Infection, Genetics and Evolution 73 (2019) 342–357

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Review

Truncated oncoproteins of retroviruses and hepatitis B virus: A lesson in contrasts

Janos Minarovits[a,*], Hans Helmut Niller[b]

[a] Faculty of Dentistry, Department of Oral Biology and Experimental Dental Research, University of Szeged, Tisza Lajos krt. 64, Szeged H-6720, Hungary
[b] Institute for Medical Microbiology, Hygiene of the University of Regensburg, D-93053 Franz-Josef-Strauß-Allee 11, Regensburg, Germany

SUMMARY

Members of the virus families *Retroviridae* and *Hepadnaviridae* use reverse transcriptase (RT) to synthesize a DNA copy of their genomic and pregenomic RNA, respectively, during the viral life cycle. A group of viruses belonging to *Retroviridae* ("acute transforming" retroviruses) as well as human hepatitis B virus (HBV), the prototype member of *Hepadnaviridae* (hepadnaviruses) are able to cause malignant neoplasms in infected hosts, due to the expression of pleiotropic "transforming proteins" encoded by the genomes of these reverse-transcribing tumor viruses. In this review we wish to compare the common and unique features of replication strategies characteristic of acute transforming retroviruses and HBV and summarize data related to the origin and evolution of their viral oncogenes either *via* transduction of cellular genes, or by accumulation of mutations in viral sequences that create a new open reading frame (overprinting). The exons of cellular genes (*proto-onc* genes or *c-onc* genes) incorporated into the genome of acute transforming retroviruses are regularly affected by deletions resulting in the expression of truncated viral oncoproteins which are frequently dysregulated compared to their cellular counterparts. These retroviral transforming proteins alter the behavior of their target cells (malignant transformation). HBx, a pleiotropic protein of HBV, regulates virus replication and contributes to hepatocarcinogenesis. In contrast to the *v-onc* genes of acute transforming retroviruses, the viral gene encoding the full-length, wild-type HBx (wtHBx) protein does not have a cellular counterpart. Mutations and deletions frequently affect, however, the HBV genome as well, resulting in the expression of truncated HBx proteins (trHBx) in liver cells. Truncated, especially C-terminal truncated variants of HBx (Ct-HBX proteins), may facilitate initiation and progression of liver carcinoma.

1. Introduction: viral tumorigenesis in a nutshell

DNA and RNA tumor viruses are associated with malignant neoplasms in humans and animals (Ewald and Swain Ewald, 2015; Plummer et al., 2016; Chang et al., 2017). Tumor virus genomes frequently carry one or several viral oncogenes coding for so called oncoproteins, i.e. pleiotropic regulators that affect the gene expression pattern, phenotype and behaviour of host cells (Vogt, 2012; Ajiro and Zheng, 2014). Non-translated regulatory RNAs are also expressed from the genome of certain tumor viruses in infected cells, contributing to the initiation and progression of malignant neoplasms (reviewed by (Grundhoff and Sullivan, 2011)). In virus-associated carcinomas, lymphomas, leukemias and sarcomas, the double stranded DNA (dsDNA) genomes of most DNA tumor viruses and the dsDNA copies generated by reverse transcription of plus-strand RNA genomes carried by RNA tumor viruses integrate into the target cell DNA (Vogt, 2012; White et al., 2014). Expression of the viral oncogenes inserted into the host cell genome not only stimulates cell growth, but may also contribute to the altered morphology and behaviour of neoplastic cells ("malignant transformation"). Furthermore, integration of viral genomes may result in host cell DNA rearrangements and chromosomal aberrations facilitating tumor initiation and progression (Robinson and Dunning Hotopp, 2014; Zhao et al., 2016).

Acute transforming RNA tumor viruses typically carry a viral oncogene (*v-onc* gene) or several *v-onc* genes and induce malignant tumors in infected hosts after a short period of latency (Vogt, 2012). In contrast, neoplasms develop only after a prolonged latency period in the hosts infected with non-acute or slow transforming RNA tumor viruses which don't harbour a viral oncogene in their genome. Non-acute RNA tumor viruses induce malignant tumors *via* so called promoter insertion and enhancer insertion. Integration of the viral dsDNA genome into the host cell genome is an important step in these processes, because it creates an opportunity for viral regulatory sequences to influence the expression of neighbouring cellular oncogenes (*c-onc* genes, also called *proto-onc* genes) that control cell proliferation. Constitutively, high level transcription of *c-onc* genes initiated at integrated viral promoters can cause dysregulated cell proliferation. Similarly, insertion of viral enhancer sequences may upregulate *c-onc* transcription even from a long distance (Tsuruyama et al., 2017) (reviewed by (Fan and Johnson, 2011)).

In case of other retroviral strains lacking a *v-onc* gene, one of the viral structural proteins may act as an oncoprotein. Expression of the envelope (env) protein encoded by the genome of Jaagsiekte sheep retrovirus (JSRV), the causative agent of ovine pulmonary

* Corresponding author.
E-mail addresses: minimicrobi@hotmail.com (J. Minarovits), Hans-Helmut.Niller@klinik.uni-regensburg.de (H.H. Niller).

https://doi.org/10.1016/j.meegid.2019.05.020
Received 18 March 2019; Received in revised form 14 May 2019; Accepted 27 May 2019
Available online 29 May 2019

adenocarcinoma, transformed rodent fibroblasts *in vitro*, partly by activating protein kinases involved in the control of cell proliferation (reviewed by (Liu and Miller, 2007)). Under experimental conditions, when introduced into the airways of immunodeficient mice *via* an adeno-associated virus (AAV) vector, JSRV env induced lung tumors (Wootton et al., 2005).

Although there is no *v-onc* gene in their genomes, the retroviruses present in the Friend virus (FV) complex are able to cause acute erythroleukemias in mice (reviewed by (Moreau-Gachelin, 2008)). One of the retroviruses in the FV complex is the replication competent F-MuLV (Friend murine leukemia virus). F-MuLV acts as a helper virus for the replication defective component called SFFV (spleen focus forming virus) (Gomez-Lucia et al., 1998; Moreau-Gachelin, 2008). Erythroleukemia develops in two phases. In the initial stage, an SFFV genome-encoded env protein, gp55, binds to the erythropoietin receptor and induces erythroblastosis, i.e. the proliferation of erythroid precursor cells. In the second stage of the disease, however, enhancer insertion, i.e. an oncogenic mechanism characteristic of non-acute RNA tumor viruses, contributes to the malignant transformation of erythroid cells. At this stage, proviral integration into the *spi-1* (SSFV proviral integration site-1) locus stimulates transcription of *spi-1* resulting in a high level of Spi-1 protein that blocks maturation of erythroid precursor cells. Additional genetic changes affecting the host cell genome may favour erythropoietin-independent proliferation of the leukemic cells (reviewed by (Moreau-Gachelin, 2008)).

In addition to structural proteins, the regulatory or accessory proteins encoded by retroviral genomes may also contribute to malignant transformation. Tax, a pleiotropic regulator encoded by the genome of human T lymphotropic virus type 1 (HTLV-1) not only enhances viral transcription and replication, but also activates cell proliferation and blocks cellular checkpoint functions and DNA repair, facilitating the development of adult T cell leukemia (reviewed by (Matsuoka and Jeang, 2011; Watanabe, 2017)).

In contrast to the majority of tumor virus infections which are accompanied with the integration of the viral genome into the host cell DNA during tumorigenesis, the genomes of oncogenic herpesviruses are typically maintained in a non-integrated, episomal form co-replicating with the cellular DNA once per cell cycle. Transcripts for latent, growth-transformation-associated oncoproteins, immune-modulator proteins and regulatory RNAs are generated from the circular, dsDNA genomes of oncogenic herpesviruses, including Epstein-Barr virus (EBV, human herpesvirus 4) and Kaposi's sarcoma-associated herpesvirus (KSHV, human herpesvirus 8). Herpesviral oncoproteins, but also the "transforming" proteins of other tumor viruses may alter the host cell epigenotype and gene expression pattern by interacting with the epigenetic regulatory machinery of the host cell (Niller et al., 2014; Minarovits et al., 2016).

Epigenetic changes, oxidative stress and modulation of intracellular signalling pathways, possibly elicited by the core protein may initiate and maintain hepatocarcinogenesis by hepatitis C virus (HCV), an RNA virus which lacks a viral oncogene (Koike, 2005; Benegiamo et al., 2012). The HCV genome does not encode a reverse transcriptase, either. Nonstructural HCV proteins may also contribute to the complex and prolonged process of hepatocarcinogenesis during persistent HCV infection which is accompanied by long-lasting chronic inflammation and liver regeneration resulting from immune-mediated cell death leading to liver fibrosis (Bartosch et al., 2009). Similar processes contribute to liver cancer development in HBV-infected patients, too (reviewed by (Seeger and Mason, 2015)).

The major mechanisms of host cell-transformation induced by RNA and DNA tumor viruses are shown in Table 1.

Reverse-transcribing tumor viruses belong to two virus families, *Retroviridae* and *Hepadnaviridae*. In this review we wish to compare the major characteristics of retroviruses and hepadnaviruses causing the development of malignant tumors, with an emphasis on the structural changes affecting their oncogenes and oncoproteins during oncogenesis and tumor progression. Changes in the anatomy of viral genes coding for virion components or non-structural virus proteins related to tumorigenesis will also be discussed. We shall concentrate on selected representatives of RNA tumor viruses, i.e. retroviruses that replicate their RNA genome *via* a dsDNA intermediate, and on human hepatitis B virus and its relatives that replicate their DNA genome *via* a reverse-transcribed viral RNA. RNA tumor viruses cause malignant diseases in a variety of vertebrate species and may target more than one host cell types, whereas chronic infection with hepatitis B virus or related viruses of the *Orthohepadnavirus* genus is associated with the development of hepatocellular carcinoma (HCC) in a subset of infected humans and rodents, respectively (reviewed by (Tennant and Gerin, 2001; Pedersen and Sorensen, 2010; Vogt, 2012; Benachenhou et al., 2013; Buendia and Neuveut, 2015)).

2. The retroviral life cycle and transduction of cellular sequences by oncogenic retroviruses

Retroviruses (family: *Retroviridae*) package two nearly identical copies of their single stranded RNA (ssRNA) genome of positive polarity into the virion. Retroviruses bind *via* their surface proteins to cellular receptors expressed either on a subset of somatic cells or on a wide variety of cell types including both somatic and germ cells (Overbaugh et al., 2001; Greenwood et al., 2018). The latter process may facilitate integration of the dsDNA copies of retroviral genomes ("proviruses", see below) into the DNA of germ cells, followed by vertical transmission ("endogenization") (Greenwood et al., 2018).

After entry of the virus into the host cell the viral reverse transcriptase (RT) associated with subviral core particles converts the ssRNA genome into a dsDNA copy which integrates into the host DNA with the help of the viral integrase enzyme. Retroviral RT initiates minus-strand DNA synthesis at the 3′-OH of a distinct cellular tRNA molecule annealed to the primer binding site located near to the 5′ end of the viral RNA genome. The minus–strand DNA serves as a template for plus-strand DNA synthesis. The dsDNA form of the retroviral genome is created by the coordinated action of the polymerase activity and RNaseH activity of RT. The retroviral dsDNA genome is flanked by identical regulatory sequences (long terminal repeats, LTRs) at both ends. Although the 5′ LTR as well as the 3′ LTR of an integrated provirus carries transcriptional promoter and terminator sequences, transcription of the retroviral genome is typically initiated at the 5′ LTR and terminated at the 3′ LTR. (For the details of the intricated process of reverse transcription see (Temin, 1993; Champoux and Schultz, 2009)) (Fig. 1). Following insertion into the host genome, the retroviral dsDNA called "provirus" is transcribed by cellular RNA polymerase II and genome-length as well as shorter, spliced viral RNAs are transported into the cytoplasm for translation. The capsid and envelope proteins are packaged into the virion together with two copies of the retroviral RNA genome, retroviral enzymes and other virion components such as cellular RNA molecules. The assembled viral particle coated by the plasma membrane leaves the cell by budding and then undergoes proteolytic processing to form mature virus (Justice and Beemon, 2013; Telesnitsky and Wolin, 2016).

The family *Retroviridae is divided into two subfamilies, of Orthoretrovirinae* (orthoretroviruses) and *Spumaretrovirinae* (spumaviruses) (Benachenhou et al., 2013). A subset of orthoretroviruses, called collectively "RNA tumor viruses", is associated with a wide variety of neoplasms in their host. Other members of *Orthoretrovirinae,* such as the human imunodeficiency virus (HIV) which belongs to the *Lentivirus* genus, cause various slowly progressing disorders, including immunodeficiency syndromes in vertebrates (Vogt, 2012; Denner, 2014; Greenwood et al., 2018). In contrast, the viruses forming the subfamily *Spumaretrovirinae* cause apathogenic infection in their natural hosts (Pinto-Santini et al., 2017).

It is a unique feature of RNA tumor viruses that their genomes could acquire, from time to time, cellular sequences derived fom proto-

Table 1
Major mechanisms of viral oncogenesis

RNA tumor viruses	DNA tumor viruses
1. Expression of oncoproteins encoded by *v-onc* genes of host cell origin; stimulation of cell growth, induction of genetic and epigenetic alterations	1. Expression of oncoproteins encoded by viral oncogenes of unknown origin or host cell origin, inhibition of cellular tumor suppressor proteins; induction of genetic and epigenetic alterations
2. Activation or dysregulation of *c-onc* genes by promoter or enhancer insertion	2. Induction of genetic instability by integration of viral DNA into the host genome
3. Expression of transforming structural proteins	3. Chronic inflammation [HBV]
4. Expression of transforming accessory (regulatory) viral proteins	
5. Chronic inflammation [HCV]	

Abbreviations: HCV, hepatitis C virus; HBV, hepatitis B virus

oncogenes (also called *c-onc* genes) (reviewed by (Sugden, 1993)). In most cases, structural alterations and mutations contributed to the conversion of the transduced cellular sequences into viral oncogenes (*v-onc* genes) (reviewed by (Vogt, 2012)). The viral oncoproteins encoded by *v-onc* genes are responsible for the tumorigenicity of "acutely transforming" retroviruses. Acquisition of proto-oncogene sequences typically coincides with the loss of replication competence due to deletions of the retroviral genome. Thus, most "acutely transforming" retroviruses are replication defective and can only be propagated in cells coinfected with a wild-type, helper retrovirus or in cell lines engineered to express the product(s) of the deleted viral genes (Mann et al., 1983; Miller and Buttimore, 1986; Vogt, 2012). The only replication competent "acutely transforming" retroviruses are, as far as we know, certain strains of Rous sarcoma virus, a chicken retrovirus carrying the *v-src* gene.

Slowly transforming, non-acute RNA tumor viruses are replication competent. In spite of the fact that their genome does not harbour a *v-onc* gene, non-acute retroviruses are able to induce late onset leukemia or lymphoma in vertebrate hosts by promoter or enhancer insertion (Fan and Johnson, 2011) (see also Introduction).

There are several scenarios regarding the transduction of cellular proto-oncogene sequences by retroviruses. Each scenario proposed two consecutive recombination events between retroviral and cellular nucleic acids. According to the "DNA model", the first integration of retroviral DNA into cellular DNA sequences could result in a 5′ retroviral - 3′ protooncogene hybrid gene (indicated by thick blue lines on the left side of Fig. 2). This would be followed by a second recombination event, i.e. by the integration of a second proviral genome (shown as a thick green bar on the right side of Fig. 2.) into the host cell DNA, 3′ from the first integration site (Goodrich and Duesberg, 1990; Schwartz et al., 1995). Thus, both recombinations would happen on the "DNA level". It was proposed that the provirus integrated 5′ from the *c-onc* gene would carry only a single LTR, where a hybrid retroviral-proto-oncogene transcript could be initiated (Fig. 2). The polyadenylation signal of the hybrid transcript would be provided, exceptionally, by the 5′ LTR of the 2nd integrated retroviral genome (Goodrich and Duesberg, 1990; Schwartz et al., 1995). Removal of cellular introns and polyadenylation of the hybrid transcript would be followed by incorporation of the hybrid RNA into viral particles along with a wild type, full-length retroviral RNA genome, generating heterozygote retroviruses (Fig. 2). Coinfection of target cells by the transducing virus and a replication competent, non-defective "helper" virus could ensure the maintenance of the transduced cellular oncogene during the infection of new target cells. Because the transduction of a *c-onc* sequence may not cause malignant transformation *per se*, it was suggested that subsequent structural changes, mutation and selection processes may result, during viral replication, in the generation of an acute transforming retrovirus with a "transforming" *v-onc* gene (Iba et al., 1984).

According to the "RNA model" for the generation of transducing retroviruses, the first recombination could also happen on the "DNA level" by integration of the proviral genome into cellular DNA sequences. The second recombination event would happen, however, during reverse transcription, i.e. RNA-dependent synthesis of a dsDNA molecule using the copackaged hybrid RNA and the wild-type retroviral RNA genome as template (Klempnauer and Bishop, 1983; Swanstrom et al., 1983; Swain and Coffin, 1992; Zhang and Temin, 1993, 1994) (Fig. 3). Template switching during minus strand DNA synthesis by reverse transcriptase, aided by short stretches of sequence homology between the copackaged RNAs could facilitate the second recombination event (reviewed by (Sugden, 1993)).

