Authors for correspondence:
Robin Fahraeus
e-mail: robin.fahraeus@inserm.fr
Alicja Sznarkowska
e-mail: alicja.sznarkowska@ug.edu.pl

Viruses, cancer and non-self recognition

Monikaben Padariya[1], Umesh Kalathiya[1], Sara Mikac[1], Katarzyna Dziubek[1], Maria C. Tovar Fernandez[1], Ewa Sroka[1], Robin Fahraeus[1,2,3,4] and Alicja Sznarkowska[1]

[1]International Centre for Cancer Vaccine Science, University of Gdansk, Kladki 24, 80-822 Gdansk, Poland
[2]Inserm UMRS1131, Institut de Génétique Moléculaire, Université Paris 7, Hôpital St. Louis, F-75010 Paris, France
[3]RECAMO, Masaryk Memorial Cancer Institute, Zluty kopec 7, 65653 Brno, Czech Republic
[4]Department of Medical Biosciences, Umeå University, Building 6M, 901 85 Umeå, Sweden

MP, 0000-0002-9132-8620; UK, 0000-0001-7757-6962; SM, 0000-0002-0591-3090; RF, 0000-0003-0402-8492; AS, 0000-0002-9461-1878

Virus–host interactions form an essential part of every aspect of life, and this review is aimed at looking at the balance between the host and persistent viruses with a focus on the immune system. The virus–host interaction is like a cat-and-mouse game and viruses have developed ingenious mechanisms to manipulate cellular pathways, most notably the major histocompatibility (MHC) class I pathway, to reside within infected cell while evading detection and destruction by the immune system. However, some of the signals sensing and responding to viral infection are derived from viruses and the fact that certain viruses can prevent the infection of others, highlights a more complex coexistence between the host and the viral microbiota. Viral immune evasion strategies also illustrate that processes whereby cells detect and present non-self genetic material to the immune system are interlinked with other cellular pathways. Immune evasion is a target also for cancer cells and a more detailed look at the interfaces between viral factors and components of the MHC class I peptide-loading complex indicates that these interfaces are also targets for cancer mutations. In terms of the immune checkpoint, however, viral and cancer strategies appear different.

1. Introduction

Viruses have been, and still are, an invaluable source for understanding various aspects of cell biology. The Rous sarcoma virus revealed viral oncogenesis while simian, human adeno (HAdV) and human papilloma (HPV) viruses paved the way for the identification of the p53 and the Retinoblastoma protein (pRB) tumour suppressors. Similarly, viruses have played an important role in our understanding of how the immune system distinguishes self from non-self, both in terms of how the innate response detects pathogen patterns and how the adaptive immune system specifically detects the presence of neoantigens on the major histocompatibility class (MHC) molecules. Different strategies employed by viruses to evade the immune system reflect the balance of specific virus–host interactions and point out distinct steps in antigen presentation and the strategies best serving each virus. Virus–host interactions are commonly seen as good versus bad, but this is an oversimplification. In fact, more recent works illustrate how viruses have adapted to the hosts' immune system and how they have helped to create it, not only as a consequence but also as a cause, pointing towards a more balanced picture that reflects the coevolutionary origin of viruses and their hosts [1–4].

There are several mechanisms in place to help the immune system distinguish between self and non-self. The detection of pathogen-associated molecules by innate cells such as dendritic (DC) macrophages results in the release of

inflammatory cytokines and interferons that promote the activation of the adaptive response or, in the case of NK cells, the elimination of stressed cells. The DCs play an important role in the interface between the innate and the adaptive immune system by presenting antigens on their MHC class I and II molecules that helps activate CD4+ and CD8+ T cells. Specific T cell receptors (TCR) on CD8+ T cells detect 8–10 amino acids peptides presented on MHC class I molecules and form a cornerstone in the immune system's capacity to distinguish self from non-self and for the elimination of virus-infected, damaged or transformed cells. Consequently, the evasion of the MHC class I pathway plays a major role in persistent viral strategies to minimize the risk of the infected host cell being detected and removed by the immune system. Similarly, cancer cells expressing neoantigens also need to evade the immune system and use similar—but also (as we will describe) different—strategies. However, the underlying processes of how cells select antigens to be presented on MHC class I molecules are still not fully understood, and a main challenge for the immune system to stay sensitive to pathogen infections or to detect mutated genes is the fact that the number of MHC class I molecules are very few compared to the number of peptides produced [5–7]. Also, each T cell expresses one TCR and the number of T cells is much lower than the number of potential antigenic peptides [8,9]. How the immune system 'deals' with this and stays sensitive to non-self is not yet fully understood.

Our limited understanding of virus–host coevolution results from the perspective that viruses are mainly disease-causing agents (run-away acute replicators) and the successful host escapes from the disease, resulting in a constant 'arms race' between viruses (predator) and their hosts (prey). However, acute replication is not the predominant lifestyle of viruses, nor is it conserved on the long-term scale, but mainly serves to explore new host possibilities [1,2,10]. The predominant successful and evolutionarily conserved viral strategy is persistence, which is the capacity of a functional or defective virus to colonize an individual host and be maintained or re-emerge at some later time [2]. It results in an intimate relationship between the virus and the host that relies on the coordinated expression of the virus and host's innate and adaptive immunity genes. Persistence tends to be genetically stable and is also highly species-and tissue-specific, demonstrated by the potential devastating effects when a virus jumps from one tissue or species to another, or when the host environment is disrupted [11–13]. Understanding the finer nuances of the persistence of a particular virus in a particular host tissue allows one to understand why a virus wreaks havoc when jumping from one host environment to another and how it affects a cellular pathway in one cell type, but not in another. Persistent functional viruses need to replicate and they need to avoid the immune system; traditionally, these two cell biological aspects have been studied separately, but a more recent example of mammalian herpes virus illustrates how viral interference with growth-promoting and antigen presentation pathways are linked [14–17].

Our close relationship with viruses is illustrated by the number of functional persistent viruses we carry and the persistence of viral remanences in our genomes. Endogenized retroviruses (ERVs) and their defectives—long and short interspersed elements (LINEs, SINEs), long terminal repeats (LTRs), Alus—constitute a majority of our 'junk' (non-protein coding) DNA [18–20]. The solo LTRs originated from full virus integrations and are present in 330 000 copies, as compared to roughly 23 000 genes [21]. The role of ERVs in the divergence of placental species is commonly accepted: Syncytins that mediate cell to cell fusion and immunosuppression during placenta formation (syncytiotrophoblast) have originated from retroviral env genes [22–24]. Each mammalian lineage has syncytins coming from distinct ERVs as a result of independent gene captures occurring separately in the genome of each lineage, making the placenta the most variable organ in mammalian species [24]. Moreover, their expression is also controlled by a complex network of ERVs (mainly their LTRs), derived from distinct retroviral integrations and expressed as non-coding RNAs [24].

Functional viruses that establish a life-long persistence (latency) in humans, such as the gamma herpesviruses Epstein–Barr virus (EBV), human cytomegalovirus (HCMV) and Kaposi's sarcoma-associated herpesvirus (KSHV) are present in the majority of the population and each virus has evolved a balance with a specific host cell type to ensure proliferation and replication without causing too much harm while at the same time preventing detection and destruction by the host's immune system. The EBV infects approximately 95% of the population and its prime resident host cell is the resting B cell [25–27]. HCMV infects 70–100% and establishes latency in CD34+ hematopoietic progenitor and myeloid lineage cells [28,29]. The proportion of people infected with KSHV is about 10–40%, with important geographic variations [30,31]. Most herpes virus infections will go unnoticed unless the host's immune system is in an unbalanced state, either via infections like HIV, or via chemically induced immune suppression in connection with, for example, organ transplantation. In such cases, the viruses can cause severe symptoms, including lymphoproliferative disorders and cancers. This illustrates how finely tuned the normal virus–host interactions are and how their cell-transforming capacities are kept in check by the immune system under normal conditions. EBV, for example, is one of the most efficient cell-transforming agents known and exposure of B cells to the virus in vitro results in the establishment of immortalized cell lines [32–34]. It is not clear how the balance between the virus's cell-transforming capacity and the immune system has evolved, and whether the immune system actively represses the cell-transforming capacity of the virus or the immortalized cells are actively suppressed. Nevertheless, it points towards a close relationship between the molecular mechanisms of viral immune escape and the regulation of cell growth and proliferation [35,36]. EBV infection prior to adolescence is non-symptomatic, but causes mononucleosis later, with increasingly severe symptoms as we get older. It is interesting that the establishment of persistent infection at a young age does not give symptoms, while many of the acute viruses cause diseases in the young, and it is worth considering that our viral microbiota helps protect from more harmful viral infections, similar to the establishment of the bacterial microbiota. In this respect, it is noteworthy that the SARS-CoV-2 (or COVID-19) gives relatively few symptoms in most younger individuals, suggesting that this virus might become the latest on a long list to form part of our virosphere. Interesting implications of viral persistence to the survival of the host population came from studies on the mouse hepatitis virus (MHV), a murine coronavirus. Mice caught in the wild are persistently colonized with an array of viruses in 80–90% of cases, including

MHV, and are perfectly healthy. However, introducing a mouse from the wild to inbred laboratory mice can have a devastating effect as the viruses of the wild mouse interferes with the reproductive biology of the uninfected colony, or may acutely infect pups, depending on the virus strategy, and cause a reproduction collapse. In this case, the persistent version is the winning outcome as it has consequences to the population survival [37,38].

2. Viruses and self-identity

The capacity of persistent, defective or functional viruses to protect or to kill cells constitutes a creative and competitive force to understand the survival and evolution of life, and it is interesting to consider how it has evolved. Exchange of information played an essential role in building the particles that eventually could form cells and replicate. These 'building blocks' were RNA consortia, later linked with amino acids or other smaller molecules to enhance their identity [39]. Once cells were formed, the exchange of genetic material remained an evolutionary driver, but with the presence of cells, the exchange took the form of virus-like particles with an active capacity to go between cells. As more advanced multicellular organisms evolved, the potential toxic effects of viral infections required a detection system that could prevent some infections while allowing those that were recognized as 'consortium members'. Interestingly, viruses actively provided the capacity of the host to detect and destroy virus-infected cells, and antiviral defences such as restriction/modification or CRISPR/Cas systems in prokaryotes, RNA interference system (very effective in e.g. *C. elegans*) or innate and adaptive immune response of higher eukaryotes all have viral origins and can be seen as systems enhancing the host identity [4]. Phylogenetic analysis indicates that the most basal members of the TCR genes are the viral *JAM/CTX/PVR* genes, also called the CTX-like family of viral receptors, that support viral interactions and often function as specific virus receptors (PVR, polio virus receptor; JAM, reovirus receptor) [40,41]. Hence, not only our 'junk' DNA has a viral origin.

