Rev. Sci. Tech. Off. Int. Epiz., 2016, **35** (1), 133-142

Viruses within animal genomes

A. de Brogniez & L. Willems

Molecular Biology (GxABT) and Molecular and Cellular Epigenetics (GIGA), University of Liege, Belgium
E-mail: Alix.deBrogniez@ulg.ac.be; Luc.willems@ulg.ac.be

Summary
Viruses and their hosts can co-evolve to reach a fragile equilibrium that allows the survival of both. An excess of pathogenicity in the absence of a reservoir would be detrimental to virus survival. A significant proportion of all animal genomes has been shaped by the insertion of viruses that subsequently became 'fossilised'. Most endogenous viruses have lost the capacity to replicate via an infectious cycle and now replicate passively. The insertion of endogenous viruses has contributed to the evolution of animal genomes, for example in the reproductive biology of mammals. However, spontaneous viral integration still occasionally occurs in a number of virus–host systems. This constitutes a potential risk to host survival but also provides an opportunity for diversification and evolution.

Keywords
Animal – Genome – Genotoxicity – Integration – Persistence – Virus.

Introduction

Viruses are biological entities that use the host cell machinery to replicate. Viruses can either integrate their genetic material into the host cell genome or replicate autonomously within cellular compartments such as the cytoplasm. For one family of RNA viruses, the *Retroviridae* (or retroviruses), chromosomal integration is an essential step of the viral replication cycle (1). Endogenous viral elements (EVEs) are evolutionary remnants of viruses that have stably integrated into the host genome and are now transmitted through the germ line (2). DNA viruses of the *Parvoviridae* (3), *Herpesviridae* (4), *Hepadnaviridae* (5) and *Adenoviridae* families (6) also occasionally undergo chromosomal integration. This review will describe the different groups of integrated viruses and the possible implications of virus integration for human and animal health.

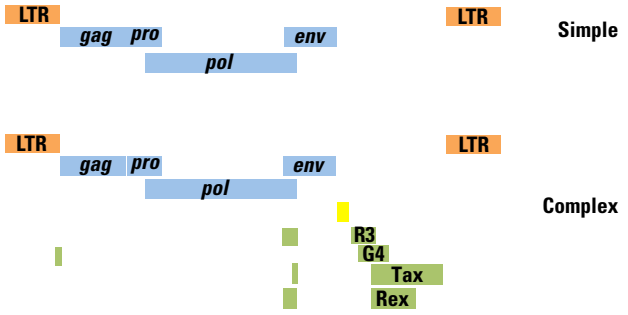
Retroviruses

Retroviruses encode a reverse transcriptase (RT) that polymerises DNA using the viral genomic RNA as a template and an integrase (IN) that inserts the proviral DNA into the host cell genome (1). Infection by retroviruses can be asymptomatic (e.g. human foamy virus) or can cause diseases such as cancer and immunodeficiency (7). Retroviruses can infect a broad variety of vertebrates from fish to humans, leading to disease, for example leukosis in cattle (8). In humans, retroviruses such as human immunodeficiency virus (HIV) induce immunosuppression (e.g. acquired immunodeficiency syndrome). Another retrovirus, human T-lymphotropic virus type 1 (HTLV1), causes adult T-cell leukaemia and neurodegenerative diseases (e.g. HTLV-associated myelopathy/tropical spastic paraparesis) (9).

Retroviruses comprise a large and diverse family of enveloped viruses, characterised by a viral particle of 80–100 nm in diameter and an RNA genome of 7–12 kilobases. The retrovirus genome is linear, unsegmented and single-stranded and has positive polarity. Retroviruses were originally classified as simple or complex based on their genomic organisation. Simple viruses contain four major coding regions: *gag*, *pro*, *pol* and *env* (Fig. 1). The *gag* gene directs the synthesis of the matrix (MA), capsid (CA) and nucleocapsid (NC) proteins that form the virion core. These proteins are cleaved from a precursor by the viral protease, encoded by the *pro* gene. *Pol* contains the genetic information for two replication enzymes: RT and IN. The viral envelope is formed by components of the host cell membrane in association with surface (SU) and transmembrane (TM) proteins encoded by the *env* gene. Complex retroviruses additionally encode a series of regulatory proteins involved in transcriptional activation (Tax), RNA processing (Rex, R3) or cellular transformation (Tax and G4) (Fig. 1).

More recently, retroviruses were subdivided into seven genera based on their evolutionary relatedness (Table I) (1).

doi: http://dx.doi.org/10.20506/rst/35.1.2423



LTR: long terminal repeat

Fig. 1
Retroviral genomic organisation
Viruses with a simple genome contain four major coding regions: *gag, pro, pol* and *env* (shown in blue). Complex viruses such as bovine leukaemia virus encode additional regulatory proteins (e.g. Tax, Rex, R3 and G4; coding regions shown in green) and small RNAs under the control of polymerase III promoter (shown in yellow)

Retroviral replication cycle

A key step in the retroviral replication cycle is the integration of viral DNA into the host cell genome. There is some evidence that genomic integration is an essential stage in the retroviral life cycle. Firstly, non-integrated retroviral DNAs do not replicate autonomously, for example as episomes (extrachromosomal circular DNA). Viral DNA integration is therefore necessary to stably maintain the virus in dividing cells. Secondly, integration is important for the efficient transcription of viral messenger and genomic RNAs.

