



# Independent and parallel lateral transfer of DNA transposons in tetrapod genomes

Peter Novick<sup>a,b</sup>, Jeremy Smith<sup>c</sup>, David Ray<sup>c</sup>, Stéphane Boissinot<sup>a,b,\*</sup>

<sup>a</sup> Department of Biology, Queens College, the City University of New York, Flushing, NY 11367, USA

<sup>b</sup> Graduate School and University Center, the City University of New York, New York, NY 10001, USA

<sup>c</sup> Department of Biology, West Virginia University, Morgantown, WV 26506, USA

## ARTICLE INFO

### Article history:

Received 22 April 2009

Received in revised form 8 July 2009

Accepted 28 August 2009

Available online 10 September 2009

Received by N. Okada

### Keywords:

Transposon

Transposase

Transposable element

Lateral transfer

Horizontal transfer

## ABSTRACT

In animals, the mode of transmission of transposable elements is generally vertical. However, recent studies have suggested that lateral transfer has occurred repeatedly in several distantly related tetrapod lineages, including mammals. Using transposons extracted from the genome of the lizard *Anolis carolinensis* as probes, we identified four novel families of hAT transposons that share extremely high similarity with elements in other genomes including several mammalian lineages (primates, chiropters, marsupials), one amphibian and one flatworm, the planarian *Schmidtea mediterranea*. The discontinuous phylogenetic distribution of these hAT families, coupled with very low synonymous divergence between species, strongly suggests that these elements were laterally transferred to these different species. This indicates that the horizontal transfer of DNA transposons in vertebrates might be more common than previously thought. Yet, it appears that the transfer of DNA transposons did not occur randomly as the same genomes have been invaded independently by different, unrelated transposon families whereas others seem to be immune to lateral transfer. This suggests that some organisms might be intrinsically more vulnerable to DNA transposon lateral transfer, possibly because of a weakened defense against transposons or because they have developed mechanisms to tolerate their impact.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

The lateral (horizontal) transfer of genetic information has had a profound impact on the evolution of unicellular organisms, in particular prokaryotes, yet the evolutionary significance of lateral gene transfer in multicellular eukaryotes remains unclear (Andersson, 2005). In eukaryotes, beside the transfer of genetic information from organelles to the nucleus, which has occurred repeatedly (Timmis et al., 2004), most cases of lateral transfer have been documented in phagotrophic protists (Andersson, 2005). In multicellular eukaryotes, lateral transfer has occurred relatively frequently in plants (Bergthorsson et al., 2003; Won and Renner, 2003) but seems exceedingly rare in fungi and animals. Most documented cases of lateral transfer in animals involve the transfer of transposable elements (TEs), usually between closely related species. The lateral transfer of TEs seems to have occurred most frequently in insects but recent studies have

demonstrated its occurrence in fish (de Boer et al., 2007) and tetrapods (Kordis and Gubensek, 1998; Pace et al., 2008).

TEs are ubiquitous in eukaryotes and have profoundly affected the size, structure and function of eukaryotic genomes. TEs have also been an important source of evolutionary novelties: exaptation of TEs within coding sequences (i.e. exonization) or as regulatory elements seems to have been relatively common (Makalowski, 2000; Nekrutenko and Li, 2001; Bejerano et al., 2006) and could have driven the evolution of some fundamental biological functions (Feschotte and Pritham, 2007; Feschotte, 2008). However, the activity of TEs can also negatively affect the fitness of their hosts (Petrov et al., 2003; Boissinot et al., 2006), either by disrupting gene function or by causing chromosomal rearrangements (Kazazian, 2004). Two main classes of TEs inhabit animal genomes: class I elements that use an RNA intermediate for their replication and class II elements, also called DNA transposons, that replicate using a cut-and-paste mechanism (for a review on transposons and their impact, see Feschotte and Pritham, 2007). DNA transposons are mobilized by a self-encoded enzyme called transposase that recognizes specific sites at the end of the element, excises the DNA and inserts it elsewhere in the genome. Although the transposition mechanism is non-replicative, DNA transposons can reach large copy number in some species due to donor site repair. As the only recognition requirements of the transposase are the terminal sequences of the elements, transposons with internal deletions can also be mobilized by the transposase encoded by complete elements. Consequently, families of autonomous DNA transposons are parasitized by a plethora of non-autonomous elements that often

**Abbreviations:** TE, Transposable element; LINE-1, Long Interspersed Nuclear Element-1; SPIN, SPace INvaders; NJ, Neighbor Joining; hAT, hobo/Activator/Tam3; ORF, Open reading Frame; AC, *Anolis carolinensis*; ML, *Myotis lucifugus*; MD, *Monodelphis domestica*; OG, *Otolemur garnettii*; SM, *Schmidtea mediterranea*; TS, *Tarsius syrichta*; MM, *Microcebus murinus*; XT, *Xenopus tropicalis*; ET, *Echinops telfairi*; PCR, Polymerase Chain Reaction; aa, amino acid; bp, base pair; MY, Million Years; dn, non-synonymous divergence; ds, synonymous divergence.

\* Corresponding author. Department of Biology, Queens College, CUNY, 65-30 Kissena Boulevard, Flushing, NY 11367, USA. Fax: +1 718 997 3445.

E-mail address: [stephane.boissinot@qc.cuny.edu](mailto:stephane.boissinot@qc.cuny.edu) (S. Boissinot).

outnumber their autonomous relatives. DNA transposons are very diverse and are represented by 15 super-families that diverged before the diversification of eukaryotic lineages (Bao et al., 2009).

The general mode of transmission of TEs in animals is vertical. Phylogenetic analyses of class I TEs, such as the LINE-1 retrotransposons, demonstrate that the evolution of these elements can be explained by a strict vertical mode of transmission (Furano, 2000; Kordis et al., 2006). The transmission of the RTE Bov-B elements from squamate reptiles to bovids constitutes at current time the only convincing exception to this model (Kordis and Gubensek, 1998). The evolution of DNA transposons is also typified by the vertical transmission of elements. However, lateral transfer of DNA transposons can occur, though rarely, and most often between closely related species. It has been unequivocally demonstrated in insects (Daniels et al., 1990; Robertson and Lampe, 1995) and fish (de Boer et al., 2007). Recently, Pace et al. (2008) described the first case of lateral transfer in mammals. The elements involved, named *SPIN*, were found in several distantly related lineages (tenrec, bush baby, murine rodents, bat, opossum, lizard and frog), yet they were absent from 27 other vertebrate genomes. The patchy phylogenetic distribution of these elements, coupled with their high sequence similarity, convincingly demonstrated that *SPIN* elements have been transferred laterally to these lineages, from a still unknown source.

During an analysis of repetitive sequences in the genome of the lizard *Anolis carolinensis*, we identified four novel families of DNA transposons that show the hallmark of lateral transfer. Surprisingly, these lateral transfers did not occur randomly in the tree of life as the same species that were hospitable to *SPIN* elements have been colonized by other distantly related families of elements. The parallel and independent invasion of the same genomes by different DNA transposons suggests that some genomes might be intrinsically more hospitable to DNA transposons (or horizontal transfer in general) than others, possibly because they lack defense against these elements or because they developed mechanisms to tolerate their impact.