The persistent state of a virus as an 'addiction module' was initially discovered in *Escherichia coli* after daughter cells that lost the P1 phage died. The P1 phage produces a stable toxin and a less stable anti-toxin, and thus, cells that lose P1, or for other reasons stop producing the anti-toxin, are killed by the stable toxin using the cell's own programmed cell death systems [42]. If another phage (such as IS2) tries to colonize a P1-carrying cell, it causes an imbalance in toxin/anti-toxin and the cell will die, and thus the concept of addiction acting as an antiviral defence [38,43,44]. The state of one genetic parasite being colonized by another (hyperparasite colonization) is common and affects the relationship of the host with other viruses [44]. The addiction concept to provide resistance to a diverse set of phages and viruses has implications for the colonization of our own microbiota, both in terms of bacteria [45–47] and also of viruses, and has interesting implications for the origin of the viral immune response in higher organisms. A self-destruction mechanism of the infected cells based on a toxin/anti-toxin system (addiction module) is observed in the white spot syndrome virus (WSSV), which is a well-characterized DNA virus causing large crashes in shrimp populations at commercial shrimp farms by producing a pro-apoptotic toxin. In some species, however, the WSSV itself produces an anti-toxin and does not cause any disease symptoms [48,49].

The adaptive immune system first appeared in jawed fish and includes recombination activating genes RAG1 and 2, immunoglobulins, T cells with their receptors (TCRs) and MHC class I and class II molecules [50–53]. There is, however, no living organism representing a gradual acquisition of components of this system, and the rapid occurrence of the adaptive immunity in jawed vertebrates has been attributed to genome duplication due to the colonization of mixed population of viruses (both DNA viruses and retroviruses) and occurrence of endogenous viruses (gypsy-like ERVs) and their defectives (Alus, LINEs, SINEs) [4,54]. In accordance with the virus addiction concept, such a large-scale colonization provides a selective pressure to resist viral competition and an ability to recognize and exclude other virus-infected self cells. Thus, it largely enhanced the self identity of the jawed fish, which now allows recognition of virus-infected self cells. Consequently, both DNA viruses and retroviruses played a role in forming the network of adaptive immune response via providing genes and their regulation, and many of the components of adaptive immunity indeed have a viral-like origin (table 1) [4,54]. For example, the human MHC class I region has 16 distinct human ERVs (HERVs) and the density of the retroelements in this region is 10 times greater than in other regions of the chromosome. These retroelements give rise to a number of non-coding RNAs (ncRNAs) that regulate the expression of various genes, including (but not restricted to) those involved in antigen processing and presentation [58,59,65,66]. One of the most widely studied is long ncRNA HCP5 (*Human Leukocyte Antigen pseudogene P5*), which evolved from an ancient HERV16 insertion (approx. 37 million years ago) and later sequestered the MHC promoter and enhancer region from an ancient *HLA* class I gene [60]. *HCP5* is an antisense retroviral transcript that interacts with regulatory micro RNAs but also binds transcription factors, enhancers and chromatin remodelling enzymes. It is involved in adaptive and innate immune responses and associates with the promotion of some autoimmune diseases and cancers.

3. The major histocompatibility class I pathway and viral interference

Peptide substrates of approximately 30 amino acids or longer need a proteasomal activity to be presented by MHC class I molecules. The TAP peptide transporter consists of TAP1 and TAP2, and is located in the membrane of the endoplasmic reticulum (ER) and brings peptides into the ER where they are further trimmed by the ERAPs1–2 aminopeptidases to 8–10 amino acid lengths before introduced to the peptide-loading complex (PLC). The PLC is linked to the TAP and includes the MHC class I molecule, B2-microglobulin, Tapasin and Erp57 that helps stabilize the complex before peptide loading and export to the cell surface via the Golgi (figure 1). Most persistent viruses have been shown to interfere with the MHC class I pathway, and perhaps this holds true for all viruses of this category. Different viruses target different components of the TAP-PLC complex. This might seem surprising and one could argue that if a virus found a clever way to interfere with a specific component of the TAP-PLC to bypass the MHC class I pathway, this would

Table 1. Viral impact on the origin and evolution of immune system.

components of innate and adaptive immune response	relation to viruses	Ref.
interferon gamma network	binding sites for IFN gamma-regulated transcription factors STAT1 and IRF1 are composed of LTR regions from endogenized gammaretrovirus family MER41 that invaded the genome of an anthropoid primate ancestor approximately 45–60 million years ago	[55]
immunoglobulin (Ig) receptors [T-cell receptors (TCRs), B-cell receptors]	the basic structure of Ig receptors is found within CTX (*cortical thymocyte marker of Xenopus*)-like family of viral receptors including junctional adhesion molecule (JAM), poliovirus receptor (PVR), coxsackie and adenovirus 5 receptor (CAR), signalling lymphocytes activation markers (SLAM)	[40,56]
RAG1/2 system	it acts like 'cut and paste' transposon—the most basal version is that of phage Mu; it also shows similarity to the RNAse H fold of retroviral integrase	[54,57]
MHC region	human MHC region is particularly dense with viral-derived elements (ERVs, Alu, SINEs, LINEs) which contribute to its evolution as e.g. recombination sites for duplication; they give rise to numerous regulatory ncRNAs (miRNAs, lncRNAs) that regulate expression of various genes	[58–64]

be copied by them all. However, if one sees the TAP-PLC as one functional unit, then this is indeed the case.

Herpes simplex (HSV) factor ICP47, the US6 of the HCMV as well as the BNLF2 of the EBV all target TAP. Interestingly, though patients with TAP deficiency syndrome do suffer from recurrent bacterial infections of the upper respiratory tract manifesting within the first 6 years of life, severe viral infections are noticeably absent and normal antibody titres against several viruses have been demonstrated in most patients [70–72]. This can perhaps be explained by more recent observations showing that viral peptides are presented to the class I pathway in a TAP-independent fashion [73–76]. Furthermore, fusing the bacterial multidrug resistance proteins A and B (TmrAB) transporter to the fragment of the transmembrane domain of TAP2 mediating interaction with tapasin restores antigen presentation in cells lacking TAP, indicating that a major role of TAP in antigen presentation is via its role in regulating PLC dynamics [77].

A question regarding peptide transport is what happens to peptides that are not loaded on the class I molecules. There is not much selection for which peptides that are transported into the ER by the TAP except for a few charged residues in the N- and C-termini. Considering the vast amount of peptide products produced by the proteasome every minute and the relative few MHC class I molecules available for peptide loading, only a very small amount of peptides can be bound to the class I molecules [78,79]. What happens to the rest? The ER is an environment sensitive to the presence of unfolded proteins and when these interact with the BiP chaperone, it triggers the unfolded protein response [80]. If antigenic peptide substrates are too small to interact with BiP, or if the charged N- and C-termini of antigenic peptides substrates is sufficient to prevent BiP recognition and is not known. Misfolded proteins are retrotranslocated out to the cytoplasm for degradation in a ubiquitin-dependent fashion by the ER-associated degradation (ERAD) system [81–83]. It is unlikely that this pathway also expels peptide substrates for the MHC pathway, and it has not been observed that such are ubiquitinated. But

as the ER membrane does not allow passive transport of peptides either, it leaves a gap in our understanding of how the majority of peptides substrates that enter the ER via the TAP are taken care of.