A third scenario, the "Sequence acquisition by reverse transcription" model, proposed that transduction of a cellular sequence by a retrovirus may happen without the involvement of a recombination step at the "DNA level" (Hajjar and Linial, 1993). Two consecutive recombinations on the "RNA level", or more precisely during RNA-dependent DNA synthesis, could possibly explain how an endogenous retrovirus could transduce *FAM8A1* (*family with sequence similarity 8 member A1*), a human gene encoding a cofactor associated with a ubiquitin ligase, during primate evolution (Jamain et al., 2001) (Fig. 4.). *FAM8A1* is not known to be involved in tumorigenesis. The frequency of host mRNA encapsidation into retroviral particles appears to be low, however, although distinct non-coding cellular RNAs are preferentially packaged into the virions of retroviruses (reviewed by (Telesnitsky and Wolin, 2016)). We would like to add that incorporation of cellular sequences into proviral DNA was also recorded during the replication of retrovirus vectors in canine cells. The data indicated that during plus-strand synthesis the reverse transcriptase possibly switched to a copackaged canine mRNA, copied a stretch of it, and then switched back to the minus strand of the viral DNA (Pulsinelli and Temin, 1991).

According to the model of Hajjar and Linial, due to the unique features of retroviral replication, even a single recombination event between viral and cellular sequences could result in the incorporation of cellular sequences into the retroviral genome (Hajjar and Linial, 1993). Such a nonhomologous recombination, possibly aided by short stretches of homology, could take place during reverse transcription of copackaged cellular and viral RNA molecules, i.e. without involving the integration of the viral DNA, and may result in the acquisition of cellular sequences by retroviral genomes. The scenario may be similar to the late steps of the "RNA model" for the transduction of proto-oncogene sequences (see Fig. 3) (Hajjar and Linial, 1993). The recombination events that shape the genomes of transducing acute RNA tumor viruses affect, in most cases, both the cellular sequences acquired and the retroviral *gag*, *pol* or *env* genes. Transduced cellular sequences may fuse to viral genes (as discussed below) and may undergo further structural changes and mutations during successive rounds of retrovirus replication (reviewed by (Correll et al., 2006).

3. Acute transforming retrovirus genomes frequently encode truncated versions of cellular proteins fused to retroviral proteins

Because the cellular sequences transduced by and inserted into retroviral genomes frequently replace viral genes or gene fragments, most of the acute transforming RNA tumor viruses are defective, i.e. they can only replicate in cells co-infected with a replication competent helper virus. Deletions occur, however, not only in the viral genome, but also at 5′ sequences, 3′ sequences or at both ends of the transduced

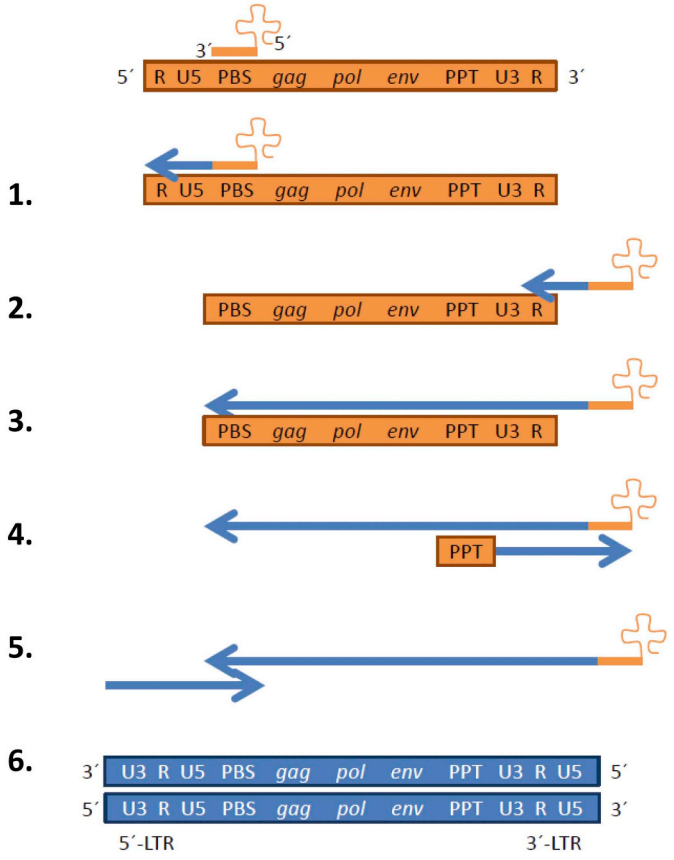

Figure 1. Reverse transcription of the retroviral ssRNA genome (Based on (Sugden, 1993; Temin, 1993; Champoux and Schultz, 2009))
Major steps
1. Initiation of minus-strand DNA synthesis by reverse transcriptase (RT)
2. Template switch 1
3. Elongation of minus-strand DNA
4. RNaseH cuts the RNA 3' from ppt; start of plus-strand DNA synthesis
5. Template switch 2 and degradation of viral RNA by RNaseH
6. Plus-strand DNA elongation and creation of long terminal repeats (LTRs) composed of U3, R and U5 sequences at both ends of the viral DNA genome which integrates into the host cell DNA (provirus). Typically, promoter and enhancer sequences located to the U3 region of the 5′ LTR control the initiation of retroviral transcription by RNA polymerase II, whereas sequences situated to the R-U5 region of the 3′LTR regulate transcriptional termination.
Symbols
a cloverleaf-like symbol indicates a cellular tRNA molecule used as a primer for minus-strand DNA synthesis by reverse transcriptase. Ochre lines depict RNA sequences, blue lines with arrows indicate viral DNA sequences. Thick ochre line: coding sequences for the retroviral genes *gag*, *pol* and *env* located to the viral RNA genome; thin ochre line: noncoding sequences in viral RNA.
Abbreviations
r: repeat region at the termini of viral RNA
R: repeat region in viral DNA
u5: unique sequence at the 5′ end of viral RNA
U5: unique 5′ region (DNA)
pbs: primer binding site in viral RNA
PBS: priner binding site in viral DNA
ppt: polypurine tract, priming plus-strand DNA synthesis, u3: unique sequence at the 3′ end of viral RNA
U3: unique 3′ region (DNA)
RT: reverse transcriptase

cellular genes that are converted into *v-onc* genes. In-frame insertion of transduced genes into viral genes results, in most cases, in the generation of fused transcripts translated as fusion proteins. The best characterised fusion proteins of acute RNA tumor viruses are listed in Table 2. Most of them are composed of a truncated retroviral Gag polyprotein fused to a truncated protein of cellular origin. Less frequently the truncated version of Env, another viral structural protein is fused to a shortened cellular protein. In rare cases, the truncated forms of the retroviral enzymes encoded by the *pol* gene may also participate in the formation of fusion oncoproteins (Table 2).

Gag sequences located at the N-terminal domain of retroviral transforming proteins play a context-dependent role in tumorigenesis. The Gag-Abl fusion protein encoded by A-MuLV (Abelson murine leukemia virus) maintained its fibroblast transforming capacity *in vitro*, even in the absence of its Gag domain (Prywes et al., 1983). The Gag sequences fused to Myc were also dispensable for fibroblast transformation by avian myelocytomatosis virus MC29 (Shaw et al., 1985). In contrast, the Gag domain of the Gag-Abl oncoprotein was indispensable for the lymphoid transforming activity of A-MuLV, because loss of the Gag domain decreased the stability of the fusion protein in lymphoid cells (Prywes et al., 1983; Prywes et al., 1985).

The genome of Hardy-Zuckerman 2 feline sarcoma virus (HZ2-FeSV) also acquired *c-abl* sequences and encodes a ΔGag-Abl-ΔPol transforming protein (Bergold et al., 1987a; Bergold et al., 1987b). Deletion of the major part of the Pol domain of the fusion protein did not affect, however, the fibroblast transforming potential of the fusion protein. In addition, when expressed from a retroviral construct infecting mice, the feline HZ2-FeSV oncoprotein with a C-terminal Pol domain induced lymphosarcomas of B cell origin *in vivo,* i.e. it transformed B-cell lineage cells similarly to the ΔGag-Abl oncoprotein of A-MuLV (Bergold et al., 1987a; Bergold et al., 1987b).

The Env-Sis fusion oncoprotein encoded by the *v-sis* gene of simian sarcoma virus (SSV) induces fibroblast transformation by an autocrine mechanism by binding to the PDGF (plateled derived growth hormone) receptor both at the cell surface and intracellularly (Betsholtz et al., 1986; Lokeshwar et al., 1990). There is a signal sequence in the N-terminal Env domain of the Env-Sis fusion protein that directs the translocation of the nascent oncoprotein through the endoplasmic reticulum (Devare et al., 1983; Hannink and Donoghue, 1984). The simian sarcoma associated virus (SSAV)-derived Env domain is important for the biological activity of the fusion protein, because removal of *v-sis* sequences coding for the signal peptide located to the Env domain interfered with cell transformation *in vitro* (Hannink and Donoghue, 1984). It is noteworthy that the genome of PI-FeSV (Parodi-Irgens feline sarcoma virus), an RNA tumor virus isolated from a cat fibrosarcoma, carried a homologue of SSV *v-sis* (Besmer et al., 1983). In this case, however, a Gag-Sis fusion protein was expressed in transformed fibroblasts (Besmer et al., 1983). Thus, expression of homologous cellular proteins fused to the C-terminal end of either truncated Env or truncated Gag induced similar changes in the behaviour of SSV and PI-FeSV infected cells, respectively.

4. Hepatitis B virus: a reverse transcribing human tumor virus that does not transduce cellular sequences

Hepatitis B virus (HBV), a reverse-transcribing small DNA virus is one of the causative agents of hepatocellular carcinoma in humans. HBV belongs to the family of *Hepadnaviridae,* and it is the type species of the genus Orthohepadnavirus. It infects human hepatocytes by binding to its unique receptor, the sodium taurocholate co-transporting polypeptide (NTCP) expressed on the surface of liver cells (Yan et al., 2012a, 2012b). Recently, another HBV receptor, hepatitis B surface antigen binding protein (SBP), was also identified (Sun et al., 2018). The restricted host range of HBV is usually attributed to the cell-type specific expression of its receptors, although the expression of liver-specific transcription factors involved in HBV replication may also contribute to HBV tropism (Tang and McLachlan, 2002).

In addition to humans, HBV variants infect apes as well (Starkman et al., 2003; Sa-Nguanmoo et al., 2009; Lyons et al., 2012). Furthermore, viruses in the *Orthohepadnavirus* genus cause liver infection in other mammalian species including New World monkeys (Woolly

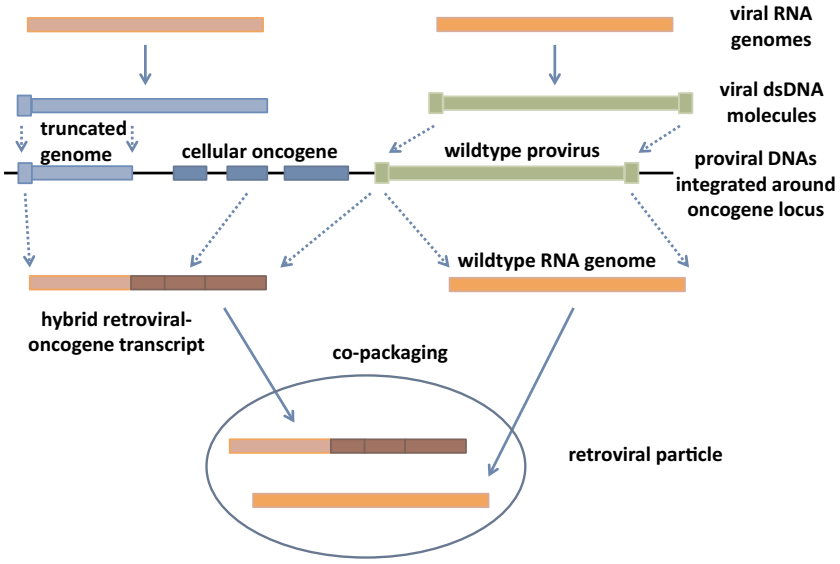

Figure 2. Transduction of cellular proto-oncogene sequences by retroviruses: The DNA model (Based on (Goodrich and Duesberg, 1990; Schwartz et al., 1995)).
On the left side of the figure integration of a truncated retroviral genome (indicated as a thick blue line) with a single LTR is shown. This integration into the host cell DNA (shown as a discontinuous thin black line) occurs 5′ from a cellular oncogene (*c-onc* gene; represented by thick horizontal blue lines alternating with thin black lines). On the right side of the figure, integration of a second, full-length, wild type provirus with two LTRs is depicted (thick green line), 3′ from the cellular oncogene. Transcripts initiated at the single LTR of the truncated provirus proceed through the *c-onc* gene and terminate at the 5′ LTR of the wild type provirus. Broken arrows indicate processing of the primary transcript yielding a hybrid retroviral-oncogene transcript co-packeged into a retroviral particle with the full length, wild type, helper retroviral genome (shown as a thick orange line) transcribed from the integrated provirus flanked by two LTRs.

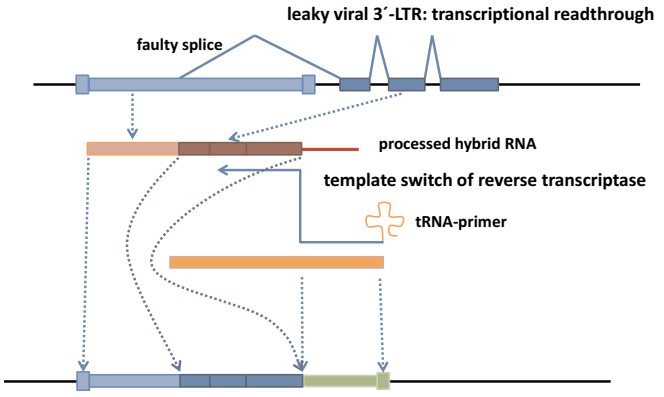

Figure 3. Transduction of cellular proto-oncogene sequences by retroviruses: The RNA model (Based on (Klempnauer and Bishop, 1983; Swanstrom et al., 1983; Swain and Coffin, 1992; Zhang and Temin, 1993)).
Integration of a retroviral genome (light blue) into the host DNA 5′ from a *c-onc* gene (exons in dark blue) may result in a processed hybrid RNA molecule (purple and dark brown). Faulty splice of the viral RNA as well as deficient termination of the viral transcript at the 3′ LTR may facilitate this process. Co-packaging of the hybrid RNA and a full length (helper) retroviral genome (ochre) into a viral particle may favor, after infection of a new host cell, non-homologous recombination during minus-strand synthesis which generates a recombinant provirus (blue and green) integrated into the host DNA (black).

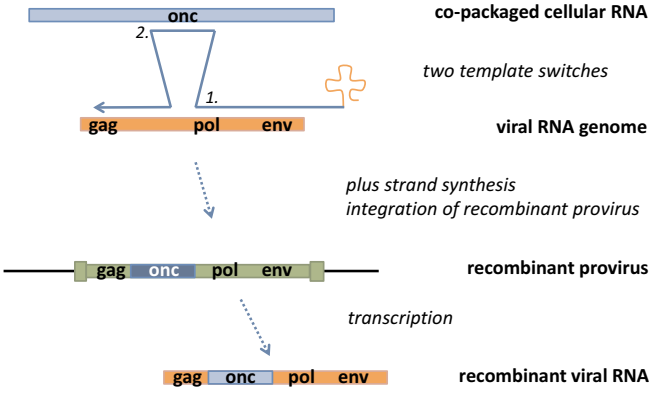

Figure 4. Transduction of cellular proto-oncogene sequences by retroviruses: The sequence acquisition by reverse transcription model (Based on (Hajjar and Linial, 1993; Jamain et al., 2001))
During reverse transcription, the minus-strand DNA synthesis (thin blue line) which started on a retroviral RNA genome (ochre) proceeds on the co-packaged cellular RNA molecule (light blue) and then switches back to the retroviral RNA genome, as indicated on the figure. Then, plus-strand DNA synthesis may generate a recombinant provirus (according to Fig. 1, not shown here). Transcription of the recombinant provirus may yield a hybrid retroviral-cellular RNA transcript (light blue and ochre). Alternatively, the second illegitimate recombination may occur during plus strand synthesis, i.e. the primer initiated at ppt (see step 4, Fig. 1) may jump to the 5′ end of the hybrid cellular-viral RNA (not shown here) to complete the synthesis of the recombinant proviral DNA.

Monkey HBV, WMHBV, Capuchin Monkey hepatitis B virus, CMHBV) and rodents (Woodchuck hepatitis virus, WHV; Ground Squirrel hepatitis virus, GSHV) (de Carvalho Dominguez Souza et al., 2018) (reviewed by (Mason, 2015)). In contrast, members of the *Avihepadnavirus* genus (prototype: Duck hepatitis virus, DHV) infect birds (reviewed by (Mason, 2015)). Although *Hepadnaviridae* are DNA viruses, they replicate their DNA genomes *via* reverse transcription of an RNA intermediate, similarly to the members of the *Caulimoviridae* family that infect plant cells (Mesnard and Carriere, 1995). The reverse transcribing DNA viruses belonging to *Caulimoviridae* and *Hepadnaviridae* are grouped together as pararetroviruses based on their replication strategy. In contrast to retroviruses, however, the polymerase proteins encoded by pararetrovirus genomes lack a functional integrase domain, and integration of the viral DNA into the host genome is not an obligatory step during pararetrovirus replication (reviewed by (Krupovic et al., 2018)).

