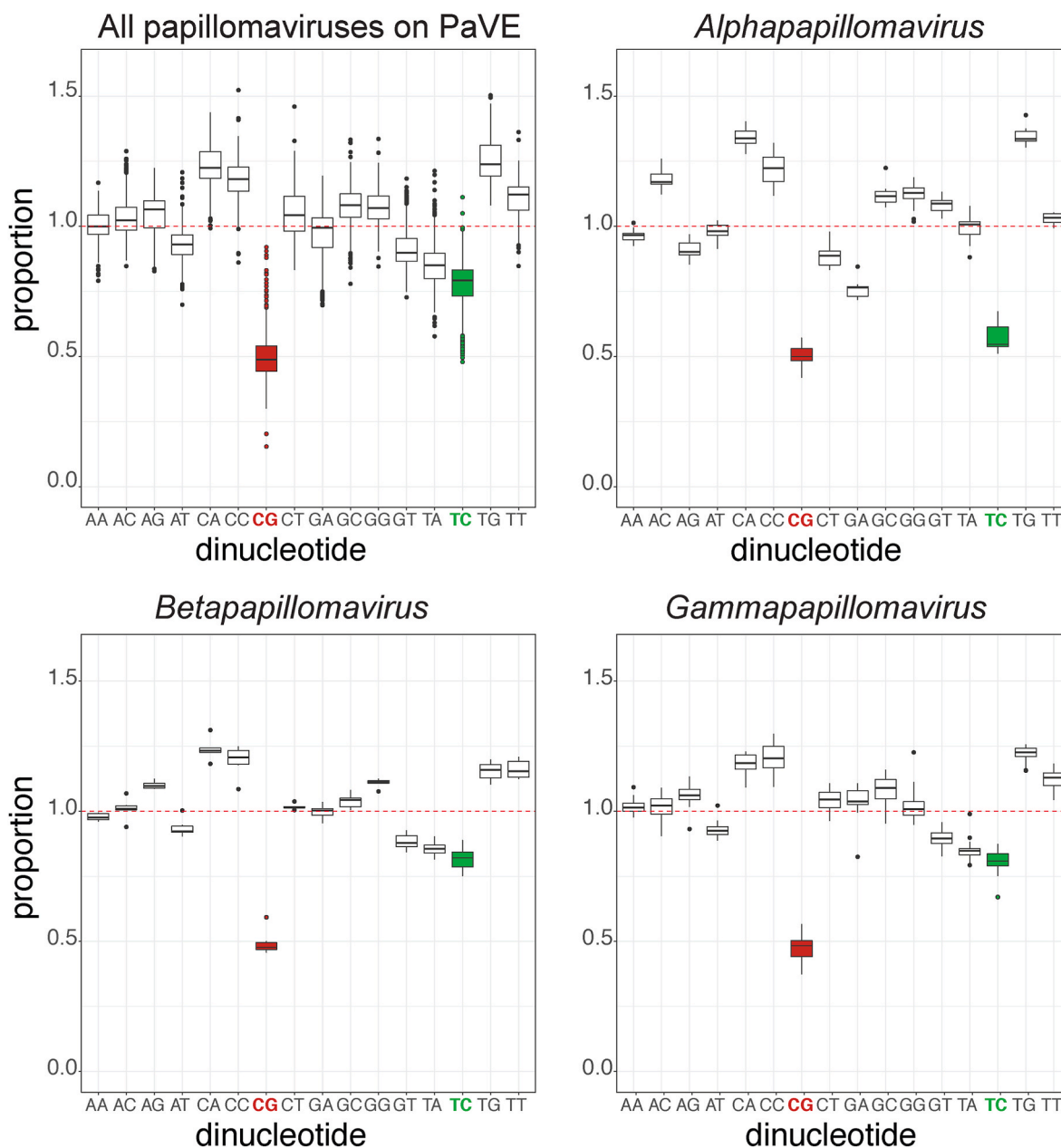


Fig. 2. Dinucleotide frequencies across papillomavirus genomes.

The observed-versus-expected (O/E) ratios of each dinucleotide in the papillomavirus genome sequences were calculated using a custom wrapper around the CompSeq program from the Emboss software suite. The red line at 1.0 indicates the ratio where a dinucleotide is seen as often as expected by chance. O/E ratios are displayed as box plots to show the minimum and maximum values (ends of the whiskers), interquartile range (length of the box), and median (line through the box). Outliers are shown as individual dots. CpG and TpC dinucleotides are highlighted in red and green, respectively. (A) All papillomaviruses in PaVE (n = 667). (B) Papillomaviruses are classified in the genus *Alphapapillomavirus* (n = 83). (C) Papillomaviruses classified in the genus *Betapapillomavirus* (n = 55). (D) Papillomaviruses classified in the genus *Gammapapillomavirus* (n = 100). A one-sample z-test was performed to obtain statistical support. In each panel, TpC and CpG were significantly less than 1 ($p < 0.01$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

cell targets such as macrophages and dendritic cells are poised to respond to these antiviral responses. Indeed, ectopic expression of APOBEC3A restricts infection with HPV16 based pseudovirions [74], and induction of APOBEC3 expression leads to hypermutation of the HPV16 genome [76]. The human APOBEC3A, 3B, and 3H preferentially target TpC-containing sequences for cytidine deamination. APOBEC3 deamination of TpC residues can result in C-to-T transitions or C-to-G transversions, leading to hypermutation and inactivation of the viral genome [77]. Unlike many other viruses, papillomaviruses have not evolved to inhibit the activity of APOBEC3. Indeed, the expression of the viral oncogenes E6 and E7 may upregulate the steady-state levels of cellular APOBEC3 proteins [77–79].

APOBEC3-induced mutations would explain the observed reduction in TpC (Fig. 2, green symbols). Warren and colleagues used computer modeling to study this reduction in TpC [67]. The authors demonstrate that the decrease in TpC is mirrored by an increase in TpG and CpA dinucleotides, suggesting that APOBEC3-driven deamination is involved in these mutations. Importantly, APOBEC3A and APOBEC3H mRNA and protein levels are significantly higher in mucosal tissues compared with cutaneous skin ([67] and Fig. 3). Papillomaviruses infect mucosal tissues (i.e., a subset of *Alphapapillomaviruses*) and display the most dramatic reduction in TpC dinucleotide ratio. These data demonstrate a correlation between APOBEC3 expression and a decrease in TpC ratio.

Interestingly, viruses in the species *Gammapapillomavirus 6* infect the cervical niche and have a reduced TpC ratio similar to what is seen in mucosal *Alphapapillomaviruses* [67]. Based on the further evolutionary analysis, the authors propose that viral replication in tissues with high APOBEC3 expression resulted in hypermutation of the ancestral infecting virus. While most of these mutated viruses were less fit and were culled from the evolutionary tree, evolution selected ancestral papillomavirus genomes with reduced target sites in their genomes. Notably, the observed depletion in TpC is most pronounced when the cytidine is in the third codon position. Therefore, the cytidines in most of the remaining TpC dinucleotides seen in extant papillomavirus genomes are

located in the first or second codon position. These observations suggest that the evolutionary loss of TpC dinucleotides likely did not change the encoded amino acids. Therefore, it seems that TpC reduction is limited by amino acid requirements, further highlighting that evolutionary selection occurs at the nucleotide sequence level with purifying selection maintaining amino acid composition and protein function. However, this also implies that APOBEC3 mutations in extant viral genomes will be non-synonymous and may impact viral fitness.

A recent paper sequenced the HPV16 genome from 5570 HPV16-infected case-control samples [80]. Because the authors used high-throughput sequencing, they could identify both major and minor allele frequencies for each sample. The authors identified rare mutations in the HPV16 genome consistent with APOBEC3 activity. These data imply that, following infection, APOBEC3 appears to further mutate a subset of the viral genomes in a cell. Interestingly, these APOBEC3 mutations were more common in HPV16(+) controls vs. HPV16(+) cervical precancer and cancer cases. The implications of these (APOBEC3 induced) mutations between cases and controls related to oncogenicity should be further examined. It will, however, be critical to account for the multiple bottlenecking events –infection, persistence, and transformation– during oncogenic progression.

