Review

Avian Endogenous Retroviruses

( birds / endogenous retroviruses / expression / evolution )

L. BORISENKO

Institute of Molecular Biology and Genetics, Ukrainian Academy of Sciences, Kiev, Ukraine

Abstract. Recent data about the diversity of AER, their structure, expression and possible ways of evolution are summarized and analysed in the present review. Additionally, the role of endogenous retroviruses in ontogenesis and pathology is discussed.

Integration into the host genome is an important step in the life cycle of retroviruses. When integration takes place in a germ cell, it gives rise to endogenous retroviruses. These viruses can be inherited along with the rest of the host genome as Mendelian genes.

Endogenous retroviruses have been reported in six classes of vertebrates (Herniou et al., 1998). Together with mammalian, avian endogenous retroviruses (AERs) are the subject of most intensive research. AERs are known to be responsible for the emergence of highly virulent virions via recombination (Venugopal, 1999). Three groups of AERs have been described to date: (1) the ev loci, (2) the endogenous avian retrovirus family (EAV), and (3) the human endogenous retrovirus type I (HERV-I)-related retroviruses. They are thought to belong to two retroviral genera: avian leukosis viruses (ALVs) and murine leukaemia viruses (MLVs). Endogenous retroviruses clustering with MLVs are widespread among vertebrates (Herniou et al., 1998) while ALVs are unique to birds.

This article reviews the structure and evolution of AERs as well as their possible role in ontogenesis and pathology.

Diversity, structure and expression

The best studied group of AERs is the chicken ev loci. This family includes a variety of complete and defective proviruses that are closely related to avian sarcoma/leukosis viruses (ASLVs) (Table 1). There are over 50 different ev loci. Twenty-two proviruses (numbered 1 through 22) have been found in White Leghorn, 7 in Brown Leghorn, 5 in Italian partridge, 12 in Rhode Island Red chickens, and 31 in the genome of meat breeds (Astrin, 1978; Hughes et al., 1979; Gudkov et al., 1981, 1986b; Chernov et al., 1984; Humphries et al., 1984; Ziemiecki et al., 1988; Sabour et al., 1992; Grunder et al., 1995; Gorbovitskaia et al., 1998). Although some evs from genomes of different chicken breeds are equivalent to each other, each breed, line or even specimen has a unique set of evs.

Several evs are capable of being expressed, while others are transcriptionally silent (Hayward et al., 1980; Baker et al., 1981; Hughes et al., 1981). Some evs encode infectious virions (ev-2, ev-7, ev-10–12, ev-14, ev-18–20) (Astrin et al., 1978; Robinson et al., 1979; Astrin, 1980; Humphries et al., 1984). Inactivation of proviruses may be due to the multiple deletions that occur in proviral structures or to the methylation of CpG dinucleotide sequences (Groudine et al., 1981; Gudkov et al., 1986a). ev-15 and ev-16 are composed exclusively of long terminal repeats (LTRs) that arise by recombinational deletion of the viral genes (Smith et al., 1984).

The EAV family found in the genomes of the Gallus species forms a heterogeneous group that shows weak similarity with evs (Tables 1, 2). Nevertheless, EAV elements may belong to the ALV genus (Dimcheff et al., 2000). We have demonstrated that the pol region of EAV-HP is more closely related to the RSV pol gene than to the pol genes of other retroviral genera (data not shown). All the EAVs are more ancient than evs and therefore defective (see next chapter). They show deletions in the env regions and, in some cases, in the gag and pol regions (Table 1). Loss of part of the env gene is also observed in some mammalian endogenous retroviruses and seems to be a common phenomenon among ancient endogenous elements (Coffin et al., 1997). This type of deletion is not specific of modern evs. Interestingly, the full-length env gene is present in E-51, though the env region coding for the surface glycoprotein (SU) is defective with multiple frameshifts (Boyce-Jacino et al., 1992).
EA V includes 5 subfamilies: EA V-0, EA V-HP, E-51, E-33, and avian retrotransposon from the chicken genome (ART-CH).

The LTRs of EA V-0 are unusually short but functional (Boyce-Jacino et al., 1989). The LTRs of E-33, E-51 and ART-CH are more closely related to each other than to the LTRs of other AERs (Table 2). This similarity may be the result of LTR recombination between several distinct AERs (Boyce-Jacino et al., 1992).