Whereas retroviral RTs perform reverse transcriptase and RNaseH activities, the corresponding polymerase enzyme (P) encoded by the partially double-stranded, relaxed circular (RC) DNA genomes of *Hepadnaviridae* has an additional function as well, called protein priming. Thus, in contrast to retroviruses that use one of the cellular tRNAs to prime reverse transcription, in case of *Hepadnaviridae* tyrosine 63 of the N-terminal domain of P protein provides the necessary 3′-OH group for minus DNA strand synthesis (Feng et al., 2013) (reviewed by (Lamontagne et al., 2016)). Hepadnaviruses are unique, because the terminal protein domain (TP domain) and the reverse transcriptase domain are part of the same polypeptide that forms a complex with the pregenomic RNA (pgRNA) pinned to the immature capsids (Lanford et al., 1997; Wang et al., 2014c). A bulge region of the epsilon encapsidation signal at the 5′ end of the pgRNA is used as a template for the synthesis of a three or four nucleotide long DNA primer which is

Table 2
Oncoproteins encoded by acute transforming retroviruses

Fusion protein	Virus	Host	Reference/review
Gag-Abl			
P160ΔGag-Abl	A-MuLV (Abelson-MuLV[a])	Mouse	(Shore et al., 2002)
P120ΔGag-Abl	A-MuLV	Mouse	(Reddy et al., 1983)
P110ΔGag-Abl-ΔPol	HZ2-FeSV (Hardy-Zuckerman 2 FeSV[b])	Cat	(Bergold et al., 1987a)
Gag-X-Akt			
P105ΔGag-X-Akt	AKT8 MuLV[a]	Mouse	(Bellacosa et al., 1991; Bellacosa et al., 1993)
Gag-Cbl			
P100ΔGag-Cbl	Cas NS-1 MuLV[a]	Mouse	(Blake et al., 1993)
Gag-Crk			
P47ΔGag-Crk	CT10 ASV (Avian Sarcoma Virus)	Chicken	(Mayer et al., 1988)
	ASV-1	Chicken	(Tsuchie et al., 1989)
Gag-Fes (Fps)			
P110ΔGag-Fes	GA-FeSV (Gardner-Arnstein FeSV[b])	Cat	(Veronese et al., 1983)
P85ΔGag-Fes	ST-FeSV (Snyder-Theilen FeSV[b])	Cat[d]	
P100ΔGag-Fes	HZ1-FeSV (Hardy-Zuckerman 1 FeSV[b])	Cat	(Snyder Jr. et al., 1984)
Gag-Fgr			
P70ΔGag-Actin-Fgr	GR-FeSV (Gardner-Rasheed FeSV[b])	Cat	(Naharro et al., 1984)
Gag-Fms			
P180ΔGag- Fms	SM-FeSV (McDonough strain of FeSV[b])	Cat	(Anderson et al., 1982)
P140ΔGag- Fms			
P120ΔGag- Fms			
Gag-Fos			
P100ΔGag-Fos	NK24 avian transforming virus	Chicken	(Nishizawa et al., 1987)
P75ΔGag-Fos-Fox	FBR murine osteosarcoma virus	Mouse	(Van Beveren et al., 1984)
Gag-Fps (Fes)			
P13 ΔGag-Fps	FSV (Fujinami sarcoma virus)	Chicken[d]	(Shibuya and Hanafusa, 1982) (Vogt et al., 1981; Huang et al., 1985)
P105ΔGag-Fps	PRCII ASV (Poultry Research Center Virus II; an Avian Sarcoma Virus)	Chicken	(Breitman et al., 1981)
	PRCIV ASV		(Wang et al., 1981)
P170ΔGag-Fps	UR1 ASV	Chicken	(Neel et al., 1982)
P150ΔGag-Fps	16L virus (an Avian Sarcoma Virus)	Chicken	(Wang et al., 2016b)
P142ΔGag-Fps	Fu-J virus (subgroup J Avian Leukosis Virus-associated acutely transforming	Chicken	
P137ΔGag-Fps	virus)	Chicken	
Gag-Jun			
P65ΔGag-Jun	ASV17	Chicken	(Bos et al., 1988)
Gag-Kit			
P80ΔGag-Kit	HZ4-FeSV (Hardy-Zuckerman 4 FeSV[b])	Cat	(Besmer et al., 1986b)
Gag-Mil (Raf)			
P100ΔGag-Mil	MH2 (avian oncovirus Mill Hill No. 2)	Chicken[e]	(Kan et al., 1984)
Gag-Myb			
P135ΔGag-Myb-Ets	E26 avian acute leukosis virus	Chicken	(Lipsick and Wang, 1999)
P48δGag-Myb-δEnv	AMV	Chicken	(Lipsick and Wang, 1999)
Gag-Myc			
P110ΔGag-Myc	MC29[c]	Chicken	(Bister and Jansen, 1986)
P90ΔGag-Myc	CMII (avian acute leukemia virus)		
P200ΔPol-Myc	OK10 (avian acute leukemia virus)		
P57δGag-Myc	OK10		
P57δGag-Myc	MH2 (avian oncovirus Mill Hill No. 2.)		
P145ΔGag-Myc	FH3 (avian acute transforming virus)	Chicken	(Chen et al., 1989)
P72ΔGag-Myc	ALV-J strain 966 (Avian Leukosis Virus subgroup J, strain 966)	Chicken	(Chesters et al., 2001)
Gag-Raf (Mil)			
P75ΔGag-Raf	MuSV3611	Mouse[e]	(Kan et al., 1984)
Gag-Ras			
P29ΔGag-Ras	RaSV (Rasheed rat sarcoma virus)	Rat	(Gonda et al., 1982)
Gag-Ros			
P68ΔGag-Ros	ASV UR2	Chicken	(Neckameyer and Wang, 1984)
Gag-Sis			
P76ΔGag-Sis	PI-FeSV (Parodi-Irgens feline sarcoma virus)	Cat[f]	(Besmer et al., 1983)
Gag-Yes			
P90ΔGag-Yes- δEnv	ASV Y73	Chicken	(Zheng et al., 1989)
Gag-Ski			
P125 Gag-Ski	SKVs (Sloan-Kettering viruses)	Chicken	(Stavnezer et al., 1989)
Env-Mos			
P41δEnv-Mos	Mo-MuSV (Moloney murine sarcoma virus	Mouse	(Donoghue, 1982; Donoghue and Hunter, 1983)
	Mo-MuSV		(Herzog et al., 1990)
(p37-v-mos, p43-v-mos)			
Env-Mpl			
P29ΔEnv-Mpl	MPL (Myeloproliferative leukemia virus)	Mouse	(Courtois et al., 1995)
Env-Rel			
P59δEnv-Rel	Rev-T (Reticuloendotheliosis virus strain T)	Turkey	(Wilhelmsen et al., 1984; Garson et al., 1990)
Env-Sis			
P28δEnv-Sis	Simian sarcoma virus	Woolly monkey[f]	(Devare et al., 1983)

(*continued on next page*)

Table 2 (*continued*)

Fusion protein	Virus	Host	Reference/review
Env-Sea			
gp155ΔEnv-Sea	S13 virus	Chicken	(Smith et al., 1989)
gp70ΔEnv-Sea	Avian erythroblastosis virus S13		(Knight et al., 1988)
Viral oncoprotein not fused to viral structural protein			
pp60-v-src	ASV (avian sarcoma virus)	Chicken	(Levinson et al., 1978)
	SJ-1, SJ-2, SJ-3 (acute transforming viruses associated with ALV-J)	Chicken	(Wang et al., 2016a)
P61-v-erbB	AEV-R (avian erythroblastosis virus, strain R)	Chicken	(Ng and Privalsky, 1986)
	AEV-R		
gp65-v-erbB	AEV-R		(Bassiri and Privalsky, 1986)
gp68-v-erbB			(Nishida et al., 1984; Bassiri and Privalsky, 1986)
	AEV-R		
gp73-v-erbB	AEV-H (strain H)	Chicken	(Ng and Privalsky, 1986)
P68-v-erbB (predicted)	AEV-H		(Yamamoto et al., 1983)
gp72-v-erbB			(Nishida et al., 1984)
P55-v-fos	FBJ murine osteosarcoma virus	Mouse	(Curran et al., 1984)
P21-v-has	Ha-MuSV rat-derived murine sarcoma retrovirus	Rat	(Andersen et al., 1981)
P21-v- kis	Ki-MuSV rat-derived murine sarcoma retrovirus	Rat	(Tsuchida et al., 1982)
P21-v-bas	BALB-MuSV Spontaneous murine (mouse) sarcoma virus		(Andersen et al., 1981; Reddy et al., 1985)
Retroviral structural protein acting as an oncoprotein			
$Pr80^{Env}$	Jaagsiekte sheep retrovirus (JSRV)	Sheep	(Hsu et al., 2015)
$P70^{Env}$			

Symbols: Δ indicates a deletion in a retroviral structural protein fused to a c-onc protein; δ indicates that due to a major deletion, only a few amino acids of a retroviral structural protein are fused to a c-onc protein

[a] MuLV: Murine Leukemia Virus

[b] FeSV: Feline Sarcoma Virus

[c] MC29: MC29 strain of Avian Myelocytomatosis Virus

[d] Gag-Fes and Gag-Fps are encoded by a cat and a chicken retrovirus, respectively; the same cellular gene was transduced by retroviruses infecting different host species

[e] Gag-Mil and Gag-Raf are encoded by a chicken and a murine retrovirus, respectively; the same cellular gene was transduced by retroviruses infecting different host species

[f] Gag-Sis and Env-Sis are encoded by a cat and a primate retrovirus, respectively; the same cellular gene was transduced by retroviruses infecting different host species

covalently linked to the polymerase (Kidd and Kidd-Ljunggren, 1996). This short nascent DNA strand and the linked P protein is transferred thereafter to the 3′ end of the terminally redundant pgRNA for the synthesis of a full-length DNA minus strand (Fig. 5). (To do) Simultaneously, the RNaseH activity of the P protein, lagging 18 nt behind the site of reverse transcription, degrades the RNA template except a terminal RNA stretch which is used as a primer for the synthesis of the DNA plus strand. In the majority of nucleocapsids primer translocation from the 5′-end to the 3′-end of the minus DNA strand template, aided by direct repeats and loop formation (circularization of the template) ensures the generation of relaxed circular DNA (rcDNA) genomes. These rcDNA genomes are not completely double stranded due to the premature termination of plus-strand synthesis (reviewed by (Seeger and Mason, 2015; Tu et al., 2017)). In a minority of nucleocapsids, however, primer translocation does not occur and the linear dslDNA genomes synthesized may relocate to the nucleus and integrate into the host cell genome. As a matter of fact, integration of HBV genomes was regularly observed in the genome of hepatocellular carcinomas as well as in noninvolved tumor-adjacent tissues of patients with liver cancer (Ding et al., 2012; Jiang et al., 2012a) (reviewed by (Honda, 2017)). Integration preferentially occurred in open chromatin regions of human liver cancer samples and HBV-infected non-tumorous liver tissues (Furuta et al., 2018). It is worthy to note that in a chimeric mouse model, HBV integration was detected in mitochondrial DNA (Furuta et al., 2018). Analysis of integration breakpoints showed the involvement of the viral enhancer as well as the X protein (HBx) and core protein coding sequence (Yang et al., 2018). Insertion of the HBV genome into the vicinity of distinct cellular genes correlated with shorter disease-free survival and overall survival of liver cancer patients (Yang et al., 2018). In patients with chronic HBV infection, integration of HBV genomes occurred both in normal-appearing hepatocytes and premalignant hepatocytes and it was associated with clonal expansion of both cell populations (Mason et al., 2010). In liver cancer cells, preferential integration of HBV genomes into the vicinity of cellular genes implicated in carcinogenesis suggested that up-regulation of distinct growth-regulators may play a role in the development of hepatocellular carcinoma (reviewed by (Hai et al., 2014; Honda, 2017)). Thus, a mechanism similar to the promoter and enhancer insertion observed in non-acute retrovirus infected cells may operate during HBV-associated carcinogenesis, too (Gilroy et al., 2016; Serrao and Engelman, 2016; Tsuruyama et al., 2017), see also: Introduction).

5. Organization of retrovirus and hepadnavirus genomes

Due to the peculiar features of reverse transcription, the ssRNA genomes of retroviruses differ from the corresponding dsDNA genomes: duplication of promoter, enhancer and polyadenylation sequences creates two identical LTRs (long terminal repeats) at the ends of the provirus which control the transcription of the retroviral genes. For a detailed description of this process involving two strand transfers, i.e. jumps of the reverse transcriptase and the associated DNA primers see (Temin, 1993; Hughes, 2015) and Fig. 2. Thus, the regulatory sequences of the provirus do not overlap with the coding sequences of retroviral proteins. In contrast, the hepadnavirus genomes are more compact: the enhancer and promoter sequences are not organized into a separate control region; instead, they overlap with the open reading frames (ORFs) of viral proteins (Torres et al., 2013). In addition, there is a partial overlap between the the four ORFs of HBV and related members of *Hepadnavirinae*. Such a compact organization of the viral genome may influence the variability and evolution of HBV and other members of the *Hepadnaviridae* family (Torres et al., 2013).

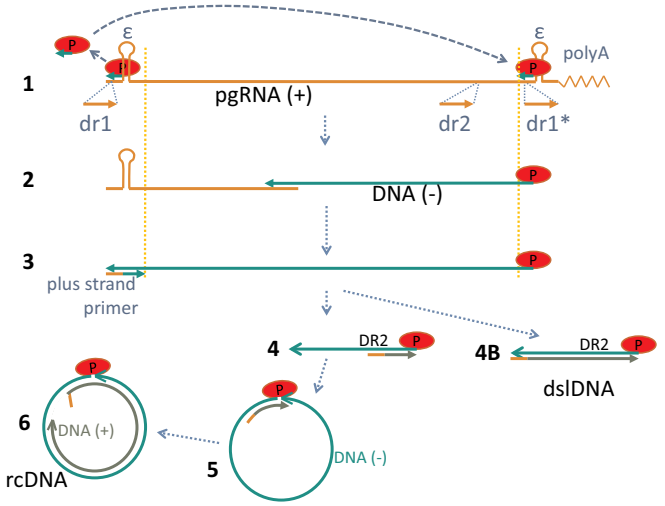

Figure 5. Reverse transcription of the pregenomic RNA of hepatitis B virus (Based on (Beck and Nassal, 2007; Seeger and Mason, 2015)).
1. Using an RNA hairpin (ε, the encapsidation signal) located to the 5′ region of the pregenomic RNA (pgRNA) as template, the polymerase (P) enzyme of HBV initiates DNA synthesis by protein priming, i.e. the first DNA nucleotide is covalently linked to the terminal protein domain of the polymerase-reverse transcriptase enzyme. After the synthesis of the first 3 nucleotides, the primer-polymerase complex is translocated (1st template switch; blue dashed line) to a repetitive sequence, dr1* which is situated to the 3′ end of the pgRNA.
2. The DNA primer is extended to the 5′ end of the pgRNA (minus-strand DNA synthesis). In parallel, the RNaseH activity of the polymerase degrades the RNA template except a repeat (dr1, direct repeat 1) at the 5′ end of the pgRNA. **3.** As a next step, dr1 is translocated and binds to the DR2 sequence of minus-strand DNA (second template switch) where it used as a primer for plus-strand DNA synthesis which continues till the end of the minus DNA strand (**4**). After circularization (**5**) aided by a repeat sequence at the end of the minus-strand DNA, the plus-strand synthesis continues but remains incomplete resulting in a double stranded, relaxed circular DNA (rcDNA) genome (**6**). If the plus strand primer fails to translocate, double stranded linear DNA (dslDNA) is generated ("in situ priming", **4B**).

6. Creation of HBx, a pleiotropic oncoprotein of orthohepadnaviruses, by overprinting

During retroviral replication reverse transcription of the viral RNA genome occurs at the initial phase of the viral life cycle creating dsDNA genomes that integrate into the host DNA (see Section 2.). In contrast, reverse transcription takes place at a late stage of hepadnavirus replication. After cell entry, and nuclear transport, the rcDNA genomes of hepadnaviruses are converted into covalently closed circular DNA (cccDNA). Transcription of the viral episomes generates viral mRNAs and pregenomic RNAs which are translated in the cytoplasm, where the pregenomic RNAs associate with the polymerase enzyme and capsid proteins. Reverse transcription generates rcDNA in the immature capsids just before the packaging of the nucleocapsids into viral envelopes and the release of mature virions (reviewed by (Seeger and Mason, 2015)). In spite of the different replication strategies of retroviruses and hepadnaviruses, they show some similarities in the arrangement and function of the viral genes (Mesnard and Carriere, 1995). The retroviral *gag* and *pol* genes code for the structural proteins of the virion and the viral replication enzymes, respectively. A third gene, *env*, encodes the glycoproteins inserted into the lipid envelope of the virus. These retroviral genes correspond to the partly overlapping open reading frames (ORFs) of hepadnavirus genomes coding for precore and core protein, polymerase (P) protein and envelope proteins (large, middle, and small surface antigen (S, HBs or HBsAg) which have analogous functions as their retroviral counterparts (Mesnard and Carriere, 1995; Lamontagne et al., 2016). There is, however, an additional, unique ORF in the genome of orthohepadnaviruses, coding for the regulatory protein HBx (also called X protein), which is absent from retroviral genomes. HBx facilitates HBV replication, but it is also regarded as an oncoprotein that plays a role in the development of hepatocellular carcinoma in humans as well as in rodents infected with other members of the *Orthohepadnavirus* genus (reviewed by (Lamontagne et al., 2016)). Unlike the viral oncoproteins of acute RNA tumor viruses, however, HBx is not specified by a *c-onc* gene acquired by the virus. Instead, it appears to be the product of a novel way of protein creation, called overprinting.