## 2. Materials and methods

The genome of *A. carolinensis* was searched for the presence of DNA transposons using the RepeatMasker ([www.repeatmasker.org](http://www.repeatmasker.org)) and PILER programs (Edgar and Myers, 2005). Additionally, we performed a BLASTX search of the Anole genome using amino acid sequences derived from the transposon library available from Repbase (v13.01). The resulting hits of at least 100 bp, with a maximum score of  $2e^{-150}$ , were extracted along with 5000 bp of flanking sequence to identify precisely the termini of each element. Elements that were recovered were aligned to each other and classified into families, and consensus sequences were constructed for each family of autonomous and non-autonomous transposons. Each consensus sequence was then screened for target site duplications, terminal inverted repeats and the presence of a transposase domain for proper classification. Sequences were manipulated and aligned using the BioEdit program (Hall, 1999).

To identify related elements in other species, nucleotide consensus sequences and their translated proteins were used as queries in a BLAST and BLASTX search of GenBank and in a BLAT search of the genomes curated on the UCSC Web page (<http://genome.ucsc.edu>). Significant hits (>80%) were collected and aligned, and consensus sequences were created and compared among species. Silent (*ds*) and replacement (*dn*) divergences between consensus sequences, as well as their ratio, were calculated using the maximum likelihood approach developed by Yang and Nielsen (2000), implemented in PAML (Yang, 2000). The mean divergence between elements and its standard deviation was calculated for each family as an estimator of its age, using Kimura 2-parameters method (Kimura, 1980). Phylogenetic analyses were performed with the neighbor-joining method (NJ) (Saitou and Nei, 1987) and maximum likelihood method using PHYML (Guindon et al., 2005). Distances and NJ trees were calculated using the MEGA 4.0 program with 1000 bootstrap replicates (bootstrap values lower than 75 are not shown) (Kumar et al., 2001). Copy numbers were estimated using BLAST or BLAT and only

**Table 1**  
Characteristics of three horizontally transferred autonomous *hAT* families and their non-autonomous relatives.

Family	Organism	TE name	Length, bp	Copy no. <sup>a</sup>	Average divergence ± S.E.	TIR
<i>hAT-1</i>	<i>Myotis lucifugus</i>	<i>hAT-HT1_ML</i>	2920	38	2.68 ± 0.17	CAGTGATGGCGAACCT
		<i>hAT-HT1N1_ML</i>	235	>5000	1.44 ± 0.51	CAGTGATGGCGAACCT
	<i>Monodelphis domestica</i>	<i>hAT-HT1_MD</i>	3002	100	8.68 ± 0.50	CYNTGATGGNNAANC
		<i>hAT-HT1N1_MD</i>	1010	>500	7.29 ± 0.42 <sup>c</sup>	CMRTGATGGSSAACCT
	<i>Anolis carolinensis</i>	<i>hAT-HT1_AC</i>	2968	6	5.82 ± 0.40	CAGTGATGGSSAACCT
		<i>hAT-HT1N1_AC</i>	585	868	4.75 ± 0.64	CARTGATGGSCAACCT
<i>hAT-2</i>	<i>Tarsius syrichta</i>	<i>hAT-HT2_TS</i>	3381	5	4.56 ± 0.32	CAGGGGTCTCTCAAACT
		<i>hAT-HT2N1_TS</i>	674	50	4.50 ± 0.55	CAGGGGTCTCTCAAACT
		<i>hAT-HT2N2_TS</i>	506	65	4.39 ± 0.40	CAGGGGTCTCTCAAACT
		<i>hAT-HT2N3_TS</i>	626	65	4.33 ± 0.36	CAGGGGTCTCTCAAACT
		<i>hAT-HT2_OG</i>	3277	23	7.81 ± 0.43	CAGGGGTCTCTCAAACT
		<i>hAT-HT2_MM</i>	2169	<10	7.23 ± 0.58	CAGGGGTCTCTCAAACT
	<i>Otolemur garnettii</i>	<i>hAT-HT2N1_MM</i>	205	>1000	5.71 ± 1.27	CAGSGTCTCTCAAACT
		<i>hAT-HT2_ML</i>	NA	<5	3.51 ± 0.72	CAGGGGTCCCCAAAC <sup>b</sup>
	<i>Microcebus murinus</i>	<i>hAT-HT2N1_ML</i>	746	27	2.16 ± 0.33	CAGGGGTCTCTCAAACT
		<i>hAT-HT2N2_ML</i>	205	>1000	1.48 ± 0.57	CAGGGGTCTCTCAAACT
	<i>Echinops telfairi</i>	<i>hAT-HT2_ET</i>	3167	16	7.55 ± 0.40	CAGGGGTCTCTCAAACT
		<i>hAT-HT2N1_ET</i>	225	>1000	4.38 ± 0.81	CAGGGGTCTCTCAAACT
	<i>M. domestica</i>	<i>hAT-HT2_MD</i>	3128	267	9.11 ± 0.37	CAGGAGTCCCCAAAC
	<i>A. carolinensis</i>	<i>hAT-HT2_AC</i>	2246	5	0.14 ± 0.07	CAGGGGTCCCCAAAC
		<i>hAT-HT2N1_AC</i>	1485	28	1.34 ± 0.31	CAGGGGTCCCCAAAC
	<i>Xenopus tropicalis</i>	<i>hAT-HT2_XT</i>	NA	<5	NA	YAGGARTCTCTCAAACT <sup>b</sup>
	<i>Schmidtea mediterranea</i>	<i>hAT-HT2_SM</i>	NA	<5	11.74 ± 1.41	NA
<i>hAT-3</i>	<i>M. lucifugus</i>	<i>hAT-HT3N1_ML</i>	326	537	0.24 ± 0.16	CAGTGRITCCCCAAA
		<i>hAT-HT3_AC</i>	2754	<25	5.62 ± 0.29	CAGTGRITCCCCAAA
	<i>A. carolinensis</i>	<i>hAT-HT3N1_AC</i>	326	344	3.98 ± 0.37	CAGTGRITCCCCAAA
		<i>hAT-HT3_SM</i>	NA	>20	12.23 ± 1.55	CAGTGTITCCCCAAA

<sup>a</sup> Due to the fact that many of these genomes are not yet fully sequenced, copy number depicts only elements already sequenced and collected from NCBI with >90% sequence identity to the consensus.