Nonetheless, it is clear that APOBEC3 is mutating individual viral genomes during infection [80]. However, the evolutionary implications of this intra-patient variation are not clear. If evolutionary selection minimized APOBEC3 recognition sites within papillomavirus genomes, we could hypothesize that further reduction of TpC dinucleotides should reduce viral fitness, either in the patient or upon transmission. Under this hypothesis, minor variants with APOBEC3 mutation signatures should be rarely seen in consensus sequences available on GenBank. This is because the reduced fitness of the minor variants does not allow these variants to outcompete the major variant. While bulk sequencing and reporting the consensus sequence, these minor variants are effectively ignored. To test the hypothesis that APOBEC3 mutations are rare in consensus extant sequences, we analyzed the genomes of ~595

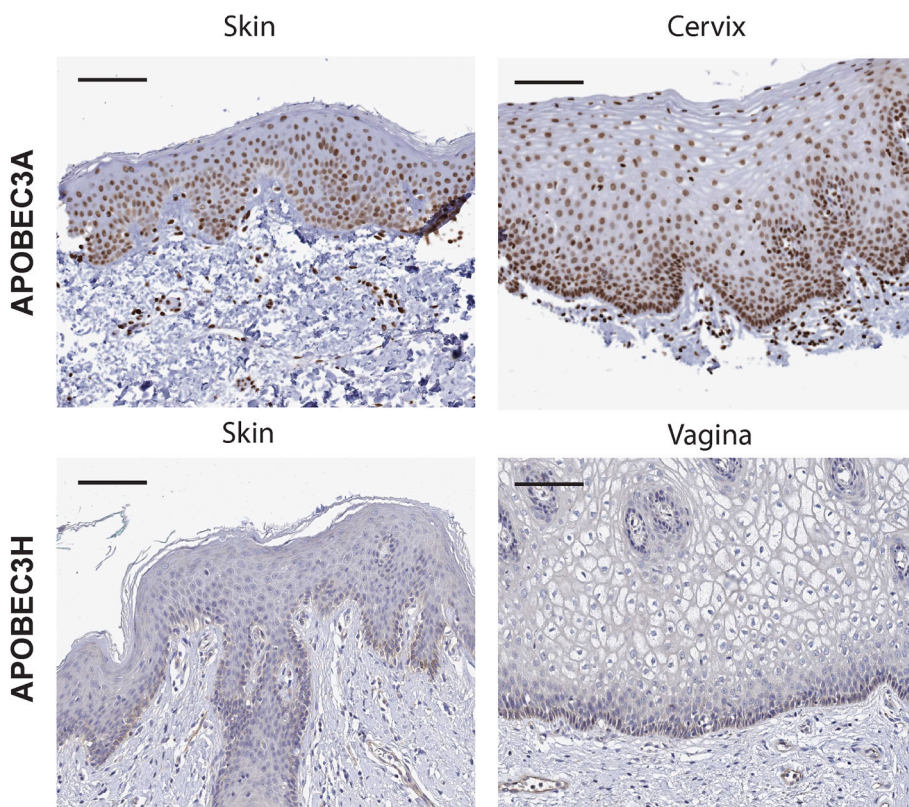


Fig. 3. Expression level of APOBEC3 in human tissues.

Images were obtained from The Human Protein Atlas, v21.proteinatlas.org. Scale bars represent 100 μ m (A) APOBEC3A (Apolipoprotein B MRNA Editing Enzyme Catalytic Subunit 3A) stained epithelia. Normal skin tissue (T-01000, Male, age 36, Patient id: 1816) and normal cervical tissue (T-83000, Female, age 36, Patient id: 1773) were both stained with antibody HPA043237. (B) APOBEC3H (Apolipoprotein B MRNA Editing Enzyme Catalytic Subunit 3H) stained epithelia. Normal vulva tissue (T-80100, Female, age 66, Patient id: 3357) and normal vaginal tissue (T-81000, Female, age 34, Patient id: 1751) were both stained with antibody HPA021492.

complete HPV16 genomes available on GenBank (Table 1). The data suggests that APOBEC3 mediated mutations of TpC sites are indeed rare in bulk, normalized sequences. In the 595 genomes, we identified 33 unique APOBEC3 mutations.

Furthermore, only seven of these mutations were synonymous. While some of these mutations were observed in ~20% of the samples, most mutations were only observed in a single sequence. Overall, the synonymous mutations were identified in more individual samples. An interesting example is a mutation in L2 (indicated with \$). This TpC dinucleotide was mutated in at least two separate instances (TTC→TTG and TTC→TTT). While the TTT mutation is silent, the TTG mutation changes the wildtype F to a variant L at this position. In line with the hypothesis that most extant TpC mutations will affect viral fitness, the silent mutation is seen in 4 times more samples than the non-synonymous mutation. Therefore, it appears that while intra-patient APOBEC3 editing alters a population of the viral genomes, the global effects of APOBEC3 mutations on HPV16 evolution may be more limited. Potentially, the bottleneck of spreading from one patient limits the evolutionary landscape available to HPV16 and other papillomavirus genomes.

The analysis in Table 1 also demonstrates that APOBEC3 is indeed editing HPV genomes and that the increased TpG proportions (Fig. 2; *Alphapapillomavirus*) are not just a mathematical effect of maintaining overall T content in the face of lowering TpC proportions.

6.2. TLR9

Papillomaviruses access mitotically active basal cells through lesions in the stratified epithelia of cutaneous or mucosal tissues. The virus is endocytosed by binding to cellular receptors, and the viral capsid disassembles, exposing the viral genome (Fig. 4) [81–87]. The exposed

viral DNA may be recognized by endosomal TLR9, resulting in a downstream inflammatory immune response [88–92].

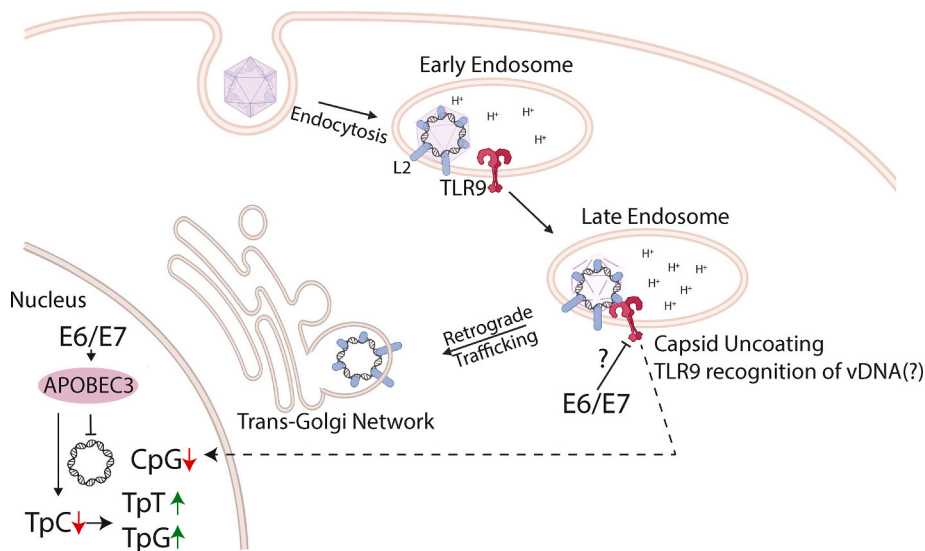
As shown in Fig. 2, in addition to a reduction in TpC, papillomavirus genomes substantially reduce CpG dinucleotides. This is in part due to the mutagenic effects of 5-methylcytosines. CpGs are frequently methylated, and the resulting 5-methylcytosines are prone to deamination leading to C→T mutations. It is likely that part of the observed CpG depletion in papillomavirus genomes can be explained by this phenomenon [13,69,70]. However, the deficit in CpG is not mirrored by an expected increase in TpG (Fig. 2). Therefore, King and colleagues tested the hypothesis that similar to TpC depletion, this CpG depletion results from an evolutionary response to antiviral effectors. King and colleagues analyzed papillomavirus genomes isolated from bats. It was previously shown that bat TLR9 is under diversifying selection in bats [68,93]. Given the co-evolution theory between PVs and their hosts, this change in TLR9's DNA recognition should induce a complementary response in the viral genome. Since the DNA recognition domain of bat TLR9 is under evolutionary selection, the authors compared papillomavirus genomes isolated from different bat species belonging to different sub-orders. The authors demonstrate that viruses infecting specific bats reduce their genomic CpG within a known TLR9 recognition motif ('ACGT'). Furthermore, the authors demonstrate that members of the genus *Alphapapillomavirus* significantly deplete TpC in the context of TCGA motifs, a known human TLR9 recognition motif [94,95,95–97].

