Karin Moelling
Int Med Microbiol , University Zürich
MPI Mol Genetics, Berlin, HPI Hamburg
moelling@molgen.mpg.de
0049 172 3274306

Thinking about Ebola viruses in times of a virus epidemic

While a devastating Ebola Virus epidemic is frightening us it may come as a surprise to learn that Ebola virus genes are present in mammalian genomes.
The DNA in mammalian cells contains many viral genomes, best known are the retroviruses, which integrate as a normal step in their viral life cycle. However other viruses integrate"illigitimately", being RNA viruses with no obligatory DNA intermediate, which normally are non integrating. One of them are **Ebola viruses.** Ebola virus genes can be traced back in the mammalian genome to about 40 to 50Mio, some even 100 Mio years ago. They are surprisingly intact, and even today express some proteins. They are related to present day`s acute viruses from Zaire (Bayli et al 2010;2010; Katzourakisand Gifford 2010). Ebola viruses are RNA-containing viruses, they should not easily have entered the genomic DNA. However a Reverse Transcriptase from retroviruses also present in mammalian cells and also in germ cells may have helped to generate a DNA copy for illigitimate integration and thereby enabled subsequent vertical transmission and inheritance. Retrotransposons such as LINE elements express Reverse transciptases and may have supplied the enzyme in trans. Recombination events could also be involved. Ebolaviruses express the nuclear protein NC in bats, in other hosts fragments of it and Viral protein vp35, opssum also part of the polymerase. **Ebola** viruses similar to Marburg viruses, another member of this filovirus family, are mainly exogenous viruses, transmitted in Sudan, Congoo and around Lake Victoria. Endogenous Ebola viruses were detected in opossum, whallaey, kangeroos, rats and in bats, guinea pigs, shrews and more than 65 Mio years ago also pigs. Ebola and Marburg viruses may have entered the genomes independently. Other relatives went extant. It came as a surprise, that filoviruses were detected in the genomes of bats in North America, in Asian primates in rodents in South America and Australia, besides swine in Reston, the only location where Ebola outbreaks have been described. What does integration of viruses into germlines and survial as fossils from many Mio ago mean?

Also Borna viruses, another RNA-containing virus with a simliar genome structure as Ebola viruses, are integrated as DNA in mammalian genome and express viral proteins in humans. **In contrast to Ebola they entered and were endogenized into the human genome.**
Bornaviruses integrated in humans, monkeys, mamosets , elephants, lemures mouse rat, birds - but not horses! Since they are absent in horse genomes one speculation goes, that horses can get Borna virus-dependent neurological diseases such as mental depressions, while humans are protected by expressing Borna virus gene products against this type of disease. We get depressions for other reasons. Integrated Bornaviruses express viral proteins thereby protecting their hosts and survived in the genomes till today. Why did Ebola viruses not get endogenized into the human genome for our protection? May be infection of humans occurred too recently, contacts to bats has become closer than before, possibly on food markets, or infection of germ cells may not be possible. In contrast to the structurally related Bornaviruses, Ebolaviruses do not replicate inside the nucleus as Bornaviruses do, therefore Bornaviruses had a higher chance to enter the hosts` DNA genomes – this is an explanation given by A. Katzourakis (personal comm.).
About 10 different viruses found their way into mammalian genomes as described by Katzourakis and Gifford (2010). Besides Ebola and Borna even circoviruses were detected harboring a rare single-stranded circular DNA. Circo viruses are frequent in pigs, chicken and

pigeon. They are rather unique by containing a circular single-stranded DNA in a small naked icosaeder.
From the published literature one can conclude, that almost all virustypes and replication strategies made it as endogenous sequences into mammalian genomes. The frequency may be underestimated because some seuqnees may no longer be recognized due to mutations. Integration occurs more frequently by viruses that establish persistant infections and replicate in the nuclesus, yet in some cases also outside.(All this is described by Katzourakis and Gifford, 2010, and Borna and Ebola also by Belyi , Levine and Skalka, 2010).

Retrovirues in mammalian genomes

Sequencing of the human genome at the beginning of our century unraveled surprises. The human genome and alomost every eukaryotic genome is full of DNA proviruses derived from retroviruses. Most of them are fossils, more or less defective, left overs from viral infections about 35 or 100 Mio years ago. Up to 45% of the human genome consists of retroviral relics with different degrees of deletions.
Integration as DNA copies must have been advantageous for a virus, this is like a safe harbor for maintenance, to be inherited and to produce progeny. There is also mutualism, because the integrated viruses also protect the cells. This is of benefit not only for the individual cell but for the entire host for survival. It is a basic principle of all viruses to provide its host cell with an antiviral defense against a similar invader. Monopoly of a virus inside a cell is advantageous for the virus to guarantee higher viral progeny than in the presence of a competitor. It is also less dangerous for the host cell. In principle, a virus infection will lead to "superinfection exclusion", or simply spoken: entrance forbidden. This term was coined for the phage-bacterial world - and holds up for the whole virus world till today.

