

HORIZONTAL TRANSFER OF NON-LTR RETROTRANSPOSONS

O.S. Novikova ¹, V. Fet ², A.G. Blinov ¹

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia, e-mail: novikova@bionet.nsc.ru;

² Marshall University, Huntington, West Virginia, USA

Horizontal transfer is a complex phenomenon usually used as explanation of phylogenetic inconsistency which can not be interpreted in terms of vertical evolution. Most examples of horizontal transfer of eukaryotic genes involve transposable elements. An intriguing feature of horizontal transfer is that its frequency differs among transposable elements' classes. Although horizontal transfer is well known for DNA transposons and LTR retrotransposons, majority of previously described cases of horizontal transfer for non-LTR retrotransposons were dismissed based on the analysis of non-LTR retrotransposon sequence change rates. Nevertheless, recent findings suggest that non-LTR retrotransposons can be horizontally transmitted. Thus, one cannot completely exclude the possibility of horizontal transfer for this group of transposable elements. We review the occurrence of horizontal transfer of non-LTR retrotransposons and examine the criteria used to infer such transfers.

Key words: non-LTR retrotransposons, horizontal transfer, evolution, Fungi, Maculinea

Introduction

Eukaryotic transposable elements (TEs) have the ability to replicate and spread in genome. They have been found virtually in all investigated eukaryotes and represent the ubiquitous components of eukaryotic genomes. For example, at least 46 % of the human genome is represented by TE-derived sequences (Bannert, Kurth, 2004) and more than 90 % of the plants' genomes can be composed by TEs (Vitte, Panaud, 2005). The relative abundance and diversity of TEs have contributed to the differences in the structure and size of eukaryotic genomes, especially in plant kingdom (Vitte, Panaud, 2005). Recent evidence suggests that TEs may provide the genome with potent agents to generate genetic and genomic plasticity (Kidwell, Lisch, 1997). TEs may have reshaped the human genome by ectopic rearrangements, by creating new genes, and by modifying and shuffling existing genes (Lander *et al.*, 2001). In some cases, TEs perform critical biological functions in their host (Kidwell, Lisch, 1997). For example, the preferential insertion of some retrotransposons in *Drosophila* at telomeric locations has removed the need for a telomerase function (Pardue *et al.*, 2005).

Two major classes of TEs are recognized (Wicker *et al.*, 2007). Class I elements use RNA-mediated mechanisms for their transposition and are called retrotransposons. Class II elements transpose through DNA-mediated mechanisms and are called DNA transposons (Feschotte, Pritham, 2007). Eukaryotic DNA transposons can be divided into three major subclasses. Elements from the first, the best-known subclass, are the classic "cut-and-paste" transposons that excise as double-stranded DNA and reinsert elsewhere in the genome. Elements from the second subclass, *Helitrons*, utilize a mechanism probably related to rolling-circle replication (Kapitonov, Jurka, 2007). For the last subclass, *Mavericks*, mechanism of transposition is not yet well understood, but that likely replicate using a self-encoded DNA polymerase (Pritham *et al.*, 2007). Both *Helitrons* and *Mavericks* most likely rely on distinct transposition mechanisms involving displacement and replication of a single-stranded DNA intermediate, respectively.

Retrotransposons transpose by the process involving transcription, reverse transcription, and integration of cDNA. Five orders of retrotransposons are recognized: LTR retrotransposons, which have

long terminal repeats (LTRs); non-LTR retrotransposons, which lack LTRs; *DIRS* retro-transposons; *Penelope*-like retrotransposons; and short interspersed nuclear elements (SINEs) (Wicker *et al.*, 2007). Non-LTR retrotransposons contain 5' and

3'-untranslated regions (UTR), two or one open reading frame (ORF), preceded by the promoter for RNA polymerase II, and are usually terminated by poly(A) tail (Fig. 1, a). All intact non-LTR retrotransposons encode a reverse transcriptase (RT),