Overprinting the +2-frame or +3-frame or both frames within an already existing +1-open reading frame (ORF) is an evolutionary process which gains coding space without enlarging a genome. Overprinting is found in cellular genomes, but due to size constraints, it is mainly used by viruses (Keese and Gibbs, 1992; Carter et al., 2013; Pavesi et al., 2013). In many cases, phylogenetic analysis combined with codon usage analysis of overlapping viral reading frames is able to discriminate between ancestral and novel, overprinted sequences with an excellent specificity and intermediate sensitivity. For example, small regulatory genes from the pX region of the deltaretrovirus HTLV-1, like Rex, p12 and p30 apparently have been generated de novo by overprinting pre-existing frames, leading Pavesi et al. to call the pX region of HTLV-1 a "gene nursery" (Pavesi et al., 2013).

An extreme and intriguing example for overprinting is Hepatitis B virus (HBV) where 50% of the viral genome is occupied by overlapping regions from which about 70% of the viral coding capacity is derived. First, the common stretch of > 1000 nts of the polymerase (P) and surface genes of HBV is the longest known overlap in animal viruses. The 5′ one-third of the overlap contains both the spacer region of the viral polymerase gene and the pre-S1 and first two thirds of the pre-S2 region of the HBs gene. A detailed codon usage analysis using a sliding window method found that within this 5′ one-third the S gene is ancestral while the P gene was overprinted. Contrary, within the 3′ two-thirds of the overlap which contains both the reverse transcriptase domain of the polymerase gene and the S domain of the HBs gene, the P gene is ancestral, while the HBs gene was overprinted (Pavesi, 2015). This puzzling dual overprinting pattern found its explanation by the discovery of the nackednaviruses. Nackednaviruses are non-enveloped fish viruses with a capsid triangulation number of 3 which are related to hepadnaviruses which exhibit a capsid triangulation number of 4. Both virus families separated from a common ancestor > 400 million years ago. The envelope of hepadnaviruses apparently emerged through the insertion of a spacer between the terminal protein (TP) domain and the reverse transcriptase (RT) domain of the P gene ("5´ one-third ancestral for HBs") and the subsequent overprinting of the RT domain ("3´ two-thirds ancestral for P") (Lauber et al., 2017).

The HBx gene which is only found in orthohepadnaviruses, but not in avihepadnaviruses and ancient endogenous avian and crocodilian HBVs is another example for the de novo origin of a gene by overprinting. A putative X-like ORF has been postulated and a functional gene product described for some avihepadnaviruses such as duck HBV (Chang et al., 2001; Schuster et al., 2002). However, there are several arguments against the existence of an X-like gene in avihepadnaviruses. All ancient endogenous HBVs and some extant avian HBVs carry multiple internal stop codons at conserved positions within the putative X-like ORF, i.e. they don´t carry an X gene (Suh et al., 2014). Contrary to mammalian HBVs where the X-ORF is essential for virus replication, the X-like ORF of duck HBV may be knocked out without any damage to virus infectivity and in vivo growth, suggesting that it might not play a physiological role in the replication cycle (Meier et al., 2003). Further, the GC content of the X/RNAseH domain-overlap is strongly elevated in mammalian HBVs, while the X-like ORF carries a moderate GC-content. Finally, while the P gene is located in the +1 frame across all HBV types, the X ORF of orthohepadnaviruses is located in the +2 frame, but the X-like ORF of duck HBV is located in the +3 frame (Suh et al., 2014). Therefore, Suh et al. argue rather convincingly that the X gene of orthohepadnaviruses arose by a duplication of the 5′-two-thirds of the

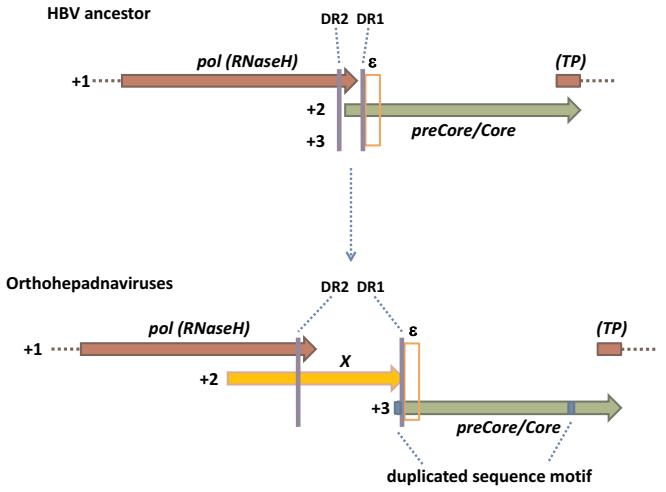

Figure 6. Generation of the HBx open reading frame by overprinting (Based on (Pavesi et al., 2013; Suh et al., 2013; Suh et al., 2014)).
A segment of a putative ancestral HBV genome is shown (top), together with the corresponding region of the extant HBV genome (bottom).
Abbreviations: *pol(RNaseH)* designates the open reading frame (ORF) for the RNaseH domain of polymerase; *(TP)* indicates a segment of the ORF for the terminal protein domain of polymerase; *preCore/Core* stands for the ORF specifying the preCore/Core protein; *X* corresponds to the ORF for the HBx protein; DR stand for direct repeat; ε, the DNA sequence corresponding to the encapsidation signal sequence of the pgRNA. Numbers (+1, +2, +3) designate open reading frames which are depicted as colored arrows. The transition from the ancestral HBV genome to the extant HBV genome (blue arrow) possibly involved a duplication of two thirds of *preCore/Core* ORF followed by a deletion affecting the 5' half of the duplicated sequence and a shift in the reading frame of *preCore/Core from* +2 to +3. Thus, reading frame +2 was liberated allowing the creation of ORF *X* for the HBx protein. In the region of the HBV genome where the *pol/RNaseH* ORF pverlaps with ORF *X*, step by step- removal of stop codons and creation of a new start codon in reading frame +2 could possibly have contributed to the *de novo* creation of a novel viral protein, HBx.

preC/C ORF, followed by a deletion of about half of the preC/C-duplicate, again followed by overprinting of the RNAse H domain of the P gene and the unoccupied +2 frame within the remaining sequence of the preC/C-duplicate (Suh et al., 2013; Suh et al., 2014) (Fig. 6).

7. Evolution of Retroviridae and Hepadnaviridae: selected aspects

Many double-stranded (ds) DNA viruses exhibit a very stable history of co-speciation with their hosts. Therefore, their divergence dates are usually inferred from the divergence dates of their hosts from their fossil records over geological time scales. On this basis, mutation rates of dsDNA viruses have been estimated at 10^{-7} to 10^{-9} nucleotide substitutions per site per year (sny) (Aiewsakun and Katzourakis, 2016). A prime example of such calculations is the family of herpesviruses (McGeoch et al., 1995; McGeoch and Gatherer, 2005). Not only DNA viruses, but also RNA viruses, such as deltaretroviruses, hantaviruses and foamyviruses have a clear and stable co-speciation history together with their hosts, although with occasional examples of cross-species transmission (reviewed in (Aiewsakun and Katzourakis, 2016; Zhang et al., 2018)). Thus, comparable to DNA viruses, RNA viruses appear to be co-speciating with surprisingly low mutation rates of 10^{-7} to 10^{-8} sny (Aiewsakun and Katzourakis, 2016).

Furthermore, paleovirological analysis of endogenous viral elements in genomes of living animals and humans reveal a surprisingly close relationship of paleoviruses with extant viral species, again confirming extremely low mutation rates over geological time scales of tens or hundreds of millions of years (Aiewsakun and Katzourakis, 2015, 2016). Phylogenetic trees of viruses can be calibrated with the help of endogenous viral elements, and the minimum age of virus families can be estimated (Holmes, 2011). A prime example for this is the hepadnavirus family which dates back at least 207 million years (Suh et al., 2013; Suh et al., 2014). With the discovery of the Nackednaviruses in diverse fish species, the precursor family of the hepadnaviruses was identified and could be dated to an age of at least 400 million years (Lauber et al., 2017). Over geological time, all lineages of retroviruses have been incorporated as endogenous retroviral elements (ERVs) into the germ lines of all vertebrate lineages. In humans ERVs comprise approximately 8% of the entire genome. A phylogenetic analysis across a diverse sample of 65 vertebrate genomes uncovered 36.000 different ERV lineages which have been very apt at host switching over time (Hayward et al., 2015). By now it became clear for the retroviruses that they have an ancient marine origin, dating back more than 450 million years, probably up to 550 million years to the early Paleozoic (Aiewsakun and Katzourakis, 2017; Hayward, 2017). All in all, co-speciation of host animals and viruses over geological times is considered as firmly established (Zhang et al., 2018). This long-term geological perspective is in concordance with phylogenomic and structural analyses which provide evidence that eukaryotic viruses are ultimately derived from prokaryotic phages (Koonin et al., 2015). Thus, viruses are probably dating back to the origin of life, and virus-like entities even further.

On the other hand, standard phylogenetic analyses of extant RNA viruses, but also DNA viruses, which are based on the time span between sampling times usually measure mutation rates between 10^{-2} and 10^{-5} sny (Sharp and Li, 1988; Jenkins et al., 2002; Davis et al., 2005; Shackelton et al., 2005; Aiewsakun and Katzourakis, 2016). Thus, when comparing short-term and long-term mutation rates, wide differences of several orders of magnitude are usually obtained. The resulting molecular clock paradoxon of RNA viruses is one of the most intriguing problems in virus evolution studies (Holmes, 2003). A comprehensive phylogenetic analysis of RNA virus mutation rates was provided by Jenkins et al. (Jenkins et al., 2002). The discrepancy between short-term and long-term mutation rates of viruses, termed time dependent rate phenomenon (TDRP), was described across almost all Baltimore virus classes by Aiewsakun and Katzourakis (Aiewsakun and Katzourakis, 2016). Interestingly, an analogy to the TDRP has been also described for living mammals compared to the fossil record of mammals (Kurten, 1959). Thus, not only viruses, but also animals are subject to a TDRP. A couple of hypotheses have been put forward to explain the molecular clock paradoxon or TDRP of viruses (Holmes, 2003; Ho et al., 2011). One obvious explanation for the discrepancy between recent molecular clock estimates for RNA virus origins and phylogenetic evidence for virus-host co-speciation over millions of years would be that mutation rates are not constant over time and may vary widely. If long-lasting periods of slow and latent virus replication would alternate with periods of faster lytic replication cycles, nucleotide substitution rates were reduced and divergence times were extended. Another explanation would be that the statistical models for calculating evolutionary distances are flawed in as yet hidden ways. A third explanation would be that both short-term and long-term estimates are true. This would imply that extant viruses are always mutating quickly and are quickly extinguished as well with only few viruses remaining over geological time scales. What remains in phylogenetic trees then would be just virus families which are related with long branching times and diversified lineages of extant viruses which in retrospect also might include many quasispecies. One implication would be that RNA viruses undergo high rates of both speciation and extinction (Holmes, 2003; Wertheim and Kosakovsky Pond, 2011).

An additional possible explanation for the molecular clock paradoxon of viruses or TDRP might be temporary endogenization and epigenetic silencing of complete viruses in the germline of animals. Endogenization might affect only a percentage of individuals of a species. Occasionally, endogenous viruses might be mobilized under specific circumstances and become the founders of extant viral lineages again. A hypothetical example could be provided by human herpesvirus

6 (HHV-6) which is carried in the germline of approximately 1% of all humans and can be inherited to the offspring while exhibiting the very low mutation rate which applies to the human genome (Pellett et al., 2012). Occasionally, an endogenous HHV-6 genome (chromosomally integrated HHV-6, ciHHV-6) can be activated and become infectious again. Cases were described where the endogenous ciHHV-6 genome of the mother was not inherited to the child, but a connatal infection occurred with the activated and mobilized ciHHV-6 genome of the mother (Hall et al., 2010; Gravel et al., 2013). If this was a more general mechanism, viruses might be latently carried within animal genomes for long periods of time and be awakened to their lytic cycle once in a while until being endogenized again. A similar proposal without demanding germline integration has been made by Lauber et al. for hepatitis B virus which is quickly integrated into the somatic genomes only hours after infecting hepatocytes (Chauhan et al., 2017; Lauber et al., 2017). Lauber et al. suggested that replicating hepatitis B viruses might be continuously replenished from the originally transmitted integrated genome copies, and the entire quasispecies, although mutationally oscillating, might overall be kept at the slow mutational rate of the host genome (Lauber et al., 2017). It is worthy to note, however, that the double-stranded linear DNA (dslDNA) genomes formed during reverse transcription as a minor form of HBV DNA, are replication incompetent, and possibly do not serve as transcriptional templates for pregenomic HBV RNA molecules during virus replication (Staprans et al., 1991; Lamontagne et al., 2016). Such dslDNA genomes may represent precursors of integrated HBV genomes. Their relationship to defective HBV genomes incorporated into virions formed with the help of replicating wild type hepatitis B viruses remains to be established. One may speculate that defective HBV particles are generated when transcripts of integrated, or episomal HBV genomes undergo splicing, followed with packaging and and reverse-transcription (Terre et al., 1991). It was observed that under experimental conditions, naturally occurring HBV deletion mutants may trans-complement each other (Okamoto et al., 1993). These phenomena may facilitate the maintenance of defective HBV genomes and virions in persistently infected individuals, as suggested by Lauber et al. (Lauber et al., 2017).

8. Intrapatient evolution of HBx generates C-terminal truncated variants of full-length, wild type HBx protein during hepatocarcinogenesis

Initial studies revealed that the full-length, “wild type” HBx protein (wtHBx), consisting of 154 amino acid residues, may play a complex role in the development of HBV-associated hepatocellular carcinoma in humans (reviewed by (Bouchard and Schneider, 2004)). In addition to wtHBx, truncated HBx variants which differed from wtHBx in their biological effects and properties, were also regularly detected in liver carcinomas (Sirma et al., 1999; Tu et al., 2001; Wang et al., 2004; Ma et al., 2008; Jiang et al., 2012b; Sung et al., 2012; Sze et al., 2013; Wang et al., 2014a) (Fig. 7).They were possibly encoded by mutated *HBx* genes generated during the integration of HBV DNA into the genome of liver cells. Integration of the HBV DNA into the host genome involves most frequently the coding regions of the viral core and HBx proteins and the introns of cellular genes, although integration within cellular repetitive elements may also occur (reviewed by (Lee et al., 2018)). It was observed that integration of the HBV genome into the host cell genome occurred more frequently in hepatocellular carcinomas (HCCs) than in adjacent normal liver tissues or pericarcinoma liver tissues (Ma et al., 2008; Sung et al., 2012; Wang et al., 2014a). Integrant-derived HBV RNAs (id-RNAs), generated by transcription of the integrated HBV genomes, could be detected both in nonmalignant liver tissues and hepatocellular carcinoma samples (Freitas et al., 2018). Integration of the HBV genome was regularly associated with deletions or mutations of the *HBx* gene, resulting in the expression of C-terminal truncated HBx (trHBx, also called ct-HBx) proteins in the majority of the HBV-positive liver tumors examined (Ma et al., 2008;

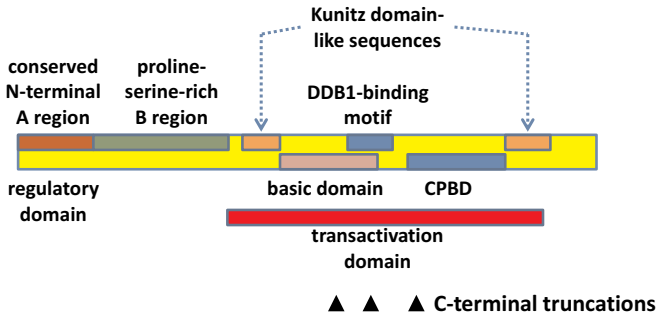

Figure 7. Structure of wtHBx (154 aa) and truncations affecting the C-terminal region (Based on (Tu et al., 2001; Slagle et al., 2015; Niller et al., 2016)). A conserved N-terminal negative regulatory domain is separated by a proline-serine rich region from the C-terminal transcription factor binding, transactivator domain which also binds to DDB1 (Damaged DNA Binding Protein 1). CPBD: cellular protein binding domain. CPBD largely overlaps with a BH3-like domain (Jiang et al., 2016). A split Kunitz-like serine protease inhibitor domain interacts with proteasome subunits. Black triangles indicate truncations affecting the C-terminal domain. Because the C-terminal transactivator domain of wtHBx is involved in apoptosis induction, truncation of the C-terminal domain may facilitate hepatocarcinogenesis by trHBx variants.

