

Review

Transposable Elements: From DNA Parasites to Architects of Metazoan Evolution

Oliver Piskurek and Daniel J. Jackson *

Courant Research Centre Geobiology, Georg-August-University of Göttingen, Goldschmidtstr. 3, Göttingen 37077, Germany; E-Mail: opiskurek@web.de

* Author to whom correspondence should be addressed; E-Mail: djacks@uni-goettingen.de; Tel.: +49-551-391-4177; Fax: +49-551-397-918.

Received: 19 May 2012; in revised form: 19 June 2012 / Accepted: 25 June 2012 /

Published: 12 July 2012

Abstract: One of the most unexpected insights that followed from the completion of the human genome a decade ago was that more than half of our DNA is derived from transposable elements (TEs). Due to advances in high throughput sequencing technologies it is now clear that TEs comprise the largest molecular class within most metazoan genomes. TEs, once categorised as "junk DNA", are now known to influence genomic structure and function by increasing the coding and non-coding genetic repertoire of the host. In this way TEs are key elements that stimulate the evolution of metazoan genomes. This review highlights several lines of TE research including the horizontal transfer of TEs through host-parasite interactions, the vertical maintenance of TEs over long periods of evolutionary time, and the direct role that TEs have played in generating morphological novelty.

Keywords: transposable element; junk DNA; molecular parasite; molecular domestication; functionalisation; exonisation; exaptation; SINE; LINE

1. Classification and Diversity of TEs

During her career Barbara McClintock discovered and described transposable elements (TEs), a class of mobile genetic elements often abundantly distributed throughout the genomes of eukaryotic organisms [1–4; reviewed in 5]. At the time, her findings were in line with the popular theory of selfish DNA in which TEs could be perceived as “genomic hitchhikers” or molecular parasites which play no

significant role in genome evolution, and provide no adaptive advantage to the host [6,7]. Ahead of her time, Barbara McClintock nonetheless suggested that TEs can indeed influence the evolution of the genome. The ways in which this has since been shown to be true are amazing. These insights have been in part due to the rise of the field of evolutionary-developmental biology (evo-devo), rapid advances in DNA sequencing technologies and the concomitant rise in the field of comparative genomics. The modern day view of TEs is that they have the potential to act as agents of evolution by increasing, rearranging and diversifying the genetic repertoire of their hosts [8–17].

With significant advances in high throughput sequencing technologies has come the democratization of genome sequencing. Following the completion of the first metazoan genome in 1998 using 'first generation' technologies [18], draft genomes of "non-model" organisms are released with increasing annual frequency. With this flood of sequence data has come the need to develop bioinformatic tools designed to detect and characterise TEs [19–21]. TEs can be broadly divided into two classes based upon their mechanism of mobilisation or transposition (however the reader should be aware that newly described classes of TEs challenge this simplistic categorisation): class I elements (retroposons) mobilise via an RNA intermediate analogous to a "copy-and-paste" mechanism where the "copy" (RNA) is biochemically distinct from the original (DNA); class II elements (DNA transposons) mobilise via a DNA-mediated mode of transposition originally known as "cut-and-paste" mechanism. TEs of both classes can be classified as autonomous or non-autonomous, based on whether or not they encode the proteins necessary for their own retrotransposition/transposition. Four types of eukaryotic class I TEs can be distinguished [22,23]: Long terminal repeat elements (LTRs); non-LTR elements such as long interspersed elements (LINEs) and non-autonomous short interspersed elements (SINEs); and two types of TE with unusual structures, namely DIRS (based on DIRS-1, characterized in *Dictyostelium*) and Penelope-like elements (PLEs). TEs of class II can be divided into three major types: type 1 elements have two terminal inverted repeats (TIRs) and are typical cut-and-paste DNA transposons which are fully excised with the help of enzyme transposase; type 2 elements are rolling-circle DNA transposons also known as Helitrons [14,24]; and type 3 elements which are self-synthesizing DNA transposons, also known as Polintons or Mavericks [25].