Together, these data demonstrate that papillomaviruses deplete CpG in the context of species-specific TLR9 recognition motifs. Since the authors detected evidence of co-evolution between TLR9 and the viral nucleotide ratio, these changes are likely due to evolutionary pressure provided by evading recognition by the TLR9 protein. Similar to what was seen as a response to APOBEC3 editing, the depletion of CpG does not appear to drastically alter the amino acid content of the viral

Table 1
APOBEC3 mutations in extant HPV16 variant genomes.

| gene (HPV16) | consensus | mutant | consensus | mutant | count | effect of mutation |
|----------------|-----------|--------|-----------|--------|-------|--------------------|
| E1 (n = 595) | CTA | TTA | L | L | 1 | synonymous |
| E1 (n = 595) | CAG | GAG | Q | E | 3 | non-synonymous |
| E1 (n = 595) | TCA | TTA | S | L | 1 | non-synonymous |
| E1 (n = 595) | CTA | TTA | L | L | 114 | synonymous |
| E1 (n = 595) | TCT | TTT | S | F | 1 | non-synonymous |
| E1 (n = 595) | TCC | TTC | S | F | 1 | non-synonymous |
| E1 (n = 595) | TCT | TTT | S | F | 1 | non-synonymous |
| E2 (n = 596) | CTT | TTT | L | F | 1 | non-synonymous |
| E2 (n = 596) | TCT | TTT | S | F | 2 | non-synonymous |
| E2 (n = 596) | CCT | GCT | P | A | 116 | non-synonymous |
| E2 (n = 596) | ATC | ATT | I | I | 2 | synonymous |
| E2 (n = 596) | CCA | TCA | P | S | 10 | non-synonymous |
| E6 (n = 596) | CAG | GAG | Q | E | 1 | non-synonymous |
| E6 (n = 596) | CCA | TCA | P | S | 1 | non-synonymous |
| E6 (n = 596) | CGG | TGG | R | W | 1 | non-synonymous |
| E6 (n = 596) | TCT | TGT | S | C | 1 | non-synonymous |
| E7 (n = 598) | CTC | CTT | L | L | 1 | synonymous |
| E7 (n = 598) | CGG | TGG | R | W | 1 | non-synonymous |
| L1 (n = 596) | GTC | GTT | V | V | 1 | synonymous |
| L1 (n = 596) | TCA | TTA | S | L | 1 | non-synonymous |
| L1 (n = 596) | CCA | GCA | P | A | 1 | non-synonymous |
| L1 (n = 596) | TCA | TTA | S | L | 1 | non-synonymous |
| L1 (n = 596) | CTA | TTA | L | L | 116 | synonymous |
| L1 (n = 596) | TCC | TTC | S | F | 1 | non-synonymous |
| L2 (n = 596) | CCC | TCC | P | S | 1 | non-synonymous |
| L2 (n = 596)# | CCA | GCA | P | A | 1 | non-synonymous |
| L2 (n = 596)# | CCA | TCA | P | S | 1 | non-synonymous |
| L2 (n = 596) | TCT | TTT | S | F | 1 | non-synonymous |
| L2 (n = 596)\$ | TTC | TTG | F | L | 46 | non-synonymous |
| L2 (n = 596)\$ | TTC | TTT | F | F | 194 | synonymous |
| L2 (n = 596) | TCA | TTA | S | L | 4 | non-synonymous |
| L2 (n = 596) | TCA | TTA | S | L | 2 | non-synonymous |
| L2 (n = 596) | CCT | TCT | P | S | 1 | non-synonymous |

HPV16 sequences were downloaded from GenBank using the PaVE database. Individual HPV16 genes were aligned, and a consensus genome was calculated. The associated amino acid was identified for each SNP that changes TpC in the consensus to TpT or TpG. The occurrence of each SNP in different genomes (count) and the effect of the mutation are indicated. # and \$ indicate SNPs at the same position in L2 that were mutated in at least two separate events.



proteins [68]. These data support the idea that papillomavirus protein sequences are evolutionarily constrained but that nucleotide sequences are evolving to escape antiviral responses shaping host-virus interactions.

7. Role of immune effectors during persistent infections

Overall, the reduction in TpC ratio suggests that papillomaviruses evolved to avoid APOBEC3-mediated editing of their genomes. Yet, papillomavirus infections upregulate the steady-state levels of APOBEC3 in infected cells. This dichotomy was recently addressed in a thoughtful review by Wallace and Munger [98]. It is well known that the viral E7 proteins, which infect the mucosal tissues, degrade pRb. pRb has been demonstrated to play critical roles in silencing repetitive elements. Therefore, degradation of pRb by E7 is expected to increase the transcription of repetitive elements [99]. Expression of Long Interspersed Element-1 (LINE1 or L1) and other repetitive elements would generate neoantigens, potentially activating the adaptive immune response. The LINE1 ORF 2 is an endonuclease that, while essential to retrotransposition, may cause excessive double-strand DNA breaks leading to senescence or apoptosis. Activation of the adaptive immune system and/or apoptosis would lead to eliminating HR-HPV-infected cells. APOBEC3s are known to restrict the expression of LINE1 and other repetitive elements [100]. Therefore, increased expression of APOBEC3 may be beneficial for HPVs, and secondary viral genome editing is minimized by reducing the available TpC sites in the genome.

Similarly, the reduction in CpG in the context of a TLR9 recognition motif suggests that TLR9 senses HPVs during infection. However, unlike APOBEC3, the viral oncogenes may reduce TLR9 levels in infected cells [101–104]. However, E6 and E7 are only expressed after the viral genome traffics to the nucleus, presumably after endosomal sensing of the viral DNA has occurred. This suggests that interfering with TLR9 activity may benefit a papillomavirus' ability to persist long-term in cells. Indeed, clearance of MmuPV1 infection in the mouse papillomavirus model (MmuPV1) is, in part, regulated by a functional MyD88 signaling pathway, suggesting a role for TLRs in the immune response against established papillomavirus infections [105]. Interestingly, other DNA viruses (e.g., Merkel cell polyomavirus, hepatitis B virus, and Epstein-Barr virus [EBV]) also interfere with TLR9 function during the maintenance phase of the infection [106–108]. This raises an interesting hypothesis that, in addition to evading TLR9 during infectious entry (by reducing the genomic CpG content), the virus benefits from interfering with the TLR9 pathway (following expression of E6 and E7).

Fig. 4. Model of infection highlighting APOBEC3 and TLR9.

The papillomavirus lifecycle risks detection by TLR9 as well as antiviral activities mediated by APOBEC3. Internalized virions enter the endolysosomal pathway triggering pH-dependent uncoating. This uncoating likely exposes the CpG dinucleotides in viral DNA to endosomal TLR9. However, evolution selected for viral genomes with depleted CpG reducing the ability of TLR9 to detect the viral DNA during entry. Interestingly, viral oncogenes have been shown to degrade TLR9 suggesting that the viral lifecycle benefits from degrading TLR9 receptors. The viral L2 protein directs retrograde trafficking to the trans-Golgi network. The viral E6/E7 proteins induce the steady-state levels of APOBEC3, which could hypermutate the viral genome leading to loss of persistent infection. However, the viral genomes evolved a reduced TpC content, thus evading spurious mutations by APOBEC3.

8. Implications and future work

In the above sections, we discussed current data suggesting that evolutionary processes shape the nucleotide sequences of papillomavirus genomes despite robust purifying selection occurring at the amino-acid/protein level. These observations likely have important implications for how we think about viral evolution and study the papillomavirus lifecycle.