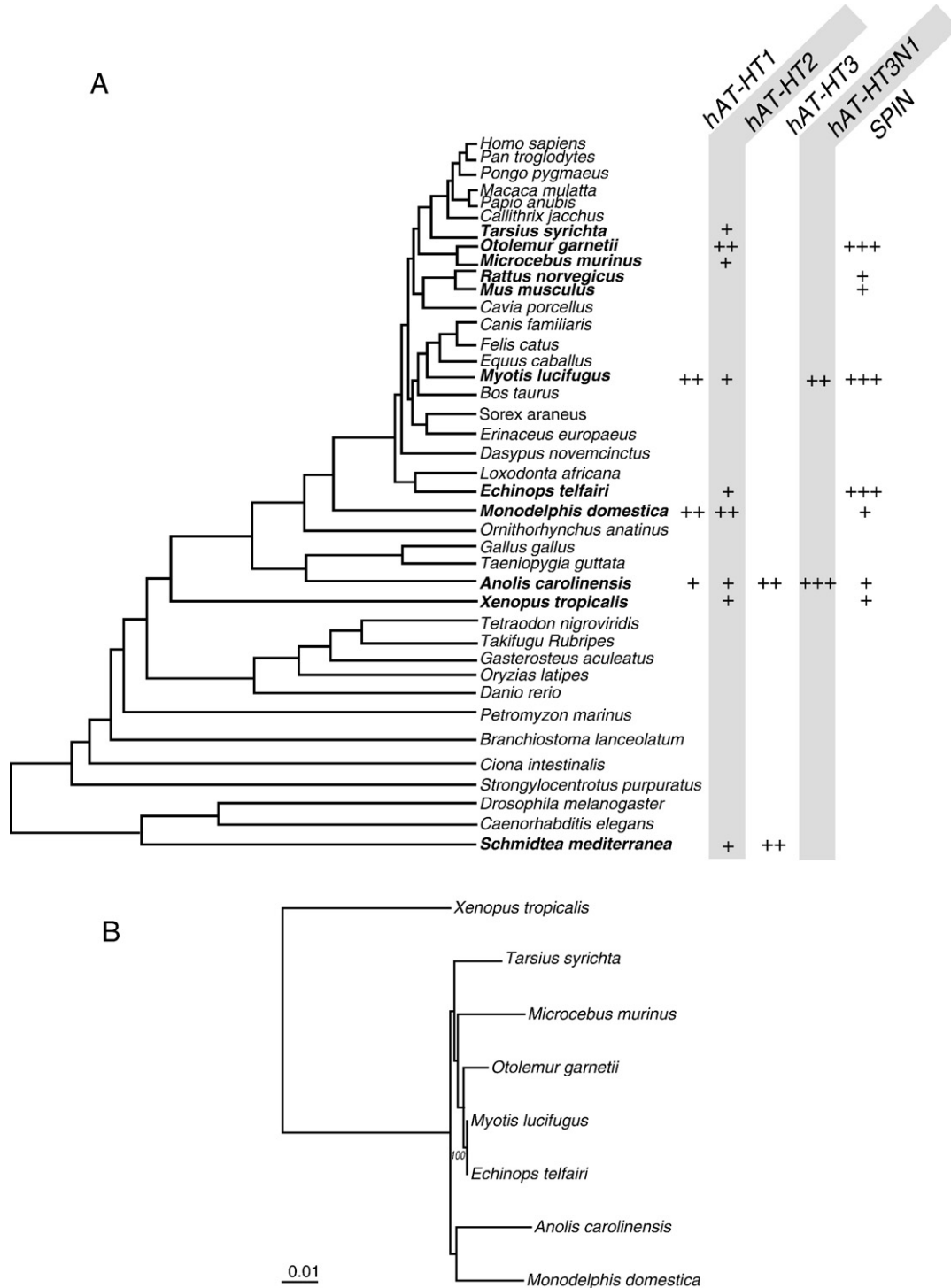
<sup>b</sup> TIR is categorized by only the 3' end of the element.

<sup>c</sup> 5' end removed from divergence calculations due to variation and subfamily structure.

elements containing >85% of the consensus query were tallied as elements representative of that family. Because some of the genome sequences are not complete, the copy number presented here are only rough estimates.

We confirmed the presence of transposon families by polymerase chain reaction (PCR). PCRs were conducted in organisms that harbor the putatively transferred transposons families and in organisms for

which we did not expect to find amplification. Twenty microliters of PCR reactions was conducted with an initial denaturing step of 94° for 5 min, followed by 30 cycles of 94°, 52° and 72° for 30 s each and ended with a 7-min extension at 72°. Primers were constructed to amplify elements in all taxa in which they are found by utilizing the most conserved regions in the ORF while minimizing ambiguous nucleotides. The primers used are as follows: hAT-HT1, forward: 5'



**Fig. 1.** (A) Distribution of five laterally transferred hAT DNA transposon families in animal species for which sufficient genome sequence is available. Species in bold contain at least one of the five families of horizontally transferred elements. The abundance of each family is symbolized as follows: + less than 10 elements, ++ more than 10 and +++ more than 1000 elements. Branch lengths are not indicative of time or genomic divergence. (B) Maximum likelihood tree based on the consensus sequence of the transposase domain of hAT-2 elements in eight of the nine species that genomic copies were found (elements in *Schmidtea mediterranea* are too fragmented for this analysis). The model of substitution (HKY+G) was determined using Modeltest (Posada and Crandall, 1998) and the tree was built under this model using PHYML with 1000 bootstrap replicates (Guindon et al., 2005). Bootstrap values less than 75 have been removed.

CTCTACAAATTTAACATCWG, reverse: 5'AATCAACYTCCWGTAAAAGT; hAT-HT2, forward: 5'TTCCTATTTTCCTTGGCAC, reverse: 5'AAGCTCTTTAAATCTVADTTG; and hAT-HT3, forward: 5'AGTTCCTCCAAATTTATCAGGAG, reverse: 5'GCAATTAAACAGAAAGATAT.

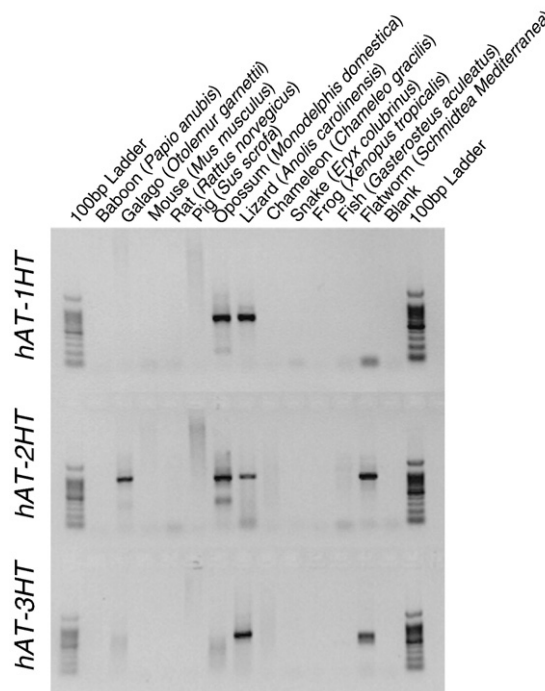
### 3. Results

The screening of the green anole genome (*A. carolinensis*) revealed the presence of 11 autonomous and 67 non-autonomous families of DNA transposons (not shown). Four of these 80 families yielded highly significant (>90%) hits in other species using the BLAST or the BLAT search engines.

#### 3.1. Four shared families of DNA transposons

We found five copies of a *hAT* (*hobo*/*Activator*/*Tam3*) element in the green anole, which are virtually identical to each other and have intact open-reading frames (ORF). This element is 2246 bp long and encodes a 602 amino acid (a.a.) long transposase (Table 1). Highly similar elements (i.e. less than 5% divergence at the DNA level) were found in eight vertebrate genomes (Fig. 1), including three primates (two strepsirrhini, *Otolemur garnettii* and *Microcebus murinus*, and the tarsier, *Tarsius syrichta*), the tenrec (*Echinops telfairi*), the little brown bat (*Myotis lucifugus*), the opossum (*Monodelphis domestica*) and the African clawed frog (*Xenopus tropicalis*). In each of these species, except the bat, we found full-length copies of the element. We named this family *hAT-HT2* to reflect its horizontal mode of transmission. Note that this family includes the previously described opossum *hAT2\_MD* family but not the *hAT2* family of bats and frog because *hAT2* families in different organisms do not form a monophyletic group. We also found fragments highly similar to *hAT-HT2* in the planarian *Schmidtea mediterranea* indicating that this transposon is not limited to vertebrate genomes. *hAT-HT2* is apparently absent from all other vertebrate and invertebrate species for which genomic data are available (>50 species). The discontinuous phylogenetic distribution of *hAT-HT2* was confirmed by PCR (Fig. 2). As expected, we amplified this family in lemur, opossum, anole and planarian but not in any other species tested. Surprisingly, we failed to amplify *hAT-HT2* in the frog, possibly because the *hAT-HT2* copies in this species are too fragmented to yield an amplification.