The EA V-HP env gene is 97% identical to the env gene of the ALV subgroup J (Smith et al., 1999). EA V-HP from domestic chicken shows deletions spanning part of the gag, the entire pol and part of the env genes (Sacco et al., 2000). By contrast, EA V-HP from grey

Table 1. Avian endogenous retroviruses (Rovigatti and Astrin, 1983; Humphries et al., 1984; Boyce-Jacino et al., 1989, 1992; Nikiforov and Gudkov, 1994; Smith et al., 1999; Sacco et al., 2000, 2001; Dimcheff et al., 2001)

<table>
<thead>
<tr>
<th>Retrovirus</th>
<th>Structure</th>
<th>Copy number per haploid genome</th>
<th>Size, kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>ev loci:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ev-1, 2</td>
<td>LTR-gag-pol-env-LTR</td>
<td>0 – 10</td>
<td>~ 7.5</td>
</tr>
<tr>
<td>ev-3</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>ev-4</td>
<td>gag-pol-LTR</td>
<td>50 - 100</td>
<td>NI</td>
</tr>
<tr>
<td>ev-5, 6</td>
<td>pol-env-LTR</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>ev-7, 8</td>
<td>LTR-gag-pol-env-LTR</td>
<td>10 - 15</td>
<td>6.8</td>
</tr>
<tr>
<td>ev-9</td>
<td>NI-gag-pol-env-LTR</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>ev-10, 11, 12</td>
<td>LTR-gag-pol-env-LTR</td>
<td>10 - 15</td>
<td>6.8</td>
</tr>
<tr>
<td>ev-13</td>
<td>NI</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>ev-14</td>
<td>LTR-gag-pol-env-LTR</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>ev-15, 16</td>
<td>NI</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>ev17, 18, 19, 20, 21, 22</td>
<td>LTR-gag-env-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>TERV</td>
<td>LTR-gag-env-LTR</td>
<td>0 – 10</td>
<td>~ 7.5</td>
</tr>
</tbody>
</table>

EAV:

<table>
<thead>
<tr>
<th>Retrovirus</th>
<th>Structure</th>
<th>Copy number per haploid genome</th>
<th>Size, kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART-CH</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>EAV-0</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>E-33</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>EAV-HP</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
<tr>
<td>E-51</td>
<td>LTR-Δgag-Δpol-Δenv-LTR</td>
<td>25 - 50</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Δ – deletion; NI – not identified

EAV includes 5 subfamilies: EAV-0, EA V-HP, E-51, E-33, and avian retrotransposon from the chicken genome (ART-CH).

The LTRs of EAV-0 are unusually short but functional (Boyce-Jacino et al., 1989). The LTRs of E-33, E-51 and ART-CH are more closely related to each other than to the LTRs of other AERs (Table 2). This similarity may be the result of LTR recombination between several distinct AERs (Boyce-Jacino et al., 1992).

The EAV-HP env gene is 97% identical to the env gene of the ALV subgroup J (Smith et al., 1999). EAV-HP from domestic chicken shows deletions spanning part of the gag, the entire pol and part of the env genes (Sacco et al., 2000). By contrast, EAV-HP from grey

Table 2. Percentages of DNA sequence similarity among AERs

<table>
<thead>
<tr>
<th>Retrovirus¹</th>
<th>% similarity to²:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ev-1</td>
</tr>
<tr>
<td>EAV-HP</td>
<td>41.9/50.4/59.7/52.0</td>
</tr>
<tr>
<td>E-33</td>
<td>48.2/NI/NI/50.9</td>
</tr>
<tr>
<td>E-51</td>
<td>53.8/NI/62.5/48.3</td>
</tr>
<tr>
<td>ART-CH</td>
<td>38.7/NI/NI/NI</td>
</tr>
</tbody>
</table>

¹GeneBank accession numbers of sequences used in the alignment: ev-1 – AY013303; EAV-HP – NC_000930; E-33 – M95190; E51 – M95189; EAV-0 – M31063 and X59844; ART-CH – L25262.