The retroviral replication cycle can be divided into several steps (shown in Fig. 2). Firstly, specific interactions between envelope glycoproteins and host cell receptors determine viral tropism. Through these interactions, the retroviral envelope attaches and then fuses with the host cell membrane, releasing the viral core into the host cell cytoplasm. After viral entry and viral particle uncoating, the RT enzyme becomes active. The primers for reverse

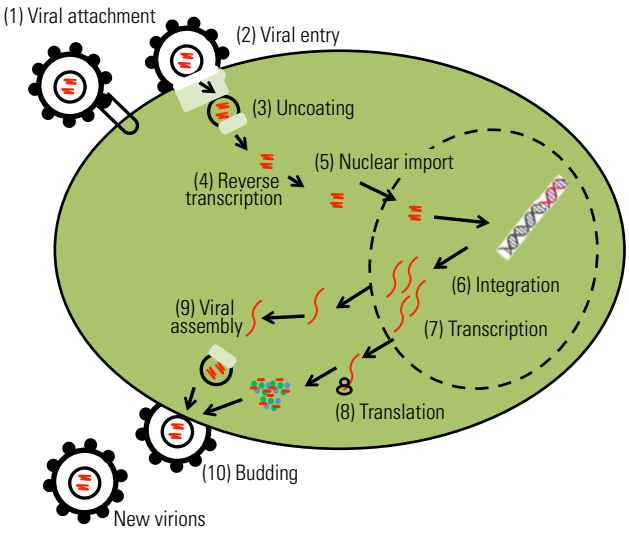


Fig. 2
Schematic representation of the retroviral replication cycle
The retrovirus enters a target cell by binding to specific receptors on the cell membrane. Once internalised, viral RNA is released from the nucleocapsid and is reverse transcribed into proviral DNA. The integrated provirus is then transcribed into genomic and messenger RNA. After translation, viral proteins are packaged along with genomic RNA into new viral particles that 'bud' off from the cell

Table I
Retrovirus classification

Genus	Examples	Viral particle morphology	Genome type
Alpharetrovirus	Rous sarcoma virus	Central, spherical core	Simple
	Avian leukosis virus		
	Mouse mammary tumor virus		
Betaretrovirus	Mason-Pfizer monkey virus	Eccentric, spherical core	Simple
	Jaagsiekte sheep viruses		
Gammaretrovirus	Murine leukaemia virus	Central, spherical core	Simple
	Feline leukaemia virus		
	Gibbon ape leukaemia virus		
Deltaretrovirus	Bovine leukaemia virus	Central, spherical core	Complex
	Human T-lymphotropic virus		
	Simian T-lymphotropic virus		
Epsilonretrovirus	Walleye dermal sarcoma virus	Cylindrical core	Complex
Lentivirus	Human immunodeficiency virus	Cylindrical core	Complex
Spumavirus	Human foamy virus	Central, spherical core	Complex

transcription of viral genomic RNA are cellular transfer RNAs (tRNAs) that hybridise with the primer-binding site of the 5′ long terminal repeat (5′LTR). The RT enzyme also has RNase H activity, which removes the parental RNA strand from the complementary DNA (cDNA). After strand transfer to the 3′LTR, strand synthesis of the proviral genome is completed. Strand synthesis is then initiated at the polypurine tract using an RNA primer that is not digested by RNase H. Double-stranded proviral DNA then forms a complex with viral (IN) and cellular proteins, enters the nucleus and integrates into the host chromosome. Once integrated, the provirus is transcribed into genomic RNA that is incorporated into viral particles and also forms viral messenger RNAs (mRNAs). The splicing of full-length viral RNA also generates mRNAs that encode structural and regulatory proteins. After their export from the nucleus, viral proteins may undergo post-translational modifications, for example SU protein glycosylation in the endoplasmic reticulum and the Golgi apparatus. Envelope glycoproteins form SU–TM oligomers, which bind to the plasma membrane. A capsid surrounding RNA dimers then assembles with the viral envelope glycoproteins to form a virion, which is released by budding from the cell membrane. The released virions can then infect other cells. The provirus is stably integrated into the host cell chromosome, where it remains and is transmitted to daughter cells by mitotic division (10, 11).

Proviral integration thus appears to be an essential stage of the retroviral life cycle and a prerequisite for virion production. Stable integration enables the provirus to regulate host gene expression and thus modulate the cell fate.

Mechanism of proviral integration into the chromosome

After reverse transcription is completed, linear double-stranded viral DNA (flanked by two LTRs) is incorporated into a preintegration complex (PIC) that migrates to the nucleus. Some retroviruses only infect dividing cells because their cDNA products can only interact with chromosomal DNA during mitosis when the nuclear membrane is disrupted (12, 13). In contrast, spumaviruses and lentiviruses can infect both dividing and resting cells because they can enter the nucleus through an active yet poorly understood mechanism (14) involving the PIC and the nuclear pore complex (15, 16).

The retroviral integration mechanism has three main steps (Fig. 3):

i) After the completion of viral DNA synthesis, both proviral ends contain 5–15 base pair extensions that are recognised by the IN protein complex. The IN enzyme cleaves the 3′ termini of both proviral DNA ends, thus eliminating the two terminal bases. The resulting 3′ hydroxyl (OH) groups are involved in proviral attachment to the host DNA. This first step is the 3′ processing reaction that forms a chemically activated viral DNA with 3′ OH radicals at the terminal ends of the viral DNA.

ii) The second step is a transfer reaction involving transesterification. Firstly, the IN protein–viral DNA complex binds to the host cell DNA and introduces an asymmetrical nick of 4–6 nucleotides in length, depending on the retroviral protein structure. In the transfer reaction, the energy released by breaking phosphodiester bonds in target DNA is used to form new bonds joining the viral 3′ ends to the target DNA.

iii) Finally, the host DNA repair machinery cleaves the protruding viral 5′ nucleotides and fills in the 4–6 base pair gap, leading to duplication of the gap nucleotide sequence surrounding the provirus. A final ligation step completes the proviral integration process.

