Arch Virol 144/1 (1999)

VDN Virology Division News

Classification of reverse transcribing elements: a discussion document

R. Hull

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

The replication of most nucleic acids is either DNA > DNA (chromosomal and viral nucleic acids) or RNA > 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 DNA. This replication uses the enzyme reverse transcriptase (RT) which was first recognised nearly 30 years ago [1, 19]. Nucleic acids which replicate by reverse transcription are termed retroelements [9, 18] and this form of replication is found for elements in higher plants and animals, fungi, insects and bacteria. Retroelements have been grouped into viral retroelements, eukaryotic chromosomal non-viral retroelements and bacterial chromosomal retroelements (Table 1) (see [8]). Retrotransposons are also known as LTR (long terminal repeat) retrotransposons and retroposons as non-LTR or poly(A) retrotransposons. There are various 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 ICTV over several years, the most recent grouping being into three separate families each containing several genera (Table 2) [15, 16]. The retrotransposons have recently been classified into two families (Table 3) [16]. However, each of these classifications is treated separately and there is no classification for the other non-viral retroelements. The suggestion being put forward here is that there is a case to bring these separate classifications together.

As well as the basic RNA > DNA replication these elements have several other features in common. The enzyme complex of active retroelements comprises reverse transcriptase (RT) and ribonuclease H (RNaseH) (except possibly in some retrons) and, in eukaryotic elements, an open reading frame encoding a nucleic acid-binding protein termed gag for retroviruses and coat protein for pararetroviruses [6]. Most elements, except hepadnaviruses, encode an aspartate proteinase and retroviruses, retrotransposons and retroposons also encode an integrase (*int*) which is absent in most, if not all, pararetroviruses and hepadnaviruses. Each of these proteins has consensus amino acid sequences, for RT (numerous publications including [4, 12, 13, 22, 23], for RNaseH (see [13]), for aspartate proteinase [13, 20], for *int* [see 13] and for gag/coat protein the cys sequence [3, 7]. These consensus sequences are taken to show a common origin for each of the four proteins

Table 1. Viral and non-viral reverse transcribing elements

Viral retroelements		I. Retroviruses (RNA in virions)	II. Pararetroviruses (DNA in virions)
Eukaryotic nonviral 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	

[a] *LTR* Long terminal repeat; *RT* reverse transcriptase; *INT* integrase. Adapted from [8]

Table 2. Classification of reverse transcribing viruses

Family *Retroviridae*	Genera	*Alpharetrovirus*	*Avian leucosis*
		Betaretrovirus	*Mouse mammary tumor virus*
		Gammaretrovirus	*Murine leukemia virus*
		Deltaretrovirus	*Bovine leukemia virus*
		Epsiloretrovirus	*Walleye dermal sarcoma virus*
		Lentivirus	*Human immunodeficiency virus 1*
		Spumavirus	*Human spumavirus*
Family *Hepadnaviridae*	Genera	*Orthohepadnavirus*	*Hepatitis B virus*
		Avihepadnavirus	*Duck hepatitis B virus*
Family *Caulimoviridae*	Genera	*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*

Table 3. Classification of retrotransposons

Family *Pseudoviridae*	Genera	*Pseudovirus*	*Saccharomyces cereviciae Ty-1 virus*
		Hemivirus	*Drosophila melanogaster copia virus*
Family *Metaviridae*	Genera	*Metavirus*	*Saccharomyces cereviciae Ty-3 virus*
		Errantivirus	*Drosophila melanogaster gypsy virus*

(see [13, 23]) and also the concept of a common replicon to which other function are added to adapt the element to the "niche" in which it exists [8].

The initial suggestion is to bring together the various previously classified families of retroelements, the retroposons and the retrons into a Class which would give the classification shown in Table 4. This Class would comprise two Orders, one (*Retrovirales*) of "viruses" potentially capable of horizontal transmission, and the other (*Retrales*) of the non-viral elements. The *Retrovirales* comprises two suborders, the *Orthoretrovirinae* which contains the retroviruses (with encapsidated RNA genomes and which involve integration in their replication) and the *Pararetrovirinae* which contains the pararetroviruses (with encapsidated DNA genomes and which replicate episomally). The lower levels of classification in this Order (family and genus) are those already accepted. The *Retrales* comprises three Suborders, the *Retrotransposinae* containing the retrotransposons which already have a classification to lower levels and the previously unclassified retroposons (*Retroposoninae*) and retrons (*Retroninae*). The latter two Suborders still require a classification structure at the lower levels.

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 has been demonstrated for several retrotransposons (e.g. yeast *Ty* elements [14]; *Drosophila copia* and *gypsy* elements [5, 10, 17]). Very recently, *env*-like coding regions have been recognised in two retrotransposons from plants [11, 21] although no particles or horizontal spread have yet been demonstrated. The potential of an env gene in these two retrotransposons poses some interesting questions as the current dogma 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 for coding an *env* gene. This raises the question as to whether retroelements which can spread horizontally, and thus are in reality viruses, should be grouped separately from retroelements which have no potential for horizontal spread.