²Sequences were aligned using program CLUSTAL_W and percentage similarity as [1-P distances]100 were calculated. First number – LTR similarity, second number – gag gene similarity, third number – pol gene similarity, fourth number – env gene similarity. NI – not identified since no data were available for comparison. Boldface numbers indicate the greatest similarity.
jungle fowl contains an intact pol gene (Sacco et al., 2001).

ART-CH has also been described as E-13 element. Since both of these clones are similar, it is reasonable to assume that E-13 and ART-CH are the same provirus integrated in different sites of the genome. ART-CH has a chimeric structure with regions homologous to all known AERs. The large deletion in the env gene of ART-CH is analogous to that seen in EAV-0 (Boyce-Jacino et al., 1992; Nikiforov and Gudkov, 1994).

Distinct combinations of homologous regions among the members of the EAV family may be the result of recombination between different retroviruses.

Some EAVs are expressed but there is no evidence for the production of proteins. ART-CH, as also murine VL30 (Adams et al., 1988) and human THE-1 (Paulson et al., 1985), may be packed and integrated using products of a helper virus (Nikiforov and Gudkov, 1994). The same process is thought to take place in the spreading of EAV-HP (Sacco et al., 2000).

In general, information about proviruses of the EAV family is incomplete: the entire proviral structure of EAV-0, E-33, E-51 is not known.

HERV-I-related endogenous retroviruses are known to be present in genomes of all vertebrates (Martin et al., 1997). Recently, we have identified a retrovirus of this group in the genome of domestic chicken (Borisenko and Rynditch, unpublished). Only genes of protease and reverse transcriptase have been sequenced for the production of proteins. ART-CH, as also murine proviruses, have been sequenced (Borisenko and Rynditch, unpublished).

Endogenous retroviruses outside of the Gallus species have not been examined as thoroughly. Recently, ALV-related gag sequences were identified in three of five families of galliform birds, indicating their ancient association with their hosts (Dimcheff et al., 2000). In addition, complete tetraonine endogenous retrovirus (TERV) has been described (Dimcheff et al., 2001).

Moreover, several AERs clustered with the MLV and spumavirus genus have been sequenced for reverse transcriptase and protease genes of members of the Passeriformes, Columbiformes, Tinamiformes and Galliformes (Martin et al., 1997, 1999; Herniou et al., 1998).

Evolution of avian endogenous retroviruses

One way of providing phylogenetic information of endogenous retroviruses is by comparing the polymorphism of the integration sites. Two or more species harbouring the same endogenous retrovirus share a common ancestor into whose germ line the integration took place (Johnson and Coffin, 1999).

The genus Gallus is composed of four species (Fig. 1). The EAV family is restricted to all Gallus species, while evs are specific for domestic chicken and its wild relative red jungle fowl only, and therefore are younger than EAV. In turn, different members of EAV seem to have infected the Gallus species at different times: E-51 and E-33 may be older than EAV-0 (Boyce-Jacino et al., 1992).

Since 5′ and 3′ LTRs were identical at the time of integration, they can serve as "evolutional clock" (Coffin et al., 1997). Based on this assumption, ancient proviruses should have differences in their LTRs. The 5′ and 3′ LTRs of ART-CH are almost identical as those of the ev loci and EAV-HP, and therefore ART-CH may be a recent component of the chicken genome. In general, the time of integration of the EAV family into the host genome remains to be determined more precisely using information about proviral distribution in all four Gallus species.

Multiple ev loci could be the result of ancient locus duplication (probably, ev-1) (Rovagatti and Astrin, 1983). Sequence mutations and deletions observed within many evs might have arisen during reverse transcription. Defects in the EAV family indirectly show the ancient origin of this group compared to the ev loci. Having accumulated numerous point mutations, deletions and insertions, none of the ancient proviruses have been found to yield infectious virions and to have exogenous counterparts. In general, the direction followed by endogenous retroviruses is that of disintegration (Tikhonenko and Gudkov, 1987).

The presence of ALV-related sequences in most galliform birds investigated leads to several conclusions. Megapodiidae and Cracidae diverged prior to the other three galliform families (Fig. 2). Since they are free of ALV-related sequences, it is reasonable to assume that association of ASLV and galliform birds occurred at the time of divergence of the Phasianidae, Numididae and Odontophoridae from the other families (50 million years ago) (Dimcheff et al., 2000).