4. Viral and cancer immune evasion

In addition to detecting and destroying virus-carrying cells, our immune system also serves to eliminate damaged or harmful cells. In both viral and cancer immune evasion, the expression of neoantigens on the MHC-I molecules plays a key role, and it could therefore be expected that the targets for viral and cancer immune evasion are similar. Indeed, downregulation of MHC-I molecules is frequent in cancers and also a common viral target. The discovery of transmembrane E3 RING-CH-type ubiquitin ligases (MARCH) came from studies on the Kaposi's sarcoma viral factors kK3 and kK5 and the equivalent cellular factors were discovered later. It is, however, not clear if a virus gave the cells MARCH, or if viruses acquired MARCH from the host. The MARCH proteins (MARCH1–10) are regulated by inflammatory signals and play important roles in various aspects of the immune system including DC differentiation. They control MHC-I and MHC-II internalization via monoubiquitination, or the degradation via lysosomal or proteasomal pathways following polyubiquitination. kK3 and kK5 targets MHC on the cell surface, whereas the MHV86 mK3 sits on the TAP molecules and wait for the nascent MHC-I to enter the ER and promotes ubiquitination and promotes ERAD. The US2 and US11 also target MHC-I for ERAD but via a slightly different mechanism. They cause retrograde transport, or dislocation, of newly synthesized HLA-I heavy chains by hijacking the cellular quality control pathway ERAD, providing it with folding-competent MHC-I to induce its rapid proteasomal degradation [84,85]. MARCH proteins are dysregulated in various cancers, suggesting that viruses and cancers share common strategies in exploiting ubiquitination pathways for immune evasion. A closer examination of the interfaces between viral
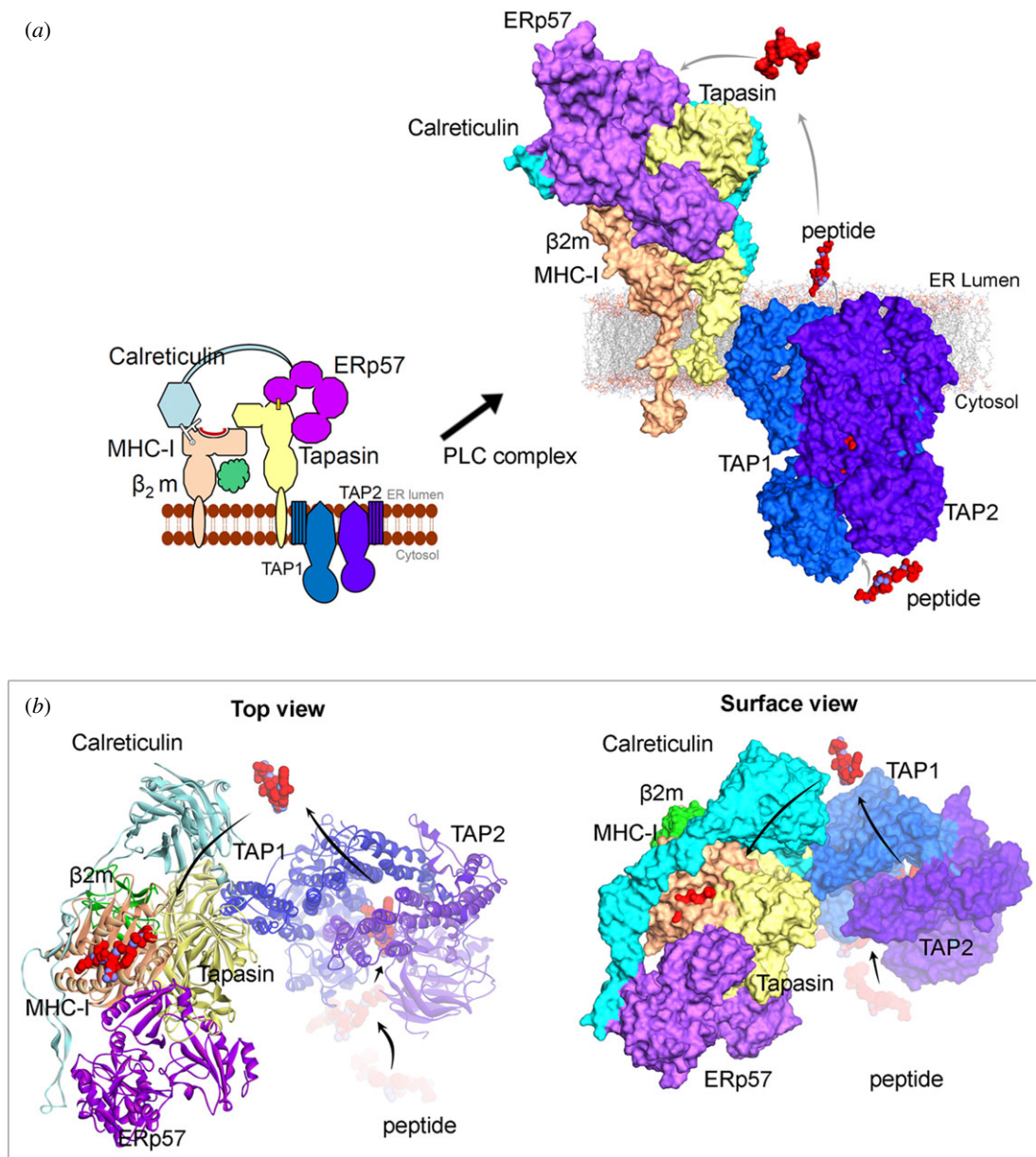
Figure 1. Different components from the peptide-loading complex (PLC) involved in the antigen presentation. (a) Representation of proteins MHC I, β2m, Calreticulin, ERp57, Tapasin, TAP1 and TAP2 and their possible protein–protein interfaces. Right panel describes the possible way of transporting peptide from cytosol-ER lumen, that could be presenting over MHC-I molecules. (b) The top surface view of the PLC, highlighting the peptide-binding cavity from MHC I molecule in the presence of a peptide bound to it. For (a) and (b), the template structures used to model PLC were PDB: 6eny [67] (MHC-I, β2m, Calreticulin, ERp57 and Tapasin) and PDB: 5u1d [68,69](TAP1 and TAP2). The three-dimensional view of the structures were prepared using BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA).

interacting proteins and cancer mutations (figures 2 and 3) further illustrate common viral and cancer strategies. The US2 binds HLA-A2 (MHC I) at the junction of the peptide interactive groove and on the MHC I surface, which allows US2 to bind independently of peptide sequence [87]. The human adenovirus E3/19 K immunomodulatory protein retains HLA-A and HLA-B, but not HLA-C, within the ER [88]. The human cytomegalovirus US3 binds to MHC-I in the ER and interferes with tapasin-dependent peptide loading [96–99]. Moreover, the molluscum contagiosum virus protein MC80 binds to tapasin, leading to its degradation and the loss of TAP proteins [100]. The TAP complex is targeted by several viral factors and the ICP47 (Infected Cell Protein 47) from the herpes simplex virus locks TAP in an inactive conformation that prevents its role in the PLC [69,90,93,94]. Similar, the HCMV-encoded US6 glycoprotein interacts with the TAP1 subunit and blocks peptide entry into the ER.

Furthermore, the ATP-driven conformational change of TAP1 required for peptide transport from cytosol to ER is the target of the US2 [93]. Peptide transport is energy-dependent, and TAP ATP binding and hydrolysis is a target for the cowpox CPXV012 protein [101]. BNLF2a from Epstein–Barr virus subsequently arrests TAP in a translocation-incompetent conformation [94,102].

Using the cancer genome atlas (TCGA) and the cBioPortal database to trace cancer mutations in different HLA types (HLA-A, HLA-B, and HLA-C; figures 1 and 2) suggests that even though there are similarities between HLA types, several mutations occur only in a particular HLA molecule and in the α3 domain and the transmembrane domains [13]. Furthermore, the HLA-A molecules have overall more residues mutated as compared to HLA-B and -C. HLA-C instead shows more mutations in the peptide-binding groove (figure 2b), and it is interesting to note that HLA-C
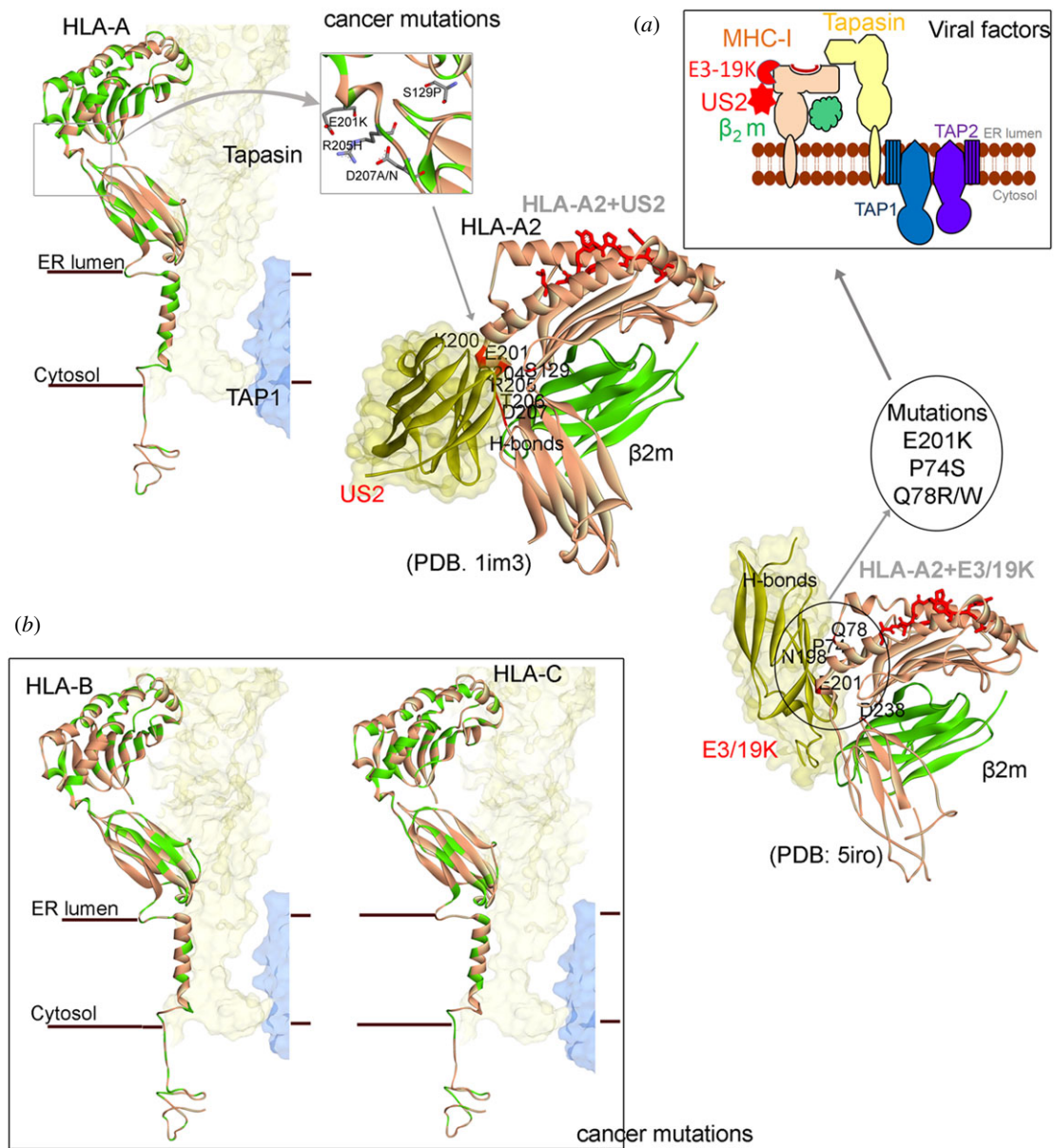
Figure 2. Viral protein binding with different regions of MHC I molecules compared with the amino acid variants (retrieved from cBioPortal [86]) from different cancer types. (*a*) Mutations in HLA-A from different cancer types (mutations are marked in green colour), compared with the viral interaction surface with HLA-A2. The crystal structure of US2 with HLA-A2 (PDB: 1im3 [87]) and E3/19 K with HLA-A2 (PDB: 5iro [88,89]). Hydrogen bond interactions between the viral proteins US2, US3 and E3/19 K with HLA-A2 are marked as red on the protein structure. (*b*) Cancer mutations marked with green colour on the HLA-B and HLA-C. Colour scheme: carbon in grey, nitrogen in blue and oxygen in red. The layover of cancer mutations on structure, and H-bond interactions were identified using BIOVIA Discovery Studio (DassaultSystèmes, BIOVIA Corp., San Diego, CA, USA).

is the major inhibitory ligand for killer immunoglobulin-like receptors (KIRs), which might help to explain why viruses and cancers target this haplotype differently. HLA-C2 is more inhibiting to NK cell activation as compared to HLA-C1 and the E3/19 K does not interact with HLA-C [88]. The TAP1/2 and β2m proteins are frequently mutated in different cancer types (figure 3). In addition, analysing the variations in Tapasin suggests that the Ig-like domain and the c-terminus region are heavily mutated. It is known that the scoop loop is of critical importance for TAPBPR-mediated stabilization of empty MHC-I client molecules [103,104], which is also found mutated in different cancer types.