8.1. Studying papillomavirus evolution

Papillomaviruses infect fish, relying on the theory of co-speciation, this implies that the ancestral papillomavirus was already infecting animals in the Palaeozoic (lasting from 540 to 250 million years ago) when fish evolved. Indeed, these dates are supported by evolutionary analyses dating the most recent common ancestor of all papillomaviruses to roughly 450 million years ago [58,109]. Despite the observation that papillomaviruses evolve slowly (roughly 5x faster than the infected host; [13]), over this timescale, much of the DNA sequence information has become saturated. Briefly, genetic saturation occurs when multiple substitutions happen at the same site in a sequence such that the apparent sequence divergence rate is lower than what has happened. The longer the time frame, the more likely this is an issue. Practically, amino acid residues are less sensitive to saturation, allowing for studying more ancient evolutionary events [110]. Indeed, the phylogenetic tree in Fig. 1A was constructed using protein sequences. However, this review argues that much of the evolutionary information is contained within the nucleotide sequences, which is lost when solely considering the encoded amino acids. A recent study by the Chen and Burk labs used trimer spectra to construct a dendrogram of the *papillomaviridae* [70]. This alignment-free approach relies on the distribution of trimers of DNA sequences and using this information to cluster similar sequences into a UPGMA-based dendrogram [111–116].

Interestingly, this approach clusters the mucosal members of the genus *Gammapapillomavirus* with the mucosal viruses belonging to the genus *Alpapapillomavirus*. Strikingly, these viruses show a similar tissue tropism and have undergone similar TpC reduction due to interactions with APOBEC3 in the mucosal tissue (see above). These relationships are not apparent using more conventional homology-based analyses. Therefore, combinations of conventional homology-based and these k-mer-based approaches may be needed to dissect the evolutionary history of the *Papillomaviridae* fully. However, it is clear that traditional dN/dS-based methods have significantly underestimated the evolutionary processes involved in shaping papillomavirus diversity and fitness.

8.2. Importance of nucleotide motif depletion

This review highlighted the importance of TpC and CpG depletion. However, Fig. 2 highlights that other dinucleotides are depleted from viral genomes. Furthermore, as illustrated by APOBEC3 in mucosal tissues, this depletion may be more apparent when considering specific clades of the phylogenetic tree or even host environmental niches. For example, the *Betapapillomavirus* and *Gammapapillomavirus* graphs in Fig. 2 show a reduction in TpA. This may be partly because the universal stop codons contain TpA. However, if this were the case, we would expect a similar decrease in TpA across all papillomaviruses. Since this is not the case, it would be interesting to explore whether the reduction in TpA results from co-evolution with antiviral responses.

Interestingly, a depletion in TpA has been linked to the observation that, when transcribed, UpA is the preferred target for ribonucleases. This was supported by the observation that TpA depletion is most robust in exons of protein-coding genes [117]. However, a computational study also suggests that the apparent TpA depletion is due to a mathematical artifact and is directly linked to the decrease in CpG [118]. Beyond dinucleotides, it is likely informative to study extended motifs.

8.3. Effects of silent mutagenesis approaches

Studying the function of viral proteins often involves mutagenesis of a specific residue within this viral protein. Since many papillomavirus genes overlap other genes, most researchers will ensure that the generated mutation does not alter the protein encoded in the next overlapping frame. However, accumulating data suggest that the nucleotide sequence, and not just the encoded protein(s), may substantially impact viral function and fitness. The data summarized in this review suggests that these mutational studies should be carefully designed and interpreted.

8.4. Intrapatient mutations and cancer progression

A recent paper by Mirabello and colleagues demonstrates a correlation between specific HPV16 SNPs and worse survival in human papillomavirus-driven oropharyngeal cancers [119]. The same authors demonstrate that APOBEC3-induced mutations in the HPV16 genome are associated with viral clearance [120]. Together with other reports, these data demonstrate that at least a subset of papillomavirus genomes acquire mutations during infection and cancer progression. The data presented in this review argues that these intrapatient mutations are not a major contributor to the current papillomavirus diversity (Table 1). Indeed, HPV16 variant lineages likely diverged between 400 and 600 thousand years ago, predating the divergence of modern *Homo sapiens* [121,122]. Importantly, papillomavirus fitness is measured by a virus' ability to persist long-term in a host while maximizing transmission to the next host. Therefore, these processes regulate the balance between nucleotide level evolution as described in this review and purifying selection at the protein level. However, the evolutionary pressures are likely very different within a patient and certainly during progression towards cancer. As such, these cancer specific SNPs may be causative of cancer progression. Therefore, studying the impact of intrapatient mutation, selection, and evolution will be necessary. High-throughput sequencing has made these types of experiments feasible [80, 123–125]. By controlling for the many bottlenecks during cancer progression, researchers may be able to identify mutations within viral genomes associated with cancer, allowing for more personalized screening of patients.

8.5. Impact of antiviral restriction factors during persistent infection

As discussed above, papillomavirus genomes are depleted in CpG to evade detection by TLR9 [68]. However, the viral oncogenes regulate the steady-state levels of this protein in infected cells [101–104]. This

suggests that, during persistent infection, when the viral DNA is mainly located in the nucleus, the virus benefits from interfering with the TLR9/MyD88/IFN pathway [126]. This is reminiscent of papillomavirus interactions with another arm of the innate immune system, cGAS/STING. cGAS recognizes cytoplasmic DNA and activates STING resulting in the activation of the interferon response. However, it was recently demonstrated that the papillomavirus genome is not sensed by cGAS during infection [126]. Yet, the papillomavirus E6 and E7 proteins have been implicated in interfering with the cGAS/STING pathway [127–131]. Together, these data suggest that papillomaviruses interfere with antiviral pathways to promote long-term infections, not just acute infection.

If papillomaviruses indeed interfere with TLR9 activity to promote long-term infections, this raises an important question about what PAMP (pathogen-associated molecular pattern) or DAMP (damage-associated molecular patterns) the cellular TLR9 may be detecting. Since papillomaviruses replicate in the nucleus of the infected cell, it is unlikely that TLR9 is detecting the viral DNA. Moreover, since the depletion of CpG dinucleotides likely further reduces TLR9 recognition, it is doubtful that the viral genome represents PAMP recognized by TLR9. This raises the possibility that persistent papillomavirus infections induce the production of DAMPs and that papillomaviruses evolved to evade the innate immune system in two complementary ways, first by escaping detection during infectious entry (e.g., by editing the genome content) into the cell and second by interfering with the pattern recognition receptors following infection.

9. Conclusions

This review discusses current evidence that papillomaviruses are evolving in response to interactions with the innate immune system. Interestingly, there is limited evidence of adaptive evolution in papillomavirus proteins. However, we propose that papillomavirus evolution primarily occurs at the nucleotide level in a way that optimizes viral fitness without altering the encoded aminoacids and thus remaining constrained by pervasive purifying selection at the protein level.

Author statement

KMK: Figure preparation and Text writing/editing EVR: Figure preparation and Text writing/editing IGT: Figure preparation and Text writing/editing KVD: Figure preparation and Text writing/editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All sequence data is available from public databases

Acknowledgments

We thank Drs. Jackson and Campos for critical feedback on the manuscript. Work in the Van Doorslaer lab is supported by grants by the Arizona Biomedical Research Centre (CTR056055), The National Institute for Dental and Craniofacial Research (NIDCR; 1R03DE030211-01), The National Institute of Allergy and Infectious Diseases (NIAID; 1R01AI165638-01A1), The American Cancer Society (RSG-22-054-01-IBCD), and a grant to the University of Arizona Cancer center (P30CA023074).