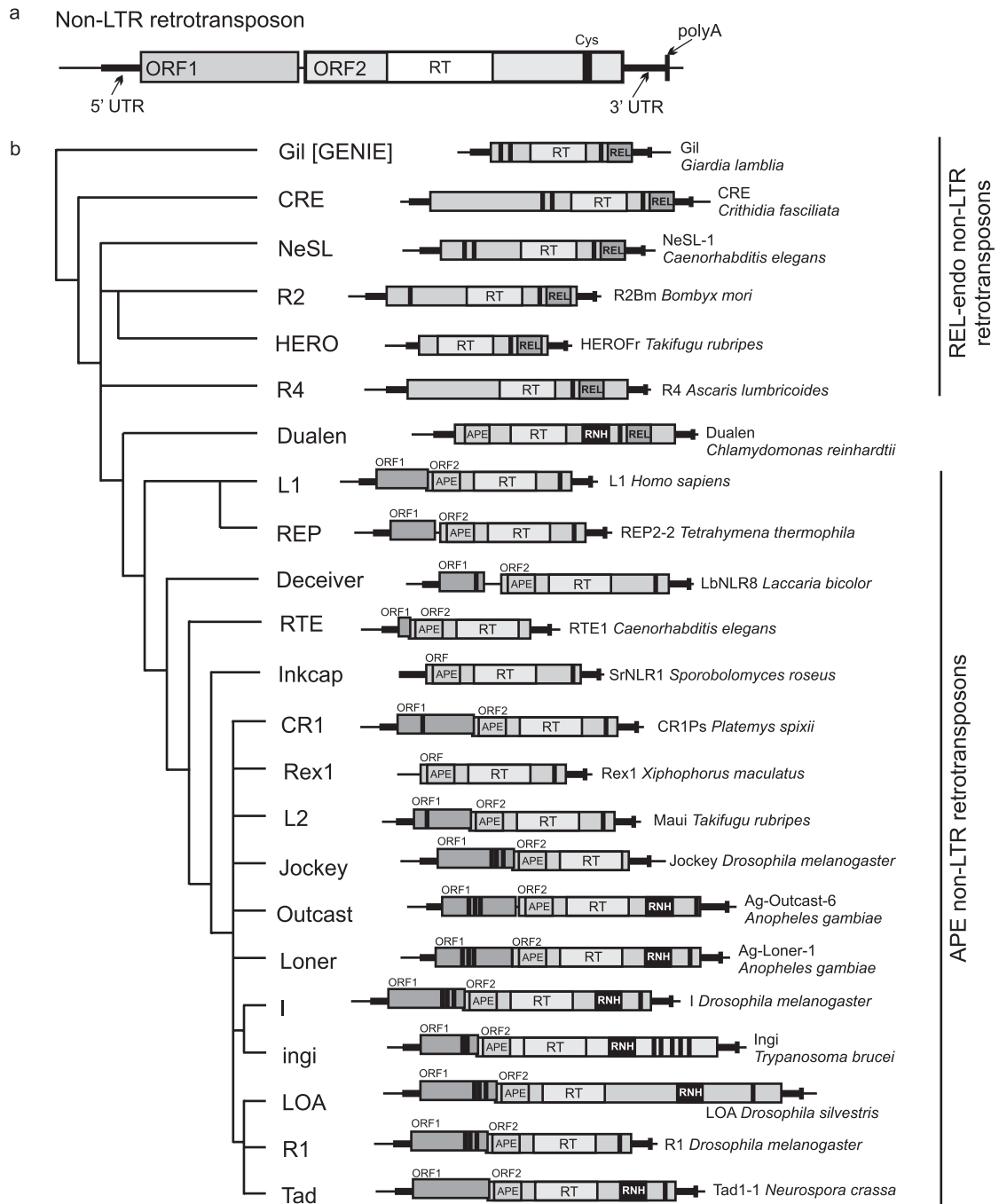


Fig. 1. Structure (a) and evolution of non-LTR retrotransposons (b).

Abbreviations: 5'UTR and 3'UTR – 5' and 3' untranslated regions; ORF – open reading frame; RT – reverse transcriptase; REL – restriction-enzyme-like endonuclease; APE – apurinic/apyrimidinic endonuclease; RNH – ribonuclease H.

which provides for their reverse transcription. Additionally, their ORFs encode a restriction-enzyme-like apurinic/apyrimidinic endonuclease, as well as a ribonuclease H (Malik *et al.*, 1999; Wicker *et al.*, 2007; Novikova *et al.*, 2008).

