Arch Virol 146/11 (2001)

VDN Virology Division News

Classifying reverse transcribing elements: a proposal and a challenge to the ICTV*

R. Hull

John Innes Centre, Norwich Research Park, Colney, Norwich, U.K.

The replication of most nucleic acids is either from DNA to DNA (chromosomal and viral nucleic acids) or from RNA to RNA (viruses and some cytoplasmic nucleic acids). However, an increasing number of nucleic acids are being found whose replication involves reverse transcription of RNA to produce DNA. This replication is driven by the enzyme reverse transcriptase (RT), which was first recognised over 30 years ago [1, 21]. Nucleic acids that replicate by reverse transcription are termed retroelements [10, 20] and this form of replication is employed by elements in higher plants, higher animals, fungi, insects and bacteria. Retroelements have been grouped into viral retroelements, eukaryotic chromosomal non-viral retroelements and bacterial chromosomal retroelements (Table 1) [9]. Retrotransposons are also known as LTR (long terminal repeat) retrotransposons and retroposons as non-LTR or poly (A) retrotransposons. There are various and separate classification systems for the viral and non-viral elements but, as these elements have many features in common, a universal classification for all retroelements should be considered.

Viral retroelements have been formally classified by the International Committee on Taxonomy of Viruses (ICTV) over several years. Recently retrotransposons have been included in the classification scheme to give a total of five families (Table 2) [17, 18, 23]. The inclusion of retrotransposons in virus classification is based primarily on similarities in their genome organizations with those of viral retroelements and on phylogenetic relationships among the reverse transcriptases. It is the encoding of a reverse transcriptase and the mechanism of replication that differentiate retroelements from all other viruses and cellular elements. Thus consideration should be given to the inclusion of other reverse transcribing elements in the classification. The suggestion being put forward here is that there is a case for classifying all these elements together.

As well as the basic RNA to DNA replication, retroelements have several other features in common. The enzyme complex of active retroelements comprises reverse

---

* Editor's footnote: This paper is a revised up-dated version of a paper published previously (Hull R, 1999, Arch Virol 144: 209–214). The ideas proposed by the author are under discussion by an *ad hoc* subcommittee of the Executive Committee of ICTV, who would welcome comments from any interested virologists on the proposals and the questions posed by Roger Hull. Please send comments to the Chair of the subcommittee (Mike Mayo) at mmayo@scri.sari.ac.uk.

Table 1. Viral and non-viral reverse transcribing elements*

	Viral retroelements		
	I. Retroviruses (RNA in virions)		II. Pararetroviruses (DNA in virions)
	Eukaryotic non-viral retroelements		
	I. Retrotransposon		II. Retroposon
LTR[a]	+		–
RT	+		+
INT	+		+
Examples	Ty, *copia* *gypsy*, Tnt1 SIRE-1, Tat-1		LINEs mitochondrial introns and plasmids
	Bacterial retroelements		
		Retron	
	LTR	–	
	RT	+	
	INT	–	
	Example	msDNA	

* Adapted from [9]

[a] LTR long terminal repeat; RT reverse transcriptase; INT integrase

transcriptase (RT), ribonuclease H (RNaseH) (except possibly for some retrons) and, in eukaryotic elements, an open reading frame that codes for a nucleic acid-binding protein, termed *gag* for retroviruses or coat protein for pararetroviruses [6]. Most elements, except hepadnaviruses, encode an aspartate proteinase, and retroviruses, retrotransposons and retroposons also code for an integrase (*int*), which is absent from most, if not all, pararetroviruses and hepadnaviruses. Each of these proteins has consensus amino acid sequences for RT (for example [4, 13, 14, 26, 27]), for RNaseH [14], for aspartate proteinase [14, 22], for *int* [14] and for gag/coat protein in the cys sequence [3, 7]. These consensus sequences are taken to show a common origin for each of the four proteins [14, 27] and also to demonstrate the concept of a common replicon to which other functions are added to adapt the element to the "niche" in which it exists [9].

The initial suggestion (Table 3) is to build from the already fixed lower taxonomy and bring the current taxa together under higher taxa. The proposal is to create two Orders, one (*Retrovirales*) of "viruses" potentially capable of horizontal transmission, and the other (*Retrales*) of the non-viral elements, and to divide these into five Suborders. The *Retrovirales* would consist of three Suborders, the *Orthoretrovirineae*, which would contain the retroviruses (with encapsidated RNA genomes and which involve integration in their replication), the *Pararetrovirineae*, which would contain the pararetroviruses (with encapsidated DNA genomes and which replicate episomally) and the *Retrotransposineae*, which would contain the retrotransposons. The lower levels of classification in this Order (family and genus) are those already accepted. The *Retrales* would consist of two Suborders, the previously unclassified retroposons (*Retroposonineae*) and retrons (*Retronineae*). The latter two Suborders still require a classification structure at the lower levels.