A majority of viral proteins interact with regions in the PLC which are highly mutated in cancers. For example, the viral factors US2 [87] and E3/19 K [88] form hydrogen bond interactions with the HLA-A2 residues (figure 2) that

are mutated in different cancer types. Similarly, UL18 viral factors bind to β2m residues that are mostly mutated in cancer types (figure 3*b*). In particular, looking at the TAP transporters from the PLC complex, the viral protein ICP47 forms h-bond interactions with four amino acids (S358, E413, Y468 and Q516) of TAP1 protein, from which two residues (S358F and Q516 K) were mutated in cancer (figure 3*a*). Conversely, TAP2 protein has formed more interactions (R210, M218, T246, E248, N250, S251, S255, N436, D432, Y422, S421 and Y477) with ICP47 compared to TAP1, and only one residue (R210 L/Q) is found mutated in cancer (figure 3). With a glance from these observations of the viral proteins (US2, E3/19 K, ICP47 and UL18) and the cancer mutations, it could be suggested that they share a common interface to interfere with the PLC (figures 2 and 3).
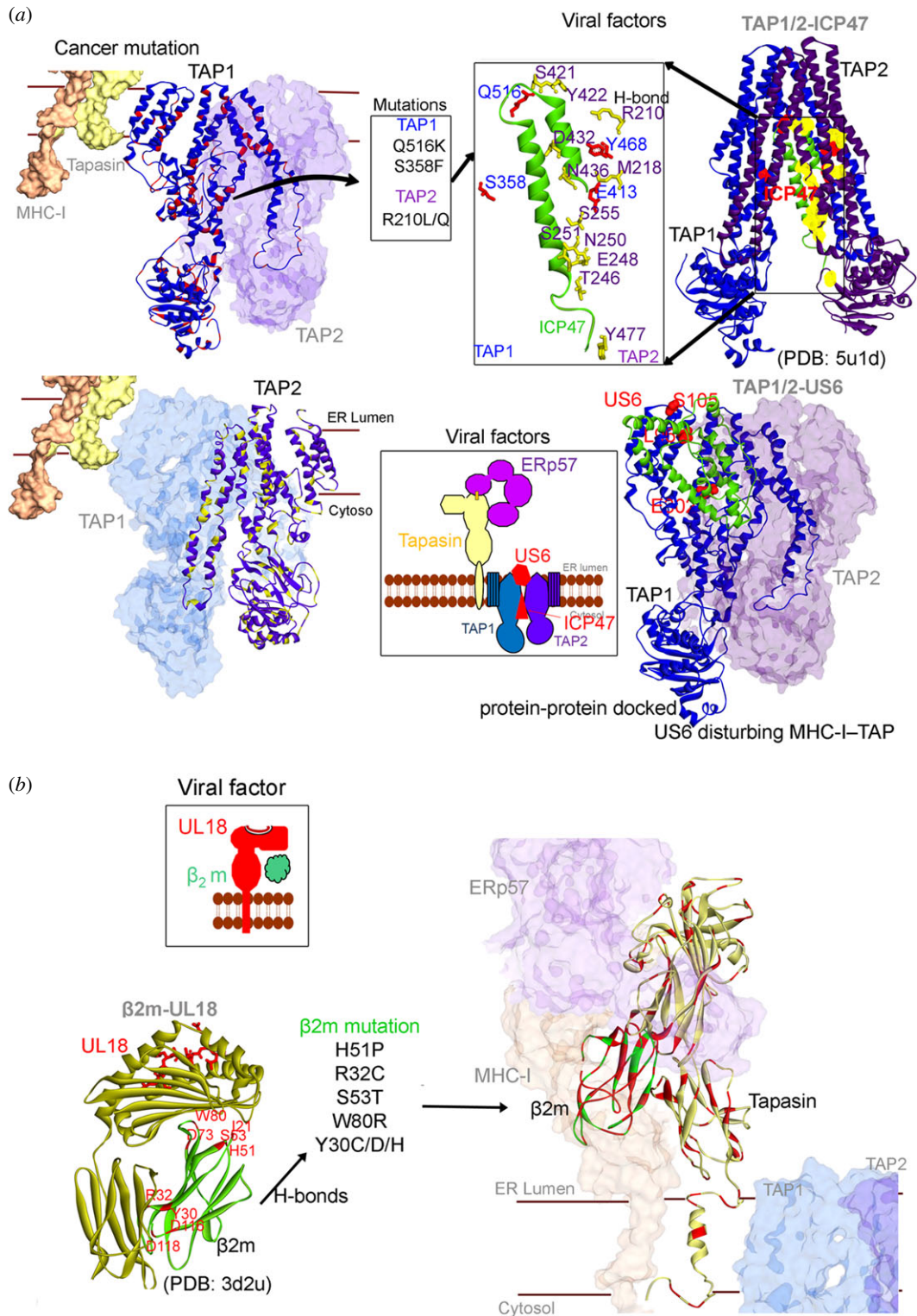
Figure 3. Comparison of cancer mutations (from cBioPortal [86]) with viral proteins interacting with the peptide-loading components. (a) TAP1 and TAP2 cancer mutations, presented as red and yellow, respectively. The viral protein ICP47 crystal structure with TAP1 and TAP2 (PDB; 5u1d [69]); residues involved in H-bond interaction with ICP47 from TAP1 and TAP2 are presented as sticks with labelled colour blue and violet, respectively. The viral protein US6 is suggested to bind the TAP1 and TAP2 blocking the peptide transport process [90]. For the US6 viral factor, its structure was modelled using homology modelling approach (SWISS-MODEL [91]) and the protein–protein docking was performed with TAP1 (using ZDOCK [92]), resulting US6-TAP1/TAP2 conformation correlates with the data available in the literature [90,93,94]. H-bond formed between US6 and TAP1 is marked in red colour. (b) The β2m and Tapasin protein cancer mutations marked in red colour. Viral protein UL18 mimicking human MHC I and interacting with β2m (PDB: 3d2u [95]) is shown, residues involved in H-bond are marked in red colour and compared with the cancer mutations. The layover of cancer mutations on structure and H-bond interactions were identified using BIOVIA Discovery Studio (DassaultSystèmes, BIOVIA Corp., San Diego, CA, USA).

5. Impact of viral infections on immune checkpoints

Recent years have seen a rapid expansion in immune-based cancer treatments, most notably using antibodies that target immune checkpoint (IC) receptors PD-1 or the co-regulatory inhibitory cytotoxic T cell receptor CTLA-4, and their ligands PD-L1/2 and CD86 and CD80, respectively. The PD-1 is expressed mainly on T- and B-lymphocytes as well as NK cells and monocytes, whereas the PD-L1 is expressed in most

tissues, including cancer cells. PD-L2 is less abundant and its role is still not completely understood [105,106]. Upon engagement by its ligand, PD-1 acts as a break to T cell cytotoxic activity and PD-1/PD-L1 and cancer cells employ different approaches to exploit this to evade the immune system. The importance of this approach for cancer immune evasion is illustrated by the use of various antibodies that block these interactions for the treatment of various cancers. Considering the importance of the ICs in cancer cell immune evasion, it would be expected that viruses would also target the immune checkpoints. However, the cellular response to acute viral infection is aimed at driving the expression of the PD-1/PD-L1 axis, perhaps to act as a rheostat to regulate the strength and duration of the immune response [107]. A direct interference with PD-1/PD-L1 axis by viral antigens has not been reported but *Herpesviridae* [108] produces a protein that mimics IL-10 and further increases PD-1/PD-L1 signalling on immune cells [109]. Furthermore, Kakizaki *et al.* [110] observed increased PD-L1 expression on monocytes that was induced by extracellular vesicles (EVs) released by HBV infected hepatocytes [110]. Prolonged and persistent PD-1/PD-L1 stimulation observed in cancer, and persistent viral infections such as the LCMV and HBV, ultimately lead to T-cell functional exhaustion and reflect a state of adapting the T cell response to chronic inflammation. Studies in a murine model for HCV (Norway rat Hepacivirus) indicated that PD-1/PD-L1 blockade implemented at the early stage of the infection indeed reduced viral load, but no significant improvement was reported as the chronic infection progressed [111]. Consequently, ICs might pose a difficult target for viruses. We do not fully understand how the IC functions and why it is not a suitable viral target.

6. The source of antigenic peptides: from immune evasion to cell cycle control

The first step in antigen presentation is, in fact, not proteasome-mediated degradation of the peptide substrate that is commonly pictured but rather the synthesis of the antigenic peptide substrate [112–114]. The different viral factors mentioned that interfere with the PLC-TAP all act in *trans*, and hence a certain amount of viral factors have to be produced in order to have an immune evasive effect. But what protects the host cell from the immune system before a sufficient level of trans-acting factors has been produced? To fill this gap in their protective repertoire, viruses can also target the production of antigenic peptide substrates [115–117].

Full-length proteins are, in fact, poor substrates for the class I pathway, while non-canonical sources are presented. For example, antigenic peptide substrates can be initiated from leucine codons in the 3′ UTR as well as from pre-spliced mRNAs [118,119]. The nature of the pre-spliced mRNA translation event is still not clear, but it can be distinguished from canonical translation [16,113]. It is not known if such pioneer translation products (PTPs) derived from pre-spliced mRNAs have other functions in the cell apart from an antigenic peptide source but it is plausible that the MHC class I pathway has its 'own' mRNA translation event [120]. The rationale for having alternative translation products rather than full-length proteins as a source of peptide substrates relates to the limited amount of MHC class I molecules as compared with the enormous amount of protein-derived degradation

products derived from full-length proteins and the need for the immune system to stay sensitive to viral antigens [120–123]. Hence, if all peptide products derived from proteasomal degradation of full-length proteins had access to the class I pathway, it would be a lottery whether non-self antigens were to be presented or not. Furthermore, if the first peptide products produced from a viral mRNA are selected for the MHC class I pathway, it ensures a rapid detection of viral infection before viral *trans*-acting factors have a chance to interfere with the class I pathway. An alternative and perhaps specific translation event for the production of antigenic peptide substrates for the MHC class I pathway might seem excessive, but if one considers the fundamental importance of the immune system in evolution, it puts things in a different perspective, and a specific translation event to help detect and eliminating virus-infected cells is perhaps not such an extreme possibility. It is possible that this non-canonical translation event is a reminiscence of a co-transcriptional translation event, similar to the prokaryotic, that became dedicated to producing antigenic peptides for the MHC class I pathway, while the canonical translation machinery evolved with its well-defined and regulated protein synthesis.