Malik *et al.* (1999) first used the term «clade» for groups of non-LTR retrotransposons, which have high phylogenetic support, share the same structural features, and are estimated to date back to the Precambrian time (older than ~570 Myr). Based on the phylogenetic analysis of RT domains, 23 diverse clades of non-LTR retrotransposons have been described (e.g. Malik *et al.*, 1999; Volff *et al.*, 2000; Biedler, Tu, 2003; Novikova *et al.*, 2008) (Fig. 1, b). The most ancestral clades of non-LTR retrotransposons (GENIE, CRE, R2, NeSL-1, and R4) contain only one ORF and show site-specific distribution in the genomes (Malik *et al.*, 1999; Malik, Eickbush, 2000). They have a restriction-enzyme-like endonuclease (REL-endo) domain. During further evolution of mobile elements, the REL-endo domain is suggested to have been substituted with an apurinic/apyrimidinic (APE) endonuclease acquired from host cells. All younger clades (L1, RTE, Tad, R1, LOA, I, Jockey, CR1, Rex1, and L2) possess the APE endonuclease domain and are called APE retrotransposons (Zingler *et al.*, 2005). The acquisition of the APE endonuclease resulted in a loss of target site specificity for all the elements (except R1 clade and some elements from L1 clade), and coincided with the origin of a second ORF in front of the RT-encoding ORF. Our knowledge of non-LTR retrotransposon diversity is likely to increase further: almost all studies of the entire genomic sequences detect new phylogenetic groups, especially those which have a limited distribution and/or are represented by a few copies per genome (e.g. Biedler, Tu, 2003; Novikova *et al.*, 2008).

Horizontal transfer of TEs

Horizontal transfer (HT) can be defined as the process by which genes can move between reproductively isolated species. It is not surprising that most examples of horizontal transfer of eukaryotic genes involve TEs (Kidwell, 1992; Hartl *et al.*, 1997). Three criteria could be used for HT event recognition. The first criterion is the inconsistency between the phylogenies of TEs

and host species. Traditionally, horizontal transfer was implied when highly similar transposable elements had been found in distantly related taxa accompanied by their discontinuous distribution, and such phenomenon could not be explained in terms of vertical inheritance (Kidwell, 1992; Hartl *et al.*, 1997). There are potential problems with application of this criterion for HT detection. Multiple TEs lineages can be present within genomes. Moreover, TEs are multicopy components of genomes. Comparisons of paralogous TE copies instead of orthologs along with varying rates of their sequence evolution are the main sources for incongruence in phylogenetic analysis, which could be misidentified as HT.

The second criterion, which seems to offer the strongest evidence, is a higher degree of observed sequence similarity for TEs than for functional genes, so-called 'slowdown effect on evolutionary rates'. Once inserted, a new copy of transposable element is presumed to evolve without functional constraints. Thus, all types of mutations should have an equal chance to be fixed. This means that the rates of synonymous and nonsynonymous nucleotide substitutions are identical, and coding regions in the sequence are disrupted by stop codons and indels (Volff *et al.*, 2000; Bensasson *et al.*, 2001). The lower than expected sequence divergence of TEs in comparison with non-mobile nuclear genes of the host species can be explained either by strong selective constraints in TE sequence coupled with a strict vertical transmission, or by horizontal transfer (Kordiš, Gubenšek, 1998; Novikova *et al.*, 2007; Novikova *et al.*, 2008). It is known that the insertion of a transposable element can alter gene expression and be selectively advantageous. However, only some of such transposable elements evolve under selective pressure (Ono *et al.*, 2001).

The third criterion of inferring HT is the discontinuous distribution of TEs among closely related taxa, i.e., presence of a TE in one lineage and its absence in a sister lineage. Such discontinuous distribution could be due to random loss of TEs, ancestral polymorphism, and independent sorting of copies into descendant species. By itself, this kind of evidence provides only weak support for HT since TE can be lost through population dynamics or ecological forces that are difficult to reconstruct (Kaplan *et al.*, 1985; Lohe *et al.*, 1995).

Horizontal transfer of non-LTR retrotransposons

While HT is well known for DNA transposons and LTR retrotransposons (Robertson, 1993; Silva, Kidwell, 2004), non-LTR retrotransposons rarely undergo HT, and their phylogenies are largely congruent to those of their hosts (Malik *et al.*, 1999). Moreover, the analysis of TEs sequence change rates led Malik *et al.* (1999) to dismiss majority of previously described HT cases. It has been suggested that non-LTR retrotransposons are rarely, if at all, horizontally transferred and their mode of transposition might be sufficient to explain this rarity since the RNA intermediate of non-LTR retrotransposons is reverse-transcribed directly into the chromosomal target site (Malik *et al.*, 1999). Nevertheless, there are a few cases of putative HT for non-LTR retrotransposons (Table 1).