Preferential integration sites

The site of proviral insertion into the host chromosome can determine the outcome of the virus infection cycle. Integration into a transcriptionally active region favours viral gene expression and exposure to the host immune response. In contrast, integration into repressed chromatin facilitates viral latency. However, retroviral integration at a site nearby or inside cellular genes can affect host transcription, leading to tumorigenesis via a mechanism called ‘insertional mutagenesis’. Differences in their integration site preferences cause retroviruses to have different genotoxic potentials (e.g. gammaretroviruses are more prone to insertional mutagenesis than lentiviruses).

Sequencing the sites flanking proviral DNA insertion has shown that retroviral integration is not random (17, 18, 19). For example, HIV integration occurs within active transcription units characterised by high G+C content, high CpG island density, high gene density, short introns and high frequencies of Alu repeats (20, 21, 22, 23). Gammaretroviruses, spumaviruses and endogenous retroviruses are preferentially integrated around transcription start sites and within CpG islands (24, 25, 26). Alpharetroviruses display only a weak preference for insertion into CpG islands and transcription units (20, 27), while betaretrovirus integration is random (28). Deltaretroviruses, which include HTLVs, simian T-cell lymphotropic viruses and bovine leukaemia virus (BLV), initially target actively transcribed regions of the genome. Deltaretroviruses are preferentially inserted into a palindromic consensus sequence with transcriptional units (29, 30). During primary infection, most infected cell clones are depleted by the host immune response (27). At this stage, the proviral load is maintained by a change in

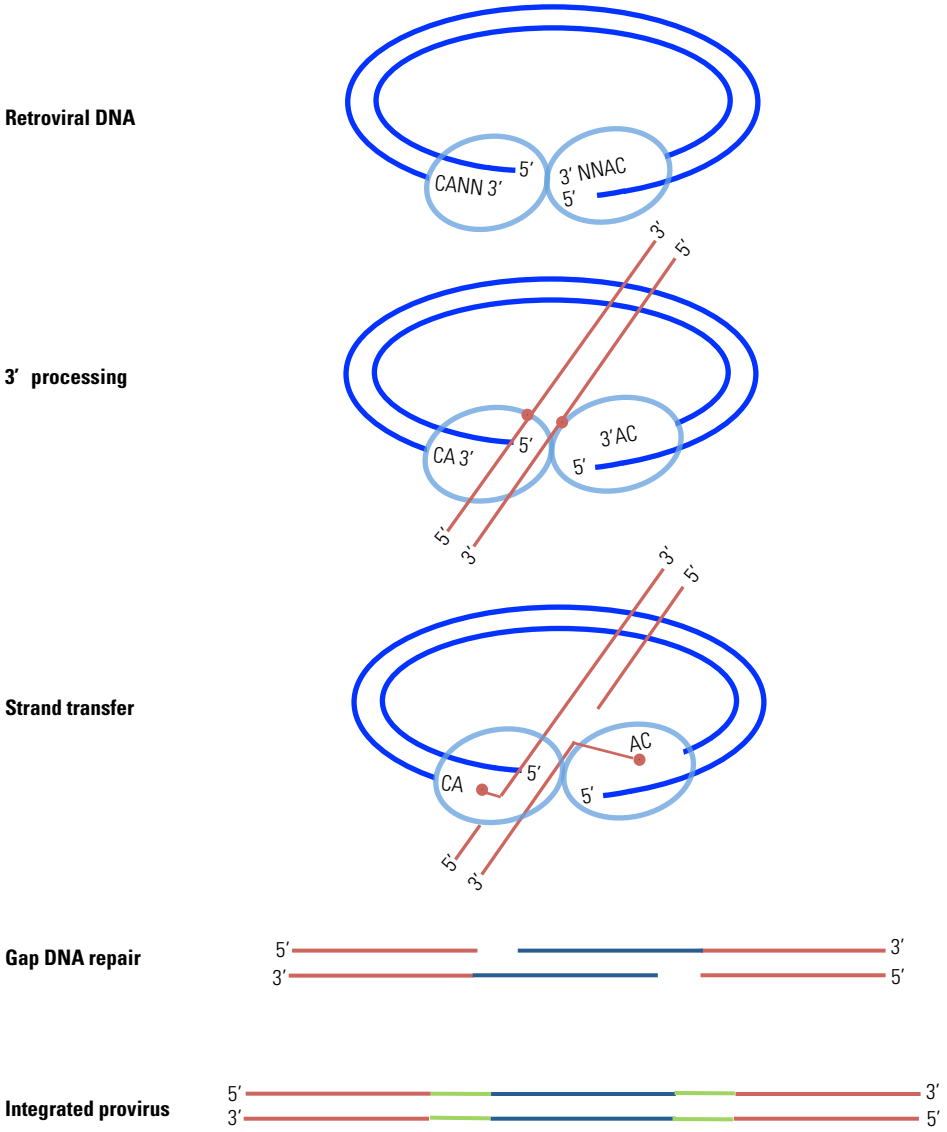


Fig. 3
Mechanism of retroviral integration
A processing reaction results in the removal of two nucleotides at the 3′ ends of the viral long terminal repeat (LTR). The 3′ hydroxyl (OH) groups perform a nucleophilic attack of the target DNA. A strand transfer reaction directed by the integrase forms two new phosphodiester bonds between the 3′ OH ends of the LTRs and the 5′ ends of target DNA. The gaps between unligated strands are repaired, resulting in short duplications of flanking DNA sequences. The integrase enzyme is represented by light blue circles; N represents any base; host DNA is shown in red; retroviral DNA is shown in dark blue; and the duplicated target site sequence is shown in green

the mode of viral propagation that favours the proliferation of pre-existing clones of infected cells. The site of viral integration thus influences the proliferative potential of infected cells. The proviral loads are therefore determined by clone abundance and limited by the host immune response.