TEs can often be recognised as genomic fossils that were once autonomously replicating elements which at some point in time experienced a deletion, inversion, or other mutation that rendered them inactive. Alternatively, a 'fossilised non-autonomous TE' can remain active as long as the enzymatic machinery required for transposition is provided by an autonomous partner. A good example of this is the LINE-SINE system. The 3' tail sequence of a SINE is identical to that of its partner LINE and is recognized by the reverse transcriptase (RT) of that LINE [26–28]. Thus, when SINEs replicate via retrotransposition they depend on the existing retrotranspositional machinery of their active LINE partners. The largest TE class within most metazoan genomes consists of LINEs and SINEs [29,30]. For example, the LINE1-*Alu*-SINE system comprises nearly 30% of the human genome [31,32]. In marsupials LINEs and their related mobilized SINEs make up nearly 40% of the genome [33,34], while LINEs/SINEs in the lizard genome (including the Bov-B LINE-Sauria SINEs system) represent about 17% of the total DNA [35–37]. Based on these observations Hua-Van *et al.* [38] discuss the concept of a struggle for survival between TE families, similar to that which occurs between species sharing the same ecological niche.

Current views of metazoan TE diversity and distribution are likely to be biased and under-representative. This is because genomes of direct relevance to human medicine or evolution, or of unusually small size are preferentially sequenced. Furthermore, comprehensive and accurate identification and annotation of the TE complement in any metazoan genome requires: (i) a high quality genome assembly and; (ii) bioinformatic tools that can recognise both conserved (typically homology based methods) and novel (*i.e.*, *de novo* detection methods) TEs. Both of these requirements are non-trivial. Continuing technical advances in sequencing and *in silico* assembly methods could be expected to eventually eliminate the first problem. Detecting truly novel, lineage specific TEs using *de novo* methods is an inherently challenging bioinformatic exercise (relative to homology based methods). However, with the sequencing of more phylogenetically representative taxa from the metazoan tree of life for evolutionary studies [39–41], and the corresponding development of tools to annotate these datasets [42–44], it can be expected that in the near future we will come to appreciate that the evolutionary histories and functions of TEs are as complex and diverse as the biological populations bearing them.

2. DNA Transposons—Horizontal Transfer Events Facilitate the Spread of TEs

Horizontal transfer (HT), the exchange of genetic material between two species that do not share an immediate ancestor-descendant relationship, of a TE from one genome to another can trigger many molecular events, which can in turn directly influence genome evolution. Furthermore, HT is an effective strategy that ensures the long-term survival of an active DNA transposon. This is because HT allows the element to evade extinction in the host which may be brought about by host repression of TE activity, or by extinction of the host lineage. The phenomenon of HT of genetic material between bacteria is well known. This mechanism can explain the abundance of insertion sequences (IS) in prokaryotic genomes [45]. However, HT events across domains of life, *i.e.*, from bacteria to eukaryotic organisms, are far less common. The genomes of bdelloid rotifers appear to have evolved for millions of years without sexual reproduction, and possess genes thought to have been acquired by HT from bacteria, fungi, and plants [46]. Concerning TEs, there is only one known HT event from a prokaryote to a eukaryote, namely the IS5-like integration from a bacterium into a bdelloid rotifer genome [47]. This HT event apparently took place recently as the TE has not increased in copy number within the bdelloid genome.

Classical examples of horizontally transferred TEs between metazoans include the P-elements in *Drosophila* [48], the Mariner transposons in insects [49], and the chromoviruses, the oldest and largest lineage of LTR elements, which were horizontally transferred into the genome of the ancestor of gnathostomes [50]. Although these and other class II TEs are well adapted to invade species via horizontal transmission [51], it has been suggested that all types of TEs may be subject to HT [52–54]. Schaack *et al.* [54] lists more than 100 cases of HTs in which class I TEs are involved, mostly LTR and non-LTR elements. A typical example for the HT of non-LTR elements is the Bov-B LINE which was initially discovered in ruminants but later shown to be ubiquitous in squamate genomes [55,56]. Given the patchy phylogenetic distribution of Bov-B LINEs among mammalian genomes it was inferred that these TEs were transferred horizontally from an ancestor of derived snakes to an ancestor

of ruminants. A potential vector for this transfer was discovered when a Bov-B-derived Sauria SINE from the snake *Echis ocellatus* was identified in a poxvirus known to infect mammals [57].