The phenomenon of the presence of different ALV age groups within the galliform lineage indicates that these retroviruses have been infecting the host genome for long periods of time. Moreover, in the periods of activity they could have undergone horizontal transmission.
The usual means of transmission of endogenous retroviruses is vertical transfer via the germ line. Recently, the evidence of horizontal transfer has been obtained (Martin et al. 1999; Dimcheff et al., 2000). Frequent horizontal transmission is indicated by the low level of correlation between the viral and host phylogenies (Martin et al., 1999). It has been shown that such transfer occurred within the ALV genus (Frisby et al., 1979; Dimcheff et al., 2000).

While endogenous ALV-related retroviruses are present throughout the Galliformes lineage only, MLV-related endogenous retroviruses are widespread within the genomes of four classes of terrestrial vertebrates, indicating their ancient germ line integration (Herniou et al., 1998).

According to Doolittle et al. (1989), exogenous retroviruses exist only a short period of time and are transmitted as endogenous elements for long periods. It does not seem unreasonable to suppose that endogenous retroviruses can generate new exogenous forms via recombination events (next chapter).

**Role of avian endogenous retroviruses**

While exogenous ASLVs cause several diseases, endogenous retroviruses are harmless for their hosts. Moreover, it seems that they do not play any role in development since chickens free of evs are normal and healthy (Astrin et al., 1979). In addition, many endogenous retroviruses are transcriptionally silent, defective, and are incapable of yielding infectious virions. Therefore, not having any function, they should mutate and lose open reading frames. The rule "the older the endogenous retrovirus, the more it has changed" seems to be true.

Nevertheless, this conclusion has not been applied to modern endogenous retroviruses with exogenous counterparts. According to the classical hypothesis explaining the conservation of endogenous retroviruses, expression of their env gene blocks the receptor utilized by exogenous retroviruses and therefore determines resistance to retroviral infection (Robinson et al., 1981).

In numerous studies such tolerance has been demonstrated (Crittenden et al., 1982; Crittenden and Fadly, 1985; Kuhnlein et al., 1993). Indeed, only modern endogenous retroviruses (ev loci) have an intact env gene, which is often lost in ancient EAVs.

The potential pathogenic effects of proviral expression are associated with recombination between exogenous and endogenous retroviruses. The possibility of such recombination has been shown for ASLV (Geryk et al., 1989), MLV (Stoye and Coffin, 1987) and feline leukaemia virus (FeLV) (Chakrabarti et al., 1994). The role of EAV-HP in the emergence of ALV-J has also been demonstrated (Venugopal, 1999). Since its isolation in 1980s ALV-J became a worldwide problem in broiler industry. It is thought that subgroup J exogenous avian leukemia virus (HPRS-103) is a product of recombination with the env gene of EAV-HP.

The integration of endogenous retroviruses during their expression cycle can also cause mutagenic effects (Coffin et al., 1997).

In addition, there is evidence that evs may play a role in the induction of autoimmune vitiligo (Sreekumar et al., 2000).

**Conclusion**

Three groups of AERs belonging to the ALV and MLV genera have been described to date. We should however bear in mind that phylogenetic relatedness of the EAV family and ALV genus is only hypothetical. The complete genomic structure of E-51, E-33, EAV-0 as well as numerous MLV-related retroviruses, several chicken ev loci and multiple ALV-related proviruses from different galliform birds remains to be characterized.

ALV-related retroviruses and galliform birds constitute an example of long co-evolution. MLV-related retroviruses are thought to co-evolve along with all classes of terrestrial vertebrates. This observation suggests that other avian orders may have their "own" groups of endogenous retroviruses acquired long ago. Studies of such birds may shed some light on the direction of retrovirus-host co-evolution.

Even if we take into account that chicken endogenous retroviruses are the most extensively studied, many questions remain. First of all, the absence of a role in development is still not conclusively established. It is known that human endogenous retroviruses have some developmental functions (Griffiths, 2001). In the context of retrovirus-host interaction, the question of provirus distribution within the host genome is also important.

**Acknowledgments**

The author thanks Dr. A. Rynditch and Dr. A.-L. Haenii for reading a draft of the manuscript, for their valuable comments and suggestions.

**References**