References

- [1] K. Van Doorslaer, Q. Tan, S. Xirasagar, S. Bandaru, V. Gopalan, Y. Mohamoud, Y. Huyen, A.A. McBride, The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis, *Nucleic Acids Res.* 41 (2013) D571–D578.
- [2] K. Van Doorslaer, Z. Li, S. Xirasagar, P. Maes, D. Kaminsky, D. Liou, Q. Sun, R. Kaur, Y. Huyen, A.A. McBride, The Papillomavirus Episteme: a major update to the papillomavirus sequence database, *Nucleic Acids Res.* 45 (2017) D499–D506.
- [3] K. Van Doorslaer, Z. Chen, H.-U. Bernard, P.K.S. Chan, R. DeSalle, J. Dillner, O. Forslund, T. Haga, A.A. McBride, L.L. Villa, R.D. Burk, Ictv Report Consortium null, ICTV virus taxonomy profile: papillomaviridae, *J. Gen. Virol.* 99 (2018) 989–990.
- [4] E.-M. de Villiers, C. Fauquet, T.R. Broker, H.-U. Bernard, H. zur Hausen, Classification of papillomaviruses, *Virology* 324 (2004) 17–27.
- [5] H.-U. Bernard, R.D. Burk, Z. Chen, K. van Doorslaer, H. zur Hausen, E.-M. de Villiers, Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments, *Virology* 401 (2010) 70–79.
- [6] F.M. Zerbin, S.G. Siddell, A.R. Mushegian, P.J. Walker, E.J. Lefkowitz, E. M. Adriaenssens, P. Alfenas-Zerbin, B.E. Dutilh, M.L. García, S. Junglen, M. Krupovic, J.H. Kuhn, A.J. Lambert, M. Łobocka, H.M. Oksanen, D. L. Robertson, L. Rubino, S. Sabanadzovic, P. Simmonds, N. Suzuki, K. Van Doorslaer, A.-M. Vandamme, A. Varsani, Differentiating between viruses and virus species by writing their names correctly, *Arch. Virol.* 167 (2022) 1231–1234.
- [7] R.L. Finnen, K.D. Erickson, X.S. Chen, R.L. Garcea, Interactions between papillomavirus L1 and L2 capsid proteins, *J. Virol.* 77 (2003) 4818–4826.
- [8] S.S. Porter, J.C. Liddle, K. Browne, D.V. Pastrana, B.A. Garcia, C.B. Buck, M. D. Weitzman, A.A. McBride, Histone modifications in papillomavirus virion minichromosomes, *mBio* 12 (2021), e03274, 20/mbio/12/1/mBio.03274-20. atom.
- [9] P.M. Larsen, L. Storgaard, S.J. Fey, Proteins present in bovine papillomavirus particles, *J. Virol.* 61 (1987) 3596–3601.
- [10] C. Johansson, S. Schwartz, Regulation of human papillomavirus gene expression by splicing and polyadenylation, *Nat. Rev. Microbiol.* 11 (2013) 239–251.
- [11] S.V. Graham, A.A.A. Faizo, Control of human papillomavirus gene expression by alternative splicing, *Virus Res.* 231 (2017) 83–95.
- [12] J. Ferguson, K. Campos-León, I. Pentland, J.D. Stockton, T. Günther, A.D. Beggs, A. Grundhoff, S. Roberts, B. Noyvert, J.L. Parish, The chromatin insulator CTCF regulates HPV18 transcript splicing and differentiation-dependent late gene expression, *PLoS Pathog.* 17 (2021), e1010032.
- [13] K. Van Doorslaer, Evolution of the papillomaviridae, *Virology* 445 (2013) 11–20.
- [14] K. Van Doorslaer, A.A. McBride, Molecular archeological evidence in support of the repeated loss of a papillomavirus gene, *Sci. Rep.* 6 (2016), 33028.
- [15] M. Bergvall, T. Melendy, J. Archambault, The E1 proteins, *Virology* 445 (2013) 35–56.
- [16] A.A. McBride, The papillomavirus E2 proteins, *Virology* 445 (2013) 57–79.
- [17] S.B. Vande Pol, A.J. Klingelutz, Papillomavirus E6 oncoproteins, *Virology* 445 (2013) 115–137.
- [18] A. Roman, K. Munger, The papillomavirus E7 proteins, *Virology* 445 (2013) 138–168.
- [19] C.A. Moody, L.A. Laimins, Human papillomavirus oncoproteins: pathways to transformation, *Nat. Rev. Cancer* 10 (2010) 550–560.
- [20] J.A. Scarth, M.R. Patterson, E.L. Morgan, A. Macdonald, The human papillomavirus oncoproteins: a review of the host pathways targeted on the road to transformation, *J. Gen. Virol.* 102 (2021).
- [21] J. Doorbar, The E4 protein; structure, function and patterns of expression, *Virology* 445 (2013) 80–98.
- [22] E. Straub, M. Dreer, J. Fertey, T. Iftner, F. Stubenrauch, The viral E8E2C repressor limits productive replication of human papillomavirus 16, *J. Virol.* 88 (2014) 937–947.
- [23] D. DiMaio, L.M. Petti, The E5 proteins, *Virology* 445 (2013) 99–114.
- [24] I.G. Bravo, A. Alonso, Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth, *J. Virol.* 78 (2004) 13613–13626.
- [25] C.B. Buck, P.M. Day, B.L. Trus, The papillomavirus major capsid protein L1, *Virology* 445 (2013) 169–174.
- [26] J.W. Wang, R.B.S. Roden, L2, the minor capsid protein of papillomavirus, *Virology* 445 (2013) 175–186.
- [27] S.K. Campos, Subcellular trafficking of the papillomavirus genome during initial infection: the remarkable abilities of minor capsid protein L2, *Viruses* 9 (2017).
- [28] M.A. Ozburn, S.K. Campos, The long and winding road: human papillomavirus entry and subcellular trafficking, *Curr Opin Virol* 50 (2021) 76–86.
- [29] Z. Chen, K. van Doorslaer, R. DeSalle, C.E. Wood, J.R. Kaplan, J.D. Wagner, R. D. Burk, Genomic diversity and interspecies host infection of alpha12 Macaca fascicularis papillomaviruses (MfPVs), *Virology* 393 (2009) 304–310.
- [30] M. Gottschling, I.G. Bravo, E. Schulz, M.A. Bracho, R. Deaville, P.D. Jepson, M.-F. V. Bresse, E. Stockfleth, I. Nindl, Modular organizations of novel cetacean papillomaviruses, *Mol. Phylogenet. Evol.* 59 (2011) 34–42.
- [31] H.-U. Bernard, Coevolution of papillomaviruses with human populations, *Trends Microbiol.* 2 (1994) 140–143.
- [32] M.S. Hafner, S.A. Nadler, Phylogenetic trees support the coevolution of parasites and their hosts, *Nature* 332 (1988) 258–259.
- [33] I.G. Bravo, A. Alonso, Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins, *Virus Gene.* 34 (2007) 249–262.
- [34] S.Y. Chan, H.U. Bernard, C.K. Ong, S.P. Chan, B. Hofmann, H. Delius, Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses, *J. Virol.* 66 (1992) 5714–5725.
- [35] S.Y. Chan, H. Delius, A.L. Halpern, H.U. Bernard, Analysis of genomic sequences of 95 papillomavirus types: uniting typing, phylogeny, and taxonomy, *J. Virol.* 69 (1995) 3074–3083.
- [36] S.Y. Chan, H.U. Bernard, M. Ratterree, T.A. Birkebak, A.J. Faras, R.S. Ostrow, Genomic diversity and evolution of papillomaviruses in rhesus monkeys, *J. Virol.* 71 (1997) 4938–4943.
- [37] S.-Y. Chan, S.-H. Chew, K. Egawa, E.-I. Grusendorf-Conen, Y. Honda, A. Rübben, K.-C. Tan, H.-U. Bernard, Phylogenetic analysis of the human papillomavirus type 2 (HPV-2), HPV-27, and HPV-57 group, which is associated with common Warts, *Virology* 239 (1997) 296–302.
- [38] S. García-Vallvé, Á. Alonso, I.G. Bravo, Papillomaviruses: different genes have different histories, *Trends Microbiol.* 13 (2005) 514–521.
- [39] M. Gottschling, A. Köhler, E. Stockfleth, I. Nindl, Phylogenetic analysis of beta-papillomaviruses as inferred from nucleotide and amino acid sequence data, *Mol. Phylogenet. Evol.* 42 (2007) 213–222.
- [40] M. Gottschling, A. Stamatakis, I. Nindl, E. Stockfleth, A. Alonso, I.G. Bravo, Multiple evolutionary mechanisms drive papillomavirus diversification, *Mol. Biol. Evol.* 24 (2007) 1242–1258.
- [41] A. Narechania, Z. Chen, R. DeSalle, R.D. Burk, Phylogenetic incongruence among oncogenic genital alpha human papillomaviruses, *J. Virol.* 79 (2005) 15503–15510.
- [42] A. Rector, H. Stevens, G. Lacave, P. Lemey, S. Mostmans, A. Salbany, M. Vos, K. Van Doorslaer, S.-J. Ghim, M. Rehtanz, G.D. Bossart, A.B. Jensen, M. Van Ranst, Genomic characterization of novel dolphin papillomaviruses provides indications for recombination within the Papillomaviridae, *Virology* 378 (2008) 151–161.
- [43] J. Kubacki, A.S. Ramsauer, C. Bachofen, C. Favrot, A. Nicolier, C. Fraefel, K. Tobler, Complete genome sequence of a Boa (*Boa constrictor*)-Specific papillomavirus type 1 isolate, *Microbiol. Resour. Anounc.* 7 (2018), e01159, 18.
- [44] C.E. Lange, C. Favrot, M. Ackermann, J. Gull, E. Vetsch, K. Tobler, Novel snake papillomavirus does not cluster with other non-mammalian papillomaviruses, *Virol. J.* 8 (2011) 436.
- [45] M. Angulo, A. Carvajal-Rodríguez, Evidence of recombination within human alpha-papillomavirus, *Virol. J.* 4 (2007) 33.
- [46] A. Carvajal-Rodríguez, Detecting recombination and diversifying selection in human alpha-papillomavirus, *Infect. Genet. Evol.* 8 (2008) 689–692.
- [47] Aaron L. Halpern, Assessing recombination in HPV, Part I, *Human Papillomavirus III* (1995) 124–138.
- [48] S.D. Shah, J. Doorbar, R.A. Goldstein, Analysis of host-parasite incongruence in papillomavirus evolution using importance sampling, *Mol. Biol. Evol.* 27 (2010) 1301–1314.
- [49] A. Varsani, E. van der Walt, L. Heath, E.P. Rybicki, A.L. Williamson, D.P. Martin, Evidence of ancient papillomavirus recombination, *J. Gen. Virol.* 87 (2006) 2527–2531.
- [50] F. Borvetó, I.G. Bravo, A. Willemsen, Papillomaviruses infecting cetaceans exhibit signs of genome adaptation following a recombination event, *Virus Evol.* 6 (2020) veaa038.
- [51] D.R. Brooks, A.L. Ferrao, The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen: guest Editorial, *J. Biogeogr.* 32 (2005) 1291–1299.
- [52] C.B. Buck, K. Van Doorslaer, A. Peretti, E.M. Geoghegan, M.J. Tisza, P. An, J. P. Katz, J.M. Pipas, A.A. McBride, A.C. Camus, A.J. McDermott, J.A. Dill, E. Delwart, T.F.F. Ng, K. Farkas, C. Austin, S. Kraberger, W. Davison, D. V. Pastrana, A. Varsani, The ancient evolutionary history of polyomaviruses, *PLoS Pathog.* 12 (2016), e1005574.
- [53] M. Gottschling, M. Goker, A. Stamatakis, O.R.P. Bininda-Emonds, I. Nindl, I. G. Bravo, Quantifying the phylodynamic forces driving papillomavirus evolution, *Mol. Biol. Evol.* 28 (2011) 2101–2113.
- [54] J.A. Kerns, M. Emerman, H.S. Malik, Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein, *PLoS Genet.* 4 (2008) e21.
- [55] M.S. Lee, P.J. Tuohy, C. Kim, K. Lichauro, H.L. Parrish, K.V. Doorslaer, M. S. Kuhns, Enhancing and Inhibitory Motifs Have Coevolved to Regulate CD4 Activity, preprint, *Immunology*, 2021.
- [56] Z. Chen, M. Terai, L. Fu, R. Herrero, R. DeSalle, R.D. Burk, Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses, *J. Virol.* 79 (2005) 7014–7023.
- [57] V.R. DeFilippis, F.J. Ayala, L.P. Villarreal, Evidence of diversifying selection in human papillomavirus type 16 E6 but not E7 oncogenes, *J. Mol. Evol.* 55 (2002) 491–499.
- [58] A. Willemsen, I.G. Bravo, Origin and evolution of papillomavirus (onco)genes and genomes, *Philos. Trans. R. Soc. B Biol. Sci.* 374 (2019), 20180303.
- [59] A. Willemsen, M. Féliz-Sánchez, I.G. Bravo, Genome plasticity in papillomaviruses and De Novo emergence of E5 oncogenes, *Genome Biol. Evol.* 11 (2019) 1602–1617.
- [60] T. Long, R.D. Burk, P.K.S. Chan, Z. Chen, Non-human primate papillomavirus E6-mediated p53 degradation reveals ancient evolutionary adaptation of carcinogenic phenotype to host niche, *PLoS Pathog.* 18 (2022), e1010444.
- [61] J. Zhou, W.J. Liu, S.W. Peng, X.Y. Sun, I. Frazer, Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability, *J. Virol.* 73 (1999) 4972–4982.