Apart from viruses and bacteria, eukaryotic parasites can also facilitate the spread of TEs across diverged host species. Houck *et al.* [58] proposed that the mite *Proctolaelaps regalis* may serve as a vector for the HT of TEs between different *Drosophila* species. Furthermore, it was recently shown that the genome of *Rhodnius prolixus*, a triatomine bug which feeds on the blood of diverse tetrapods, harbours four DNA transposon families which are also present in the genomes of the bug's preferred hosts [54,59]. Finally, the hookworm *Ancylostoma caninum*, a parasite of dogs that is frequently detected in the human small intestine, harbours a mariner-like DNA transposon (bandit) that is phylogenetically related to the human Hsmar DNA transposon, suggesting that a HT of bandit may have taken place between hookworm parasites and mammalian hosts [60]. The complexity of TE evolution is highlighted in these cases where the TE and the host genome co-evolve in parallel with the co-evolutionary arms race of the parasite and host [61]. In order to fully understand how the genomic TE complement of a given eukaryotic organism evolved, it is therefore often necessary to understand the ecology of any host-parasite interactions that organism may have. Nonetheless, because TEs can become functionally relevant for the host genome (see Section 4 below), the exchange of such genetic material between species, regardless of how it is delivered, can have a striking impact on genome evolution.

3. Retroposons—TEs as Molecular Markers to Infer Phylogenetic Relationships

While horizontally transferred DNA transposons can provide information concerning the ecological interactions of host species, retroposons are usually vertically inherited and can therefore provide information concerning the phylogenetic relationships of species. Using SINE/LINE partners to infer phylogenetic relationships by parsimony is a powerful method when multiple retroposon insertions show the same phylogenetic pattern [62,63]. This method simply treats an insertion at a specific genomic location as a derived character state, while the lack of an insertion at an orthologous locus is regarded as the ancestral state. SINEs and LINEs are suitable for phylogenetic studies for several reasons: they insert almost randomly into genomic DNA; most copies are non-autonomous; they exist in large copy numbers; and their transfer usually occurs vertically. For more than a decade, SINEs and LINEs have been successfully employed as molecular markers, particularly by the Okada group [64–69] and Schmitz and colleagues [34,70–73]. It is also known that retroposons can reveal rapid radiation (incomplete lineage sorting) events as shown recently for the origin of placental mammals [69,72]. In this case retroposons were not fixed in the ancestral population before the separation of lineages (incomplete lineage sorting), and therefore cannot be used as phylogenetic markers. Such situations can nonetheless provide insight into the evolutionary and geological history of the three placental lineages (Afrotheria, Xenarthra, and Boreotheria). These lineages divided nearly simultaneously in parallel with the division of continents that lead to isolated Africa, South America, and Laurasia. Such TE insertion polymorphisms are not informative for phylogenetic analyses, but they do provide an efficient tool for the identification of rapid speciation events [63,74]. For example, recent high-throughput sequencing and comparison of human genomes revealed extensive variation in LINE1

content [42], illustrating that such TEs can not only be used as phylogenetic markers but are a major source of individual genomic variation.