Full-length or partial consensus sequences were constructed for each species (except for *S. mediterranea* because most elements were fractionated into small pieces) and compared to each other. The synonymous divergence (*ds*) between transposase genes, that solely reflects the neutral history of the sequence, is relatively low, <7% for all inter-specific comparisons (Table 2), considering how distantly related these species are. For instance, the *ds* between *Anolis* and the little brown bat is about 5%, although reptiles and mammals diverged about 300 MY ago (Donoghue and Benton, 2007). Similarly among mammals, the *ds* between the opossum and the other mammals is remarkably low (<5%) considering that marsupials and eutherians split 120 MY ago (Donoghue and Benton, 2007). Such a low level of silent divergence between distantly related lineages is not consistent with a vertical mode of transmission of the *hAT-HT2* transposons and, instead, strongly suggests it has been laterally transferred. A phylogenetic analysis based on *hAT-HT2* consensus sequences shows a star-like phylogeny, with extremely short internal branches (Fig. 1B), also supporting the lateral transfer of *hAT-HT2*. The tree also suggests that the lateral transfer likely originated from a single common source. The length of the terminal branches indicates that *hAT-HT2* elements have resided in each genome long enough to accumulate nucleotide substitutions since they were laterally transferred. Interestingly, transposases differ at many non-synonymous positions, which is surprising considering that the integrity of the transposase is critical for the mobilization of DNA transposons. We



**Fig. 2.** PCR amplification of three horizontally transferred *hAT* families. Using the same primers, *hAT-1HT* amplifies in the lizard and the opossum, *hAT-2HT* in the lemur, opossum, lizard and flatworm and *hAT-3HT* in only the lizard and the flatworm. As *hAT-2HT* elements in frog are extremely fragmented, we were unable to amplify this family. In all other organisms screened, we repeatedly failed to amplify any of the three *hAT* families.

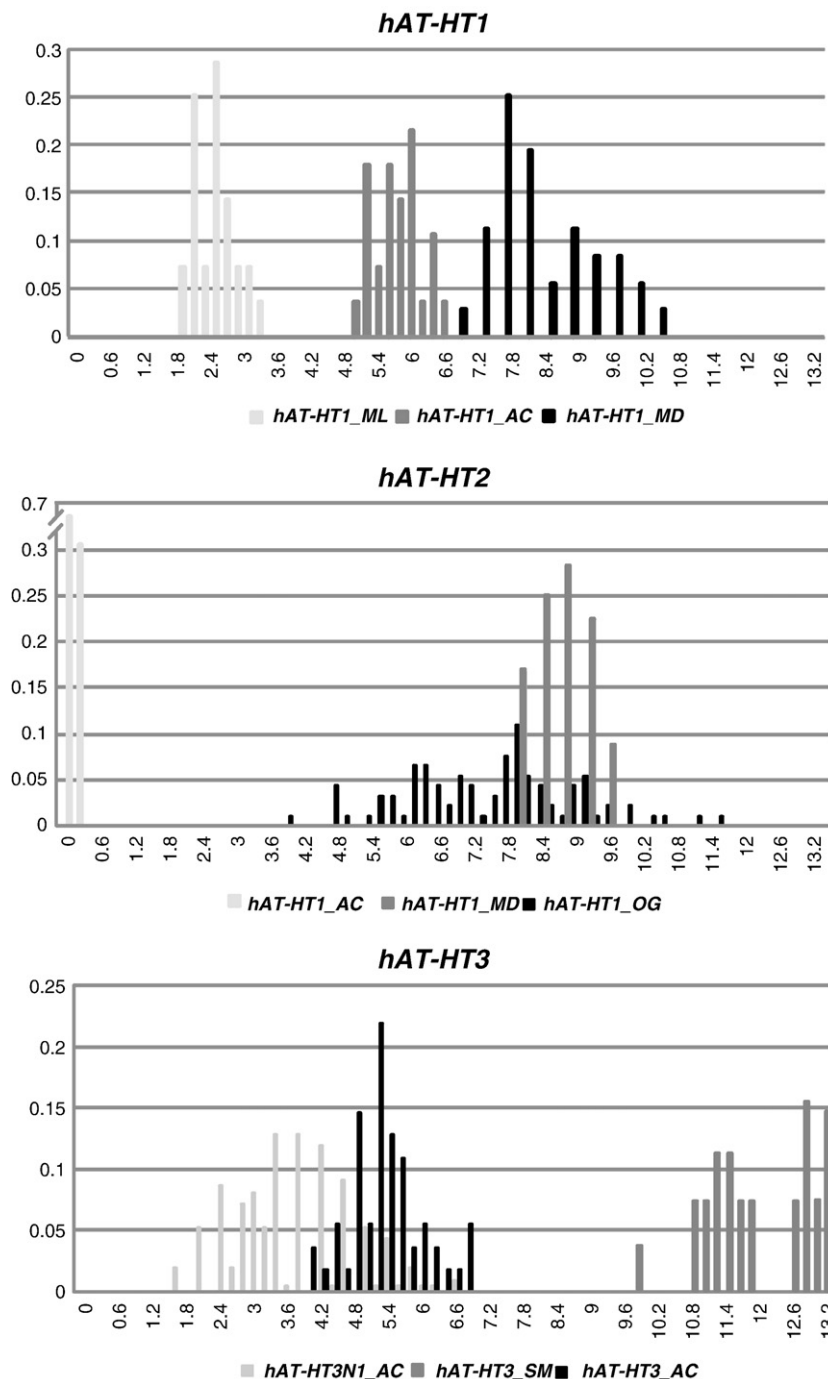
calculated the ratio *dn/ds* between transposase consensus sequences and we found that the values of the ratio are remarkably high for a protein coding sequence, although only a few were higher than 1 (Table 2). These high *dn/ds* values suggest that the transposase gene evolved neutrally, which is consistent with the very definite, and relatively short lifespan of *hAT-HT2* in the anole and in the opossum (Fig. 3). However, the distribution of pairwise divergence in *O. garnettii* (the bush baby; Fig. 3) suggests that *hAT-HT2* has persisted in this species for about 18MY (assuming a mutation rate in prosimians of ~0.2% per MY; Liu et al., 2003). It is very unlikely that *hAT-HT2* would have persisted so long in this genome if the transposase was evolving neutrally. Instead, the high *dn/ds* values could result from the action of positive selection, i.e. selection in favor of amino acid changes. Under this scenario, the colonization of the naïve bush baby genome by *hAT-HT2* could have been followed by a phase of adaptation of the transposon to its new genomic environment, possibly in response to host defense.