One of the most reliable cases of HT for non-LTR retrotransposons was described for *CR1B* elements from CR1 clade (Novikova *et al.*, 2007). Initially, TE sequences with extremely high similarity were identified in genomes of silkworm *Bombyx mori* (family Bombycidae) and *Maculinea* large blue butterflies (family Lycaenidae). Further comparative and phylogenetic analysis provided more evidence for possible HT. Analysis of closely related species showed that *CR1B* element presented only in *Maculinea* butterflies and in both studied Bombycidae moths (*Bombyx mori* and *Oberthueria*

caeca) (Fig. 2, a). All three criteria were satisfied including unexpectedly high similarity between elements, phylogenetic inconsistency, and discontinuous distribution.

Along with three listed criteria, the 'divergence-versus-age' analysis can be used for HT inference (Kordiš, Gubenšek, 1998; Malik *et al.*, 1999). It includes the comparison of divergence rates between the enzymatic domains of the TEs with the host divergence time estimates. This method is widely used for non-LTR retrotransposon analysis (Kordiš, Gubenšek, 1998; Malik *et al.*, 1999; Novikova *et al.*, 2007; Novikova *et al.*, 2008). Amino acid sequence distances between the reverse transcriptase (RT) domains of various non-LTR retrotransposon clades (R1, R2, Jockey, I, CR1, RTE, Tad and L1) are plotted against estimates of host divergence time (Fig. 3). The last common ancestor of lepidopteran superfamilies Papilionoidea (*Maculinea* spp.) and Bombycoidea (*Bombyx mori*) has been estimated to exist 140 million years ago (Mya) (Gaunt, Miles, 2002). The point for *BmCR1B* (from *B. mori*) versus *MteCR1B* (from *Maculinea teleius*) fell markedly below all other comparisons. It means that the rate of evolution is much lower than expected for *BmCR1B* and *MteCR1B* elements. The slowdown effect on evolutionary rates can be explained by HT event (Kordiš, Gubenšek, 1998; Malik *et al.*, 1999; Župunski *et al.*, 2001; Novikova *et al.*, 2007; Novikova *et al.*, 2008). The HT of *CR1B* elements between lepidopteran taxa is an example of the

Table 1

Putative cases of horizontal transfer of non-LTR retrotransposons

Clade: elements	Direction of transfer	Time estimate (Mya*)	References
Jockey: diverse elements	Insects: between Drosophilidae	~ 5–12	Sánchez-Gracia <i>et al.</i> , 2005
RTE: <i>Bta/Vma</i>	Vertebrates: from an ancestral snake lineage (Boidae) to the ancestor of ruminant mammals	~ 40–50	Kordiš, Gubenšek, 1998 Župunski <i>et al.</i> , 2001
Rex1: <i>Bab/Ana & Bab/Anj</i>	Vertebrates: from Percomorpha fishes to eels (<i>Anguilla</i>)	less than 90	Volff <i>et al.</i> , 2000
CR1: <i>BmCR1B/MteCR1B</i>	Insects: from Bombycidae moths to <i>Maculinea</i> butterflies (Lycaenidae)	~ 5–10	Novikova <i>et al.</i> , 2007
Tad: <i>AorNLR2 & AniNLR2/FoNLR5 & CgNLR3</i>	Fungi: from Sordaryomycetes (<i>Chetomium globosum</i> and <i>Fusarium oxysporum</i>) to Eurotiomycetes (<i>Aspergillus oryzae</i> and <i>A. niger</i>)	less than 10	Novikova <i>et al.</i> , 2008

* Mya – million years ago.