Yin et al., 2013). Both wtHBx and its C-terminal truncated variants (trHBx) may contribute to HCC development. trHBx proteins differ, however, from wtHBx in their biological activities, because the C-terminal truncations and mutations frequently affect the transactivation domain identified in wtHBx (reviewed by (Kew, 2011; Niller et al., 2016)). It was observed that wtHBx activated the expression of pro-apoptotic genes and blocked cell proliferation, whereas trHBx variants defective in transactivation failed to induce apoptosis and failed to block cell proliferation (Sirma et al., 1999; Tu et al., 2001; Al-Anazi et al., 2018). C-terminal truncation may influence the binding of HBx to cellular proteins resulting in altered expression patterns of protein coding genes as well as microRNA genes in cells expressing wtHBx, trHBx or both (Neuveut et al., 2010; Yip et al., 2011; Ng et al., 2016; Guerrieri et al., 2017; Niu et al., 2017; Wang et al., 2017). The exact timing of wtHBx and trHBx expression in HBV-infected individuals remains to be elucidated. They may act either in different cell types, i.e. in phenotyically normal liver cells and hepatocellular carcinoma cells of the same patient, suggesting that they are involved in consecutive stages of hepatocarcinogenesis, or they are co-expressed in the same cell, where they may act in concert to reprogram the gene expression pattern of the HBV-infected hepatocyte (reviewed by (Niller et al., 2016)).

A phenomenon similar to the truncation of HBx, the putative “transforming protein” of HBV, was regularly observed during the generation of retroviral oncoproteins, due to the deletion of 5′ sequences, 3′ sequences, or both, of the transduced cellular proto-onc genes (reviewed by (Pedersen and Sorensen, 2010); Table 3). Truncations and mutations in viral oncoproteins frequently remove or alter the negative regulatory domains of the corresponding c-onc proteins, resulting in constitutive and increased enzyme activity, loss of enzyme activity, altered interactions with partner proteins and loss of tumor suppressor or trans-repressor activity (Table 3).

9. C-terminal truncations affecting the cellular adapter protein Crk and the viral regulatory protein HBx

Out of the numerous retroviral v-onc proteins we would like to briefly overview the properties of v-Crk, a transforming protein expressed in chicken cells infected by the avian sarcoma viruses CT10 and ASV-1 (Mayer et al., 1988; Tsuchie et al., 1989). v-Crk is a C-terminally truncated version of c-Crk, a cellular “adapter” protein which – similarly to wtHBx – does not have any enzymatic activity. c-Crk promotes

Table 3
Primary structure of retroviral oncoproteins compared to the protein products encoded by the corresponding cellular *proto-onc* genes – Selected examples

Protein	Consequence of truncation/Note	Reference/Review
Truncated oncoproteins		
N-terminal truncation		
P120ΔGag-Abl	Constitutive activation of tyrosine kinase	(Srinivasan et al., 1982; Zou and Calame, 1999)
P100ΔGag-Fos	Unknown	(Nishizawa et al., 1987)
P105ΔGag-Fps	Unknown	(Huang et al., 1984)
P65ΔGag-Jun	Loss of phosphorylation by JNK1 and 2[a]; misregulation of target genes?, interference with c-Jun?	(Nishimura and Vogt, 1988; Vogt, 2001)
C-terminal truncation		
P100ΔGag-Cbl	Loss of E3 ubiquitin ligase activity; Loss of tumor suppressor activity	(Blake et al., 1991; Blake et al., 1993; Swaminathan and Tsygankov, 2006)
P47ΔGag-Crk	Loss of the C-terminal SH3 domain may alter the interactome of the protein; loss of the autoinhibitory linker region may increase the interactions of the SH2 domain with tyrosine-phosphorylated proteins	(Reichman et al., 1992; Bell and Park, 2012)
P120ΔGag- Fms	Constitutive activation of tyrosine kinase; loss of negative regulation by c-Cbl	(Besmer et al., 1986a; Mancini et al., 2002)
P55-v-fos	Loss of *trans*-repression activity	(van Straaten et al., 1983; Curran et al., 1984; Kamata and Holt, 1992)
P90 ΔGag-Yes- δEnv	Unknown	(Sudol et al., 1988; Zheng et al., 1989)
Both N-terminal and C-terminal truncation in the viral oncoprotein		
gp65-v-erbB gp68-v-erbB	Constitutive activation of tyrosine kinase	(Bassiri and Privalsky, 1986; Ng and Privalsky, 1986)
P70ΔGag-Actin-Fgr	Constitutive activation of tyrosine kinase	(Naharro et al., 1984; Cooper et al., 1986; Katamine et al., 1988)
P75ΔGag-Fos-Fox	Loss of *trans*-repression activity	(Kamata and Holt, 1992)
P80ΔGag-Kit	Constitutive activation of tyrosine kinase	(Besmer et al., 1986a; Yarden et al., 1987)
Unaltered or mutated viral oncoprotein fused to retroviral structural protein		
P105ΔGag-X-Akt	Myristylated fusion protein[b]; the cellularly derived part of v-Akt is identical to c-Akt	(Ahmed et al., 1993; Bellacosa et al., 1993)
P130 ΔGag-Fps	Amino acid substitutions in v-Fps; no truncation in v-Fps compared to c-Fps	(Shibuya and Hanafusa, 1982; Huang et al., 1985)
P110 ΔGag-Myc	No major structural changes in c-Myc	(Bister and Jansen, 1986)
P90 ΔGag-Myc P200 ΔPol-Myc P57 δGag-Myc	Mutations are not necessary for oncogenicity	(Patschinsky et al., 1986)
P145 ΔGag-Myc		(Chen et al., 1989)
P72 ΔGag-Myc		(Chesters et al., 2001)
Mutated viral oncoprotein not fused to retroviral structural protein		
P21-v-has	Activation (constitutive GTP binding) due to two amino acid changes of the GTPase protein	(Ruta et al., 1986)
P21-v- kis	Activation (constitutive GTP binding) due to two amino acid changes of the GTPase protein	(Tsuchida et al., 1982; Iritani et al., 1986; Tsuchida et al., 2016)
pp60-v-src	Loss of Tyr527 involved in negative regulation; increased tyrosine kinase activity; increased transforming ability	(Levinson et al., 1978; Takeya and Hanafusa, 1983; Cooper et al., 1986; Laudano and Buchanan, 1986)

Symbols: Δ indicates a deletion in a retroviral structural protein fused to a c-onc protein; δ indicates that due to a major deletion, only a few amino acids of a retroviral structural protein are fused to a c-onc protein

[a] JNK1 & 2: Jun N-terminal kinase 1 & 2; ¶SH2 & SH3 domain:Src homology domain 2 & 3.

[b] The N-terminus of the ΔGag region is modified.

protein-protein interactions *via* its SH2 and SH3 domains (Src homology domain 2 and 3) (Sriram and Birge, 2010) (Fig. 8). The chicken c-Crk protein corresponding to the CrkII protein of mammals, contains a modular SH2 domain and two SH3 domains (SH3N and SH3C). In contrast, the v-Crk protein expressed in chicken cells transformed by the avian sarcoma viruses CT10 or ASV-1 lacks the C-terminal SH3C domain (Reichman et al., 1992) (reviewed by (Birge et al., 2009)). The N-terminal region of the retroviral v-Crk fusion protein is formed by a truncated viral Gag sequence whereas the C-terminal part of v-Crk consists of a truncated c-Crk sequence which lost its carboxy teminal amino acid residues. The C-terminal truncation may contribute to the enhanced levels of proteins phosphorylated on tyrosine residues in *v-crk* transformed cells, because the loss of the negative regulatory elements located to the SH3C domain and the neighbouring inter-SH3 region apparently facilitates both the binding of the SH2 domain to tyrosine-phosphorylated, activated receptor protein kinases and enhances the phosphorylation of their substrates associated with the SH3N domain of the Crk "connector" protein (Sriram and Birge, 2010; Bell and Park, 2012). Constitutive activation of the growth-promoting PI3K-Akt signal transduction pathway by v-Crk plays an important role in the transformation of CT10 virus infected chicken embryo fibroblasts (Akagi et al., 2002).

One of the protein-protein interactons facilitated by v-Crk is the binding of the p85 regulatory subunit of phosphoinositide 3-kinase (PI3K) to focal adhesion kinase (FAK) which results in FAK phosphorylation by the Src family tyrosine kinase(s) associated with the SH2 domain of v-Crk. Phosphatidylinositol (3,4,5)-trisphosphate (PIP_3) produced by the activated PI3K attracts Akt to the cell membrane, initiating further signaling events by the activated Akt serine threonine kinase that modulates a series of proteins regulating cell growth and motiliy. Increased protein phosphorylation contributes to malignant transformation of the CT10 and ASV1 infected cells, and both viral and cellular Crk proteins contribute to increased cell motility, invasion and metastasis formation during tumor progression (reviewed by (Rodrigues et al., 2005; Zheng et al., 2010; Tsuda and Tanaka, 2012)).

We would like to emphasize that various cellular Crk proteins play a prominent role in the development of oral, head and neck, mammary and non-small cell lung carcinomas in humans (Yamada et al., 2011; Fathers et al., 2012; Yanagi et al., 2012; Mortazavi et al., 2015).

Similarly to the deletion in the cellular *Crk* sequences transduced by retroviruses, deletions of the *HBx* gene in HBV infected liver cells typically also affect the carboxy-terminal region of the encoded protein. Such a C-terminal truncation significantly affects the properties of the pleiotropic HBx protein and alters the behaviour of HBV- infected

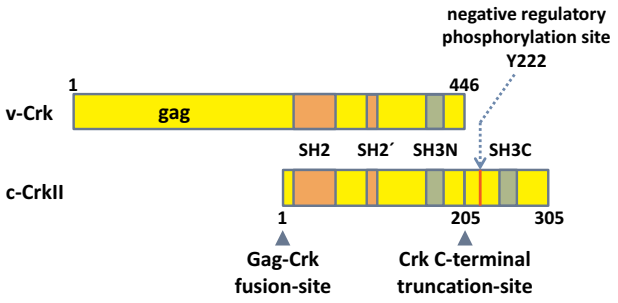

Figure 8. Structure of v-Crk, a retroviral adaptor protein and its cellular counterpart (Based on (Reichman et al., 1992; Birge et al., 2009; Sriram and Birge, 2010)).
v-Crk (top), the transforming protein of avian sarcoma viruses CT10 and ASV-1 is a fusion protein composed of a truncated retroviral Gag domain and a C-terminally truncated cellular Crk domain. The corresponding cellular adaptor protein, c-Crk, promotes protein-protein interactions *via* its SH2 and SH3 domains (Src homology domain 2 and 3). Due to the C-terminal truncation, SH3C, a SH3 domain present in the cellular protein, is lacking in v-Crk. A neighbouring tyrosine residue involved in negative regulation of the interactions of c-Crk with its protein partners is also missing in v-Crk. These structural changes may contribute to the enhanced binding of v-Crk, *via* its SH2 domain to tyrosine-phosphorylated, activated receptor protein kinases and enhances the phosphorylation of their substrates associated with the SH3N domain of v-Crk. Constitutive signaling by the activated cellular tyrosine kinases may result in malignant transformation of CT10 and ASV-1 infected cells.

hepatocytes, too (Fig. 7). Interacting with a series of cellular proteins, wtHBx activates various signaling pathways resulting in the stimulation of HBV replication (reviewed by (Niller et al., 2016)). In parallel, as indicated by several experiments, wtHBx may inhibit cell growth and induce apoptosis via its C-terminal transactivation domain. Integration of the viral dslDNA genomes frequently associates, however, with deletions or mutations of the HBx coding sequence resulting in the appearance of C-terminal truncated HBx proteins which lack a functional transactivator domain. Accordingly, trHBx stimulates cell growth, motility, invasion and metastasis formation, similarly to v-Crk. C-terminal truncation possibly alters not only the structure but also the interactome of both Crk and HBx, resulting in regulatory changes in the target cells and malignant transformation.

HBx is regarded as a hybrid type intrinsically unstructured protein, composed of a disordered N-terminal region and a folded C-terminal domain (Lee et al., 2012). The partly unstructured character of HBx may facilitate its interactions with a series of cellular proteins. Such promiscuous interactions may permit activation of the PI3K/Akt pathway by HBx, similarly to v-Crk which has well characterised SH2 and SH3 domains (Akagi et al., 2002; Riggins et al., 2003; Wang et al., 2014b; Rawat and Bouchard, 2015). Stimulation of Akt by HBx may enhance HBV replication in infected hepatocytes and may inhibit, in parallel, apoptosis induction by the wild type HBx protein which carries a C-terminal “death motif“ (also called BH3 motif) (Rawat and Bouchard, 2015; Jiang et al., 2016). How C-terminal truncations of HBx affect its interaction with the PI3K/Akt pathway is unknown at present. One may speculate, however, that the loss of the pro-apoptotic C-terminal domain is sufficient by itself to shift the balance from programmed cell death to cell proliferation in HBV infected cells (Ma et al., 2008). Thus, transforming proteins with properties similar to those of the oncoproteins of acute RNA tumor viruses are regularly generated during intrapatient evolution of HBV.

Recently, the degradation of the cellular restriction factor of extrachromosomal DNA transcription Smc5/6 (structural maintenance of chromosomes) as a specific function of HBx was revealed. By interacting with DDB1-containing E3 ubiquitin ligase, HBx targets Smc5/6 for degradation. The degradation of Smc5/6 enables HBx to stimulate transcription from intracellular HBV genomes. It is unknown at present whether Smc5/6 serves as a host cell restriction factor also for dsDNA viruses other than HBV (Decorsiere et al., 2016).

10. Outlook

We have presented a comprehensive overview of the origins of viral oncoproteins of two classes of reverse-transcribing tumor viruses, i.e. the acute transforming retroviruses and human hepatitis B virus, and we have highlighted the mechanisms of truncated viral oncoproteins in the malignant transformation of eukaryotic cells. The research on viral oncoproteins has contributed a great deal to our understanding of their counterparts, the cellular protooncogenes. As far as protooncogenes or their fusions with other genes are involved in human tumorigenesis, oncoproteins serve as target structures for anti-tumoral therapeutic approaches a long time. Since the early 2000s, an ever expanding list of small molecule kinase inhibitors and antibodies has been approved for the treatment of diverse cancers and hematological diseases (Roskoski Jr., 2019). HBx which is not derived from a cellular gene may also be accessible for antiviral treatment with small molecule inhibitors. An example is given by Ghosh et al. (2017), where a chemical library screen yielded the pharmacologically active thiourea compound IR415 which inhibited the HBx activity which is directed against the anti-viral Dicer endoribonuclease, and thereby strongly inhibited HBV replication in cell culture (Ghosh et al., 2017). Because truncation of full-length, wtHBx apparently facilitates hepatocarcinogenesis, we speculate that the development of small molecule inhibitors targeting trHBx molecules may also be a promising therapeutic approach. The list of small-molecule oncoprotein inhibitors and anti-oncoprotein antibodies is certainly bound to be enlarged and improved in the near future.

Acknowledgement

The support of the grant GINOP-2.3.2-15-2016-00011 by the European Regional Development Fund to a project led by JM is kindly acknowledged. The grant was managed by European Union Economic Development and Innovation Operational Programme of the Ministry of National Economy, National Research, Development and Innovation Office, Budapest, Hungary.

References

Ahmed, N.N., Franke, T.F., Bellacosa, A., Datta, K., Gonzalez-Portal, M.E., Taguchi, T., Testa, J.R., Tsichlis, P.N., 1993. The proteins encoded by c-akt and v-akt differ in post-translational modification, subcellular localization and oncogenic potential. Oncogene 8, 1957–1963.
Aiewsakun, P., Katzourakis, A., 2015. Endogenous viruses: connecting recent and ancient viral evolution. Virology 479-480, 26–37.
Aiewsakun, P., Katzourakis, A., 2016. Time-dependent rate phenomenon in viruses. J Virol 90, 7184–7195.
Aiewsakun, P., Katzourakis, A., 2017. Marine origin of retroviruses in the early Palaeozoic Era. Nat Commun 8, 13954.
Ajiro, M., Zheng, Z.M., 2014. Oncogenes and RNA splicing of human tumor viruses. Emerg Microbes Infect 3, e63.
Akagi, T., Murata, K., Shishido, T., Hanafusa, H., 2002. v-Crk activates the phosphoinositide 3-kinase/AKT pathway by utilizing focal adhesion kinase and H-Ras. Mol Cell Biol 22, 7015–7023.
Al-Anazi, M.R., Nazir, N., Colak, D., Al-Ahdal, M.N., Al-Qahtani, A.A., 2018. Deletion and functional analysis of hepatitis B virus X protein: evidence for an effect on cell cycle regulators. Cell Physiol Biochem 49, 1987–1998.
Andersen, P.R., Devare, S.G., Tronick, S.R., Ellis, R.W., Aaronson, S.A., Scolnick, E.M., 1981. Generation of BALB-MuSV and Ha-MuSC by type C virus transduction of homologous transforming genes from different species. Cell 26, 129–134.
Anderson, S.J., Furth, M., Wolff, L., Ruscetti, S.K., Sherr, C.J., 1982. Monoclonal antibodies to the transformation-specific glycoprotein encoded by the feline retroviral oncogene v-fms. J Virol 44, 696–702.
Bartosch, B., Thimme, R., Blum, H.E., Zoulim, F., 2009. Hepatitis C virus-induced hepatocarcinogenesis. J Hepatol 51, 810–820.
Bassiri, M., Privalsky, M.L., 1986. Mutagenesis of the avian erythroblastosis virus erbB coding region: an intact extracellular domain is not required for oncogenic transformation. J Virol 59, 525–530.
Beck, J., Nassal, M., 2007. Hepatitis B virus replication. World J. Gastroenterol. 13, 48–64.
Bell, E.S., Park, M., 2012. Models of crk adaptor proteins in cancer. Genes Cancer 3,