The average divergence between *hAT-HT2* copies is 0.14% in the anole indicating that these copies have been inserted very recently in this genome. In other species, *hAT-HT2* insertions are a little older, the divergence between elements ranging from 11.74% in the planarian to 3.51% in the bat (Table 1). Assuming that variations in mutation rate are small among species, differences in pairwise divergence suggest that the lateral transfer of *hAT-HT2* occurred at different times in different lineages or that the elements remained dormant in some species until they began amplifying in their host genome. *hAT-HT2* is found in low copy numbers in most species (<50 copies), except in the tenrec and in the opossum where this element is found in more than 200 copies. This family is also responsible for the mobility of non-autonomous elements in most lineages (Table 1). We identified one non-autonomous family related to *hAT-HT2* in *Microcebus*, anole and tenrec, at least two non-autonomous families in the bat and at least three in the tarsier. These non-autonomous families have amplified extremely successfully and largely outnumber their autonomous progenitors.

**Table 2**Summary of non-synonymous (*dn*) and synonymous (*ds*) substitutions of the four families of autonomously replicating horizontally transferred *hAT* families.

	dn						ds						dn/ds							
hAT-HT1	Md	MI	Ac				Md	MI	Ac				Md	MI	Ac					
Monodelphis domestica (Md)																				
Myotis lucifugus (MI)	0.0983						0.3002													
Anolis carolinensis (Ac)	0.0157	0.0932					0.0315	0.3201					0.3276							
													0.4986	0.2912						
hAT-HT2	Ts	MI	Mm	Md	Og	Et	Ts	MI	Mm	Md	Og	Et	Ac	Ts	MI	Mm	Md	Og	Et	Ac
Tarsius syrichta (Ts)																				
M. lucifugus (MI)	0.0444						0.0299							1.48495						
Microcebus murinus (Mm)	0.0298	0.0527					0.0205	0.0371						1.45366	1.42049					
M. domestica (Md)	0.0337	0.0432	0.0407				0.0407	0.0493	0.0428					0.8280	0.87627	0.95093				
Otolemur garnettii (Og)	0.0191	0.0364	0.0270	0.0230			0.0171	0.0206	0.0194	0.0420				1.1170	1.7670	1.39175	0.54762			
Echinops telfairi (Et)	0.0147	0.0342	0.0225	0.0208	0.0065		0.0102	0.0229	0.0124	0.0347	0.0068			1.44118	1.49345	1.81452	0.59942	0.95588		
A. carolinensis (Ac)	0.0229	0.0363	0.0308	0.0257	0.0179	0.0135	0.0450	0.0525	0.0470	0.0608	0.0388	0.0412		0.50889	0.69143	0.65532	0.4227	0.46134	0.32767	
hAT-HT3	Ac						Ac							Ac						
Schmidtea mediterranea (Sm)	0.0098						0.0027							3.6360						
SPIN	Xt	Ac	Og	Et	MI	Md	Xt	Ac	Og	Et	MI	Md	Xt	Ac	Og	Et	MI	Md		
Xenopus tropicalis (Xt)																				
A. carolinensis (Ac)	0.1002						0.1198							0.8369						
O. garnettii (Og)	0.0580	0.0529					0.0366	0.0862						1.5848	0.6138					
E. telfairi (Et)	0.0541	0.0513	0.0101				0.0304	0.0987	0.0111					1.7789	0.5201	0.9017				
M. lucifugus (MI)	0.0578	0.0514	0.0118	0.0067			0.0304	0.0931	0.0067	0.0044				1.9037	0.5519	1.7662	1.5286			
M. domestica (Md)	0.0997	0.0952	0.0606	0.0581	0.0558		0.2537	0.2688	0.2323	0.2425	0.2370			0.3931	0.3542	0.2608	0.2395	0.2356		

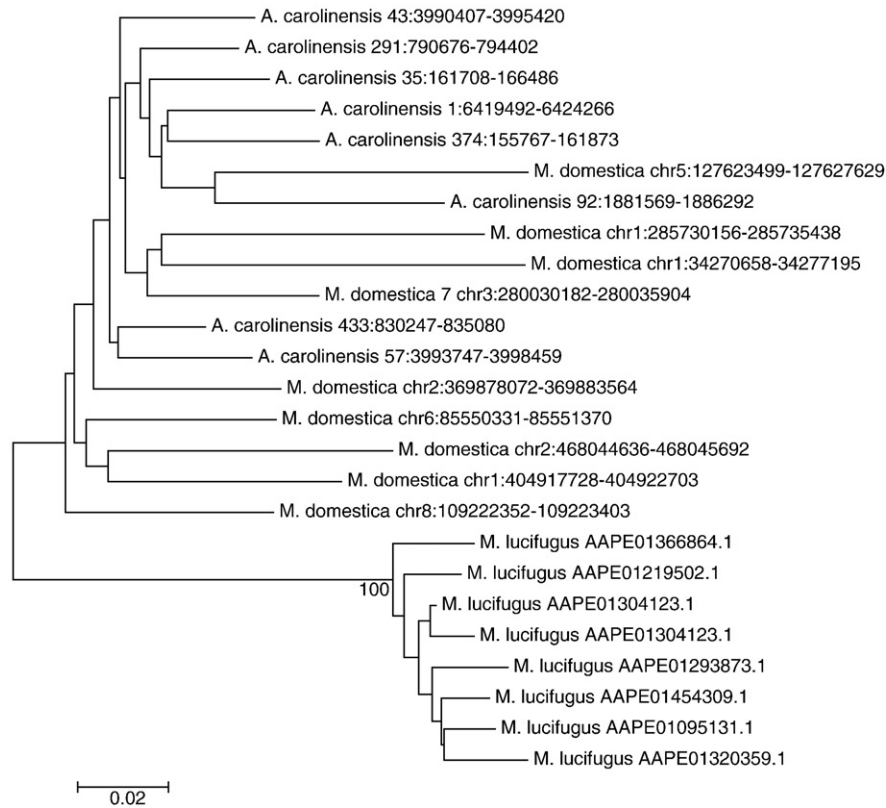




**Fig. 3.** Pairwise divergence distribution of *hAT-HT1*, *hAT-HT2* and *hAT-HT3* families. Abbreviations: *Anolis carolinensis* (AC), *Myotis lucifugus* (ML), *Monodelphis domestica* (MD), *Otolemur garnettii* (OG) and *Schmidtea mediterranea* (SM). Genomic elements were collected and at least 1000 bp across the element was used to calculate pairwise divergence.