- [62] A.E. Firth, C.M. Brown, Detecting overlapping coding sequences in virus genomes, *BMC Bioinf.* 7 (2006) 75.
- [63] L.A. Shackelton, C.R. Parrish, E.C. Holmes, Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses, *J. Mol. Evol.* 62 (2006) 551–563.
- [64] P. Schmid, W.A. Flegel, Codon usage in vertebrates is associated with a low risk of acquiring nonsense mutations, *J. Transl. Med.* 9 (2011) 87.
- [65] I.G. Bravo, M. Müller, Codon usage in papillomavirus genes: practical and functional aspects, *Papillomavirus Rep.* 16 (2005) 63–72.
- [66] M.D. Daugherty, H.S. Malik, Rules of engagement: molecular insights from host-virus arms races, *Annu. Rev. Genet.* 46 (2012) 677–700.
- [67] C.J. Warren, K. Van Doorslaer, A. Pandey, J.M. Espinosa, D. Pyeon, Role of the host restriction factor APOBEC3 on papillomavirus evolution, *Virus Evol.* 1 (2015).
- [68] K. King, B.B. Larsen, S. Gryseels, C. Richet, S. Kraberger, R. Jackson, M. Worobey, J.S. Harrison, A. Varsani, K.V. Doorslaer, Co-evolutionary analysis suggests a role for TLR9 in papillomavirus restriction, preprint. *Microbiology* (2021).
- [69] M. Upadhyay, P. Vivekanandan, Depletion of CpG dinucleotides in papillomaviruses and polyomaviruses: a role for divergent evolutionary pressures, *PLoS One* 10 (2015), e0142368.
- [70] Z. Chen, F. Utro, D. Platt, R. DeSalle, L. Parida, P.K.S. Chan, R.D. Burk, K-mer analyses reveal different evolutionary histories of alpha, beta, and gamma papillomaviruses, *Int. J. Mol. Sci.* 22 (2021) 9657.
- [71] J.-P. Vartanian, D. Guétard, M. Henry, S. Wain-Hobson, Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions, *Science* 320 (2008) 230–233.
- [72] R. Suspène, M.-M. Aynaud, S. Koch, D. Pasdeloup, M. Labetoulle, B. Gaertner, J.-P. Vartanian, A. Meyerhans, S. Wain-Hobson, Genetic editing of herpes simplex virus 1 and Epstein-Barr herpesvirus genomes by Human APOBEC3 Cytidine deaminases in culture and *in vivo*, *J. Virol.* 85 (2011) 7594–7602.
- [73] N. Minkah, K. Chavez, P. Shah, T. MacCarthy, H. Chen, N. Landau, L.T. Krug, Host restriction of murine gammaherpesvirus 68 replication by human APOBEC3 cytidine deaminases but not murine APOBEC3, *Virology* 454–455 (2014) 215–226.
- [74] C.J. Warren, T. Xu, K. Guo, L.M. Griffin, J.A. Westrich, D. Lee, P.F. Lambert, M. L. Santiago, D. Pyeon, APOBEC3A functions as a restriction factor of human papillomavirus, *J. Virol.* 89 (2015) 688–702.
- [75] S. Stavrou, S.R. Ross, APOBEC3 proteins in viral immunity, *J. Immunol.* 195 (2015) 4565–4570.
- [76] Z. Wang, K. Wakae, K. Kitamura, S. Aoyama, G. Liu, M. Koura, A.M. Monjurul, I. Kukimoto, M. Muramatsu, APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation, *J. Virol.* 88 (2014) 1308–1317.
- [77] C.J. Warren, J.A. Westrich, K.V. Doorslaer, D. Pyeon, Roles of APOBEC3A and APOBEC3B in human papillomavirus infection and disease progression, *Viruses* 9 (2017).
- [78] V.C. Vieira, B. Leonard, E.A. White, G.J. Starrett, N.A. Temiz, L.D. Lorenz, D. Lee, M.A. Soares, P.F. Lambert, P.M. Howley, R.S. Harris, Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B, *mBio* 5 (2014), e02234, 14.
- [79] J.A. Westrich, C.J. Warren, M.J. Klausner, K. Guo, C.-W. Liu, M.L. Santiago, D. Pyeon, Human papillomavirus 16 E7 stabilizes APOBEC3A protein by inhibiting Cullin 2-dependent protein degradation, *J. Virol.* 92 (2018), e01318, 17.
- [80] L. Mirabello, M. Yeager, K. Yu, G.M. Clifford, Y. Xiao, B. Zhu, M. Cullen, J. F. Boland, N. Wentzensen, C.W. Nelson, T. Raine-Bennett, Z. Chen, S. Bass, L. Song, Q. Yang, M. Steinberg, L. Burdett, M. Dean, D. Roberson, J. Mitchell, T. Lorey, S. Franceschi, P.E. Castle, J. Walker, R. Zuna, A.R. Kreimer, D. C. Beachler, A. Hildesheim, P. Gonzalez, C. Porras, R.D. Burk, M. Schiffman, HPV16 E7 genetic conservation is critical to carcinogenesis, *Cell* 170 (2017) 1164–1174.e6.
- [81] P. Aksoy, E.Y. Gottschalk, P.I. Meneses, HPV entry into cells, *Mutat. Res. Rev. Mutat. Res.* 772 (2017) 13–22.
- [82] P.M. Day, M. Schelhaas, Concepts of papillomavirus entry into host cells, *Curr Opin Virol* 4 (2014) 24–31.
- [83] S. DiGiuseppe, W. Luszczek, T.R. Keiffer, M. Bienkowska-Haba, L.G.M. Guion, M. J. Sapp, Incoming human papillomavirus type 16 genome resides in a vesicular compartment throughout mitosis, *Proc. Natl. Acad. Sci. USA* 113 (2016) 6289–6294.
- [84] R.M. Richards, D.R. Lowy, J.T. Schiller, P.M. Day, Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 1522–1527.
- [85] C. Cerqueira, P. Samperio Ventayol, C. Voageley, M. Schelhaas, Kallikrein-8 proteolytically processes human papillomaviruses in the extracellular space to facilitate entry into host cells, *J. Virol.* 89 (2015) 7038–7052.
- [86] P.M. Day, A.S. Weisberg, C.D. Thompson, M.M. Hughes, Y.Y. Pang, D.R. Lowy, J. T. Schiller, Human papillomavirus 16 capsids mediate nuclear entry during infection, *J. Virol.* 93 (2019), e00454, 19, /jvi/93/15/JVI.00454-19.atom.
- [87] M. Schelhaas, B. Shah, M. Holzer, P. Blattmann, L. Kühling, P.M. Day, J. T. Schiller, A. Helenius, Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis, *PLoS Pathog.* 8 (2012), e1002657.
- [88] A.G. Bowie, L. Unterholzner, Viral evasion and subversion of pattern-recognition receptor signalling, *Nat. Rev. Immunol.* 8 (2008) 911–922.
- [89] M.R. Thompson, J.J. Kaminski, E.A. Kurt-Jones, K.A. Fitzgerald, Pattern recognition receptors and the innate immune response to viral infection, *Viruses* 3 (2011) 920–940.
- [90] T. Kawai, S. Akira, Innate immune recognition of viral infection, *Nat. Immunol.* 7 (2006) 131–137.