Table 2. Classification of reverse transcribing viruses*

Family: *Retroviridae*	
Genera	Type species
Alpharetrovirus	*Avian leukosis virus*
Betaretrovirus	*Mouse mammary tumor virus*
Gammaretrovirus	*Murine leukemia virus*
Deltaretrovirus	*Bovine leukemia virus*
Epsiloretrovirus	*Walleye dermal sarcoma virus*
Lentivirus	*Human immunodeficiency virus 1*
Spumavirus	*Chimpanzee foamy virus*
Family: *Hepadnaviridae*	
Genera	Type species
Orthohepadnavirus	*Hepatitis B virus*
Avihepadnavirus	*Duck hepatitis B virus*
Family: *Caulimoviridae*	
Genera	Type species
Badnavirus	*Commelina yellow mottle virus*
Caulimovirus	*Cauliflower mosaic virus*
"Rice tungro bacilliform-like viruses"	*Rice tungro bacilliform virus*
"Soybean chlorotic mottle-like viruses"	*Soybean chlorotic mottle virus*
"Cassava vein mosaic-like viruses"	*Cassava vein mosaic virus*
"Petunia vein clearing-like viruses"	*Petunia vein clearing virus*
Family: *Pseudoviridae*	
Genera	Type species
Pseudovirus	*Saccharomyces cerevisiae Ty-1 virus*
Hemivirus	*Drosophila melanogaster copia virus*
Family: *Metaviridae*	
Genera	Type species
Metavirus	*Saccharomyces cerevisiae Ty-3 virus*
Errantivirus	*Drosophila melanogaster gypsy virus*

* from van Regenmortel *et al.* [23]

This classification will most probably need further refining to allow for some recent findings and suggestions:

1. Although the genome organizations and replication mechanisms of retroviruses and retrotransposons have many features in common, it was considered that the main difference between them was that the former encoded a protein, *env*, which is involved in the formation of virus particles and in horizontal spread. It was considered that retrotransposons and retroposons could transpose within an individual cell and pass vertically but could not infect another host by horizontal transmission. However, particles and possible horizontal spread have been demonstrated for several retrotransposons (e.g. yeast *Ty* elements [15]; *copia* and *gypsy* elements of *Drosophila* [5, 11, 19]). Recently, *env*-like coding regions have been recognised in retrotransposons from plants [reviewed in 16] although no retrotransposon particles have been detected and horizontal spread has not yet been demonstrated. The potential presence of an *env* gene in these plant retrotransposons poses some interesting questions as the current dogma

Table 3. Proposed classification of reverse transcribing elements

Order: *Retrovirales*	
Suborder: *Orthoretrovirineae*	
Family: *Retroviridae*	
Genera	Type species
Alpharetrovirus	*Avian leukosis virus*
Betaretrovirus	*Mouse mammary tumor virus*
Gammaretrovirus	*Murine leukemia virus*
Deltaretrovirus	*Bovine leukemia virus*
Epsiloretrovirus	*Walleye dermal sarcoma virus*
Lentivirus	*Human immunodeficiency virus 1*
Spumavirus	*Chimpanzee foamy virus*
Suborder: *Pararetrovirineae*	
Family: *Hepadnaviridae*	
Genera	Type species
Orthohepadnavirus	*Hepatitis B virus*
Avihepadnavirus	*Duck hepatitis B virus*
Family: *Caulimoviridae*	
Genera	Type species
Badnavirus	*Commelina yellow mottle virus*
Caulimovirus	*Cauliflower mosaic virus*
"Rice tungro bacilliform-like viruses"	*Rice tungro bacilliform virus*
"Soybean chlorotic mottle-like viruses"	*Soybean chlorotic mottle virus*
"Cassava vein mosaic-like viruses"	*Cassava vein mosaic virus*
"Petunia vein clearing-like viruses"	*Petunia vein clearing virus*
Suborder: *Retrotransposineae*	
Family: *Pseudoviridae*	
Genera	Type species
Pseudovirus	*Saccharomyces cerevisiae Ty-1 virus*
Hemivirus	*Drosophila melanogaster copia virus*
Family: *Metaviridae*	
Genera	Type species
Metavirus	*Saccharomyces cerevisiae Ty-3 virus*
Errantivirus	*Drosophila melanogaster gypsy virus*
Order: *Retrales*	
Suborder: *Retroposineae*	
Suborder: *Retronineae*	

is that cell-to-cell movement within plants and horizontal spread between plants does not require the membrane-associated features of the *env* gene product. However, the majority of sequenced retrotransposons lack any potential to code for an *env* gene. This lack raises the question of whether or not retroelements that can spread horizontally, and thus are in reality viruses, should be grouped separately from retroelements that have no potential for horizontal spread (see comments below on the definition of a virus).