Endogenous viruses protect their host.

Some examples may illustrate the protective effect of an integrated or endogenized virus - best studied for retroviruses. Virologists know that one cannot easily superinfect a cell with the same virus, designated as viral interference. Protection inside cells was developed by monkeys as defense against a superinfecting virus. They express the TRIM1alpha protein directed against the Gag protein of an incoming virus. Even certain honey bees defend themselves by expressing a structural protein of an endogenized virus against superinfection (the Israeli acute paralysis virus, an RNA virus. Ref, see Belyi et al. 2010). Also a phage can protect its host against superinfection by the same phage: a retrophage expressing an revers transcriptase. The phage does not even have a name but its host is the well-known bacteria *Bordetella pertussis*, causing whooping caughs.
Endogenization of a retroviral DNA as a protective measure can be witnessed in an actively ongoing reality show with Koalas in Australia: these animals were threatened from going extinct and were therefore isolated on an island about 100 years ago to recover - yet the opposite happened. They contracted a retroviral infection by a gibbon-ape leukemia virus and many of them died. The survivors exhibited two kinds of integrated retroviral sequences, some from de novo infections at variable integration sites and - to everybodies surprise - also at stable integration sites in all cells (Tarlington et al , 2008). This must have been caused by fixation through germline infection and inheritance. The process of retroviral endogenization happened in the Koalas in less than 100 years.

Viruses and anitviral defense are related.

Retroviruses are ideal for creating antiviral defense, because they are cellular genes immediately after infection, reverse transcription and integration and undergo mutations by an error-prone RT during each round of replication. This will lead to genetic diversity and allow selection of survivors. it is a fast way for a cell to develop a defense system even a more general one, not only against identical invaders. Thus retroviruses as DNA- proviruses generated immunity not only against their colleague viruses - but are probably the basis for a general defense system. Thus, our genome and that of many other animal species have accumulated retroviruses as antiviral defense system. Retroviruses created immune systems. Thus retroviral sequences in our human

genome protected us against superinfections. With time, when a related virus did not show up as possible invader, loss of selective pressure resulted in accumulation of mutations or deletions and finally loss.

Do we need to worry?

Similarly illigitimate DNA versions of any other virus which becomes integrated will protect the cell against superinfection. Viruses may then still replicate but not cause diseases. This is the case with Ebola viruses. They were endogenized in bats, where they replicate without causing diseases but may be transmitted to humans. The geography of bats carrying endogenized Ebola sequences is very surprising and indicates that the virus was not confined to areas in Africa. What about bats in other places in the world? What about today? Bats are a species with more than 1000 types. Ebola was also found endogenized in swines. Nobody can predict what this may mean or whether there are today potential dangers associated.

References:

Endogenous viral elements in animal genomes.
Katzourakis A, Gifford RJ.
PLoS Genet. 2010 Nov 18;6(11):e1001191
Genome Res. 2006 Dec;16(12):1548-56. Epub 2006 Oct 31.

Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes.
Belyi VA, Levine AJ, Skalka AM.
PLoS Pathog. 2010 Jul 29;6(7):e1001030.

Sequences from ancestral single-stranded DNA viruses in vertebrate genomes: the parvoviridae and circoviridae are more than 40 to 50 million years old.
Belyi VA, Levine AJ, Skalka AM.
J Virol. 2010 Dec;84(23):12458-62.

Initial sequencing and analysis of the human genome.
Lander ES et al Nature. 2001 Feb 15;409(6822):860-921

Identification of the regions of Fv1 necessary for murine leukemia virus restriction.
Bishop KN, Bock M, Towers G, Stoye JP.
J Virol. 2001 Jun;75(11):5182-8.

Genome Res. 2006 Dec;16(12):1548-56. Epub 2006 Oct 31.
Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements.
Dewannieux M[1], Harper F, Richaud A, Letzelter C, Ribet D, Pierron G, Heidmann T.

Biology and evolution of the endogenous koala retrovirus.
Tarlinton R, Meers J, Young P.
Cell Mol Life Sci. 2008 Nov;65(21):3413-21.