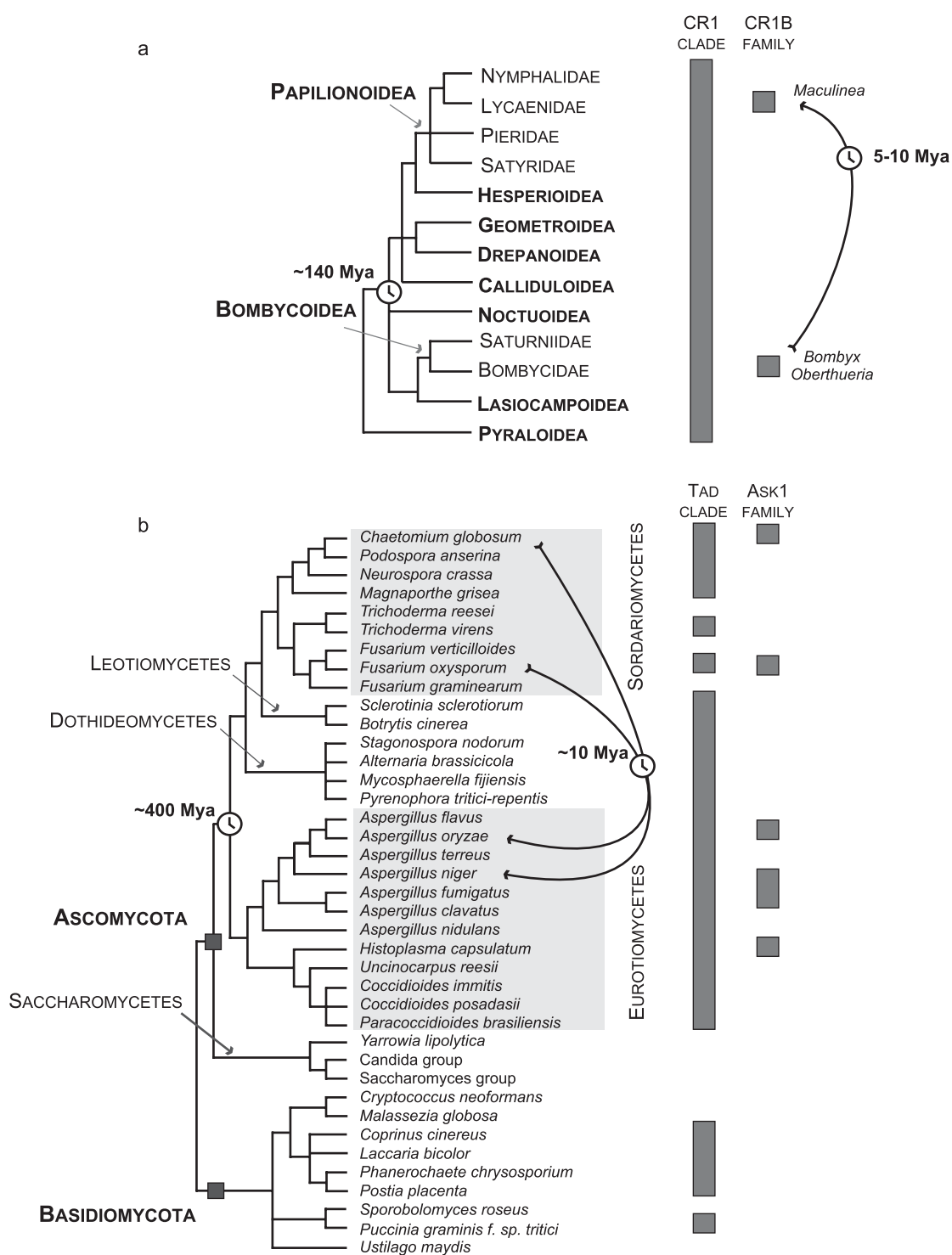


Fig. 2. Horizontal transfers of non-LTR retrotransposons HT of CR1 elements between *Maculinea* butterflies and *Bombyx* moths, dated 5–10 Mya (a), and HT of TAD elements between Sordariomycetes (*Chaetomium globosum* and *Fusarium oxysporum*) to Eurotiomycetes (*Aspergillus oryzae* and *A. niger*) probably occurred 10 Mya (b). Phylogeny of superfamilies and families of butterfly suborder Ditrysia based on Tree of Life (<http://tolweb.org/>). Evolutionary tree sequenced fungal genomes is represented according to Fitzpatrick *et al.* (2006) with minor modifications.