2. Many of the retrotransposons and retroposons of plants have mutations or deletions that are considered to render them incapable of replication. Although some may possibly have replication functions restored by complementation from another retroelement, it is generally thought that most are essentially "fossil" remains of earlier transposition events. Should a classification of retroelements distinguish between potentially active and inactive elements?
3. Phylogenetic analyses of the amino acid sequences of retrotransposon reverse transcriptases reveal two distinct groupings: the *Ty1/copia* group and the *Ty3/gypsy* group (see [2, 27]). Of these groups, the animal-infecting enveloped retroviruses are more closely related to the *Ty3/gypsy* group, as is the element from *Arabidopsis thaliana* that contains the env gene [25]; another plant retrotransposon with an env gene, SIRE-1 [12], belongs to the *Ty1/copia* group. All plant *Ty1/copia* group elements have a (–) strand DNA synthesis primer-binding site complementary to $tRNA^{met}_{i}$, whereas the animal *Ty1/copia* group elements have a range of primer tRNAs. Should these different groupings be recognised in a classification?

This classification will raise various challenges for the ICTV:

1. Definition of a virus: The ICTV does not formally define a virus. In the ICTV 7th Report [23] a virus is described as "*an elementary biosystem that possesses some of the properties of living systems such as having a genome and being able to adapt to changing environments. However, viruses cannot capture and store free energy and they are not functionally active outside their host cells*". For classifying any organisms, it would seem essential to have a definition of what one is classifying. A suggestion is the definition given by Hull [8]: "*A virus is a set of one or more nucleic acid template molecules, either DNA or RNA but not both, normally encased in a protective coat or coats of protein or lipoprotein, that is able to organize its own replication only within suitable host cells. Within such cells, virus replication is (i) dependent on the host's protein-synthesizing machinery, (ii) organized from pools of the required materials rather than by binary fission, (iii) located at sites that are not separated from the host cell contents by a lipoprotein bilayer membrane, and (iv) continually giving rise to variants through various kinds of change in the viral nucleic acid*."
2. Retroposons and retrons would fit with the ICTV description of a virus but not with the generally accepted image of a virus, which has implications for the potential for horizontal spread. If one adopts a more specific definition of a virus, there could be problems in incorporating retroposons and retrons in a formal classification. However, as far as I know there is no organization that is involved in classifying such non-viral elements. The similarities in genome organization, replication cycle and in the conserved sequences themselves blur the distinction between viral and non-viral retroelements. Thus, there is a strong case for the ICTV being involved in their classification, and the challenge is to do so.
3. Most biological classification systems have taxa at higher levels than Order but the ICTV has not yet adopted any such taxa. As noted above, the results of sequence analyses point to a common evolutionary origin for the major genes involved in the reverse transcription method of nucleic acid replication. This opens the possibility for the ICTV to consider adopting higher taxa such that the two Orders in this proposed classification of retroelements would be grouped as a Class, the *Retroelementopsida.*

4. On a more specific taxonomic issue, the ICTV use the suffix *-virinae* for sub-families (23). The suffix *-ineae* is used in the subordinal name in botanical classification (Article 17 of the International Code of Botanical Nomenclature; reference [24]). There is potential for confusion here that needs to be resolved.