341–352.
Bellacosa, A., Testa, J.R., Staal, S.P., Tsichlis, P.N., 1991. A retroviral oncogene, akt, encoding a serine-threonine kinase containing an SH2-like region. Science 254, 274–277.
Bellacosa, A., Franke, T.F., Gonzalez-Portal, M.E., Datta, K., Taguchi, T., Gardner, J., Cheng, J.Q., Testa, J.R., Tsichlis, P.N., 1993. Structure, expression and chromosomal mapping of c-akt: relationship to v-akt and its implications. Oncogene 8, 745–754.
Benachenhou, F., Sperber, G.O., Bongcam-Rudloff, E., Andersson, G., Boeke, J.D., Blomberg, J., 2013. Conserved structure and inferred evolutionary history of long terminal repeats (LTRs). Mob DNA 4, 5.
Benegiamo, G., Vinciguerra, M., Mazzoccoli, G., Piepoli, A., Andriulli, A., Pazienza, V., 2012. DNA methyltransferases 1 and 3b expression in Huh-7 cells expressing HCV core protein of different genotypes. Dig Dis Sci 57, 1598–1603.
Bergold, P.J., Blumenthal, J.A., D'andrea, E., Snyder, H.W., Lederman, L., Silverstone, A., Nguyen, H., Besmer, P., 1987a. Nucleic acid sequence and oncogenic properties of the HZ2 feline sarcoma virus v-abl insert. J Virol 61, 1193–1202.
Bergold, P.J., Wang, J.Y., Hardy Jr., W.D., Littau, V., Johnson, E., Besmer, P., 1987b. Structure and origins of the HZ2-feline sarcoma virus. Virology 158, 320–329.
Besmer, P., Snyder Jr., H.W., Murphy, J.E., Hardy Jr., W.D., Parodi, A., 1983. The Parodi-Irgens feline sarcoma virus and simian sarcoma virus have homologous oncogenes, but in different contexts of the viral genomes. J Virol 46, 606–613.
Besmer, P., Lader, E., George, P.C., Bergold, P.J., Qiu, F.H., Zuckerman, E.E., Hardy, W.D., 1986a. A new acute transforming feline retrovirus with fms homology specifies a C-terminally truncated version of the c-fms protein that is different from SM-feline sarcoma virus v-fms protein. J Virol 60, 194–203.
Besmer, P., Murphy, J.E., George, P.C., Qiu, F.H., Bergold, P.J., Lederman, L., Snyder Jr., H.W., Brodeur, D., Zuckerman, E.E., Hardy, W.D., 1986b. A new acute transforming feline retrovirus and relationship of its oncogene v-kit with the protein kinase gene family. Nature 320, 415–421.
Betsholtz, C., Johnsson, A., Heldin, C.H., Westermark, B., 1986. Efficient reversion of simian sarcoma virus-transformation and inhibition of growth factor-induced mitogenesis by suramin. Proc Natl Acad Sci USA 83, 6440–6444.
Birge, R.B., Kalodimos, C., Inagaki, F., Tanaka, S., 2009. Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. Cell Commun Signal 7, 13.
Bister, K., Jansen, H.W., 1986. Oncogenes in retroviruses and cells: biochemistry and molecular genetics. Adv Cancer Res 47, 99–188.
Blake, T.J., Shapiro, M., Morse 3rd, H.C., Langdon, W.Y., 1991. The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was generated by a large truncation encompassing a proline-rich domain and a leucine zipper-like motif. Oncogene 6, 653–657.
Blake, T.J., Heath, K.G., Langdon, W.Y., 1993. The truncation that generated the v-cbl oncogene reveals an ability for nuclear transport, DNA binding and acute transformation. EMBO J 12, 2017–2026.
Bos, T.J., Bohmann, D., Tsuchie, H., Tjian, R., Vogt, P.K., 1988. v-jun encodes a nuclear protein with enhancer binding properties of AP-1. Cell 52, 705–712.
Bouchard, M.J., Schneider, R.J., 2004. The enigmatic X gene of hepatitis B virus. J Virol 78, 12725–12734.
Breitman, M.L., Hirano, A., Wong, T., Vogt, P.K., 1981. Characteristics of avian sarcoma virus strain PRCIV and comparison with strain PRCII-p. Virology 114, 451–462.
Buendia, M.A., Neuveut, C., 2015. Hepatocellular carcinoma. Cold Spring Harb Perspect Med 5, a021444.
Carter, J.J., Daugherty, M.D., Qi, X., Bheda-Malge, A., Wipf, G.C., Robinson, K., Roman, A., Malik, H.S., Galloway, D.A., 2013. Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. Proc Natl Acad Sci U S A 110, 12744–12749.
Champoux, J.J., Schultz, S.J., 2009. Ribonuclease H: properties, substrate specificity and roles in retroviral reverse transcription. FEBS J 276, 1506–1516.
Chang, S.F., Netter, H.J., Hildt, E., Schuster, R., Schaefer, S., Hsu, Y.C., Rang, A., Will, H., 2001. Duck hepatitis B virus expresses a regulatory HBx-like protein from a hidden open reading frame. J Virol 75, 161–170.
Chang, Y., Moore, P.S., Weiss, R.A., 2017. Human oncogenic viruses: nature and discovery. Philos Trans R Soc Lond B Biol Sci 372.
Chauhan, R., Churchill, N.D., Mulrooney-Cousins, P.M., Michalak, T.I., 2017. Initial sites of hepadnavirus integration into host genome in human hepatocytes and in the woodchuck model of hepatitis B-associated hepatocellular carcinoma. Oncogenesis 6, e317.
Chen, C., Biegalke, B.J., Eisenman, R.N., Linial, M.L., 1989. FH3, a v-myc avian retrovirus with limited transforming ability. J Virol 63, 5092–5100.
Chesters, P.M., Howes, K., Mckay, J.C., Payne, L.N., Venugopal, K., 2001. Acutely transforming avian leukosis virus subgroup J strain 966: defective genome encodes a 72-kilodalton Gag-Myc fusion protein. J Virol 75, 4219–4225.
Cooper, J.A., Gould, K.L., Cartwright, C.A., Hunter, T., 1986. Tyr527 is phosphorylated in pp60c-src: implications for regulation. Science 231, 1431–1434.
Correll, P.H., Paulson, R.F., Wei, X., 2006. Molecular regulation of receptor tyrosine kinases in hematopoietic malignancies. Gene 374, 26–38.
Courtois, G., Benit, L., Mikaeloff, Y., Pauchard, M., Charon, M., Varlet, P., Gisselbrecht, S., 1995. Constitutive activation of a variant of the env-mpl oncogene product by disulfide-linked homodimerization. J Virol 69, 2794–2800.
Curran, T., Miller, A.D., Zokas, L., Verma, I.M., 1984. Viral and cellular fos proteins: a comparative analysis. Cell 36, 259–268.
Davis, P.L., Holmes, E.C., Larrous, F., Van Der Poel, W.H., Tjornehoj, K., Alonso, W.J., Bourhy, H., 2005. Phylogeography, population dynamics, and molecular evolution of European bat lyssaviruses. J Virol 79, 10487–10497.
De Carvalho Dominguez Souza, B.F., Konig, A., Rasche, A., De Oliveira Carneiro, I., Stephan, N., Corman, V.M., Roppert, P.L., Goldmann, N., Kepper, R., Muller, S.F., Volker, C., De Souza, A.J.S., Gomes-Gouvea, M.S., Moreira-Soto, A., Stocker, A., Nassal, M., Franke, C.R., Rebello Pinho, J.R., Soares, M., Geyer, J., Lemey, P., Drosten, C., Netto, E.M., Glebe, D., Drexler, J.F., 2018. A novel hepatitis B virus species discovered in capuchin monkeys sheds new light on the evolution of primate hepadnaviruses. J Hepatol 68, 1114–1122.
Decorsiere, A., Mueller, H., Van Breugel, P.C., Abdul, F., Gerossier, L., Beran, R.K., Livingston, C.M., Niu, C., Fletcher, S.P., Hantz, O., Strubin, M., 2016. Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature 531, 386–389.
Denner, J., 2014. The transmembrane proteins contribute to immunodeficiencies induced by HIV-1 and other retroviruses. AIDS 28, 1081–1090.
Devare, S.G., Reddy, E.P., Law, J.D., Robbins, K.C., Aaronson, S.A., 1983. Nucleotide sequence of the simian sarcoma virus genome: demonstration that its acquired cellular sequences encode the transforming gene product p28sis. Proc Natl Acad Sci U S A 80, 731–735.
Ding, D., Lou, X., Hua, D., Yu, W., Li, L., Wang, J., Gao, F., Zhao, N., Ren, G., Li, L., Lin, B., 2012. Recurrent targeted genes of hepatitis B virus in the liver cancer genomes identified by a next-generation sequencing-based approach. PLoS Genet 8, e1003065.
Donoghue, D.J., 1982. Demonstration of biological activity and nucleotide sequence of an in vitro synthesized clone of the Moloney murine sarcoma virus mos gene. J Virol 42, 538–546.
Donoghue, D.J., Hunter, T., 1983. Recombinational junctions of variants of Moloney murine sarcoma virus: generation and divergence of a mammalian transforming gene. J Virol 45, 607–617.
Ewald, P.W., Swain Ewald, H.A., 2015. Infection and cancer in multicellular organisms. Philos Trans R Soc Lond B Biol Sci 370.
Fan, H., Johnson, C., 2011. Insertional oncogenesis by non-acute retroviruses: implications for gene therapy. Viruses 3, 398–422.
Fathers, K.E., Bell, E.S., Rajadurai, C.V., Cory, S., Zhao, H., Mourskaia, A., Zuo, D., Madore, J., Monast, A., Mes-Masson, A.M., Grosset, A.A., Gaboury, L., Hallet, M., Siegel, P., Park, M., 2012. Crk adaptor proteins act as key signaling integrators for breast tumorigenesis. Breast Cancer Res 14, R74.
Feng, H., Chen, P., Zhao, F., Nassal, M., Hu, K., 2013. Evidence for multiple distinct interactions between hepatitis B virus P protein and its cognate RNA encapsidation signal during initiation of reverse transcription. PLoS One 8, e72798.
Freitas, N., Lukash, T., Gunewardena, S., Chappell, B., Slagle, B.L., Gudima, S.O., 2018. Relative abundance of integrant-derived viral RNAs in infected tissues harvested from chronic hepatitis B virus carriers. J Virol 92.
Furuta, M., Tanaka, H., Shiraishi, Y., Unida, T., Imamura, M., Fujimoto, A., Fujita, M., Sasaki-Oku, A., Maejima, K., Nakano, K., Kawakami, Y., Arihiro, K., Aikata, H., Ueno, M., Hayami, S., Ariizumi, S.I., Yamamoto, M., Gotoh, K., Ohdan, H., Yamaue, H., Miyano, S., Chayama, K., Nakagawa, H., 2018. Characterization of HBV integration patterns and timing in liver cancer and HBV-infected livers. Oncotarget 9, 25075–25088.
Garson, K., Percival, H., Kang, C.Y., 1990. The N-terminal env-derived amino acids of v-rel are required for full transforming activity. Virology 177, 106–115.
Ghosh, S., Kaushik, A., Khurana, S., Varshney, A., Singh, A.K., Dahiya, P., Thakur, J.K., Sarin, S.K., Gupta, D., Malhotra, P., Mukherjee, S.K., Bhatnagar, R.K., 2017. An RNAi-based high-throughput screening assay to identify small molecule inhibitors of hepatitis B virus replication. J Biol Chem 292, 12577–12588.
Gilroy, K.L., Terry, A., Naseer, A., De Ridder, J., Allahyar, A., Wang, W., Carpenter, E., Mason, A., Wong, G.K., Cameron, E.R., Kilbey, A., Neil, J.C., 2016. Gamma-retrovirus integration marks cell type-specific cancer genes: a novel profiling tool in cancer genomics. PLoS One 11, e0154070.
Gomez-Lucia, E., Zhi, Y., Nabavi, M., Zhang, W., Kabat, D., Hoatlin, M.E., 1998. An array of novel murine spleen focus-forming viruses that activate the erythropoietin receptor. J Virol 72, 3742–3750.
Gonda, M.A., Young, H.A., Elser, J.E., Rasheed, S., Talmadge, C.B., Nagashima, K., Li, C.C., Gilden, R.V., 1982. Molecular cloning, genomic analysis, and biological properties of rat leukemia virus and the onc sequences of Rasheed rat sarcoma virus. J Virol 44, 520–529.
Goodrich, D.W., Duesberg, P.H., 1990. Evidence that retroviral transduction is mediated by DNA not by RNA. Proc Natl Acad Sci U S A 87, 3604–3608.
Gravel, A., Hall, C.B., Flamand, L., 2013. Sequence analysis of transplacentally acquired human herpesvirus 6 DNA is consistent with transmission of a chromosomally integrated reactivated virus. J Infect Dis 207, 1585–1589.
Greenwood, A.D., Ishida, Y., O'brien, S.P., Roca, A.L., Eiden, M.V., 2018. Transmission, evolution, and endogenization: lessons learned from recent retroviral invasions. Microbiol Mol Biol Rev 82.
Grundhoff, A., Sullivan, C.S., 2011. Virus-encoded microRNAs. Virology 411, 325–343.
Guerrieri, F., Belloni, L., D'andrea, D., Pediconi, N., Le Pera, L., Testoni, B., Scisciani, C., Floriot, O., Zoulim, F., Tramontano, A., Levrero, M., 2017. Genome-wide identification of direct HBx genomic targets. BMC Genomics 18, 184.
Hai, H., Tamori, A., Kawada, N., 2014. Role of hepatitis B virus DNA integration in human hepatocarcinogenesis. World J Gastroenterol 20, 6236–6243.
Hajjar, A.M., Linial, M.L., 1993. A model system for nonhomologous recombination between retroviral and cellular RNA. J Virol 67, 3845–3853.
Hall, C.B., Caserta, M.T., Schnabel, K.C., Shelley, L.M., Carnahan, J.A., Marino, A.S., Yoo, C., Lofthus, G.K., 2010. Transplacental congenital human herpesvirus 6 infection caused by maternal chromosomally integrated virus. J Infect Dis 201, 505–507.
Hannink, M., Donoghue, D.J., 1984. Requirement for a signal sequence in biological expression of the v-sis oncogene. Science 226, 1197–1199.
Hayward, A., 2017. Origin of the retroviruses: when, where, and how? Curr Opin Virol 25, 23–27.
Hayward, A., Cornwallis, C.K., Jern, P., 2015. Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. Proc Natl Acad Sci U S A 112, 464–469.
Herzog, N.K., Nash, M., Ramagli, L.S., Arlinghaus, R.B., 1990. v-mos protein produced by