Accurate quantification by high-throughput sequencing has demonstrated that integration site preference is neither specific to the host cell nor dependent on the viral entry route. Different models have been proposed to explain the integration site preference of retroviruses (31). It is possible that the chromatin structure (relaxed or condensed) influences the accessibility of target DNA sequences to the PIC. The cell cycle phase may also modulate site-specific integration; moreover, it is possible that for each retroviral genus cellular proteins act as tethering factors that simultaneously interact with specific chromatin sites and PICs.

Consequences of viral integration

Viral infection can have opposing outcomes on the cell fate, i.e. proliferation or death (32). Indeed, strand breaks

generated during viral integration activate the DNA repair machinery, which can be overloaded by an excess of infection events. In contrast, apoptosis may also result from the deregulation of host cell metabolism (33).

If the cell survives infection, the virus can be either actively expressed and modulate the cell transcriptome in *trans* or silenced and thus enter latency. Some retroviruses encode viral oncoproteins that directly inhibit cell death or stimulate cell proliferation (34). Cell fate therefore depends on the viral activation state. For example, HIV infection of resting or activated lymphocytes leads to pyroptosis or caspase-9 dependent apoptosis, respectively. Integration can also lead to insertional mutagenesis that disrupts genes or alters cellular transcription in *cis*. Viral integration at sites near to oncogenes or tumour suppressor genes can also lead to oncogenesis (35).

Another outcome of viral integration is transcriptional silencing and latency to form 'viral reservoirs' that escape host immunity. Several factors influence entry into latency: the availability of cellular transcription factors, the expression of viral proteins that activate transcription and the integration site in the host cell genome. Integration into transcriptionally inactive heterochromatin is reported to favour viral latency (36). Thus, the formation of viral reservoirs can be controlled by epigenetic mechanisms such as lysine acetylation of histones (4, 37).

Endogenous viral elements

Viruses other than retroviruses have also become integrated into host cell genomes. EVEs are retrotransposons that evolved from viruses and are vertically transmitted through germ cells (38). Although they are mostly latent, EVEs can be reactivated by specific environmental conditions (e.g. irradiation), leading to the production of new infectious viral particles. The uncontrolled spread of these viruses can be hazardous to health (39). Nevertheless, EVEs are essential drivers of evolution that act by promoting genomic diversity.

The RepBase database has compiled EVEs into six families containing more than 200 consensus sequences (40, 41). LTR-containing members include simple retrotransposons and endogenous retroviruses. Non-LTR retroelements comprise long and short interspersed nuclear elements (known as LINEs and SINEs, respectively) (42). These retrotransposons and their evolutionary remnants constitute a large proportion (45%) of the human genome (43). Besides retroviruses (44), EVEs are also derived from RNA viruses; for example, the flavivirus-related endogenous element has been integrated into the *Aedes* mosquito genome (45). Other examples include EVEs derived from bornaviruses (46), filoviruses (47), orthomyxoviruses (48), reoviruses and rhabdoviruses (48). Even small DNA viruses such as circoviruses and parvoviruses have entered animal genomes during evolution (3, 49). In paleovirology, EVEs derived from germ-line integration events that occurred millions of years ago can provide historical information about ancient viruses (48). The analysis of EVEs gives an estimate of the time since divergence of orthologue-containing host species groups and thus provides insight into long-term viral evolution.

Like retroviruses, EVEs are generally latent but can be reactivated by epigenetic mechanisms such as cytosine methylation at CpG dinucleotides or acetylation of histone lysines (50). Although reintegration of activated EVEs can pose a threat to the host genome, the resulting mutations can provide a selective advantage during host evolution. For example, the syncytin protein essential for placental morphogenesis in mammals is derived from a retroviral envelope protein (51, 52, 53). In humans, syncytin is encoded by an endogenous retrovirus ERVWE1 on chromosome 7.

The mechanism for stable integration of a non-retroviral RNA element into chromosomal DNA has been demonstrated for lymphocytic choriomeningitis virus (LCMV) (31, 54). LCMV is transcribed in the cytoplasm by the virally encoded RNA-dependent RNA polymerase (RdRp). During reverse transcription, LCMV RNA recombines with intracisternal A-type particle (IAP) elements, and the resulting cDNA undergoes random integration (55).

Integration of DNA viruses

DNA viruses can sometimes be integrated into the host genome without the need for a processing step such as reverse transcription. The genome of DNA viruses can translocate into the nucleus and persist in an episomal form. Integration events into host chromosomes have occasionally been reported for viruses of different families: *Parvoviridae*, *Herpesviridae*, *Hepadnaviridae* and *Adenoviridae*.

Adeno-associated virus type 2 (AAV2) is a widespread non-pathogenic virus of the *Parvoviridae* family that requires a helper virus, such as adenovirus, human papillomavirus or herpes simplex virus, to complete its replicative cycle. In the absence of a helper virus, AAV integrates at specific regions of the host genome known as AAV integration site (AAVS) loci, e.g. AAVS1 on chromosome 19, AAVS2 on chromosome 5 and AAVS3 on chromosome 3 (56). Site-specific AAV-2 integration into non-repetitive elements located in gene-dense regions involves two specific host genome sequences: the terminal resolution site (TRS) and the Rep-binding site (RBS) (57, 58). TRS and RBS are

predicted to recombine with homologous sequences of the viral genome. Thus, AAVS1 provides a safe site for AAV integration.

Although herpesviruses persist as double-stranded DNA episomes in the nucleus, they are occasionally integrated into the host cell genome. Stable integration of Epstein-Barr virus (EBV) ensures its long-term persistence (59, 60, 61). EBV integration is non-random: heterochromatin regions are preferentially targeted (59, 62). Herpesvirus 6 (HHV-6) also undergoes chromosomal integration (ciHHV-6), probably by homologous recombination between viral and cellular telomeric sequences (4, 63).