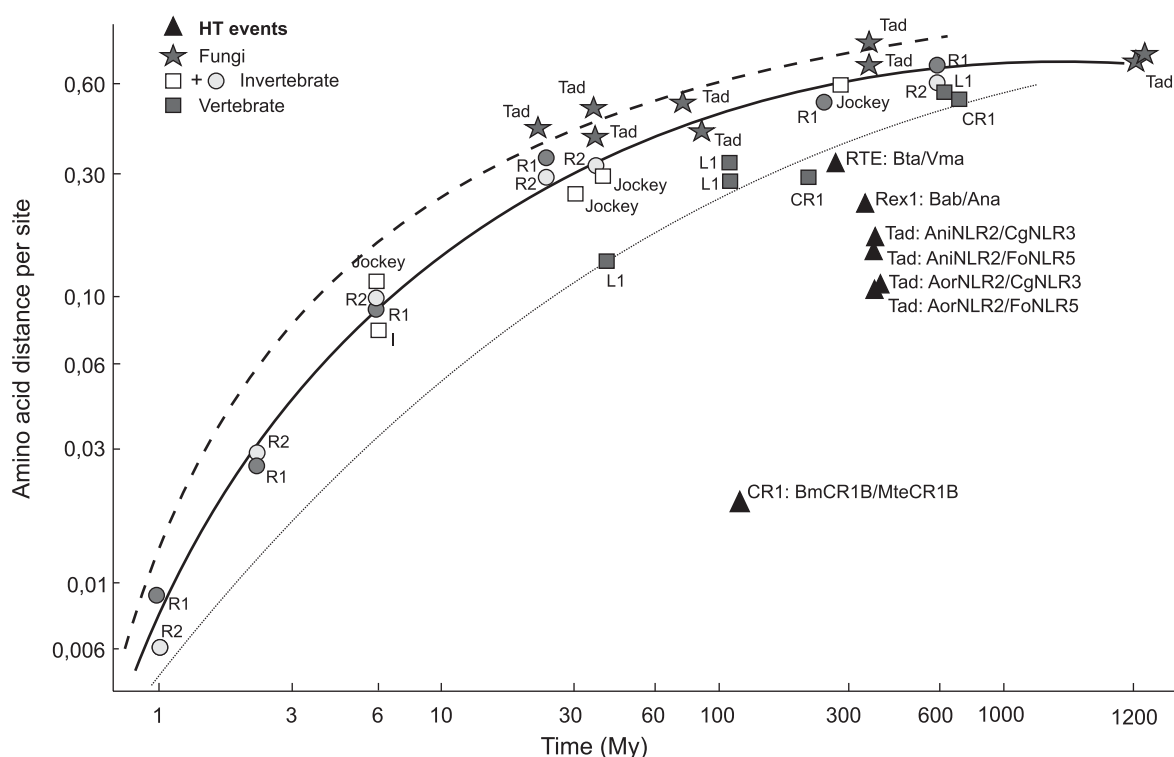


Fig. 3. Divergence-versus-age analysis of non-LTR retrotransposons.

Amino acid sequence distances were calculated from the sequences of the complete reverse transcriptase (RT) domain. The curves for arthropods and vertebrates are reproduced from Malik *et al.* (1999); the curve for fungi is reproduced based on Novikova *et al.* (2008). Putative HT events are shown.

most recently occurred transmission of non-LTR retrotransposons between reproductively isolated species that belong to the taxa separated more than 100 million years ago.

Horizontal transfer is a complex phenomenon which has only indirect evidence. Moreover, it is impossible to investigate HT directly in a laboratory; all known HT cases between eukaryotes were discovered accidentally during phylogenetic studies (e.g. Kordiš, Gubenšek, 1998; Volff *et al.*, 2000; Novikova *et al.*, 2007; Novikova *et al.*, 2008). For example, HT of *CR1B* elements was identified along with the analysis of non-LTR retrotransposons diversity in *Maculinea* genomes. Bioinformatic approaches can greatly expand opportunities to observe the distribution and evolutionary history of non-LTR retrotransposons in various eukaryotic taxa and to identify novel cases of HT. The most suitable eukaryotic group for such investigations proved to be Fungi. More than 100 fungal genomes are available for analysis in various genomic databases and the number of projects is

constantly increasing (e.g. Galagan *et al.*, 2005). Fungi have small genomes, usually with limited amounts of repetitive DNA (including non-LTR retrotransposons) that makes the genomic analysis faster and easier.