References

1. Baltimore D (1970) RNA-dependent DNA polymerase in virions of RNA tumor viruses. Nature 226: 1209–1211
2. Boeke JD, Corces VG (1989) Transcription and reverse transcription or retrotransposons. Annu Rev Microbiol 43: 403–434
3. Covey SN (1986) Amino acid sequence homology in *gag* region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res 14: 623–633
4. Doolittle RF, Feng D-F, Johnson MS, McClure MA (1989) Origins and evolutionary relationships of retroviruses. Q Rev Biol 64: 1–30
5. Emori Y, Shiba T, Kanaya S, Inouye S, Yuki S, Saigo K (1985) The nucleotide sequence of *copia* and *copia*-related RNA in *Drosophila* virus-like particles. Nature 315: 773–776
6. Fuetterer J, Hohn T (1987) Involvement of nucleocapsids in reverse transcription: a general phenomenon? Trends Biochem Sci 12: 92–95
7. Henderson LE, Copeland TD, Oroszlan S (1981). Primary structure of low molecular weight nucleic acid-binding proteins of murine leukemia viruses. J Biol Chem 256: 8400–8406
8. Hull R (2001). Matthew's Plant Virology. Academic Press, London
9. Hull R, Covey, SN (1996) Retroelements: propagation and adaptation. Virus Genes 11: 105–118
10. Hull R, Will H (1989) Molecular biology of viral and nonviral retroelements. Trends Genet 5: 357–359
11. Kim A, Terzian C, Santamaria P, Pelisson A, Prud'homme N, Bucheton A (1994) Retroviruses in invertebrates: the *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. Proc Natl Acad Sci USA 91: 1285–1289
12. Laten HM, Majumdar A, Gaucher EA (1998) *SIRE-1*, a *copia/Ty1*-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc Natl Acad Sci USA 95: 6897–6902
13. Li MD, Bronson DL, Lemke TD, Faras AJ (1995) Phylogenetic analysis of 55 retroelements on the basis of the nucleotide and product amino acid sequences of the *pol* gene. Mol Biol Evol 12: 657–670
14. McClure MA (1991) Evolution of retroposons by acquisition or deletion of retrovirus-like genes. Mol Biol Evol 8: 835–856
15. Mellor J, Malim MH, Gull K, Tuite MF, McReady S, Dibbayan T, Kingsman SM, Kingsman AJ (1985) Reverse transcriptase activity and Ty RNA are associated with virus-like particles in yeast. Nature 318: 583–586
16. Peterson-Burch BD, Wright DA, Laten HM, Voytas DF (2000) Retroviruses in plants? Trends Genet 16: 151–152
17. Pringle CR (1998) The universal system of virus taxonomy of the International Committee on Virus Taxonomy (ICTV) including new proposals ratified since publication of the Sixth ICTV Report in 1995. Arch Virol 143: 203–210
18. Pringle CR (1998) Virus taxonomy – San Diego 1998. Arch Virol 143: 1449–1459
19. Song SU, Gerasimova T, Kurkulos M, Boeke JD, Corces VG (1994) An *env*-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. Genes Dev 8: 2046–2057
20. Temin HM (1989) Retrons in bacteria. Nature 339: 254–255
21. Temin HM, Mizutani S (1970) RNA-directed DNA polymerase in virions of Rous sarcoma virus. Nature 226: 1211–1213
22. Toh H, Ono M, Saigo K, Miyata T (1985) Retroviral protease-like sequence in the yeast retrotransposon Ty1. Nature 351: 691
23. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens E, Estes M, Lemon S, Maniloff J, Mayo MA, McGeoch D, Pringle CR, Wickner RB (eds) (2000) Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, New York San Diego
24. Voss EG, Burdet HM, Chaloner WG, Demoulin P, Hiepko P, McNiell J, Meikle RD, Nicolson DH, Rollins RC, Silva PC, Greuter W (1983) International Code of Botanical Nomenclature. Bohn, Scheltema and Holkema, Utrecht and Dr. W. Junk, The Hague

25. Wright DA, Voytas DF (1998) Potential retroviruses in plants: Tat1 is related to a group of *Arabidopsis thaliana* Ty3/*gypsy* retrotransposons that encode envelope-like proteins. Genetics 149: 703–715
26. Xiong Y, Eickbush TH (1988) Similarity of reverse transcriptase-like sequences of viruses, transposable elements and mitochondrial introns. Mol Biol Evol 5: 675–690
27. Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9: 3353–3362

Author's address: Dr. R. Hull, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, U.K. e-mail: Roger.Hull@bbsrc.ac.uk.

Verleger: Springer-Verlag KG, Sachsenplatz 4–6, A-1201 Wien. – Herausgeber: Dr. M. H. V. Van Regenmortel, École Supérieure de Biotechnologie de Strasbourg (ESBS), Parc d'Innovation, Boulevard Sébastian Brandt, F-67400 Illkirch, France. – Redaktion: Sachsenplatz 4–6, A-1201 Wien. – Satz und Umbruch: Thomson Press (India) Ltd., New Delhi. – Offsetdruck: Adolf Holzhausens Nachfolger, Holzhausenplatz 1, A-1140 Wien. – Herstellungsort: Wien. – Printed in Austria.