A second autonomous *hAT* family, *hAT-HT1*, is shared only among the anole, the brown bat and the opossum (Figs. 1 and 2). The *hAT-HT1* family includes the opossum *hAT1\_MD* and bat *Myotis\_hAT1* families described in Repbase. In each of these species, we were able to recover full-length elements and to reconstruct complete consensus sequences. The divergence between the anole and opossum *hAT-HT1* is extremely low, in particular at synonymous sites (3.15%), and is suggestive of lateral transfer. The *ds* between the anole and the bat is much higher (about 30%), suggesting a distinct origin for the *hAT-HT1* family in the bat. The distributions of pairwise divergence among elements (Fig. 3) barely overlap, indicating that the *hAT-HT1* family amplified at different time in these three species. A phylogenetic

analysis of genomic copies of *hAT-HT1* elements (Fig. 4) shows that the anole and opossum elements form an unresolved polytomy, suggesting a single source for the lateral transfer of *hAT-HT1* into these two species. In contrast, bat elements form a clearly distinct monophyletic group indicating that bat *hAT-HT1* might be coming from a different source or that the elements were transferred at different times. This independent origin is also supported by the observation that bat *hAT-HT1* elements have very different (non-homologous) termini than the anole and opossum elements (Fig. 5). *hAT-HT1* transposases also differ remarkably at non-synonymous sites, although the values of the *dn/ds* ratio were not as high as the values for *hAT-HT2* (Table 2). *hAT-HT1* elements are found in very low



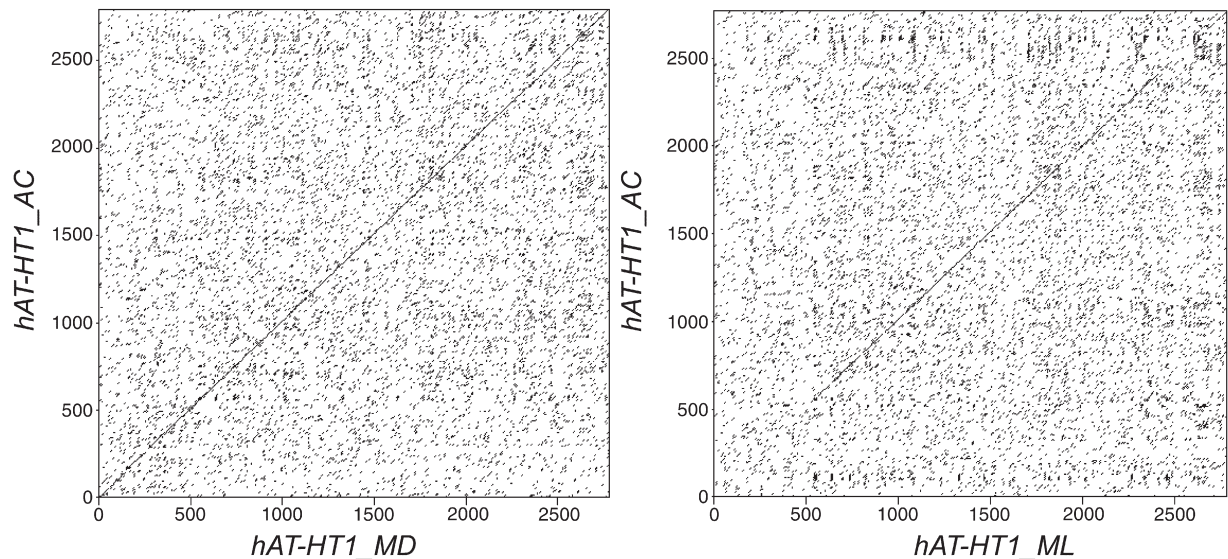
**Fig. 4.** Neighbor-joining phylogeny of genomic copies of *hAT-HT1* elements from the three species known to foster this family, *Monodelphis domestica*, *Anolis carolinensis* and *Myotis lucifugus*. The tree is based on a 1000-bp long alignment of the transposase domain. The robustness of the tree was assessed with 1000 bootstrap replicates. Only bootstrap values higher than 75 are shown. Accession numbers for each element are delineated on each branch (chromosome or scaffold locus from the UCSC Web site for the opossum and lizard respectively, or its NCBI accession number for the bat).

copy number in the anole (about 8) but have been successful at colonizing the bat and opossum genomes, with more than 100 copies in both species.

A third autonomous *hAT* family, *hAT-HT3*, is shared by the anole and the planarian *S. mediterranea* (Figs. 1 and 2). We recovered a number of full-length elements in the anole but only fragments in the planarian, yet we were able to construct full-length *hAT-HT3* consensus for both species. The *ds* value between the anole and the

planarian consensus is the lowest reported in this study, with a value of 0.27%, suggesting a lateral transfer from the very same source or directly from a planarian species to the anole (because the *hAT-HT3* has apparently resided longer in the planarian than in the lizard).

Finally, we found in anole several families of non-autonomous elements related to *hAT-HT3*. One of these non-autonomous elements, *hAT-HT3N1*, is also found in the bat. The anole and the bat *hAT-HT3N1* elements are very similar and diverge by only 0.65%, a value best



**Fig. 5.** Dot plot comparison of the complete *hAT-HT1* consensus sequences of the anole to the gray short-tailed opossum, *Monodelphis domestica* (left), and to the little brown bat, *Myotis lucifugus* (right).

explained by a lateral transfer of this non-autonomous element. The transfer and successful amplification of the non-autonomous element is surprising because we failed to find in the bat genome an autonomous element similar to the *hAT-HT3* family. It is plausible that, following its transfer, *hAT-HT3N1* recruited another autonomous element for its mobilization. We cannot exclude that the number of autonomous *hAT-HT3* elements in the bat genome is very small and that they reside in genomic regions that have yet to be fully sequenced. It is also possible that the autonomous element was so recently transferred that it is still polymorphic in little brown bat populations and is missing from the individual sequenced.

### 3.2. Evidence for lateral transfer

The evidence we presented above unequivocally demonstrates that the presence in distantly related genomes of three autonomous and one non-autonomous *hAT* transposons results from lateral transfer. The first line of evidence comes from the phylogenetically discontinuous distribution of the families. If these elements had been vertically transmitted from a common ancestor, they would have been independently lost from other lineages. In fact, we would need to assume as many as 14, 8, 9 and 9 independent losses to account for the phylogenetic distribution of the *hAT-HT2*, *hAT-HT1*, *hAT-HT3* and *hAT-HT3N1* families, respectively. We would also expect to find the remnants of ancient elements in the primates, tenrec, bat and opossum genomes. TEs readily accumulate in mammalian genomes where they remain as DNA fossils (Gilbert and Labuda, 1999; Bejerano et al., 2006). Given the relatively slow mutation rate of mammals, ancient elements remain detectable long after their insertion. In the case of these families, we failed to find ancient copies in any of these genomes suggesting these families are recent colonizers of mammalian genomes. This reasoning applies only to mammals because TEs, especially long ones, do not accumulate in the anole genome, which is characterized by a high turnover of inserts (Novick et al., 2009). Thus, we do not expect to find ancient elements in this genome. The second line of evidence in favor of lateral transfer comes from the very low divergence between consensus sequences, especially at silent sites. Silent divergences faithfully reflect the evolutionary history of sequences because synonymous mutations are virtually invisible to natural selection. For all autonomous families, we report levels of silent divergence considerably lower than expected for such ancient lineages. For instance, lophotrochozoans and vertebrates separated more than 800 MYA, yet the *ds* between the anole and the planarian *hAT-HT3* consensus is 0.27%. Similarly among mammals, the *ds* values between *hAT-HT2* consensus sequences were all lower than 5%, although chiropters diverged from primates 65 MY ago and eutherians split from marsupials more than 120 MY ago (Donoghue and Benton, 2007). The third line of evidence comes from phylogenetic analyses. Phylogenetic trees showed a striking lack of structure, for *hAT-HT2* and for part of the *hAT-HT1* tree. This lack of structure is not consistent with a vertical mode of transmission which yields bifurcating phylogenies that mimic the phylogeny of the host species. The phylogenies we obtained are best explained by the invasion of different genomes from a single or a small number of sources.