Indeed, the genomic analysis of large set of fungal genomes revealed a possible case of HT of Tad-like non-LTR retroelements between Eurotiomycetes and Sordariomycetes, which is estimated to have taken place less than 10 Mya (Fig. 2, b). The time since divergence of Eurotiomycetes and Sordariomycetes is estimated between 310 Myr and 670 Myr (Berbee, Taylor, 1993; Heckman *et al.*, 2001). The oldest well-documented ascomycete fossils are found in the 400-Myr-old Rhynie chert (Taylor *et al.*, 1999). Based on this finding, it was suggested that 400 Myr for the Eurotiomycetes and Sordariomycetes divergence would seem to provide a conservative date estimate; even earlier dates could be expected (Kasuga *et al.*, 2002). Elements from Eurotiomycetes *Aspergillus niger* and *A. oryzae* (*AniNLR2* and *AorNLR2*) shared higher

similarity with elements from Sordariomycetes *Chaetomium globosum* and *Fusarium oxysporum* (CgNLR3 and FoNLR5) than with other non-LTR retrotransposons from *Aspergillus*; this cannot be explained solely by vertical evolution (Novikova *et al.*, 2008). Divergence-versus-age analysis as well as comparative analysis of evolutionary rates for functional genes and described elements confirmed the presence of slowdown effect on evolutionary rates, which always accompanies possible HTs (Fig. 3).

The criterion of discontinuous distribution cannot be applied in this case since all investigated fungal species differ in number and diversity of non-LTR retrotransposon clades presented in their genome (Novikova *et al.*, 2008). There are several main processes which could affect the copy number and diversity of non-LTR retrotransposons in fungal genomes: stochastic loss of non-LTR retrotransposons; burst of retrotransposition; the limitation of copy number increase by natural selection, which removes deleterious insertions; horizontal transfer; passive and active inactivation of repetitive sequences; self-regulation of transposition (decrease of transposition rate when the copy number increases) (e.g. Hua-Van *et al.*, 2005; Le Rouzic, Capy, 2005). Complex interactions between these processes lead to the formation of a unique repertoire of non-LTR retrotransposons in each fungal species. Although the criterion of “patchy” distribution does not work for fungi, the two strongest criteria were satisfied; therefore, the demonstrated inconsistency can be due to a possible HT.

Conclusions

Horizontal transfer (HT) of non-LTR retrotransposons appears to be not just an artifact of analysis as suggested earlier (Malik *et al.*, 1999). Non-LTR retrotransposons, in fact, can be horizontally transmitted, although it is a very rare event in comparison to other types of transposable elements. The actual mechanisms of horizontal transfer are still unknown for eukaryotic TEs since it is not possible to show experimentally how the HT can occur. Parasites, symbionts, bacteria, or viruses all could be suggested as potential vectors for horizontal transfer.

References

- Bannert N., Kurth R. Retroelements and the human genome: new perspectives on an old relation // Proc. Natl Acad. Sci. USA. 2004. V. 101. (Suppl. 2). P. 14572–14579.
- Bensasson D., Petrov D.A., Zhang D.X. *et al.* Genomic gigantism: DNA loss is slow in mountain grasshoppers // Mol. Biol. Evol. 2001. V. 18. P. 246–253.
- Berbee M.L., Taylor J.W. Dating the evolutionary radiations of the true fungi // Can. J. Bot. 1993. V. 71. P. 1114–1127.
- Biedler J., Tu Z. Non-LTR retrotransposons in the African malaria mosquito, *Anopheles gambiae*: unprecedented diversity and evidence of recent activity // Mol. Biol. Evol. 2003. V. 20. P. 1811–1825.
- Feschotte C., Pritham E.J. DNA transposons and the evolution of eukaryotic genomes // Annu. Rev. Genet. 2007. V. 41. P. 331–368.
- Fitzpatrick D.A., Logue M.E., Stajich J.E., Butler G. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis // BMC Evol. Biol. 2006. V. 6. P. 99.
- Galagan J.E., Henn M.R., Ma L.J. *et al.* Genomics of the fungal kingdom: insights into eukaryotic biology // Genome Res. 2005. V. 15. P. 1620–1631.
- Gaunt M.W., Miles M.A. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks // Mol. Biol. Evol. 2002. V. 19. P. 748–761.
- Hartl D.L., Lohe A.R., Lozovskaya E.R. Modern thoughts on an ancient marinere: function, evolution, regulation // Annu. Rev. Genet. 1997. V. 31. P. 337–358.
- Heckman D.S., Geiser D.M., Eidell B.R. *et al.* Molecular evidence for the early colonization of land by fungi and plants // Science. 2001. V. 293. P. 1129–1133.
- Hua-Van A., Le Rouzic A., Maisonhaute C., Capy P. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences // Cytogenet. Genome Res. 2005. V. 110. P. 426–440.
- Kapitonov V.V., Jurka J. Helitrons on a roll: eukaryotic rolling-circle transposons // Trends Genet. 2007. V. 23. P. 521–529.
- Kaplan N., Darden T., Langley C.H. Evolution and extinction of transposable elements in Mendelian populations // Genetics. 1985. V. 109. P. 459–480.
- Kasuga T., White T.J., Taylor J.W. Estimation of nucleotide substitution rates in Eurotiomycete fungi // Mol. Biol. Evol. 2002. V. 19. P. 2318–2324.
- Kidwell M.G. Horizontal transfer // Curr. Opin. Genet. Dev. 1992. V. 2. P. 868–873.
- Kidwell M.G., Lisch D. Transposable elements as sources of variation in animals and plants // Proc. Natl Acad. Sci. USA. 1997. V. 94. P. 7704–7711.