- [91] C.L. Gupta, S. Akhtar, A. Wayne, N.R. Pandey, N. Pathak, P. Bajpai, Cross talk between Leishmania donovani CpG DNA and Toll-like receptor 9: an immunoinformatics approach, *Biochem. Biophys. Res. Commun.* 459 (2015) 424–429.
- [92] J. Thompson, A. Iwasaki, Toll-like receptors regulation of viral infection and disease, *Adv. Drug Deliv. Rev.* 60 (2008) 786–794.
- [93] M. Escalera-Zamudio, M.L. Zepeda-Mendoza, E. Loza-Rubio, E. Rojas-Anaya, M. L. Méndez-Ojeda, C.F. Arias, A.D. Greenwood, The evolution of bat nucleic acid-sensing Toll-like receptors, *Mol. Ecol.* 24 (2015) 5899–5909.
- [94] J. Pohar, D. Lainšček, R. Fukui, C. Yamamoto, K. Miyake, R. Jerala, M. Benčina, Species-specific minimal sequence motif for oligodeoxyribonucleotides activating mouse TLR9, *J. Immunol.* 195 (2015) 4396–4405.
- [95] A.M. Krieg, A.-K. Yi, S. Matson, T.J. Waldschmidt, G.A. Bishop, R. Teasdale, G. A. Koretzky, D.M. Klinman, CpG motifs in bacterial DNA trigger direct B-cell activation, *Nature* 374 (1995) 546–549.
- [96] A.K. Yi, M. Chang, D.W. Peckham, A.M. Krieg, R.F. Ashman, CpG oligodeoxyribonucleotides rescue mature spleen B cells from spontaneous apoptosis and promote cell cycle entry, *J. Immunol. Baltim. Md* 1950 160 (1998) 5898–5906.
- [97] G. Hartmann, A.M. Krieg, Mechanism and function of a newly identified CpG DNA motif in human primary B cells, *J. Immunol.* 164 (2000) 944–953.
- [98] N.A. Wallace, K. Münger, The curious case of APOBEC3 activation by cancer-associated human papillomaviruses, *PLoS Pathog.* 14 (2018), e1006717.
- [99] C.A. Ishak, A.E. Marshall, D.T. Passos, C.R. White, S.J. Kim, M.J. Cecchini, S. Ferwati, W.A. MacDonald, C.J. Howlett, I.D. Welch, S.M. Rubin, M.R.W. Mann, F.A. Dick, An RB-EZH2 complex mediates silencing of repetitive DNA sequences, *Mol. Cell.* 64 (2016) 1074–1087.
- [100] A. Koito, T. Ikeda, Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases, *Front. Microbiol.* 4 (2013).
- [101] U.A. Hasan, C. Zannetti, P. Parroche, N. Goutagny, M. Malfroy, G. Roblot, C. Carreira, I. Hussain, M. Müller, J. Taylor-Papadimitriou, D. Picard, B.S. Sylla, G. Trinchieri, R. Medzhitov, M. Tommasino, The Human papillomavirus type 16 E7 oncoprotein induces a transcriptional repressor complex on the Toll-like receptor 9 promoter, *J. Exp. Med.* 210 (2013) 1369–1387.
- [102] U.A. Hasan, E. Bates, F. Takeshita, A. Biliato, R. Accardi, V. Bouvard, M. Mansour, I. Vincent, L. Gissmann, T. Iftner, M. Sideri, F. Stubenrauch, M. Tommasino, TLR9 expression and function is abolished by the cervical cancer-associated human papillomavirus type 16, *J. Immunol.* 178 (2007) 3186–3197.
- [103] L. Pacini, C. Savini, R. Ghittoni, D. Saidj, J. Lamartine, U.A. Hasan, R. Accardi, M. Tommasino, Downregulation of toll-like receptor 9 expression by beta human papillomavirus 38 and implications for cell cycle control, *J. Virol.* 89 (2015) 11396–11405.
- [104] L. Pacini, M.G. Ceraolo, A. Venuti, G. Melita, U.A. Hasan, R. Accardi, M. Tommasino, UV radiation activates toll-like receptor 9 expression in primary human keratinocytes, an event inhibited by human papillomavirus 38 E6 and E7 oncoproteins, *J. Virol.* 91 (2017).
- [105] C. Scagnolari, F. Cannella, A. Pierangeli, R. Mellinger Pilgrim, G. Antonelli, D. Rowley, M. Wong, S. Best, D. Xing, R.B.S. Roden, R. Viscidi, Insights into the role of innate immunity in cervicovaginal papillomavirus infection from studies using gene-deficient mice, *J. Virol.* 94 (2020), e00087, 20.
- [106] I. Fathallah, P. Parroche, H. Gruffat, C. Zannetti, H. Johansson, J. Yue, E. Manet, M. Tommasino, B.S. Sylla, U.A. Hasan, EBV latent membrane protein 1 is a negative regulator of TLR9, *J. Immunol. Baltim. Md* 1950 185 (2010) 6439–6447.
- [107] I.E. Vincent, C. Zannetti, J. Lucifora, H. Norder, U. Protzer, P. Hainaut, F. Zoulim, M. Tommasino, C. Trépo, U. Hasan, I. Chemin, Hepatitis B virus impairs TLR9 expression and function in plasmacytoid dendritic cells, *PLoS One* 6 (2011), e26315.
- [108] N. Shahzad, M. Shuda, T. Gheit, H.J. Kwun, I. Cornet, D. Saidj, C. Zannetti, U. Hasan, Y. Chang, P.S. Moore, R. Accardi, M. Tommasino, The T antigen locus of Merkel cell polyomavirus downregulates human Toll-like receptor 9 expression, *J. Virol.* 87 (2013) 13009–13019.
- [109] K. Van Doorslaer, V. Ruoppolo, A. Schmidt, A. Lescroël, D. Jongsojmit, M. Elrod, S. Kraberger, D. Stainton, K.M. Dugger, G. Ballard, D.G. Ainsley, A. Varsani, Unique genome organization of non-mammalian papillomaviruses provides insights into the evolution of viral early proteins, *Virus Evol.* 3 (2017) vex027.
- [110] K.M. King, K. Van Doorslaer, Building (viral) phylogenetic trees using a maximum likelihood approach, *Curr. Protoc. Microbiol.* 51 (2018) e63.
- [111] N. Pornputtapong, D.A. Acheampong, P. Patumcharoenpol, P. Jenjaroenpun, T. Wongsurawat, S.-R. Jun, S. Yongkiettrakul, N. Chokesajjawatee, I. Nookaew, KITSUNE: a tool for identifying empirically optimal K-mer length for alignment-free phylogenomic analysis, *Front. Bioeng. Biotechnol.* 8 (2020), 556413.
- [112] G. Bernard, C.X. Chan, M.A. Ragan, Alignment-free microbial phylogenomics under scenarios of sequence divergence, genome rearrangement and lateral genetic transfer, *Sci. Rep.* 6 (2016), 28970.
- [113] G. Bernard, C.X. Chan, Y. Chan, X.-Y. Chua, Y. Cong, J.M. Hogan, S.R. Maetschke, M.A. Ragan, Alignment-free inference of hierarchical and reticulate phylogenomic relationships, *Briefings Bioinf.* 20 (2019) 426–435.
- [114] S.V. Thankachan, S.P. Chockalingam, Y. Liu, A. Krishnan, S. Aluru, A greedy alignment-free distance estimator for phylogenetic inference, *BMC Bioinf.* 18 (2017) 238.