One would expect that viruses would have a strategy to counteract an alternative source of antigenic peptides and it appears that one approach is to minimize the translation of the viral mRNAs using cis-acting mechanisms [124]. This is the case for the EBNA1 of the EBV, which is the only viral antigen expressed in resting B cells and in EBV-carrying Burkitt's lymphomas, and compounds that stimulates EBNA1 translation also increase the production of antigenic peptides from the EBNA1 message [124]. The LANA1 of the KSHV is using a similar *cis*-acting mechanisms but it is not yet known if other viruses are also using a similar strategy [116,125,126].

Apart from evading the immune system, latent viruses can also control the proliferation of the host cell. Despite being two essential aspects for viruses, there are yet few examples where these two concepts are shown to be interlinked. Simian, adeno and human papilloma viruses express proteins that interact with the retinoblastoma protein (pRb) and the p53 tumour suppressors [127–129]. The interaction with pRb by E6, E1A or large T antigen releases the E2F1 transcription factors and promotes cell growth and proliferation via a number of downstream target genes including cyclins and c-myc [130,131]. The latent herpes viruses do not seem to express proteins that bind pRb, but more recent works on EBV suggest that the virus is using another mechanism to induce E2F1 expression [14]. Two EBNA1 transgenic animal models show an inverse correlation with tumour phenotype and EBNA1 expression [132,133]. Thus, the less EBNA1 expression, the more tumours, which is counterintuitive, as it would be expected that more of an oncogene would result in more tumours. It turns out that it is not the EBNA1 protein that is oncogenic—it is the *EBNA1* mRNA. Under conditions when EBNA1 suppresses its own mRNA translation it activates E2F1 synthesis. The signalling pathway mediating this link requires the PI3 Kδ, and inhibition of PI3 Kδ prevents the expression of E2F1 and its downstream target c-myc, and suppresses growth of EBNA1-induced tumour cells [14]. Hence, by suppressing its own synthesis, EBNA1 evades MHC class I antigen presentation, while at the same time, targeting the pRb tumour suppressor pathways. This example illustrates how two key viral targets for a successful latent virus–host interaction have co-evolved.

7. Non-protein-mediated control of the host cell

As the immune system uses peptides to distinguish between self and non-self it raises the question to what extent viruses use non-coding RNAs to manipulate the host cell environment to minimize the production of antigenic peptides and acquire stable infection. All human DNA viruses use ncRNAs that interact with nucleotides and proteins, and regulate expression of both cellular and viral genes (both individual genes and gene sets) via both induction and interference with gene expression, serving as scaffolds to ribonucleoproteins (RNPs) or as guides in target recognition within the genome. First reports on the virus expressing miRNAs came from studies on EBV in 2004 [134], and to date 44 EBV-encoded miRNAs produced have been described [135]. EBV miRNAs downregulate the viral genes such as the viral DNA polymerase *BALF5*, and the early EBV genes *BZLF1* and *BRLF1* to keep the viral levels in check and to prevent the transition to the lytic state [136]. Viral miRNAs also interfere with the host immunity by targeting pattern recognition receptors such as RIG-I (*retinoic acid-inducible protein 1*) [137] and IFN gamma-STAT1 pathway [138] thus suppressing interferon signalling and the antiviral response. They also downregulate cytokines and chemokines production and block the antigen presentation and subsequent T cell responses. They are also engaged in the control of cell proliferation and apoptosis by targeting pro-apoptotic genes such as *BAD, BIM, BAX, MAP3K5, PUMA* and capase-3 for inhibition of apoptosis.

A different aspect of using RNAs to manipulate the host comes from the concept of quasispecies that are diverse RNA species derived from RNA viruses due to RNA polymerase infidelity. These can form groups of interactive variants that are crucial for the viral population to attain fitness and to persist in the host. This interesting concept was initially thought of as an error-based selection of the fittest (master) type replicator from the spectrum of mutants. More recent models indicate that quasispecies is a result of a collective viral evolution and not as separate species that together contribute to the characteristics of the population [139–141]. For example, in poliovirus pathogenesis, the generation of diversity (not errors) and cooperation between quasispecies is key to attain disease-associated viral fitness, and a virus carrying a high-fidelity polymerase generates less genomic diversity and is less capable to adapt to adverse growth conditions [142]. HCV quasispecies form swarms of variants that both complement and interfere with the replication of the whole population. Preclusions between different HCV quasispecies populations (clades) have been observed *in vivo* after blood donation and liver transplants of HCV-infected donors to the HCV-infected recipients [143–145]. Only one HCV strain would survive the mixing of quasispecies clades, indicating that the selection occurs on the population level. The joining of RNA species to form a collective occurs via RNA stem-loops structure interactions. Thus, RNA stem-loop structures provide molecular identity to the RNA species, and individual membership is therefore dynamic, but must be coherent with the group [146].

8. Conclusion

Viruses skilfully evade both innate and acquired immune responses by targeting factors within pathways distinguishing self from non-self, but at the same time viruses have contributed to mechanisms whereby cells sense viral infections. This apparent contradiction warrants a deeper look at virus–host interactions, what each partner is trying to achieve and why in some cases they coexist in harmony but in others they do not. Viruses are the most abundant entities on the planet and all species live in their corresponding virosphere. While viruses both cooperate and compete to infect their respective hosts, they also provide ready network solutions in order to enhance the host's self identity. An exchange of genetic material between the two, in case of persistent viruses, has evolved to become a fine-tuned balance of coexistence dating back to the origin of that host. And when the immune system is altered, or when the virus jumps species or cell type, the consequences can be severe. The high number of epigenomic viruses and the accumulation of virus-derived information in the 'non-coding DNA', as well as in the genes, and the role of these elements in regulating network functions is hard evidence of this evolutionary complexity.

Viral persistence can be seen as consequential to the survival of the colonized population, and P1-infected *E. coli* or wild mouse carrying persistent MHV are examples where viruses outcompete the noncolonized bacteria or mice. The addiction model in which viruses have different host- and cell-type-specific mechanisms to set the persistence helps in better understanding the antiviral immunity processes, their origin and what happens when we are targeted by viruses we are not used to. The connection of these mechanisms to cell proliferation and differentiation underlies the long-term coevolution of human and these viruses.

References

1. Villarreal LP. 2005 *Viruses and the evolution of life*. Washington, DC: ASM Press.

2. Villarreal LP. 2006 How viruses shape the tree of life. *Future Virol.* **1**, 587–595. (doi:10.2217/17460794.1.5.587)

3. Villarreal LP, Witzany G. 2010 Viruses are essential agents within the roots and stem of the tree of life.

J Theor. Biol. **262**, 698–710. (doi:10.1016/j.jtbi. 2009.10.014)

4. Villarreal LP. 2011 Viral ancestors of antiviral systems. *Viruses* **3**, 1933–1958. (doi:10.3390/v3101933)

5. Yewdell JW, Reits E, Neefjes J. 2003 Making sense of mass destruction: quantitating MHC class I antigen presentation. *Nat. Rev. Immunol.* **3**, 952–961. (doi:10.1038/nri1250)

6. Princiotta M, Finzi D, Qian S, Gibbs J, Schuchmann S, Buttgereit F, Bennink JR, Yewdell JW. 2003 Quantitating protein synthesis, degradation, and endogenous antigen processing. *Immunity* **18**, 343–354. (doi:10.1016/S1074-7613(03)00051-7)

7. Wei J *et al.* 2017 Varied role of ubiquitylation in generating MHC class I peptide ligands. *J. Immunol.* **198**, 3835–3845. (doi:10.4049/jimmunol.1602122)

8. Marrack P, Scott-Browne J, Dai S, Gapin L, Kappler J. 2008 Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu. Rev. Immunol.* **26**, 171–203. (doi:10.1146/annurev.immunol.26.021607.090421)

9. Yin L, Scott-Browne J, Kappler J, Gapin L, Marrack P. 2012 T cells and their eons-old obsession with MHC. *Immunol. Rev.* **250**, 49–60. (doi:10.1111/imr.12004)

10. Villarreal LP, Defilippis VR, Gottlieb KA. 2000 Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology* **272**, 1–6. (doi:10.1006/viro.2000.0381)

11. Forte E, Zhang Z, Thorp EB, Hummel M. 2020 Cytomegalovirus latency and reactivation: an intricate interplay with the host immune response. *Front. Cell. Inf. Microbiol.* **10**, 130. (doi:10.3389/fcimb.2020.00130)

12. Cesarman E. 2014 Gammaherpesviruses and lymphoproliferative disorders. *Annu. Rev. Pathol. Mech. Dis.* **9**, 349–372. (doi:10.1146/annurev-pathol-012513-104656)

13. Das C, Hansen KC, Tyler JKLMS. 2017 The virome in host health and disease Ken. *Physiol. Behav.* **176**, 139–148. (doi:10.1016/j.physbeh.2017.03.040)

14. Gnanasundram SV, Pyndiah S, Daskalogianni C, Armfield K, Nylander K, Wilson JB, Fåhraeus R. 2017 PI3 Kδ activates E2F1 synthesis in response to mRNA translation stress. *Nat. Commun.* **8**, 2103. (doi:10.1038/s41467-017-02282-w)

15. Gnanasundram SV, Fåhraeus R. 2018 Translation stress regulates ribosome synthesis and cell proliferation. *Int. J. Mol. Sci.* **19**, 3757. (doi:10.3390/ijms19123757)

16. Prado Martins R, Fahraeus R. 2017 A matter of maturity: the impact of pre-mRNA processing in gene expression and antigen presentation. *Int. J. Biochem. Cell Biol.* **91**, 203–211. (doi:10.1016/j.biocel.2017.05.023)

17. Seliger B, Ritz U, Ferrone S. 2006 Molecular mechanisms of HLA class I antigen abnormalities following viral infection and transformation. *Int. J. Cancer* **118**, 129–138. (doi:10.1002/ijc.21312)