About one million people die each year from hepatitis B virus (HBV)-induced hepatocellular carcinoma, liver cirrhosis and other complications (64). Although usually present in an episomal form, HBV can integrate into the host chromosome during acute infection by a mechanism similar to retroviral insertion. HBV integration preferentially targets genes involved in cell survival, proliferation and oncogenesis in transcriptionally active regions (65, 66, 67).

Although adenoviruses replicate extrachromosomally in the nucleus, integration occasionally occurs by homologous recombination and non-homologous end joining (6, 68).

Conclusion

Animal genomes contain a large proportion of endogenous viruses that are generally unable to undergo an infectious cycle but replicate passively via mitotic division. Of the exogenous viruses, retroviruses are designed to integrate into the host genome and therefore constitute a potential risk of mutation. Other types of viruses with either RNA or DNA genomes can also occasionally be integrated into the host genome.

Acknowledgements

The work was supported by the 'Fonds National de la Recherche Scientifique' (FNRS), the Interuniversity Attraction Poles BELVIR programme (initiated by the Belgian Science Policy Office), the Belgian Foundation against Cancer, the 'Centre anticancéreux près l'Université de Liège' (ULg), the 'AgricultureIsLife' project of Gembloux Agro-Bio Tech, and the ULg 'Fonds Spéciaux pour la Recherche'. AdB is currently a postdoctoral researcher at the Télévie. LW is research director of the FNRS. The authors thank Arsène Burny for his helpful comments.

Les virus présents dans les génomes animaux

A. de Brogniez & L. Willems

Résumé

Les virus et leurs hôtes ont démontré leur capacité à co-évoluer pour atteindre le fragile équilibre qui assure leur survie mutuelle. Une pathogénicité excessive sans réservoir disponible peut compromettre la survie d'un virus. Un nombre proportionnellement significatif de génomes animaux ont vu leur structure modifiée par l'insertion de virus qui se sont par la suite « fossilisés ». La plupart des virus endogènes ayant perdu leur aptitude à se répliquer via le déclenchement d'un cycle infectieux, leur réplication s'effectue désormais de manière passive. L'insertion de virus endogènes dans les génomes animaux a contribué à les faire évoluer, comme l'illustre la biologie de la reproduction des mammifères. Néanmoins, des intégrations spontanées de virus continuent de se produire ponctuellement dans certains systèmes virus-hôtes. Elles représentent un risque potentiel pour la survie de l'hôte mais ouvrent également de nouvelles perspectives de diversification et d'évolution.

Mots-clés

Animal – Génome – Génotoxicité – Intégration – Persistance – Virus.

Virus dentro de genomas animales

A. de Brogniez & L. Willems

Resumen
Los virus y sus anfitriones pueden coevolucionar hasta alcanzar un frágil equilibrio que permite la supervivencia de ambos. A falta de un reservorio, una patogenicidad excesiva resultaría perjudicial para la supervivencia del virus. Hay una proporción importante de todos los genomas animales en cuya configuración ha intervenido la inserción de virus, ulteriormente «fosilizados» en el genoma. La mayoría de los virus endógenos han perdido la capacidad de replicarse por medio de un ciclo infeccioso y se replican ahora de forma pasiva. La inserción de virus endógenos ha contribuido a la evolución de los genomas animales, por ejemplo en la biología reproductiva de los mamíferos. No obstante, en muchos sistemas virus-anfitrión se sigue dando ocasionalmente una integración vírica espontánea, lo que supone a la vez un posible riesgo para la supervivencia del anfitrión y una oportunidad de diversificación y evolución.

Palabras clave
Animal – Genoma – Genotoxicidad – Integración – Persistencia – Virus.

References

1. Coffin J.M., Hughes S.H. & Varmus H.E. (1997). – The interactions of retroviruses and their hosts. *In* Retroviruses (J.M. Coffin, S.H. Hughes & H.E. Varmus, eds). Cold Spring Harbor, New York.

2. Holmes E.C. (2011). – The evolution of endogenous viral elements. *Cell Host Microbe*, **10** (4), 368–377. doi:10.1016/j.chom.2011.09.002.

3. Belyi V.A., Levine A.J. & Skalka A.M. (2010). – Sequences from ancestral single-stranded DNA viruses in vertebrate genomes: the *Parvoviridae* and *Circoviridae* are more than 40 to 50 million years old. *J. Virol.*, **84** (23), 12458–12462. doi:10.1128/JVI.01789-10.

4. Arbuckle J.H., Medveczky M.M., Luka J., Hadley S.H., Luegmayr A., Ablashi D., Lund T.C., Tolar J., De Meirleir K., Montoya J.G., Komaroff A.L., Ambros P.F. & Medveczky P.G. (2010). – The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes *in vivo* and *in vitro*. *Proc. Natl Acad. Sci. USA*, **107** (12), 5563–5568. doi:10.1073/pnas.0913586107.

5. Murakami Y., Minami M., Daimon Y. & Okanoue T. (2004). – Hepatitis B virus DNA in liver, serum, and peripheral blood mononuclear cells after the clearance of serum hepatitis B virus surface antigen. *J. Med. Virol.*, **72** (2), 203–214. doi:10.1002/jmv.10547.

6. Stephen S.L., Sivanandam V.G. & Kochanek S. (2008). – Homologous and heterologous recombination between adenovirus vector DNA and chromosomal DNA. *J. Gene Med.*, **10** (11), 1176–1189. doi:10.1002/jgm.1246.

7. Gallo R.C. (2015). – Developing a successful HIV vaccine. *J. Infect. Dis.*, **212** (Suppl. 1), S40–S41. doi:10.1093/infdis/jiv069.