2. Many of the retrotransposons and retroposons of plants have mutations or deletions which are considered to render them incapable of replication. Although possibly some may 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 retrotransposon reverse transcriptase amino acid sequences reveal two distinct groupings: the *Ty1/copia* group and the *Ty3/gypsy* group (see [2, 23]). The animal enveloped retroviruses are more closely related to the *Ty3/gypsy* group as is the element from Arabidopsis thaliana that contains the env gene [21]. The other recently

Table 4. Proposed classification of reverse transcribing elements

Class: *Retroelementopsida*	
Order: *Retrovirales*	
Suborder: *Orthoretrovirinae*	
Family: *Retroviridae*	
Genera	
Alpharetrovirus	*Avian leucosis 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	*Human spumavirus*
Suborder: *Pararetrovirinae*	
Family: *Hepadnaviridae*	
Genera	
Orthohepadnavirus	*Hepatitis B virus*
Avihepadnavirus	*Duck hepatitis B virus*
Family: *Caulimoviridae*	
Genera	
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*
Order: *Retrales*	
Suborder: *Retrotransposinae*	
Family: *Pseudoviridae*	
Genera	
Pseudovirus	*Saccharomyces cereviciae Ty-1 virus*
Hemivirus	*Drosophila melanogaster copia virus*
Family: *Metaviridae*	
Genera	
Metavirus	*Saccharomyces cereviciae Ty-3 virus*
Errantivirus	*Drosophila melanogaster gypsy virus*
Suborder: *Retroposinae*	
Suborder: *Retroninae*	

described plant retrotransposon with an env gene, SIRE-1 [11], 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 *Ty3/gypsy* elements and retroviruses have a range of primer tRNAs. Thus, should these different groupings be recognised in a classification?

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, Covey SN (1996) Retroelements: propagation and adaptation. Virus Genes 11: 105–118
9. Hull R, Will H (1989) Molecular biology of viral and nonviral retroelements. Trends Genet 5: 357–359
10. Kim A, Terzian C, Santamaria P, Pelisson A, Purdíhomme 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
11. 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
12. 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
13. McClure MA (1991) Evolution of retroposons by acquisition or deletion of retrovirus-like genes. Mol Biol Evol 8: 835–856
14. 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
15. 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
16. Pringle CR (1998) Virus taxonomy – San Diego 1998. Arch Virol 143: 1449–1459
17. 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
18. Temin HM (1989) Retrons in bacteria. Nature 339: 254–255
19. Temin HM, Mizutani S (1970) RNA-directed DNA polymerase in virions of Rous sarcoma virus. Nature 226: 1211–1213
20. Toh H, Ono M, Saigo K, Miyata T (1985) Retroviral protease-like sequence in the yeast retrotransposon Ty1. Nature 351: 691
21. 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
22. Xiong Y, Eickbush TH (1988) Similarity of reverse transcriptase-like sequences of viruses, transposable elements and mitochondrial introns Mol Biol Evol 5: 675–690
23. 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.

Editor's Note

Professor Hull is intending to convene an *ad hoc* meeting at the International Congress of Virology in Sydney in August to discuss his proposals for the introduction of higher taxa to unify into a common structure the existing ICTV-ratified classifications of reverse tran-

scribing agents. At present these classifications extend only to the family level [*vide* Pringle CR (1999) Virus taxonomy 1999. Arch Virol 144/2, in preparation]. Anyone interested in participating in this meeting should contact Professor Hull. Subsequently the recommendations of this meeting will be considered by the relevant Sub-committees and the Executive Committee of the ICTV and further action taken if considered appropriate.

In the foregoing article Professor Hull has employed the ending "… *virinae*" to designate the taxonomic category "suborder". This conflicts with the current version of the International Code of Virus Classification and Nomenclature [*vide* Mayo MA, Horzinek MC (1998) A revised version of the International Code of Virus Classification and Nomenclature. Arch Virol 143: 1645–1654]. In virus taxonomy the ending "… *virinae*" has already been assigned to the taxon "subfamily". If Professor Hull's proposals are accepted, it will be necessary to assign a different ending to the taxon suborder. Since many of the conventions of virus taxonomy are derived from zoological nomenclature, the ending "… *oidea*" could have been substituted directly in this article, but for its similarity to the ending "… *viroidea*", already adopted for viroid families. A decision on the nomenclature of suborders will be the responsibility of the ICTV, if Professor Hull's proposals are adopted.

C. P.

Verleger: Springer-Verlag KG, Sachsenplatz 4–6, A-1201 Wien. – Herausgeber: Dr. F. A. Murphy, School of Veterinary Medicine, University of California Davis, Davis, CA 95616, U.S.A. – Redaktion: Sachsenplatz 4–6, A-1201 Wien. – Satz und Umbruch: Thomson Press (India) Ltd., New Delhi. – Offsetdruck: Adolf Holzhausens Nachfolger, Kandlgasse 19–21, A-1070 Wien. – Herstellungsort: Wien. – Printed in Austria.