18. Volff JN. 2006 Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays* **28**, 913–922. (doi:10.1002/bies.20452)

19. Mattick JS. 2009 Deconstructing the dogma: a new view of the evolution and genetic programming of complex organisms. *Ann. NY Acad. Sci.* **1178**, 29–46. (doi:10.1111/j.1749-6632.2009.04991.x)

20. Villarreal LP, Witzany G. 2013 The DNA habitat and its RNA inhabitants: at the dawn of RNA sociology. *Genomics Insights* **6**, 1–12. (doi:10.4137/GEI.S11490)

21. Moelling K. 2013 What contemporary viruses tell us about evolution: a personal view. *Arch. Virol.* **158**, 1833–1848. (doi:10.1007/s00705-013-1679-6)

22. Blond J-L, Besème F, Duret L, Bouton O, Bedin F, Perron H, Mandrand B, Mallet F. 1999 Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. *J. Virol.* **73**, 1175–1185. (doi:10.1128/JVI.73.2.1175-1185.1999)

23. Haig D. 2012 Retroviruses and the placenta. *Curr. Biol.* **22**, R609–R613. (doi:10.1016/j.cub.2012.06.002)

24. Villarreal LP. 2016 Viruses and the placenta: the essential virus first view. *APMIS. Blackwell Munksgaard* **124**, 20–30. (doi:10.1111/apm.12485)

25. Borza CM, Hutt-Fletcher LM. 2002 Alternate replication in B cells and epithelial cells switches tropism of Epstein-Barr virus. *Nat. Med.* **8**, 594–599. (doi:10.1038/nm0602-594)

26. Fries KL, Sculley TB, Webster-Cyriaque J, Rajadurai P, Sadler RH, Raab-Traub N. 1997 Identification of a novel protein encoded by the BamHI A region of the Epstein-Barr virus. *J. Virol.* **71**, 2765–2771. (doi:10.1128/JVI.71.4.2765-2771.1997)

27. Neves M, Marinho-Dias J, Ribeiro J, Sousa H. 2017 Epstein–Barr virus strains and variations: geographic or disease-specific variants? *J. Med. Virol.* **89**, 373–387. (doi:10.1002/jmv.24633)

28. Geisler J, Touma J, Rahbar A, Söderberg-Nauclér C, Vetvik K. 2019 A review of the potential role of human cytomegalovirus (HCMV) infections in breast cancer carcinogenesis and abnormal immunity. *Cancers* **11**, 1842. (doi:10.3390/cancers11121842)

29. Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleesschauwer B, Griffiths P. 2019 Estimation of the worldwide seroprevalence of cytomegalovirus: a systematic review and meta-analysis. *Rev. Med. Virol.* **29**, e2034. (doi:10.1002/rmv.2034)

30. Mesri EA, Cesarman E, Boshoff C. 2010 Kaposi's sarcoma and its associated herpesvirus. Nature Reviews Cancer. *Nat. Rev. Cancer* **10**, 707–719. (doi:10.1038/nrc2888)

31. Hayward GS. 1999 KSHV strains: the origins and global spread of the virus. *Semin Cancer Biol.* **9**, 187–199. (doi:10.1006/scbi.1998.0116)

32. Pope JH. 1967 Establishment of cell lines from peripheral leucocytes in infectious mononucleosis. *Nature* **216**, 810–811. (doi:10.1038/216810a0)

33. Henle W. 1968 Evidence for viruses in acute leukemia and Burkitt's tumor. *Cancer* **21**, 580–586. (doi:10.1002/1097-0142(196804)21:4<580::AID-CNCR2820210406>3.0.CO;2-Q)

34. Mrozek-Gorska P, Buschle A, Pich D, Schwarzmayr T, Fechtner R, Scialdone A, Hammerschmidt W. 2019 Epstein–Barr virus reprograms human B

35. Shannon-Lowe C, Rickinson AB, Bell AI. 2017 Epstein-barr virus-associated lymphomas. *Phil.Trans. R. Soc. B Biol. Sci.* **372**, 20160271. (doi:10.1098/rstb.2016.0271)

36. Shannon-Lowe C, Rickinson A. 2019 The global landscape of EBV-associated tumors. *Front. Oncol.* **9**, 713. (doi:10.3389/fonc.2019.00713)

37. Homberger FR. 1997 Enterotropic mouse hepatitis virus. *Lab. Anim.* **31**, 97–115. (doi:10.1258/002367797780600189)

38. Villarreal LP. 2009 Persistence pays: how viruses promote host group survival. *Curr. Opin. Microbiol.* **12**, 467–472. (doi:10.1016/j.mib.2009.06.014)

39. Villarreal LP, Witzany G. 2019 That is life: communicating RNA networks from viruses and cells in continuous interaction. *Ann. NY Acad. Sci.* **1447**, 5–20. (doi:10.1111/nyas.14040)

40. Chrétien I, Marcuz A, Courtet M, Katevuo K, Vainio O, Heath JK, White SJ, Du Pasquier L. 1998 CTX, a *Xenopus* thymocyte receptor, defines a molecular family conserved throughout vertebrates. *Eur. J. Immunol.* **28**, 4094–4104. (doi:10.1002/(SICI)1521-4141(199812)28:12<4094::AID-IMMU4094>3.0.CO;2-2)

41. Du Pasquier L, Zucchetti I, De Santis R. 2004 Immunoglobulin superfamily receptors in protochordates: before RAG time. *Immunol. Rev.* **198**, 233–248. (doi:10.1111/j.0105-2896.2004.00122.x)

42. Lehnherr H, Maguin E, Jafri S, Yarmolinsky MB. 1993 Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J. Mol. Biol.* **233**, 414–428. (doi:10.1006/jmbi.1993.1521)

43. Villarreal LP. 2009 *Origin of group identity: viruses, addiction, and cooperation*. Berlin, Germany: Springer.

44. Villarreal LP. 2016 Persistent virus and addiction modules: an engine of symbiosis. *Curr. Opin. Microbiol.* **31**, 70–79. (doi:10.1016/j.mib.2016.03.005)

45. Roach DR, Leung CY, Henry M, Morello E, Singh D, Di Santo JP, Weitz JS, Debarbieux L. 2017 Synergy between the host immune system and bacteriophage is essential for successful phage therapy against an acute respiratory pathogen. *Cell Host Microbe* **22**, 38–47.e4. (doi:10.1016/j.chom.2017.06.018)

46. Ott SJ *et al.* 2017 Efficacy of sterile fecal filtrate transfer for treating patients with clostridium difficile infection. *Gastroenterology* **152**, 799–811.e7. (doi:10.1053/j.gastro.2016.11.010)

47. Maurice CF. 2019 Considering the other half of the gut microbiome: bacteriophages. *mSystems* **4**, e00102-19. (doi:10.1128/mSystems.00102-19)

48. Khadijah S, Neo SY, Hossain MS, Miller LD, Mathavan S, Kwang J. 2003 Identification of white spot syndrome virus latency-related genes in specific-pathogen-free shrimps by use of a

microarray. *J. Virol.* **77**, 10 162–10 167. (doi:10.1128/JVI.77.18.10162-10167.2003)

49. Wang Z, Hu L, Yi G, Xu H, Qi Y, Yao L. 2004 ORF390 of white spot syndrome virus genome is identified as a novel anti-apoptosis gene. *Biochem. Biophys. Res. Commun.* **325**, 899–907. (doi:10.1016/j.bbrc.2004.09.224)

50. Flajnik MF. 1994 Primitive vertebrate immunity: what is the evolutionary derivative of molecules that define the adaptive immune system? *Ciba Foundation symp.* **186**, 224–232. discussion 233-6.

51. Flajnik MF, Kasahara M. 2001 Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity* **15**, 351–362. (doi:10.1016/S1074-7613(01)00198-4)

52. Ohta Y, Goetz W, Hossain MZ, Nonaka M, Flajnik MF. 2006 Ancestral organization of the MHC revealed in the amphibian *Xenopus*. *J. Immunol.* **176**, 3674–3685. (doi:10.4049/jimmunol.176.6.3674)

53. Flajnik MF. 2014 Re-evaluation of the immunological big bang. *Curr. Biol.* **24**, R1060–R1065. (doi:10.1016/j.cub.2014.09.070)

54. Villarreal LP. 2009 The source of self: genetic parasites and the origin of adaptive immunity. *Ann. NY Acad. Sci.* **1178**, 194–232. (doi:10.1111/j.1749-6632.2009.05020.x)

55. Chuong EB, Elde NC, Feschotte C. 2016 Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**, 1083–1087. (doi:10.1126/science.aad5497)

56. Du Pasquier L. 2004 Speculations on the origin of the vertebrate immune system. *Immunol. Lett.* **92**, 3–9. (doi:10.1016/j.imlet.2003.10.012)

57. Dreyfus DH, Jones JF, Gelfand EW. 1999 Asymmetric DDE (D35E)-like sequences in the RAG proteins: implications for V(D)J recombination and retroviral pathogenesis. *Med. Hypotheses* **52**, 545–549. (doi:10.1054/mehy.1997.0691)

58. Clark PM, Chitnis N, Shieh M, Kamoun M, Johnson FB, Monos D. 2018 Novel and haplotype specific microRNAs encoded by the major histocompatibility complex. *Sci. Rep.* **8**, 3832. (doi:10.1038/s41598-018-19427-6)

59. Chitnis N, Clark PM, Kamoun M, Stolle C, Johnson FB, Monos DS. 2017 An expanded role for HLA genes: HLA-B encodes a microRNA that regulates IgA and other immune response transcripts. *Front. Immunol.* **8**, 583. (doi:10.3389/fimmu.2017.00583)

60. Kulski JK. 2019 Long noncoding RNA HCP5, a hybrid HLA class I endogenous retroviral gene: structure, expression, and disease associations. *Cells* **8**, 480. (doi:10.3390/cells8050480)

61. Gaudieri S, Giles KM, Kulski JK, Dawkins RL. 1997 Duplication and polymorphism in the MHC: Alu generated diversity and polymorphism within the PERB11 gene family. *Hereditas* **127**, 37–46. (doi:10.1111/j.1601-5223.1997.00037.x)

62. Kulski JK, Gaudieri S, Bellgard M, Balmer L, Giles K, Inoko H, Dawkins RL. 1997 The evolution of MHC diversity by segmental duplication and transposition of retroelements. *J. Mol. Evol.* **45**, 599–609. (doi:10.1007/PL00006264)