8. Rodríguez S., Florins A., Gillet N., de Brogniez A., Sánchez-Alcaraz M.T., Boxus M., Boulanger F., Gutiérrez G., Trono K., Alvarez I., Vagnoni L. & Willems L. (2011). – Preventive and therapeutic strategies for bovine leukemia virus: lessons for HTLV. *Viruses*, **3** (7), 1210–1248. doi:10.3390/v3071210.

9. Kchour G., Rezaee R., Farid R., Ghantous A., Rafatpanah H., Tarhini M., Kooshyar M.M., El Hajj H., Berry F., Mortada M., Nasser R., Shirdel A., Dassouki Z., Ezzedine M., Rahimi H., Ghavamzadeh A., de The H., Hermine O., Mahmoudi M. & Bazarbachi A. (2013). – The combination of arsenic, interferon-alpha, and zidovudine restores an 'immunocompetent-like' cytokine expression profile in patients with adult T-cell leukemia lymphoma. *Retrovirology*, **10**, 91. doi:10.1186/1742-4690-10-91.

10. Florins A., Gillet N., Asquith B., Boxus M., Burteau C., Twizere J.C., Urbain P., Vandermeers F., Debacq C., Sanchez-Alcaraz M.T., Schwartz-Cornil I., Kerkhofs P., Jean G., Thewis A., Hay J., Mortreux F., Wattel E., Reichert M., Burny A., Kettmann R., Bangham C. & Willems L. (2007). – Cell dynamics and immune response to BLV infection: a unifying model. *Front. Biosci*, **12**, 1520–1531. doi:10.2741/2165.

11. Friedrich B.M., Dziuba N., Li G., Endsley M.A., Murray J.L. & Ferguson M.R. (2011). – Host factors mediating HIV-1 replication. *Virus Res.*, **161** (2), 101–114. doi:10.1016/j.virusres.2011.08.001.

12. Lewis P.F. & Emerman M. (1994). – Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.,* **68** (1), 510–516.

13. Roe T., Reynolds T.C., Yu G. & Brown P.O. (1993). – Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.*, **12** (5), 2099–2108.

14. Suzuki Y. & Craigie R. (2007). – The road to chromatin - nuclear entry of retroviruses. *Nat. Rev. Microbiol.*, **5** (3), 187–196. doi:10.1038/nrmicro1579.

15. Ao Z., Danappa Jayappa K., Wang B., Zheng Y., Kung S., Rassart E., Depping R., Kohler M., Cohen E.A. & Yao X. (2010). – Importin alpha3 interacts with HIV-1 integrase and contributes to HIV-1 nuclear import and replication. *J. Virol.*, **84** (17), 8650–8663. doi:10.1128/JVI.00508-10.

16. Christ F., Thys W., De Rijck J., Gijsbers R., Albanese A., Arosio D., Emiliani S., Rain J.C., Benarous R., Cereseto A. & Debyser Z. (2008). – Transportin-SR2 imports HIV into the nucleus. *Curr. Biol.*, **18** (16), 1192–1202. doi:10.1016/j.cub.2008.07.079.

17. Ciuffi A. (2008). – Mechanisms governing lentivirus integration site selection. *Curr. Gene Ther.*, **8** (6), 419–429. doi:10.2174/156652308786848021.

18. Ciuffi A., Ronen K., Brady T., Malani N., Wang G., Berry C.C. & Bushman F.D. (2009). – Methods for integration site distribution analyses in animal cell genomes. *Methods*, **47** (4), 261–268. doi:10.1016/j.ymeth.2008.10.028.

19. Bushman F., Lewinski M., Ciuffi A., Barr S., Leipzig J., Hannenhalli S. & Hoffmann C. (2005). – Genome-wide analysis of retroviral DNA integration. *Nat. Rev. Microbiol.*, **3**, 848–858. doi:10.1038/nrmicro1263.

20. Derse D., Crise B., Li Y., Princler G., Lum N., Stewart C., McGrath C.F., Hughes S.H., Munroe D.J. & Wu X. (2007). – Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J. Virol.,* **81** (12), 6731–6741. doi:10.1128/JVI.02752-06.

21. Brady T., Roth S.L., Malani N., Wang G.P., Berry C.C., Leboulch P., Hacein-Bey-Abina S., Cavazzana-Calvo M., Papapetrou E.P., Sadelain M., Savilahti H. & Bushman F.D. (2011). – A method to sequence and quantify DNA integration for monitoring outcome in gene therapy. *Nucleic Acids Res.*, **39** (11), e72. doi:10.1093/nar/gkr140.

22. Roth S.L., Malani N. & Bushman F.D. (2011). – Gammaretroviral integration into nucleosomal target DNA *in vivo*. *J. Virol.*, **85** (14), 7393–7401. doi:10.1128/JVI.00635-11.

23. Mitchell R.S., Beitzel B.F., Schroder A.R., Shinn P., Chen H., Berry C.C., Ecker J.R. & Bushman F.D. (2004). – Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.*, **2**, E234. doi:10.1371/journal.pbio.0020234.

24. Kim H.H., van den Heuvel A.P., Schmidt J.W. & Ross S.R. (2011). – Novel common integration sites targeted by mouse mammary tumor virus insertion in mammary tumors have oncogenic activity. *PloS ONE*, **6**, e27425. doi:10.1371/journal.pone.0027425.

25. Brady T., Lee Y.N., Ronen K., Malani N., Berry C.C., Bieniasz P.D. & Bushman F.D. (2009). – Integration target site selection by a resurrected human endogenous retrovirus. *Genes Dev.*, **23**, 633–642. doi:10.1101/gad.1762309.

26. Wu X., Li Y., Crise B. & Burgess S.M. (2003). – Transcription start regions in the human genome are favored targets for MLV integration. *Science*, **300** (5626), 1749–1751. doi:10.1126/science.1083413.