- Kordiš D., Gubenšek F. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes // *Proc. Natl Acad. Sci. USA*. 1998. V. 95. P. 10704–10709.
- Lander E.S., Linton L.M., Birren B. *et al.* Initial sequencing and analysis of the human genome // *Nature*. 2001. V. 409. P. 860–921.
- Le Rouzic A., Capy P. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift // *Genetics*. 2005. V. 169. P. 1033–1043.
- Lohe A.R., Moriyama E.N., Lidholm D.A., Hartl D.L. Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements // *Mol. Biol. Evol.* 1995. V. 12. P. 62–72.
- Malik H.S., Burke W.D., Eickbush T.H. The age and evolution of non-LTR retrotransposable elements // *Mol. Biol. Evol.* 1999. V. 16. P. 793–805.
- Malik H.S., Eickbush T.H. NeSL-1, an ancient lineage of site-specific non-LTR retrotransposons from *Caenorhabditis elegans* // *Genetics*. 2000. V. 154. P. 193–203.
- Novikova O., Fet V., Blinov A. Non-LTR retrotransposons in fungi // *Funct. Integr. Genomics*. 2008. doi:10.1007/s10142-008-0093-8.
- Novikova O., Sliwińska E., Fet V. *et al.* CR1 clade of non-LTR retrotransposons from *Maculinea* butterflies (Lepidoptera: Lycaenidae): evidence for recent horizontal transmission // *BMC Evol. Biol.* 2007. V. 7. P. 93.
- Ono R., Kobayashi S., Wagatsuma H. *et al.* A retro-transposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21 // *Genomics*. 2001. V. 73. P. 232–237.
- Pardue M.L., Rashkova S., Casacuberta E. *et al.* Two retrotransposons maintain telomeres in *Drosophila* // *Chromosome Res.* 2005. V. 13. P. 443–453.
- Pritham E.J., Putliwala T., Feschotte C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses // *Gene*. 2007. V. 390. P. 3–17.
- Robertson H.M. The mariner transposable element is widespread in insects // *Nature*. 1993. V. 362. P. 241–245.
- Sánchez-Gracia A., Maside X., Charlesworth B. High rate of horizontal transfer of transposable elements in *Drosophila* // *Trends Genet.* 2005. V. 21. P. 200–203.
- Silva J.C., Kidwell M.G. Evolution of P elements in natural populations of *Drosophila willistoni* and *D. sturtevantii* // *Genetics*. 2004. V. 168. P. 1323–1335.
- Taylor T.N., Hass H., Kerp H. The oldest fossil ascomycetes // *Nature*. 1999. V. 399. P. 648.
- Vitte C., Panaud O. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model // *Cytogenet. Genome Res.* 2005. V. 110. P. 91–107.
- Volff J.N., Körting C., Schartl M. Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes // *Mol. Biol. Evol.* 2000. V. 17. P. 1673–1684.
- Wicker T., Sabot F., Hua-Van A. *et al.* A unified classification system for eukaryotic transposable elements // *Nat. Rev. Genet.* 2007. V. 8. P. 973–982.
- Zingler N., Weichenrieder O., Schumann G.G. APE-type non-LTR retrotransposons: determinants involved in target site recognition // *Cytogenet. Genome Res.* 2005. V. 110. P. 250–268.
- Župunski V., Gubenšek F., Kordiš D. Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons // *Mol. Biol. Evol.* 2001. V. 18. P. 1849–1863.