63. Kulski JK, Gaudieri S, Inoko H, Dawkins RL. 1999 Comparison between two human endogenous retrovirus (HERV)-rich regions within the major histocompatibility complex. *J. Mol. Evol.* **48**, 675–683. (doi:10.1007/PL00006511)

64. Kulski JK, Gaudieri S, Martin A, Dawkins RL. 1999 Coevolution of PERB11 (MIC) and HLA class I genes with HERV-16 and retroelements by extended genomic duplication. *J. Mol. Evol.* **49**, 84–97. (doi:10.1007/PL00006537)

65. Olgun G, Sahin O, Tastan O. 2018 Discovering lncRNA mediated sponge interactions in breast cancer molecular subtypes. *BMC Genomics* **19**, 1–2. (doi:10.1186/s12864-018-5006-1)

66. Teng H, Wang P, Xue Y, Liu X, Ma J, Cai H, Xi Z, Li Z, Liu Y. 2016 Role of HCP5-miR-139-RUNX1 feedback loop in regulating malignant behavior of Glioma cells. *Mol. Ther.* **24**, 1806–1822. (doi:10.1038/mt.2016.103)

67. Blees A, Januliene D, Hofmann T, Koller N, Schmidt C, Trowitzsch S, Moeller A, Tampé R. 2017 Structure of the human MHC-I peptide-loading complex. *Nature* **551**, 525–528. (doi:10.1038/nature24627)

68. Eggensperger S, Tampé R. 2015 The transporter associated with antigen processing: a key player in adaptive immunity. *Biol. Chem.* **396**, 1059–1072. (doi:10.1515/hsz-2014-0320)

69. Oldham ML, Grigorieff N, Chen J. 2016 Structure of the transporter associated with antigen processing trapped by herpes simplex virus. *Elife* **5**, e21829. (doi:10.7554/eLife.21829)

70. De La Salle H et al. 1994 Homozygous human TAP peptide transporter mutation in HLA class I deficiency. *Science* **265**, 237–241. (doi:10.1126/science.7517574)

71. Teisserenc H, Schmitt W, Blake N, Dunbar R, Gadola S, Gross WL, Exley A, Cerundolo V. 1997 A case of primary immunodeficiency due to a defect of the major histocompatibility gene complex class I processing and presentation pathway. *Immunol. Lett.* **57**, 183–187. (doi:10.1016/S0165-2478(97)00072-2)

72. Gadola SD, Moins-Teisserenc HT, Trowsdale J, Gross WL, Cerundolo V. 2000 TAP deficiency syndrome. *Clin. Exp. Immunol.* **121**, 173–178. (doi:10.1046/j.1365-2249.2000.01264.x)

73. De La Salle H et al. 1997 Human peptide transporter deficiency: importance of HLA-B in the presentation of TAP-independent EBV antigens. *J. Immunol.* **158**, 4555–4563.

74. Lautscham G, Mayrhofer S, Taylor G, Haigh T, Leese A, Rickinson A, Blake N. 2001 Processing of a multiple membrane spanning Epstein-Barr virus protein for CD8+ T cell recognition reveals a proteasome-dependent, transporter associated with antigen processing-independent pathway. *J. Exp. Med.* **194**, 1053–1068. (doi:10.1084/jem.194.8.1053)

75. Lautscham G, Rickinson A, Blake N. 2003 TAP-independent antigen presentation on MHC class I molecules: lessons from Epstein-Barr virus. *Microbes Infection* **5**, 291–299. (doi:10.1016/S1286-4579(03)00031-5)

76. Del Val M, Antón LC, Ramos M, Muñoz-Abad V, Campos-Sánchez E. 2020 Endogenous TAP-independent MHC-I antigen presentation: not just the ER lumen. *Curr. Opin. Immunol.* **64**, 9–14. (doi:10.1016/j.coi.2019.12.003)

77. Nöll A et al. 2017 Crystal structure and mechanistic basis of a functional homolog of the antigen transporter TAP. *Proc. Natl Acad. Sci. USA* **114**, E438–E447. (doi:10.1073/pnas.1620009114)

78. Neefjes J, Marlieke LMJ, Petra P, Oddmund B. 2011 Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836. (doi:10.1038/nri3084)

79. Garstka MA et al. 2015 The first step of peptide selection in antigen presentation by MHC class I molecules. *Proc. Natl Acad. Sci. USA* **112**, 1505–1510. (doi:10.1073/pnas.1416543112)

80. Gardner BM, Pincus D, Gotthardt K, Gallagher CM, Walter P. 2013 Endoplasmic reticulum stress sensing in the unfolded protein response. *Cold Spring Harb. Perspect. Biol.* **5**, 1–15. (doi:10.1101/cshperspect.a013169)

81. Hoseki J, Ushioda R, Nagata K. 2010 Mechanism and components of endoplasmic reticulum-associated degradation. *J. Biochem.* **147**, 19–25. (doi:10.1093/jb/mvp194)

82. Lopata A, Kniss A, Löhr F, Rogov VV, Dötsch V. 2020 Ubiquitination in the erad process. *Int. J. Mol. Sci.* **21**, 1–21. (doi:10.3390/ijms21155369)

83. Nakatsukasa K, Brodsky JL. 2008 The recognition and retrotranslocation of misfolded proteins from the endoplasmic reticulum. *Traffic* **9**, 861–870. (doi:10.1111/j.1600-0854.2008.00729.x)

84. Van den Boomen DJH, Lehner PJ. 2015 Identifying the ERAD ubiquitin E3 ligases for viral and cellular targeting of MHC class I. *Mol. Immunol.* **68**, 106–111. (doi:10.1016/j.molimm.2015.07.005)

85. van de Weijer ML, Schuren ABC, Boomen DJH, Mulder A, Claas FHJ, Lehner PJ, Lebbink RJ, Wiertz EJHJ. 2017 Multiple E2 ubiquitin-conjugating enzymes regulate human cytomegalovirus US2-mediated immunoreceptor downregulation. *J. Cell Sci.* **130**, 2883–2892. (doi:10.1242/jcs.206839)

86. Cerami E et al. 2012 The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404. (doi:10.1158/2159-8290.CD-12-0095)

87. Gewurz BE, Gaudet R, Tortorella D, Wang EW, Ploegh HL, Wiley DC. 2001 Antigen presentation subverted: structure of the human cytomegalovirus protein US2 bound to the class I molecule HLA-A2. *Proc. Natl Acad. Sci. USA* **98**, 6794–6799. (doi:10.1073/pnas.121172898)

88. Li L, Santarsiero BD, Bouvier M. 2016 Structure of the adenovirus type 4 (Species E) E3-19 K/HLA-A2 complex reveals species-specific features in MHC class I recognition. *J. Immunol.* **197**, 1399–1407. (doi:10.4049/jimmunol.1600541)

89. Sester M, Ruszics Z, Mackley E, Burgert H-G. 2013 The transmembrane domain of the adenovirus E3/19 K protein acts as an endoplasmic reticulum retention signal and contributes to intracellular

sequestration of major histocompatibility complex class I molecules. *J. Virol.* **87**, 6104–6117. (doi:10.1128/JVI.03391-12)

90. Berry R, Watson GM, Jonjic S, Degli-Esposti MA, Rossjohn J. 2020 Modulation of innate and adaptive immunity by cytomegaloviruses. *Nat. Rev. Immunol.* **20**, 113–127. (doi:10.1038/s41577-019-0225-5)

91. Waterhouse A *et al.* 2018 SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303. (doi:10.1093/nar/gky427)

92. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. 2014 ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **30**, 1771–1773. (doi:10.1093/bioinformatics/btu097)

93. Parcej D, Tampé R. 2010 ABC proteins in antigen translocation and viral inhibition. *Nat. Chem. Biol.* **6**, 572–580. (doi:10.1038/nchembio.410)

94. Ahn K, Gruhler A, Galocha B, Jones TR, Wiertz EJHJ, Ploegh HL, Peterson PA, Yang Y, Früh K. 1997 The ER-luminal domain of the HCMV glycoprotein US6 inhibits peptide translocation by TAP. *Immunity* **6**, 613–621. (doi:10.1016/S1074-7613(00)80349-0)

95. Yang Z, Bjorkman PJ. 2008 Structure of UL18, a peptide-binding viral MHC mimic, bound to a host inhibitory receptor. *Proc. Natl Acad. Sci. USA* **105**, 10 095–10 100. (doi:10.1073/pnas.0804551105)

96. Lehner PJ, Karttunen JT, Wilkinson GWG, Cresswell P. 1997 The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation. *Proc. Natl Acad. Sci. USA* **94**, 6904–6909. (doi:10.1073/pnas.94.13.6904)

97. Verweij MC, Horst D, Griffin BD, Luteijn RD, Davison AJ, Ressing ME, Wiertz EJ. 2015 Viral inhibition of the transporter associated with antigen processing (TAP): a striking example of functional convergent evolution. *PLoS Pathogens* **11**, e1004743. (doi:10.1371/journal.ppat.1004743)

98. Hewitt EW. 2003 The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology* **110**, 163–169. (doi:10.1046/j.1365-2567.2003.01738.x)

99. Park B, Kim Y, Shin J, Lee S, Cho K, Früh K, Lee S, Ahn K. 2004 Human cytomegalovirus inhibits tapasin-dependent peptide loading and optimization of the MHC class I peptide cargo for immune evasion. *Immunity* **20**, 71–85. (doi:10.1016/S1074-7613(03)00355-8)

100. Harvey IB, Wang X, Fremont DH. 2019 Molluscum contagiosum virus MC80 sabstages MHC-I antigen presentation by targeting tapasin for ER-associated degradation. *PLOS Pathogens* **15**, e1007711.