27. Gillet N.A., Gutiérrez G., Rodriguez S.M., de Brogniez A., Renotte N., Alvarez I., Trono K. & Willems L. (2013). – Massive depletion of bovine leukemia virus proviral clones located in genomic transcriptionally active sites during primary infection. *PLoS Pathog.*, **9**, e1003687. doi:10.1371/journal.ppat.1003687.

28. Faschinger A., Rouault F., Sollner J., Lukas A., Salmons B., Gunzburg W.H. & Indik S. (2008). – Mouse mammary tumor virus integration site selection in human and mouse genomes. *J. Virol.,* **82** (3), 1360–1367. doi:10.1128/JVI.02098-07.

29. Wu X., Li Y., Crise B., Burgess S.M. & Munroe D.J. (2005). – Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *J. Virol.,* **79** (8), 5211–5214. doi:10.1128/JVI.79.8.5211-5214.2005.

30. Holman A.G. & Coffin J.M. (2005). – Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proc. Natl Acad. Sci. USA*, **102** (17), 6103–6107. doi:10.1073/pnas.0501646102.

31. Desfarges S. & Ciuffi A. (2010). – Retroviral integration site selection. *Viruses*, **2** (1), 111–130. doi:10.3390/v2010111.

32. Roulston A., Marcellus R.C. & Branton P.E. (1999). – Viruses and apoptosis. *Annu. Rev. Microbiol.*, **53**, 577–628. doi:10.1146/annurev.micro.53.1.577.

33. Kerkhofs P., Adam E., Droogmans L., Portetelle D., Mammerickx M., Burny A., Kettmann R. & Willems L. (1996). Cellular pathways involved in the *ex vivo* expression of bovine leukemia virus. *J. Virol.*, **70** (4), 2170–2177.

34. Twizere J.C., Kruys V., Lefèbvre L., Vanderplasschen A., Collete D., Debacq C., Lai W.S., Jauniaux J.C., Bernstein L., Semmes O.J., Burny A., Blackshear P.J., Kettmann R. & Willems L. (2003). – Interaction of retroviral tax oncoproteins with tristetraprolin and regulation of TNF-α expression. *J. Natl Cancer Inst.*, **95** (24), 1846–1859. doi:10.1093/jnci/djg118.

35. Witzany G. (2012). – Viruses: essential agents of life. Springer, Dordrecht, 428 pp. doi:10.1007/978-94-007-4899-6.

36. Arbuckle J.H. & Medveczky P.G. (2011). – The molecular biology of human herpesvirus-6 latency and telomere integration. *Microbes Infect.*, **13** (8–9), 731–741. doi:10.1016/j.micinf.2011.03.006.

37. Achachi A., Florins A., Gillet N., Debacq C., Urbain P., Manfouo Foutsop G., Vandermeers F., Jasik A., Reichert M., Kerkhofs P., Lagneaux L., Burny A., Kettmann R. & Willems L. (2005). – Valproate activates bovine leukemia virus gene expression, triggers apoptosis and induces leukemia/lymphoma regression *in vivo*. *Proc. Natl Acad. Sci. USA*, **102** (29), 10309–10314. doi:10.1073/pnas.0504248102.

38. Jaenisch R. (1976). – Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc. Natl Acad. Sci. USA*, **73** (4), 1260–1264. doi:10.1073/pnas.73.4.1260.

39. Rowe H.M., Jakobsson J., Mesnard D., Rougemont J., Reynard S., Aktas T., Maillard P.V., Layard-Liesching H., Verp S., Marquis J., Spitz F., Constam D.B. & Trono D. (2010). – KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*, **463**, 237–240. doi:10.1038/nature08674.

40. Blomberg J., Benachenhou F., Blikstad V., Sperber G. & Mayer J. (2009). – Classification and nomenclature of endogenous retroviral sequences (ERVs): problems and recommendations. *Gene*, **448** (2), 115–123. doi:10.1016/j.gene.2009.06.007.

41. Jurka J., Kapitonov V.V., Pavlicek A., Klonowski P., Kohany O. & Walichiewicz J. (2005). – Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.*, **110** (1–4), 462–467. doi:10.1159/000084979.

42. Goodier J.L. & Kazazian H.H. Jr (2008). – Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell*, **135** (1), 23–35. doi:10.1016/j.cell.2008.09.022.

43. Jern P. & Coffin J.M. (2008). – Effects of retroviruses on host genome function. *Annu. Rev. Genet.*, **42**, 709–732. doi:10.1146/annurev.genet.42.110807.091501.

44. Benveniste R.E. & Todaro G.J. (1974). – Evolution of C-type viral genes: inheritance of exogenously acquired viral genes. *Nature*, **252**, 456–459. doi:10.1038/252456a0.

45. Crochu S., Cook S., Attoui H., Charrel R.N., De Chesse R., Belhouchet M., Lemasson J.J., de Micco P. & de Lamballerie X. (2004). – Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of *Aedes* spp. mosquitoes. *J. Gen. Virol.*, **85** (7), 1971–1980. doi:10.1099/vir.0.79850-0.

46. Horie M., Honda T., Suzuki Y., Kobayashi Y., Daito T., Oshida T., Ikuta K., Jern P., Gojobori T., Coffin J.M. & Tomonaga K. (2010). – Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature*, **463**, 84–87. doi:10.1038/nature08695.

47. Belyi V.A., Levine A.J. & Skalka A.M. (2010). – Unexpected inheritance: multiple integrations of ancient Bornavirus and Ebolavirus/Marburgvirus sequences in vertebrate genomes. *PLoS Pathog.*, **6** (7), e1001030. doi:10.1371/journal.ppat.1001030.

48. Katzourakis A. & Gifford R.J. (2010). – Endogenous viral elements in animal genomes. *PLoS Genet.*, **6** (11), e1001191. doi:10.1371/journal.pgen.1001191.