## 4. Discussion

We identified four new families of DNA transposons that have been horizontally transferred in tetrapods. Because these families represent distinct monophyletic groups (Fig. 6), they signify independent events of lateral transfer. The discovery of these four cases results from searching other genomes for the presence of transposons discovered in *A. carolinensis*. It is therefore biased towards elements found in this species and certainly underestimates the frequency of transposon lateral transfer in tetrapods. The only case previously reported was the *SPIN* family, which was shared between bush baby,

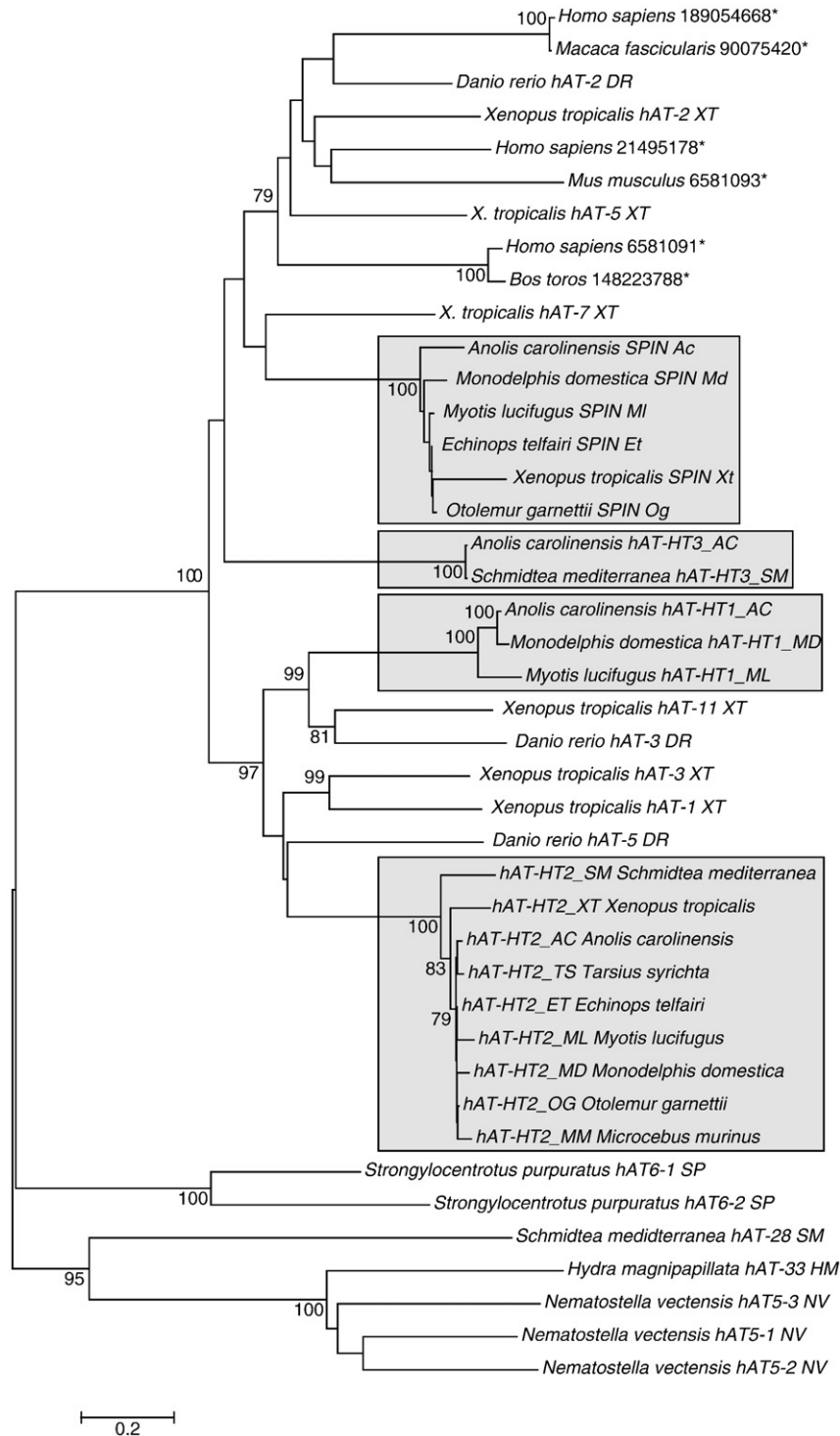
bats, murine rodents, marsupials, anole and frog (Pace et al., 2008). Thus, at least five events of lateral transfer have occurred during the evolution of tetrapods and we suspect many more will be discovered. Lateral transfer of transposons had previously been described in animals but, in most cases, was between closely related species and was mostly limited to insects. Our analysis suggests that the transfer of DNA transposons might be much more common in vertebrates than previously thought. In fact, it probably happens more often than our data suggest because we detected only the successful events of lateral transfer.

This has important implications for the genomic evolution of mammals. Until the discovery of recent transposon activity in the bat *M. lucifugus* (Ray et al., 2007, 2008), it was believed that DNA transposons were extinct in mammals and had no impact on the evolution of this group. However, lateral transfer provides a means for the re-colonization of genomes and we cannot exclude that DNA transposons are active in many species for which genomic data are not currently available. The successful amplification of laterally transferred DNA transposons can have a significant impact on genomic diversity. Altogether the five events of lateral transfer discussed here are responsible for the amplification of approximately 1300, 57,000 and 100,000 elements in the anole, the bat and the tenrec genomes, respectively. As the amplification of DNA transposons can trigger the evolution of new genes or even new biological functions by providing coding and regulatory motifs on which natural selection can act (reviewed in Feschotte and Pritham, 2007), it is plausible that some insertions have been recruited by their host to perform a new function.

The evidence in favor of lateral transfer is overwhelming, yet our data provide no information about the source of the transfer and its mechanism(s). The main problem is that in amniotes (reptiles and mammals) the germ line and the early stages of embryogenesis are sequestered within the gonads and the genital tract of the female. In species that have free living unprotected eggs and/or external fertilization, such as the frog and the planarian, it is easier to imagine a mechanism of transfer of genetic material directly in the germ line. In amniotes, the lateral transfer of transposons implies the existence of an agent that is mediating the transfer to the different species. It was recently proposed that poxviruses could mediate lateral transfer of TEs because a snake retroposon was found in the genome of a rodent poxvirus (Piskurek and Okada, 2007), raising the possibility that this virus could transfer the reptilian element to its rodent host. Poxviruses are good candidates for mediating lateral transfer due to their low host and cell specificity. However, we are not aware of any poxvirus (or any other viruses) with a range of hosts broad enough to explain the phylogenetic distribution of DNA transposons, from planarian to primates. The geographic distributions and ecology of the host species differ drastically: tarsiers are found in the Philippines, the bush baby and *Xenopus* frog in Africa, lemurs and tenrecs in Madagascar, the little brown bat and the green anole in North America, the opossum *M. domestica* in South America and the planarian *S. mediterranea* around the Mediterranean. Not only do these species inhabit different geographic areas, but their center of origin and their evolution took place on different continents. Therefore, the source of the transposons must have (or had if it is now extinct) a very wide, possibly global, geographic distribution. A majority of the species that experienced lateral transfer have an insectivorous diet (anole, bats). It is therefore tempting to speculate that a transfer of transposons occurred from prey to predators. However, none of the laterally transferred elements described here were found in the insect genomes currently available, although we can not exclude that these transposons are present in an insect species that has not yet been sequenced.