101. Luteijn RD *et al.* 2014 Cowpox virus protein CPXV012 eludes CTLs by blocking ATP binding to TAP. *J. Immunol.* **193**, 1578–1589. (doi:10.4049/jimmunol.1400964)

102. Wycisk AI *et al.* 2011 Epstein-barr viral BNLF2a protein hijacks the tail-anchored protein insertion machinery to block antigen processing by the transport complex TAP. *J. Biol. Chem.* **286**, 41 402–41 412. (doi:10.1074/jbc.M111.237784)

103. Sagert L, Hennig F, Thomas C, Tampé R. 2020 A loop structure allows TAPBPR to exert its dual function as MHC I chaperone and peptide editor. *Elife* **9**, e55326. (doi:10.7554/eLife.55326)

104. Padariya M, Kalathiya U, Houston DR, Alfaro JA. 2020 Recognition dynamics of cancer mutations on the ERp57-tapasin interface. *Cancers* (*Basel*) **12**, 737. (doi:10.3390/cancers12030737)

105. Zamani MR, Aslani S, Salmaninejad A, Javan MR, Rezaei N. 2016 PD-1/PD-L and autoimmunity: a growing relationship. *Cell Immunol.* **310**, 27–41. (doi:10.1016/j.cellimm.2016.09.009)

106. Bardhan K, Anagnostou T, Boussiotis VA. 2016 The PD1: PD-L1/2 pathway from discovery to clinical implementation. *Front. Immunol.* **7**, 550. (doi:10.3389/fimmu.2016.00550)

107. Schönrich G, Raftery MJ. 2019 The PD-1/PD-L1 axis and virus infections: a delicate balance. *Front. Cell. Inf. Microbiol.* **9**, 207. (doi:10.3389/fcimb.2019.00207)

108. Raftery MJ, Wieland D, Gronewald S, Kraus AA, Giese T, Schönrich G. 2004 Shaping phenotype, function, and survival of dendritic cells by cytomegalovirus-encoded IL-10. *J. Immunol.* **173**, 3383–3391. (doi:10.4049/jimmunol.173.5.3383)

109. Selenko-Gebauer N *et al.* 2003 B7-H1 (programmed death-1 ligand) on dendritic cells is involved in the induction and maintenance of T cell anergy. *J. Immunol.* **170**, 3637–3644. (doi:10.4049/jimmunol.170.7.3637)

110. Kakizaki M, Yamamoto Y, Yabuta S, Kurosaki N, Kagawa T, Kotani A. 2018 The immunological function of extracellular vesicles in hepatitis B virus-infected hepatocytes. *PLoS ONE* **13**, 1–13. (doi:10.1371/journal.pone.0205886)

111. Billerbeck E *et al.* 2017 Mouse models of acute and chronic hepacivirus infection. *Science* **357**, 204–208. (doi:10.1126/science.aal1962)

112. Boon T, Van Pel A. 1989 T cell-recognized antigenic peptides derived from the cellular genome are not protein degradation products but can be generated directly by transcription and translation of short subgenic regions. A hypothesis. *Immunogenetics* **29**, 75–79. (doi:10.1007/BF00395854)

113. Apcher S, Daskalogianni C, Lejeune F, Manoury B, Imhoos G, Heslop L, Fahraeus R. 2011 Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation. *Proc. Natl Acad. Sci. USA* **108**, 11 572–11 577. (doi:10.1073/pnas.1104104108)

114. Anton LC, Yewdell JW. 2014 Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors. *J. Leukoc. Biol.* **95**, 551–562. (doi:10.1189/jlb.1113599)

115. Bennett NJ, May JS, Stevenson PG. 2005 Gamma-herpesvirus latency requires T cell evasion during episome maintenance. *PLoS Biol.* **3**, 0638–0649. (doi:10.1371/journal.pbio.0030120)

116. Kwun HJ, da Silva SR, Shah IM, Blake N, Moore PS, Chang Y. 2007 Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mimics Epstein-Barr virus EBNA1 immune evasion through central repeat domain effects on protein processing. *J. Virol.* **81**, 8225–8235. (doi:10.1128/JVI.00411-07)

117. Apcher S, Daskalogianni C, Manoury B, Fåhraeus R. 2010 Epstein barr virus-encoded EBNA1 interference with MHC class I antigen presentation reveals a close correlation between mRNA translation initiation and antigen presentation. *PLoS Pathog.* **6**, 1–14. (doi:10.1371/journal.ppat.1001151)

118. Schwab SR, Shugart JA, Horng T, Malarkannan S, Shastri N. 2004 Unanticipated antigens: translation initiation at CUG with leucine. *PLoS Biol.* **2**, 1774–1784. (doi:10.1371/journal.pbio.0020366)

119. Starck SR, Jiang V, Pavon-Eternod M, Prasad S, McCarthy B, Pan T, Shastri N. 2012 Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**, 1719–1723. (doi:10.1126/science.1220270)

120. Apcher S, Millot G, Daskalogianni C, Scherl A, Manoury B, Fahraeus R. 2013 Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. *Proc. Natl Acad. Sci.* **110**, 17 951–17 956. (doi:10.1073/pnas.1309956110)

121. Yewdell JW, Dersh D, Fahraeus R. 2019 Peptide channeling: the key to MHC class I immunosurveillance? *Trends Cell Biol.* **29**, 929–939. (doi:10.1016/j.tcb.2019.09.004)

122. Yewdell JW, Anton LC, Bennink JR. 1996 Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J. Immunol.* **157**, 1823–1826.

123. Khan S, de Giuli R, Schmidtke G, Bruns M, Buchmeier M, van den Broek M, Groettrup M. 2001 Cutting edge: neosynthesis is required for the presentation of a T cell epitope from a long-lived viral protein. *J. Immunol.* **167**, 4801–4804. (doi:10.4049/jimmunol.167.9.4801)

124. Yili Y, Manoury B, Fahraeus R. 2003 Self-inhibition of synthesis and antigen presentation by epstein-barr virus-encoded EBNA1. *Science* **301**, 1371–1374. (doi:10.1126/science.1088902)

125. Zaldumbide A, Ossevoort M, Wiertz EJHJ, Hoeben RC. 2007 In cis inhibition of antigen processing by the latency-associated nuclear antigen I of Kaposi sarcoma herpes virus. *Mol. Immunol.* **44**, 1352–1360. (doi:10.1016/j.molimm.2006.05.012)

126. Su L, Liao Q, Wu Y, Chen X. 2011 Kaposi's sarcoma-associated herpesvirus-encoded LANA down-regulates IL-22R1 expression through a cis-acting element within the promoter region. *PLoS ONE* **6**, e19106. (doi:10.1371/journal.pone.0019106)

127. Whyte P, Buchkovich KJ, Horowitz JM, Friend SH, Raybuck M, Weinberg RA, Harlow E. 1988 Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* **334**, 124–129. (doi:10.1038/334124a0)

128. Phelps WC, Yee C, Munger K, Howley PM. 1988 The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar

to those of adenovirus E1A. *Cell* **53**, 539–547. (doi:10.1016/0092-8674(88)90570-3)

129. Felsani A, Mileo AM, Paggi MG. 2006 Retinoblastoma family proteins as key targets of the small DNA virus oncoproteins. *Oncogene* **25**, 5277–5285. (doi:10.1038/sj.onc.1209621)

130. Chen HZ, Tsai SY, Leone G. 2009 Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat. Rev. Cancer* **9**, 785–797. (doi:10.1038/nrc2696)

131. Roussel MF, Davis JN, Cleveland JL, Ghysdael J, Hiebert SW. 1994 Dual control of MYC expression through a single DNA binding site trageted by ETS family proteins and E2F-1. *Oncogene* **9**, 405–415.

132. Wilson JB, Bell JL, Levine AJ. 1996 Expression of Epstein-Barr virus nuclear antigen-1 induced B cell neoplasia in transgenic mice. *EMBO J.* **15**, 3117–3126. (doi:10.1002/j.1460-2075.1996.tb00674.x)

133. Kang MS, Lu H, Yasui T, Sharpe A, Warren H, Cahir-McFarland E, Bronson R, Hung SC, Kieff E. 2005 Epstein-Barr virus nuclear antigen-1 does not induce lymphoma in transgenic FVB mice. *Proc. Natl Acad. Sci. USA* **102**, 820–825. (doi:10.1073/pnas.0408774102)

134. Pfeffer S *et al.* 2004 Identification of virus-encoded microRNAs. *Science* **304**, 734–736. (doi:10.1126/science.1096781)

135. Wang M, Yu F, Wu W, Wang Y, Ding H, Qian L. 2018 Epstein-Barr virus-encoded microRNAs as regulators in host immune responses. *Int. J. Biol. Sci.* **14**, 565–576. (doi:10.7150/ijbs.24562)

136. Brady CS *et al.* 2000 Multiple mechanisms underlie HLA dysregulation in cervical cancer. *Tissue Antigens* **55**, 401–411. (doi:10.1034/j.1399-0039.2000.550502.x)

137. Lu Y *et al.* 2017 Epstein-Barr virus miR-BART6-3p inhibits the RIG-I pathway. *J. Innate Immun.* **9**, 574–586. (doi:10.1159/000479749)

138. Huang WT, Lin CW. 2014 EBV-encoded miR-BART20-5p and miR-BART8 inhibit the IFN-γ-STAT1 pathway associated with disease progression in nasal NK-cell lymphoma. *Am. J. Pathol.* **184**, 1185–1197. (doi:10.1016/j.ajpath.2013.12.024)

139. Ojosnegros S, Perales C, Mas A, Domingo E. 2011 Quasispecies as a matter of fact: viruses and beyond., virus research. *Virus Res.* **162**, 203–215. (doi:10.1016/j.virusres.2011.09.018)

140. Domingo E, Sheldon J, Perales C. 2012 Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* **76**, 159–216. (doi:10.1128/MMBR.05023-11)

141. Domingo E, Perales C. 2019 Viral quasispecies. *PLoS Genet.* **15**, e1008271. (doi:10.1371/journal.pgen.1008271)

142. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006 Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348. (doi:10.1038/nature04388)

143. Fan X, Lang DM, Xu Y, Lyra AC, Yusim K, Everhart JE, Korber BT, Perelson AS, Di Bisceglie AM. 2003 Liver transplantation with hepatitis C virus-infected graft: interaction between donor and recipient viral strains. *Hepatology* **38**, 25–33. (doi:10.1053/jhep.2003.50264)

144. Evans MJ, Rice CM, Goff SP. 2004 Genetic interactions between hepatitis C virus replicons. *J. Virol.* **78**, 12 085–12 089. (doi:10.1128/JVI.78.21.12085-12089.2004)

145. Ramírez S *et al.* 2010 Hepatitis C virus superinfection of liver grafts: a detailed analysis of early exclusion of non-dominant virus strains. *J. Gen. Virol.* **91**, 1183–1188. (doi:10.1099/vir.0.018929-0)

146. Villarreal LP. 2015 Force for ancient and recent life: viral and stem-loop RNA consortia promote life. *Ann. NY Acad. Sci.* **1341**, 25–34. (doi:10.1111/nyas.12565)