in vitro translation has protein kinase activity. J Virol 64, 3093–3096.
Ho, S.Y., Lanfear, R., Bromham, L., Phillips, M.J., Soubrier, J., Rodrigo, A.G., Cooper, A., 2011. Time-dependent rates of molecular evolution. Mol Ecol 20, 3087–3101.
Holmes, E.C., 2003. Molecular clocks and the puzzle of RNA virus origins. J Virol 77, 3893–3897.
Holmes, E.C., 2011. The evolution of endogenous viral elements. Cell Host Microbe 10, 368–377.
Honda, T., 2017. Potential links between Hepadnavirus and Bornavirus sequences in the host genome and cancer. Front Microbiol 8, 2537.
Hsu, T., Phung, A., Choe, K., Kim, J.W., Fan, H., 2015. Role for a Zinc Finger Protein (Zfp111) in transformation of 208F rat fibroblasts by Jaagsiekte sheep retrovirus envelope protein. J Virol 89, 10453–10466.
Huang, C.C., Hammond, C., Bishop, J.M., 1984. Nucleotide sequence of v-fps in the PRCII strain of avian sarcoma virus. J Virol 50, 125–131.
Huang, C.C., Hammond, C., Bishop, J.M., 1985. Nucleotide sequence and topography of chicken c-fps: Genesis of a retroviral oncogene encoding a tyrosine-specific protein kinase. J Mol Biol 181, 175–186.
Hughes, S.H., 2015. Reverse Transcription of Retroviruses and LTR Retrotransposons. Microbiol Spectr 3, MDNA3–0027-2014.
Iba, H., Takeya, T., Cross, F.R., Hanafusa, T., Hanafusa, H., 1984. Rous sarcoma virus variants that carry the cellular src gene instead of the viral src gene cannot transform chicken embryo fibroblasts. Proc Natl Acad Sci U S A 81, 4424–4428.
Iritani, A., Katayama, N., Tahira, T., Hayashi, K., Tsuchida, N., 1986. Nucleotide sequence of exon I of the rat c-K-ras gene. Bull Tokyo Med Dent Univ 33, 35–40.
Jamain, S., Girondot, M., Leroy, P., Clergue, M., Quach, H., Fellous, M., Bourgeron, T., 2001. Transduction of the human gene FAM8A1 by endogenous retrovirus during primate evolution. Genomics 78, 38–45.
Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54, 156–165.
Jiang, S., Yang, Z., Li, W., Li, X., Wang, Y., Zhang, J., Xu, C., Chen, P.J., Hou, J., Mccrae, M.A., Chen, X., Zhuang, H., Lu, F., 2012a. Re-evaluation of the carcinogenic significance of hepatitis B virus integration in hepatocarcinogenesis. PLoS One 7, e40363.
Jiang, Z., Jhunjhunwala, S., Liu, J., Haverty, P.M., Kennemer, M.I., Guan, Y., Lee, W., Carnevali, P., Stinson, J., Johnson, S., Diao, J., Yeung, S., Jubb, A., Ye, W., Wu, T.D., Kapadia, S.B., De Sauvage, F.J., Gentleman, R.C., Stern, H.M., Seshagiri, S., Pant, K.P., Modrusan, Z., Ballinger, D.G., Zhang, Z., 2012b. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. Genome Res 22, 593–601.
Jiang, T., Liu, M., Wu, J., Shi, Y., 2016. Structural and biochemical analysis of Bcl-2 interaction with the hepatitis B virus protein HBx. Proc Natl Acad Sci U S A 113, 2074–2079.
Justice, J.T., Beemon, K.L., 2013. Avian retroviral replication. Curr Opin Virol 3, 664–669.
Kamata, N., Holt, J.T., 1992. Inhibitory effect of myristylation on transrepression by FBR (Gag-Fos) protein. Mol Cell Biol 12, 876–882.
Kan, N.C., Flordellis, C.S., Mark, G.E., Duesberg, P.H., Papas, T.S., 1984. A common onc gene sequence transduced by avian carcinoma virus MH2 and by murine sarcoma virus 3611. Science 223, 813–816.
Katamine, S., Notario, V., Rao, C.D., Miki, T., Cheah, M.S., Tronick, S.R., Robbins, K.C., 1988. Primary structure of the human fgr proto-oncogene product p55c-fgr. Mol Cell Biol 8, 259–266.
Keese, P.K., Gibbs, A., 1992. Origins of genes: “big bang” or continuous creation? Proc Natl Acad Sci U S A 89, 9489–9493.
Kew, M.C., 2011. Hepatitis B virus x protein in the pathogenesis of hepatitis B virus-induced hepatocellular carcinoma. J Gastroenterol Hepatol 26, 144–152 Suppl 1.
Kidd, A.H., Kidd-Ljunggren, K., 1996. A revised secondary structure model for the 3'-end of hepatitis B virus pregenomic RNA. Nucleic Acids Res 24, 3295–3301.
Klempnauer, K.H., Bishop, J.M., 1983. Transduction of c-myb into avian myeloblastosis virus: locating points of recombination within the cellular gene. J Virol 48, 565–572.
Knight, J., Zenke, M., Disela, C., Kowenz, E., Vogt, P., Engel, J.D., Hayman, M.J., Beug, H., 1988. Temperature-sensitive v-sea transformed erythroblasts: a model system to study gene expression during erythroid differentiation. Genes Dev 2, 247–258.
Koike, K., 2005. Molecular basis of hepatitis C virus-associated hepatocarcinogenesis: lessons from animal model studies. Clin Gastroenterol Hepatol 3, S132–S135.
Koonin, E.V., Dolja, V.V., Krupovic, M., 2015. Origins and evolution of viruses of eukaryotes: The ultimate modularity. Virology 479-480, 2–25.
Krupovic, M., Blomberg, J., Coffin, J.M., Dasgupta, I., Fan, H., Geering, A.D., Gifford, R., Harrach, B., Hull, R., Johnson, W., Kreuze, J.F., Lindemann, D., Llorens, C., Lockhart, B., Mayer, J., Muller, E., Olszewski, N.E., Pappu, H.R., Pooggin, M.M., Richert-Poggeler, K.R., Sabanadzovic, S., Sanfacon, H., Schoelz, J.E., Seal, S., Stavolone, L., Stoye, J.P., Teycheney, P.Y., Tristem, M., Koonin, E.V., Kuhn, J.H., 2018. Ortervirales: new virus order unifying five families of reverse-transcribing viruses. J Virol 92.
Kurten, B., 1959. Rates of evolution in fossil mammals. Cold Spring Harb Symp Quant Biol 24, 205–215.
Lamontagne, R.J., Bagga, S., Bouchard, M.J., 2016. Hepatitis B virus molecular biology and pathogenesis. Hepatoma Res 2, 163–186.
Lanford, R.E., Notvall, L., Lee, H., Beames, B., 1997. Transcomplementation of nucleotide priming and reverse transcription between independently expressed TP and RT domains of the hepatitis B virus reverse transcriptase. J Virol 71, 2996–3004.
Lauber, C., Seitz, S., Mattei, S., Suh, A., Beck, J., Herstein, J., Borold, J., Salzburger, W., Kaderali, L., Briggs, J.a.G., Bartenschlager, R., 2017. Deciphering the origin and evolution of hepatitis B viruses by means of a family of non-enveloped fish viruses. Cell Host Microbe 22, 387–399 e386.
Laudano, A.P., Buchanan, J.M., 1986. Phosphorylation of tyrosine in the carboxyl-terminal tryptic peptide of pp60c-src. Proc Natl Acad Sci U S A 83, 892–896.
Lee, S.H., Cha, E.J., Lim, J.E., Kwon, S.H., Kim, D.H., Cho, H., Han, K.H., 2012. Structural characterization of an intrinsically unfolded mini-HBX protein from hepatitis B virus. Mol Cells 34, 165–169.
Lee, W.Y., Bachtiar, M., Choo, C.C.S., Lee, C.G., 2018. Comprehensive review of Hepatitis B Virus-associated hepatocellular carcinoma research through text mining and big data analytics. Biol Rev Camb Philos Soc. 94, 353–357.
Levinson, A.D., Oppermann, H., Levintow, L., Varmus, H.E., Bishop, J.M., 1978. Evidence that the transforming gene of avian sarcoma virus encodes a protein kinase associated with a phosphoprotein. Cell 15, 561–572.
Lipsick, J.S., Wang, D.M., 1999. Transformation by v-Myb. Oncogene 18, 3047–3055.
Liu, S.L., Miller, A.D., 2007. Oncogenic transformation by the jaagsiekte sheep retrovirus envelope protein. Oncogene 26, 789–801.
Lokeshwar, V.B., Huang, S.S., Huang, J.S., 1990. Intracellular turnover, novel secretion, and mitogenically active intracellular forms of v-sis gene product in simian sarcoma virus-transformed cells. Implications for intracellular loop autocrine transformation. J Biol Chem 265, 1665–1675.
Lyons, S., Sharp, C., Lebreton, M., Djoko, C.F., Kiyang, J.A., Lankester, F., Bibila, T.G., Tamoufe, U., Fair, J., Wolfe, N.D., Simmonds, P., 2012. Species association of hepatitis B virus (HBV) in non-human apes; evidence for recombination between gorilla and chimpanzee variants. PLoS One 7, e33430.
Ma, N.F., Lau, S.H., Hu, L., Xie, D., Wu, J., Yang, J., Wang, Y., Wu, M.C., Fung, J., Bai, X., Tzang, C.H., Fu, L., Yang, M., Su, Y.A., Guan, X.Y., 2008. COOH-terminal truncated HBV X protein plays key role in hepatocarcinogenesis. Clin Cancer Res 14, 5061–5068.
Mancini, A., Koch, A., Wilms, R., Tamura, T., 2002. c-Cbl associates directly with the C-terminal tail of the receptor for the macrophage colony-stimulating factor, c-Fms, and down-modulates this receptor but not the viral oncogene v-Fms. J Biol Chem 277, 14635–14640.
Mann, R., Mulligan, R.C., Baltimore, D., 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33, 153–159.
Mason, W.S., 2015. Animal models and the molecular biology of hepadnavirus infection. Cold Spring Harb Perspect Med 5.
Mason, W.S., Liu, C., Aldrich, C.E., Litwin, S., Yeh, M.M., 2010. Clonal expansion of normal-appearing human hepatocytes during chronic hepatitis B virus infection. J Virol 84, 8308–8315.
Matsuoka, M., Jeang, K.T., 2011. Human T-cell leukemia virus type 1 (HTLV-1) and leukemic transformation: viral infectivity, Tax, HBZ and therapy. Oncogene 30, 1379–1389.
Mayer, B.J., Hamaguchi, M., Hanafusa, H., 1988. Characterization of p47gag-crk, a novel oncogene product with sequence similarity to a putative modulatory domain of protein-tyrosine kinases and phospholipase C. Cold Spring Harb Symp Quant Biol 53 (Pt 2), 907–914.
Mcgeoch, D.J., Gatherer, D., 2005. Integrating reptilian herpesviruses into the family herpesviridae. J Virol 79, 725–731.
Mcgeoch, D.J., Cook, S., Dolan, A., Jamieson, F.E., Telford, E.A., 1995. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. J Mol Biol 247, 443–458.
Meier, P., Scougall, C.A., Will, H., Burrell, C.J., Jilbert, A.R., 2003. A duck hepatitis B virus strain with a knockout mutation in the putative X ORF shows similar infectivity and in vivo growth characteristics to wild-type virus. Virology 317, 291–298.
Mesnard, J.M., Carriere, C., 1995. Comparison of packaging strategy in retroviruses and pararetroviruses. Virology 213, 1–6.
Miller, A.D., Buttimore, C., 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. Mol Cell Biol 6, 2895–2902.
Minarovits, J., Demcsak, A., Banati, F., Niller, H.H., 2016. Epigenetic dysregulation in virus-associated neoplasms. Adv Exp Med Biol 879, 71–90.
Moreau-Gachelin, F., 2008. Multi-stage Friend murine erythroleukemia: molecular insights into oncogenic cooperation. Retrovirology 5, 99.
Mortazavi, F., Lu, J., Phan, R., Lewis, M., Trinidad, K., Aljilani, A., Pezeshkpour, G., Tamanoi, F., 2015. Significance of KRAS/PAK1/Crk pathway in non-small cell lung cancer oncogenesis. BMC Cancer 15, 381.
Naharro, G., Robbins, K.C., Reddy, E.P., 1984. Gene product of v-fgr onc: hybrid protein containing a portion of actin and a tyrosine-specific protein kinase. Science 223, 63–66.
Neckameyer, W.S., Wang, L.H., 1984. Molecular cloning and characterization of avian sarcoma virus UR2 and comparison of its transforming sequence with those of other avian sarcoma viruses. J Virol 50, 914–921.
Neel, B.G., Wang, L.H., Mathey-Prevot, B., Hanafusa, T., Hanafusa, H., Hayward, W.S., 1982. Isolation of 16L virus: a rapidly transforming sarcoma virus from an avian leukosis virus-induced sarcoma. Proc Natl Acad Sci U S A 79, 5088–5092.
Neuveut, C., Wei, Y., Buendia, M.A., 2010. Mechanisms of HBV-related hepatocarcinogenesis. J Hepatol 52, 594–604.
Ng, M., Privalsky, M.L., 1986. Structural domains of the avian erythroblastosis virus erbB protein required for fibroblast transformation: dissection by in-frame insertional mutagenesis. J Virol 58, 542–553.
Ng, K.Y., Chai, S., Tong, M., Guan, X.Y., Lin, C.H., Ching, Y.P., Xie, D., Cheng, A.S., Ma, S., 2016. C-terminal truncated hepatitis B virus X protein promotes hepatocellular carcinogenesis through induction of cancer and stem cell-like properties. Oncotarget 7, 24005–24017.
Niller, H.H., Szenthe, K., Minarovits, J., 2014. Epstein-Barr virus-host cell interactions: an epigenetic dialog? Front Genet 5, 367.
Niller, H.H., Ay, E., Banati, F., Demcsak, A., Takacs, M., Minarovits, J., 2016. Wild type HBx and truncated HBx: Pleiotropic regulators driving sequential genetic and epigenetic steps of hepatocarcinogenesis and progression of HBV-associated neoplasms. Rev Med Virol 26, 57–73.