Although the phylogenetic distribution of the five transposon families is discontinuous, it is not random. For instance, the anole shares four families with the bat, three with the opossum and three with the planarian *S. mediterranea* but none with the other 24





**Fig. 6.** Neighbor-joining phylogeny of 35 sequences consisting of 600 amino acids of the 3' end of the hAT transposase domain in vertebrates. In addition to the 14 sequences included in this study (shaded in gray), six mammalian sequences were included from the NCBI database (protein accession numbers are available next to the scientific name and indicated by an asterisk) as well as 15 sequences retrieved from Repbase (SPIN elements are also shaded in gray). As an outgroup, we have included seven sequences from four species of invertebrates including *Schmidtea mediterranea* (also retrieved from Repbase). Bootstrap values less than 75 have been removed from the tree.

vertebrates and more than 25 invertebrates for which genomic data are available. This pattern suggests that some genomes have a higher chance of lateral transfer or are more hospitable to transposons than others or both. It is plausible that genomes that are vulnerable to lateral transfer have a weakened resistance to transposon invasion and, at first, lose control of the amplification whereas genomes with a strong response never let the transposons amplify in their genome. Our data suggest that a mechanism of resistance might exist in some

species. When consensus transposase sequences are compared, we found high values for the ratio  $dn/ds$ , with a number of values higher than 1. Values of  $dn/ds$  higher than 1 indicate that selection in favor of amino acid changes is acting on a gene. One can wonder why a gene as indispensable to the fitness of the transposon would evolve rapidly at the amino acid level. A possibility is that the amplification of the transposon could trigger the evolution of a response by the host, leading to an arms race between the transposon and its host. This

evolutionary race could cause a rapid evolution of the transposase and of the host factor responsible for its control. The role of a genomic conflict has previously been proposed to explain the rapid evolution of the first open-reading frame of the LINE-1 retrotransposon in primates (Boissinot and Furano, 2001; Khan et al., 2006) and of transposase genes in bacterial DNA transposons (Petersen et al., 2007).

## Acknowledgments

This research was supported by PSC-CUNY grant 61542-00-39 to Stéphane Boissinot. We would like to thank Dr. Peter Reddien (MIT), Dr. Michael Steiper (Hunter College), Dr. Mark Batzer (Louisiana State University) and Doug Dellecave from D&J's Reptiles for their tissue or DNA contributions. Molecular experiments were conducted in part with equipment from the Core Facility for Imaging, Cellular and Molecular Biology at Queens College.

## References

- Andersson, J.O., 2005. Lateral gene transfer in eukaryotes. *Cell. Mol. Life Sci.* 62, 1182–1197.
- Bao, W., Jurka, M.G., Kapitonov, V.V., Jurka, J., 2009. New superfamilies of eukaryotic DNA transposons and their internal divisions. *Mol. Biol. Evol.* 26, 983–993.
- Bejerano, G., et al., 2006. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 441, 87–90.
- Bergthorsson, U., Adams, K.L., Thomason, B., Palmer, J.D., 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424, 197–201.
- Boissinot, S., Furano, A.V., 2001. Adaptive evolution in LINE-1 retrotransposons. *Mol. Biol. Evol.* 18, 2186–2194.
- Boissinot, S., Davis, J., Entezam, A., Petrov, D., Furano, A.V., 2006. Fitness cost of LINE-1 (L1) activity in humans. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9590–9594.
- Daniels, S.B., Peterson, K.R., Strausbaugh, L.D., Kidwell, M.G., Chovnick, A., 1990. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* 124, 339–355.
- de Boer, J.G., Yazawa, R., Davidson, W.S., Koop, B.F., 2007. Bursts and horizontal evolution of DNA transposons in the speciation of pseudotetraploid salmonids. *BMC Genomics* 8, 422.
- Donoghue, P.C.J., Benton, M.J., 2007. Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol. Evol.* 22, 424–431.
- Edgar, R.C., Myers, E.W., 2005. PILER: identification and classification of genomic repeats. *Bioinformatics* 21 (Suppl. 1), i152–i158.
- Feschotte, C., 2008. Transposable elements and the evolution of regulatory networks. *Nat. Rev., Genet.* 9, 397–405.
- Feschotte, C., Pritham, E.J., 2007. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* 41, 331–368.
- Furano, A.V., 2000. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog. Nucleic Acid Res. Mol. Biol.* 64, 255–294.
- Gilbert, N., Labuda, D., 1999. CORE-SINEs: eukaryotic short interspersed retroposing elements with common sequence motifs. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2869–2874.
- Guindon, S., Lethiec, P., Duroux, P., Gascuel, O., 2005. PHYML online—a Web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, W557–W559.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Kazazian, H.H., 2004. Mobile elements: drivers of genome evolution. *Science* 303, 1626–1632.
- Khan, H., Smit, A., Boissinot, S., 2006. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res.* 16, 78–87.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120.
- Kordis, D., Gubensek, F., 1998. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10704–10709.
- Kordis, D., Lovsin, N., Gubensek, F., 2006. Phylogenomic analysis of the L1 retrotransposons in Deuterostomia. *Syst. Biol.* 55, 886–901.
- Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: Molecular Evolutionary Genetics Analysis Software. Arizona State University, Tempe, Arizona, USA.
- Liu, G., et al., 2003. Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome. *Genome Res.* 13, 358–368.
- Makalowski, W., 2000. Genomic scrap yard: how genomes utilize all that junk. *Gene* 259, 61–67.
- Nekrutenko, A., Li, W.H., 2001. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* 17, 619–621.
- Novick, P.A., Basta, H., Floumanhaft, M., McClure, M.A., Boissinot, S., 2009. The evolutionary dynamics of autonomous non-LTR retrotransposons in the lizard *Anolis carolinensis* shows more similarity to fish than mammals. *Mol. Biol. Evol.* 26, 1811–1822.
- Pace, J.K., Gilbert, C., Clark, M.S., Feschotte, C., 2008. Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17023–17028.
- Petersen, L., Bollback, J.P., Dimmic, M., Hubisz, M., Nielsen, R., 2007. Genes under positive selection in *Escherichia coli*. *Genome Res.* 17, 1336–1343.
- Petrov, D., Aminetzach, Y.T., Davis, J.C., Bensasson, D., Hirsh, A.E., 2003. Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol. Biol. Evol.* 20, 880–892.
- Piskurek, O., Okada, N., 2007. Poxviruses as possible vectors for horizontal transfer of retrotransposons from reptiles to mammals. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12046–12051.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Ray, D.A., Pagan, H.J., Thompson, M.L., Stevens, R.D., 2007. Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*. *Mol. Biol. Evol.* 24, 632–639.
- Ray, D.A., et al., 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res.* 18, 717–728.
- Robertson, H.M., Lampe, D.J., 1995. Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. *Mol. Biol. Evol.* 12, 850–862.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Timmis, J.N., Ayliffe, M.A., Huang, C.Y., Martin, W., 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev., Genet.* 5, 123–135.
- Won, H., Renner, S.S., 2003. Horizontal gene transfer from flowering plants to Gnetum. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10824–10829.
- Yang, Z., 2000. PAML (Phylogenetic Analysis by Maximum Likelihood) version 3.0. University College, London.
- Yang, Z., Nielsen, R., 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17, 32–43.