49. Kapoor A., Simmonds P. & Lipkin W.I. (2010). – Discovery and characterization of mammalian endogenous parvoviruses. *J. Virol.*, **84** (24), 12628–12635. doi:10.1128/JVI.01732-10.

50. Feng S., Jacobsen S.E. & Reik W. (2010). – Epigenetic reprogramming in plant and animal development. *Science*, **330** (6004), 622–627. doi:10.1126/science.1190614.

51. Rawn S.M. & Cross J.C. (2008). – The evolution, regulation, and function of placenta-specific genes. *Annu. Rev. Cell Dev. Biol.*, **24**, 159–181. doi:10.1146/annurev.cellbio.24.110707.175418.

52. Blaise S., de Parseval N., Benit L. & Heidmann T. (2003). – Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc. Natl Acad. Sci. USA*, **100** (22), 13013–13018. doi:10.1073/pnas.2132646100.

53. Dupressoir A., Vernochet C., Bawa O., Harper F., Pierron G., Opolon P. & Heidmann T. (2009). – Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proc. Natl Acad. Sci. USA*, **106** (29), 12127–12132. doi:10.1073/pnas.0902925106.

54. Klenerman P., Hengartner H. & Zinkernagel R.M. (1997). – A non-retroviral RNA virus persists in DNA form. *Nature*, **390**, 298–301. doi:10.1038/36876.

55. Geuking M.B., Weber J., Dewannieux M., Gorelik E., Heidmann T., Hengartner H., Zinkernagel R.M. & Hangartner L. (2009). – Recombination of retrotransposon and exogenous RNA virus results in nonretroviral cDNA integration. *Science*, **323** (5912), 393–396. doi:10.1126/science.1167375.

56. Hüser D., Gogol-Döring A., Lutter T., Weger S., Winter K., Hammer E.M., Cathomen T., Reinert K. & Heilbronn R. (2010). – Integration preferences of wildtype AAV-2 for consensus rep-binding sites at numerous loci in the human genome. *PLoS Pathog.*, **6** (7), e1000985. doi:10.1371/journal.ppat.1000985.

57. Brister J.R. & Muzyczka N. (1999). – Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. *J. Virol.,* **73** (11), 9325–9336.

58. McCarty D.M., Pereira D.J., Zolotukhin I., Zhou X., Ryan J.H. & Muzyczka N. (1994). – Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. *J. Virol.*, **68** (8), 4988–4997.

59. Gao J., Luo X., Tang K., Li X. & Li G. (2006). – Epstein-Barr virus integrates frequently into chromosome 4q, 2q, 1q and 7q of Burkitt's lymphoma cell line (Raji). *J. Virol. Meth.*, **136** (1–2), 193–199. doi:10.1016/j.jviromet.2006.05.013.

60. Hurley E.A., Klaman L.D., Agger S., Lawrence J.B. & Thorley-Lawson D.A. (1991). – The prototypical Epstein-Barr virus-transformed lymphoblastoid cell line IB4 is an unusual variant containing integrated but no episomal viral DNA. *J. Virol.,* **65** (7), 3958–3963.

61. Lestou V.S., De Braekeleer M., Strehl S., Ott G., Gadner H. & Ambros P.F. (1993). – Non-random integration of Epstein-Barr virus in lymphoblastoid cell lines. *Genes Chromosomes Cancer*, **8** (1), 38–48. doi:10.1002/gcc.2870080108.

62. Takakuwa T., Luo W.J., Ham M.F., Sakane-Ishikawa F., Wada N. & Aozasa K. (2004). – Integration of Epstein-Barr virus into chromosome 6q15 of Burkitt lymphoma cell line (Raji) induces loss of BACH2 expression. *Am. J. Pathol.*, **164** (3), 967–974. doi:10.1016/S0002-9440(10)63184-7.

63. Nacheva E.P., Ward K.N., Brazma D., Virgili A., Howard J., Leong H.N. & Clark D.A. (2008). – Human herpesvirus 6 integrates within telomeric regions as evidenced by five different chromosomal sites. *J. Med. Virol.*, **80** (11), 1952–1958. doi:10.1002/jmv.21299.

64. Neuveut C., Wei Y. & Buendia M.A. (2010). – Mechanisms of HBV-related hepatocarcinogenesis. *J. Hepatol.*, **52** (4), 594–604. doi:10.1016/j.jhep.2009.10.033.

65. Murakami Y., Saigo K., Takashima H., Minami M., Okanoue T., Brechot C. & Paterlini-Brechot P. (2005). – Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. *Gut*, **54** (8), 1162–1168. doi:10.1136/gut.2004.054452.

66. Tamori A., Nishiguchi S., Shiomi S., Hayashi T., Kobayashi S., Habu D., Takeda T., Seki S., Hirohashi K., Tanaka H. & Kubo S. (2005). – Hepatitis B virus DNA integration in hepatocellular carcinoma after interferon-induced disappearance of hepatitis C virus. *Am. J. Gastroenterol.*, **100**, 1748–1753. doi:10.1111/j.1572-0241.2005.41914.x.

67. Ferber M.J., Montoya D.P., Yu C., Aderca I., McGee A., Thorland E.C., Nagorney D.M., Gostout B.S., Burgart L.J., Boix L., Bruix J., McMahon B.J., Cheung T.H., Chung T.K., Wong Y.F., Smith D.I. & Roberts L.R. (2003). – Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. *Oncogene*, **22**, 3813–3820. doi:10.1038/sj.onc.1206528.

68. Stephen S.L., Montini E., Sivanandam V.G., Al-Dhalimy M., Kestler H.A., Finegold M., Grompe M. & Kochanek S. (2010). – Chromosomal integration of adenoviral vector DNA *in vivo*. *J. Virol.*, **84** (19), 9987–9994. doi:10.1128/JVI.00751-10