Nishida, T., Sakamoto, S., Yamamoto, T., Hayman, M., Kawai, S., Toyoshima, K., 1984. Comparison of genome structures among three different strains of avian erythroblastosis virus. Gan 75, 325–333.
Nishimura, T., Vogt, P.K., 1988. The avian cellular homolog of the oncogene jun. Oncogene 3, 659–663.
Nishizawa, M., Goto, N., Kawai, S., 1987. An avian transforming retrovirus isolated from a nephroblastoma that carries the fos gene as the oncogene. J Virol 61, 3733–3740.
Niu, Y., Xu, M., Slagle, B.L., Huang, H., Li, S., Guo, G.L., Shi, G., Qin, W., Xie, W., 2017. Farnesoid X receptor ablation sensitizes mice to hepatitis b virus X protein-induced hepatocarcinogenesis. Hepatology 65, 893–906.
Okamoto, H., Wang, Y., Tanaka, T., Machida, A., Miyakawa, Y., Mayumi, M., 1993. Trans-complementation among naturally occurring deletion mutants of hepatitis B virus and integrated viral DNA for the production of viral particles with mutant genomes in hepatoma cell lines. J Gen Virol 74, 407–414 Pt 3.
Overbaugh, J., Miller, A.D., Eiden, M.V., 2001. Receptors and entry cofactors for retroviruses include single and multiple transmembrane-spanning proteins as well as newly described glycophosphatidylinositol-anchored and secreted proteins. Microbiol Mol Biol Rev 65, 371–389 table of contents.
Patschinsky, T., Jansen, H.W., Blocker, H., Frank, R., Bister, K., 1986. Structure and transforming function of transduced mutant alleles of the chicken c-myc gene. J Virol 59, 341–353.
Pavesi, A., 2015. Different patterns of codon usage in the overlapping polymerase and surface genes of hepatitis B virus suggest a de novo origin by modular evolution. J Gen Virol 96, 3577–3586.
Pavesi, A., Magiorkinis, G., Karlin, D.G., 2013. Viral proteins originated de novo by overprinting can be identified by codon usage: application to the "gene nursery" of Deltaretroviruses. PLoS Comput Biol 9, e1003162.
Pedersen, F.S., Sorensen, A.B., 2010. Pathogenesis of oncoviral infections. In: Kurth, R., Bannert, N. (Eds.), Retroviruses: Molecular Biology, Genomics and Pathogenesis. Caister Academic Press, Norfolk, pp. 237–267.
Pellett, P.E., Ablashi, D.V., Ambros, P.F., Agut, H., Caserta, M.T., Descamps, V., Flamand, L., Gautheret-Dejean, A., Hall, C.B., Kamble, R.T., Kuehl, U., Lassner, D., Lautenschlager, I., Loomis, K.S., Luppi, M., Lusso, P., Medveczky, P.G., Montoya, J.G., Mori, Y., Ogata, M., Pritchett, J.C., Rogez, S., Seto, E., Ward, K.N., Yoshikawa, T., Razonable, R.R., 2012. Chromosomally integrated human herpesvirus 6: questions and answers. Rev Med Virol 22, 144–155.
Pinto-Santini, D.M., Stenbak, C.R., Linial, M.L., 2017. Foamy virus zoonotic infections. Retrovirology 14, 55.
Plummer, M., De Martel, C., Vignat, J., Ferlay, J., Bray, F., Franceschi, S., 2016. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health 4, e609–e616.
Prywes, R., Foulkes, J.G., Rosenberg, N., Baltimore, D., 1983. Sequences of the A-MuLV protein needed for fibroblast and lymphoid cell transformation. Cell 34, 569–579.
Prywes, R., Hoag, J., Rosenberg, N., Baltimore, D., 1985. Protein stabilization explains the gag requirement for transformation of lymphoid cells by Abelson murine leukemia virus. J Virol 54, 123–132.
Pulsinelli, G.A., Temin, H.M., 1991. Characterization of large deletions occurring during a single round of retrovirus vector replication: novel deletion mechanism involving errors in strand transfer. J Virol 65, 4786–4797.
Rawat, S., Bouchard, M.J., 2015. The hepatitis B virus (HBV) HBx protein activates AKT to simultaneously regulate HBV replication and hepatocyte survival. J Virol 89, 999–1012.
Reddy, E.P., Smith, M.J., Srinivasan, A., 1983. Nucleotide sequence of Abelson murine leukemia virus genome: structural similarity of its transforming gene product to other onc gene products with tyrosine-specific kinase activity. Proc Natl Acad Sci U S A 80, 3623–3627.
Reddy, E.P., Lipman, D., Andersen, P.R., Tronick, S.R., Aaronson, S.A., 1985. Nucleotide sequence analysis of the BALB/c murine sarcoma virus transforming gene. J Virol 53, 984–987.
Reichman, C.T., Mayer, B.J., Keshav, S., Hanafusa, H., 1992. The product of the cellular crk gene consists primarily of SH2 and SH3 regions. Cell Growth Differ 3, 451–460.
Riggins, R.B., Deberry, R.M., Toosarvandani, M.D., Bouton, A.H., 2003. Src-dependent association of Cas and p85 phosphatidylinositol 3'-kinase in v-crk-transformed cells. Mol Cancer Res 1, 428–437.
Robinson, K.M., Dunning Hotopp, J.C., 2014. Mobile elements and viral integrations prompt considerations for bacterial DNA integration as a novel carcinogen. Cancer Lett 352, 137–144.
Rodrigues, S.P., Fathers, K.E., Chan, G., Zuo, D., Halwani, F., Meterissian, S., Park, M., 2005. CrkI and CrkII function as key signaling integrators for migration and invasion of cancer cells. Mol Cancer Res 3, 183–194.
Roskoski Jr., R., 2019. Properties of FDA-approved small molecule protein kinase inhibitors. Pharmacol Res 144, 19–50.
Ruta, M., Wolford, R., Dhar, R., Defeo-Jones, D., Ellis, R.W., Scolnick, E.M., 1986. Nucleotide sequence of the two rat cellular rasH genes. Mol Cell Biol 6, 1706–1710.
Sa-Nguanmoo, P., Rianthavorn, P., Amornsawadwattana, S., Poovorawan, Y., 2009. Hepatitis B virus infection in non-human primates. Acta Virol 53, 73–82.
Schuster, R., Hildt, E., Chang, S.F., Terradillos, O., Pollicino, T., Lanford, R., Gerlich, W.H., Will, H., Schaefer, S., 2002. Conserved transactivating and pro-apoptotic functions of hepadnaviral X protein in ortho- and avihepadnaviruses. Oncogene 21, 6606–6613.
Schwartz, J.R., Duesberg, S., Duesberg, P.H., 1995. DNA recombination is sufficient for retroviral transduction. Proc Natl Acad Sci U S A 92, 2460–2464.
Seeger, C., Mason, W.S., 2015. Molecular biology of hepatitis B virus infection. Virology 479-480, 672–686.
Serrao, E., Engelman, A.N., 2016. Sites of retroviral DNA integration: from basic research to clinical applications. Crit Rev Biochem Mol Biol 51, 26–42.
Shackelton, L.A., Parrish, C.R., Truyen, U., Holmes, E.C., 2005. High rate of viral evolution associated with the emergence of carnivore parvovirus. Proc Natl Acad Sci U S A 102, 379–384.
Sharp, P.M., Li, W.H., 1988. Understanding the origins of AIDS viruses. Nature 336, 315.
Shaw, J., Hayman, M.J., Enrietto, P.J., 1985. Analysis of a deleted MC29 provirus: gag sequences are not required for fibroblast transformation. J Virol 56, 943–950.
Shibuya, M., Hanafusa, H., 1982. Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses. Cell 30, 787–795.
Shore, S.K., Tantravahi, R.V., Reddy, E.P., 2002. Transforming pathways activated by the v-Abl tyrosine kinase. Oncogene 21, 8568–8576.
Sirma, H., Giannini, C., Poussin, K., Paterlini, P., Kremsdorf, D., Brechot, C., 1999. Hepatitis B virus X mutants, present in hepatocellular carcinoma tissue abrogate both the antiproliferative and transactivation effects of HBx. Oncogene 18, 4848–4859.
Slagle, B.L., Andrisani, O.M., Bouchard, M.J., Lee, C.G., Ou, J.H., Siddiqui, A., 2015. Technical standards for hepatitis B virus X protein (HBx) research. Hepatology 61, 1416–1424.
Smith, D.R., Vogt, P.K., Hayman, M.J., 1989. The v-sea oncogene of avian erythroblastosis retrovirus S13: another member of the protein-tyrosine kinase gene family. Proc Natl Acad Sci U S A 86, 5291–5295.
Snyder Jr., H.W., Singhal, M.C., Zuckerman, E.E., Hardy Jr., W.D., 1984. Isolation of a new feline sarcoma virus (HZ1-FeSV): biochemical and immunological characterization of its translation product. Virology 132, 205–210.
Srinivasan, A., Dunn, C.Y., Yuasa, Y., Devare, S.G., Reddy, E.P., Aaronson, S.A., 1982. Abelson murine leukemia virus: structural requirements for transforming gene function. Proc Natl Acad Sci U S A 79, 5508–5512.
Sriram, G., Birge, R.B., 2010. Emerging roles for crk in human cancer. Genes Cancer 1, 1132–1139.
Staprans, S., Loeb, D.D., Ganem, D., 1991. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. J Virol 65, 1255–1262.
Starkman, S.E., Macdonald, D.M., Lewis, J.C., Holmes, E.C., Simmonds, P., 2003. Geographic and species association of hepatitis B virus genotypes in non-human primates. Virology 314, 381–393.
Stavnezer, E., Brodeur, D., Brennan, L.A., 1989. The v-ski oncogene encodes a truncated set of c-ski coding exons with limited sequence and structural relatedness to v-myc. Mol Cell Biol 9, 4038–4045.
Sudol, M., Kieswetter, C., Zhao, Y.H., Dorai, T., Wang, L.H., Hanafusa, H., 1988. Nucleotide sequence of a cDNA for the chick yes proto-oncogene: comparison with the viral yes gene. Nucleic Acids Res 16, 9876.
Sugden, B., 1993. How some retroviruses got their oncogenes. Trends Biochem Sci 18, 233–235.
Suh, A., Brosius, J., Schmitz, J., Kriegs, J.O., 2013. The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses. Nat Commun 4, 1791.
Suh, A., Weber, C.C., Kehlmaier, C., Braun, E.L., Green, R.E., Fritz, U., Ray, D.A., Ellegren, H., 2014. Early mesozoic coexistence of amniotes and hepadnaviridae. PLoS Genet 10, e1004559.
Sun, Y., Wang, S., Yi, Y., Zhang, J., Duan, Z., Yuan, K., Liu, W., Li, J., Zhu, Y., 2018. The hepatitis B surface antigen binding protein: an immunoglobulin G constant region-like protein that interacts with HBV envelop proteins and mediates HBV entry. Front Cell Infect Microbiol 8, 338.
Sung, W.K., Zheng, H., Li, S., Chen, R., Liu, X., Li, Y., Lee, N.P., Lee, W.H., Ariyaratne, P.N., Tennakoon, C., Mulawadi, F.H., Wong, K.F., Liu, A.M., Poon, R.T., Fan, S.T., Chan, K.L., Gong, Z., Hu, Y., Lin, Z., Wang, G., Zhang, Q., Barber, T.D., Chou, W.C., Aggarwal, A., Hao, K., Zhou, W., Zhang, C., Hardwick, J., Buser, C., Xu, J., Kan, Z., Dai, H., Mao, M., Reinhard, C., Wang, J., Luk, J.M., 2012. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet 44, 765–769.
Swain, A., Coffin, J.M., 1992. Mechanism of transduction by retroviruses. Science 255, 841–845.
Swaminathan, G., Tsygankov, A.Y., 2006. The Cbl family proteins: ring leaders in regulation of cell signaling. J Cell Physiol 209, 21–43.
Swanstrom, R., Parker, R.C., Varmus, H.E., Bishop, J.M., 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc Natl Acad Sci U S A 80, 2519–2523.
Sze, K.M., Chu, G.K., Lee, J.M., Ng, I.O., 2013. C-terminal truncated hepatitis B virus x protein is associated with metastasis and enhances invasiveness by C-Jun/matrix metalloproteinase protein 10 activation in hepatocellular carcinoma. Hepatology 57, 131–139.
Takeya, T., Hanafusa, H., 1983. Structure and sequence of the cellular gene homologous to the RSV src gene and the mechanism for generating the transforming virus. Cell 32, 881–890.
Tang, H., Mclachlan, A., 2002. Avian and Mammalian hepadnaviruses have distinct transcription factor requirements for viral replication. J Virol 76, 7468–7472.
Telesnitsky, A., Wolin, S.L., 2016. The host RNAs in retroviral particles. Viruses 8.
Temin, H.M., 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc Natl Acad Sci U S A 90, 6900–6903.
Tennant, B.C., Gerin, J.L., 2001. The woodchuck model of hepatitis B virus infection. ILAR J 42, 89–102.
Terre, S., Petit, M.A., Brechot, C., 1991. Defective hepatitis B virus particles are generated by packaging and reverse transcription of spliced viral RNAs in vivo. J Virol 65, 5539–5543.
Torres, C., Fernandez, M.D., Flichman, D.M., Campos, R.H., Mbayed, V.A., 2013. Influence of overlapping genes on the evolution of human hepatitis B virus. Virology 441, 40–48.
Tsuchida, N., Ryder, T., Ohtsubo, E., 1982. Nucleotide sequence of the oncogene encoding

the p21 transforming protein of Kirsten murine sarcoma virus. Science 217, 937–939.
Tsuchida, N., Murugan, A.K., Grieco, M., 2016. Kirsten Ras* oncogene: significance of its discovery in human cancer research. Oncotarget 7, 46717–46733.
Tsuchie, H., Chang, C.H., Yoshida, M., Vogt, P.K., 1989. A newly isolated avian sarcoma virus, ASV-1, carries the crk oncogene. Oncogene 4, 1281–1284.
Tsuda, M., Tanaka, S., 2012. Roles for crk in cancer metastasis and invasion. Genes Cancer 3, 334–340.
Tsuruyama, T., Hiratsuka, T., Yamada, N., 2017. Hotspots of MLV integration in the hematopoietic tumor genome. Oncogene 36, 1169–1175.
Tu, H., Bonura, C., Giannini, C., Mouly, H., Soussan, P., Kew, M., Paterlini-Brechot, P., Brechot, C., Kremsdorf, D., 2001. Biological impact of natural COOH-terminal deletions of hepatitis B virus X protein in hepatocellular carcinoma tissues. Cancer Res 61, 7803–7810.
Tu, T., Budzinska, M.A., Shackel, N.A., Urban, S., 2017. HBV DNA integration: molecular mechanisms and clinical implications. Viruses 9.
Van Beveren, C., Enami, S., Curran, T., Verma, I.M., 1984. FBR murine osteosarcoma virus. II. Nucleotide sequence of the provirus reveals that the genome contains sequences acquired from two cellular genes. Virology 135, 229–243.
Van Straaten, F., Muller, R., Curran, T., Van Beveren, C., Verma, I.M., 1983. Complete nucleotide sequence of a human c-onc gene: deduced amino acid sequence of the human c-fos protein. Proc Natl Acad Sci U S A 80, 3183–3187.
Veronese, F., Kelloff, G.J., Reynolds Jr., F.H., Hill, R.W., Stephenson, J.R., 1983. Monoclonal antibodies to feline sarcoma virus gag and fes gene translational products. J Cell Biochem 21, 9–18.
Vogt, P.K., 2001. Jun, the oncoprotein. Oncogene 20, 2365–2377.
Vogt, P.K., 2012. Retroviral oncogenes: a historical primer. Nat Rev Cancer 12, 639–648.
Vogt, P.K., Neil, J.C., Moscovici, C., Breitman, M.L., 1981. PRCII, a representative of a new class of avian sarcoma viruses. Haematol Blood Transfus 26, 424–428.
Wang, L.H., Feldman, R., Shibuya, M., Hanafusa, H., Notter, M.F., Balduzzi, P.C., 1981. Genetic structure, transforming sequence, and gene product of avian sarcoma virus UR1. J Virol 40, 258–267.
Wang, Y., Lau, S.H., Sham, J.S., Wu, M.C., Wang, T., Guan, X.Y., 2004. Characterization of HBV integrants in 14 hepatocellular carcinomas: association of truncated X gene and hepatocellular carcinogenesis. Oncogene 23, 142–148.
Wang, D., Cai, H., Yu, W.B., Yu, L., 2014a. Identification of hepatitis B virus X gene variants between hepatocellular carcinoma tissues and pericarcinoma liver tissues in Eastern China. Int J Clin Exp Pathol 7, 5988–5996.
Wang, H.Y., Yang, S.L., Liang, H.F., Li, C.H., 2014b. HBx protein promotes oval cell proliferation by up-regulation of cyclin D1 via activation of the MEK/ERK and PI3K/Akt pathways. Int J Mol Sci 15, 3507–3518.
Wang, J.C., Nickens, D.G., Lentz, T.B., Loeb, D.D., Zlotnick, A., 2014c. Encapsidated hepatitis B virus reverse transcriptase is poised on an ordered RNA lattice. Proc Natl Acad Sci U S A 111, 11329–11334.
Wang, Y., Fang, L., Li, J., Li, Y., Cui, S., Sun, X., Chang, S., Zhao, P., Cui, Z., 2016a. Rescue of avian leukosis subgroup-J-associated acutely transforming viruses carrying different lengths of the v-fps oncogene and analysis of their tumorigenicity. Arch Virol 161, 3473–3481.
Wang, Y., Li, J., Li, Y., Fang, L., Sun, X., Chang, S., Zhao, P., Cui, Z., 2016b. Identification of ALV-J associated acutely transforming virus Fu-J carrying complete v-fps oncogene. Virus Genes 52, 365–371.
Wang, Z., Wu, Z., Huang, P., 2017. The function of miRNAs in hepatocarcinogenesis induced by hepatitis B virus X protein (Review). Oncol Rep 38, 652–664.
Watanabe, T., 2017. Adult T-cell leukemia: molecular basis for clonal expansion and transformation of HTLV-1-infected T cells. Blood 129, 1071–1081.
Wertheim, J.O., Kosakovsky Pond, S.L., 2011. Purifying selection can obscure the ancient age of viral lineages. Mol Biol Evol 28, 3355–3365.
White, M.K., Pagano, J.S., Khalili, K., 2014. Viruses and human cancers: a long road of discovery of molecular paradigms. Clin Microbiol Rev 27, 463–481.
Wilhelmsen, K.C., Eggleton, K., Temin, H.M., 1984. Nucleic acid sequences of the oncogene v-rel in reticuloendotheliosis virus strain T and its cellular homolog, the proto-oncogene c-rel. J Virol 52, 172–182.
Wootton, S.K., Halbert, C.L., Miller, A.D., 2005. Sheep retrovirus structural protein induces lung tumours. Nature 434, 904–907.
Yamada, S., Yanamoto, S., Kawasaki, G., Rokutanda, S., Yonezawa, H., Kawakita, A., Nemoto, T.K., 2011. Overexpression of CRKII increases migration and invasive potential in oral squamous cell carcinoma. Cancer Lett 303, 84–91.
Yamamoto, T., Hihara, H., Nishida, T., Kawai, S., Toyoshima, K., 1983. A new avian erythroblastosis virus, AEV-H, carries erbB gene responsible for the induction of both erythroblastosis and sarcomas. Cell 34, 225–232.
Yan, H., Zhong, G., Xu, G., He, W., Jing, Z., Gao, Z., Huang, Y., Qi, Y., Peng, B., Wang, H., Fu, L., Song, M., Chen, P., Gao, W., Ren, B., Sun, Y., Cai, T., Feng, X., Sui, J., Li, W., 2012a. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. Elife 1, e00049.
Yan, H., Zhong, G., Xu, G., He, W., Jing, Z., Gao, Z., Huang, Y., Qi, Y., Peng, B., Wang, H., Fu, L., Song, M., Chen, P., Gao, W., Ren, B., Sun, Y., Cai, T., Feng, X., Sui, J., Li, W., 2012b. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. Elife 3.
Yanagi, H., Wang, L., Nishihara, H., Kimura, T., Tanino, M., Yanagi, T., Fukuda, S., Tanaka, S., 2012. CRKL plays a pivotal role in tumorigenesis of head and neck squamous cell carcinoma through the regulation of cell adhesion. Biochem Biophys Res Commun 418, 104–109.
Yang, L., Ye, S., Zhao, X., Ji, L., Zhang, Y., Zhou, P., Sun, J., Guan, Y., Han, Y., Ni, C., Hu, X., Liu, W., Wang, H., Zhou, B., Huang, J., 2018. Molecular characterization of HBV DNA integration in patients with hepatitis and hepatocellular carcinoma. J Cancer 9, 3225–3235.
Yarden, Y., Kuang, W.J., Yang-Feng, T., Coussens, L., Munemitsu, S., Dull, T.J., Chen, E., Schlessinger, J., Francke, U., Ullrich, A., 1987. Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. EMBO J 6, 3341–3351.
Yin, J., Li, N., Han, Y., Xue, J., Deng, Y., Shi, J., Guo, W., Zhang, H., Wang, H., Cheng, S., Cao, G., 2013. Effect of antiviral treatment with nucleotide/nucleoside analogs on postoperative prognosis of hepatitis B virus-related hepatocellular carcinoma: a two-stage longitudinal clinical study. J Clin Oncol 31, 3647–3655.
Yip, W.K., Cheng, A.S., Zhu, R., Lung, R.W., Tsang, D.P., Lau, S.S., Chen, Y., Sung, J.G., Lai, P.B., Ng, E.K., Yu, J., Wong, N., To, K.F, Wong, V.W., Sung, J.J., Chan, H.L., 2011. Carboxyl-terminal truncated HBx regulates a distinct microRNA transcription program in hepatocellular carcinoma development. PLoS One 6, e22888.
Zhang, J., Temin, H.M., 1993. 3' junctions of oncogene-virus sequences and the mechanisms for formation of highly oncogenic retroviruses. J Virol 67, 1747–1751.
Zhang, J., Temin, H.M., 1994. Retrovirus recombination depends on the length of sequence identity and is not error prone. J Virol 68, 2409–2414.
Zhang, Y.Z., Wu, W.C., Shi, M., Holmes, E.C., 2018. The diversity, evolution and origins of vertebrate RNA viruses. Curr Opin Virol 31, 9–16.
Zhao, J.W., Fang, F., Guo, Y., Zhu, T.L., Yu, Y.Y., Kong, F.F., Han, L.F., Chen, D.S., Li, F., 2016. HPV16 integration probably contributes to cervical oncogenesis through interrupting tumor suppressor genes and inducing chromosome instability. J Exp Clin Cancer Res 35, 180.
Zheng, X.F., Podell, S., Sefton, B.M., Kaplan, P.L., 1989. The sequence of chicken c-yes and p61c-yes. Oncogene 4, 99–104.
Zheng, J., Machida, K., Antoku, S., Ng, K.Y., Claffey, K.P., Mayer, B.J., 2010. Proteins that bind the Src homology 3 domain of CrkI have distinct roles in Crk transformation. Oncogene 29, 6378–6389.
Zou, X., Calame, K., 1999. Signaling pathways activated by oncogenic forms of Abl tyrosine kinase. J Biol Chem 274, 18141–18144.