- [115] Q. Zhang, S.-R. Jun, M. Leuze, D. Ussery, I. Nookaew, Viral phylogenomics using an alignment-free method: a three-step approach to determine optimal length of k-mer, *Sci. Rep.* 7 (2017), 40712.
- [116] J. Choi, S.-H. Kim, Whole-proteome tree of life suggests a deep burst of organism diversity, *Proc. Natl. Acad. Sci. USA* 117 (2020) 3678–3686.
- [117] E. Beutler, T. Gelbart, J.H. Han, J.A. Koziol, B. Beutler, Evolution of the genome and the genetic code: selection at the dinucleotide level by methylation and polyribonucleotide cleavage, *Proc. Natl. Acad. Sci. U. S. A.* 86 (1989) 192–196.
- [118] L. Duret, N. Galtier, The covariation between Tpa deficiency, CpG deficiency, and G+C content of human isochores is due to a mathematical artifact, *Mol. Biol. Evol.* 17 (2000) 1620–1625.
- [119] K.A. Lang Kuhs, D.L. Faden, L. Chen, D.K. Smith, M. Pinheiro, C.B. Wood, S. Davis, M. Yeager, J.F. Boland, M. Cullen, M. Steinberg, S. Bass, X. Wang, P. Liu, M. Mehrad, T. Tucker, J.S. Lewis, R.L. Ferris, L. Mirabello, Genetic variation within the human papillomavirus type 16 genome is associated with oropharyngeal cancer prognosis, *Ann. Oncol.* 33 (2022) 638–648.
- [120] B. Zhu, Y. Xiao, M. Yeager, G. Clifford, N. Wentzensen, M. Cullen, J.F. Boland, S. Bass, M.K. Steinberg, T. Raine-Bennett, D. Lee, R.D. Burk, M. Pinheiro, L. Song, M. Dean, C.W. Nelson, L. Burdett, K. Yu, D. Roberson, T. Lorey, S. Franceschi, P. E. Castle, J. Walker, R. Zuna, M. Schiffman, L. Mirabello, Mutations in the HPV16 genome induced by APOBEC3 are associated with viral clearance, *Nat. Commun.* 11 (2020) 886.
- [121] V.N. Pimenoff, C.M. de Oliveira, I.G. Bravo, Transmission between archaic and modern human ancestors during the evolution of the oncogenic human papillomavirus 16, *Mol. Biol. Evol.* 34 (2017) 4–19.
- [122] Z. Chen, R. DeSalle, M. Schiffman, R. Herrero, C.E. Wood, J.C. Ruiz, G.M. Clifford, P.K.S. Chan, R.D. Burk, Niche adaptation and viral transmission of human papillomaviruses from archaic hominins to modern humans, *PLoS Pathog.* 14 (2018), e1007352.
- [123] C.A. Mangold, M.M. Rathbun, D.W. Renner, C.V. Kuny, M.L. Szpara, Viral infection of human neurons triggers strain-specific differences in host neuronal and viral transcriptomes, *PLoS Pathog.* 17 (2021), e1009441.
- [124] M.M. Rathbun, M.L. Szpara, A holistic perspective on herpes simplex virus (HSV) ecology and evolution, p. 27–57, in: *Advances in Virus Research*, Elsevier, 2021.
- [125] M.M. Rathbun, M.M. Shipley, C.D. Bowen, S. Selke, A. Wald, C. Johnston, M. L. Szpara, Comparison of herpes simplex virus 1 genomic diversity between adult sexual transmission partners with genital infection, *PLoS Pathog.* 18 (2022), e1010437.
- [126] B.L. Uhrlorn, R. Jackson, S. Li, S.M. Bratton, K. Van Doorslaer, S.K. Campos, Vesicular trafficking permits evasion of cGAS/STING surveillance during initial human papillomavirus infection, *PLoS Pathog.* 16 (2020), e1009028.
- [127] L. Lau, E.E. Gray, R.L. Brunette, D.B. Stetson, DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway, *Science* 350 (2015) 568–571.
- [128] X. Luo, C.R. Donnelly, W. Gong, B.R. Heath, Y. Hao, L.A. Donnelly, T. Moghbeli, Y.S. Tan, X. Lin, E. Bellile, B.A. Kansy, T.E. Carey, J.C. Brenner, L. Cheng, P. J. Polverini, M.A. Morgan, H. Wen, M.E. Prince, R.L. Ferris, Y. Xie, S. Young, G. T. Wolf, Q. Chen, Y.L. Lei, HPV16 drives cancer immune escape via NLRX1-mediated degradation of STING, *J. Clin. Invest.* (2019), <https://doi.org/10.1172/JCI129497>.
- [129] L.V. Ronco, A.Y. Karpova, M. Vidal, P.M. Howley, Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity, *Genes Dev.* 12 (1998) 2061–2072.
- [130] I. Tobey, B. Uhrlorn, R. Jackson, K. King, S. Campos, K. Van Doorslaer, Human Papillomavirus Oncogenes Limit cGAMP Production, N/A, 2022. Unpublished, In preparation.
- [131] E. Gusho, L.A. Laimins, Human papillomaviruses sensitize cells to DNA damage induced apoptosis by targeting the innate immune sensor cGAS, *PLoS Pathog.* 18 (2022), e1010725.