4. TEs Are a Source of Novel Genetic Material

TEs are often associated with genome expansions and increases in genomic diversity (e.g. 85% of the maize genome, 57% of the *Hydra* genome, and 45–69% of the human genome are composed of TEs) [31,75–77]. In many cases the insertion of a TE has either no impact (typical for insertions into non-functional DNA regions), or deleterious effects as a result of disrupting coding DNA or gene regulatory regions such as promoters. TE insertions can also cause sequence inversions, duplications or deletion events and are therefore a potential source of genetic diseases [78,79]. Deleterious TE activity can also be brought about by the ectopic recombination of non-homologous regions of a chromosome. Such rearrangements during meiosis can produce unviable gametes, and it has been recently shown in *Drosophila* that selection against such events is the major force driving TE population dynamics [80].

SINEs, LINEs and other transposed sequences can however positively influence the host genome in many ways. For example, Jordan *et al.* [81] analysed promoter regions in the human genome and found that almost 25% contain TE-derived sequences associated with transcriptional regulation, while Nekrutenko and Li [82] examined 13,799 human genes and found 533 genes associated with TE insertions, of which these were mostly SINEs (~40%) and LINEs (~27%). As a specific example of this, Lunyak *et al.* [83] showed that tissue-specific transcription of a SINE sequence in the murine growth hormone locus is required for the establishment of functional chromatin domains, which in turn permit gene activation. It is now clear that TE insertions into the untranslated regions of genes are frequently associated with alternative splicing events (exonizations), and the *de novo* generation of exons [84–87]. In addition to exonization, TE insertions are also known to deliver novel introns. Indeed TE mediated intron insertions are thought to be responsible for much of the wide-spread intron gains observed in mammalian genomes [88]. TEs that evolve into novel protein coding sequences by exonization might subsequently acquire a function in a process called exaptation [8,89]. Several criteria can be used to detect whether such domestication of a TE has taken place: (i) evidence of TE fixation in a population; (ii) the presence of an intact open reading frame and splice sites; (iii) the presence of orthologous TE sequences in several species. Comparative analyses of mammalian genomes have revealed that only 1.5–2.0% of the human genome constitute protein coding genes, while up to 5% of the genome consists of conserved non-coding elements (CNEs) [90]. Two recent studies characterized certain SINEs which make up this population of CNEs: Bejerano *et al.* [12] identified the living fossil LF-SINE in the coelacanth, while Nishihara *et al.* [66] found homologous AmnSINE1 members in amniotes. Some of the exapted SINE copies, which accumulated mutations over evolutionary time, form ultra-conserved enhancers, such as the LF-SINE locus 0.5 Mb upstream of the neuro-developmental gene *ISL1* [12]. Recently, the genome of the marsupial *Monodelphis domestica* revealed that at least 16% of eutherian-specific CNEs are derived from TEs [33]. It has been suggested that phylogenetically conserved SINEs with conserved domains, such as the CORE-SINEs found in bilaterians [91], V-SINEs in vertebrates [92], DeuSINEs in deuterostomes [66], and CephSINEs in cephalopods [93], can act as functional modules that enhance gene expression. Some

recent studies have indeed experimentally demonstrated this to be true [15,16,94,95]. Interestingly, these examples of TE exaptation are often associated with morphological innovations. Santangelo *et al.* [94] identified an ancient exaptation of a CORE-SINE that has remained under purifying selection in all mammalian orders for the last 170 million years. This CORE-SINE locus, as well as the recently identified exapted LTR locus in placental mammals [95], function as *cis*-regulatory elements that regulate hormone activity of the proopiomelanocortin gene. Sasaki *et al.* [15] also used a transgenic mouse system to show that the AmnSINE1 loci AS071 and AS021 were exapted about 300 million years ago in a common ancestor of reptiles, birds, and mammals to a role in forebrain development. It has been proposed that this biological innovation allowed these lineages to more readily adapt to the low oxygen concentrations (~ 10%) that predominated after the Permian-Triassic mass extinction 250 million years ago [96]. Another striking example of the ability of TEs to influence the evolution of morphological novelties is that of the MER20 locus (hAT-Charlie family DNA transposon) which possibly contributed to the origin of a novel gene regulatory network dedicated to pregnancy in placental mammals [97]. These examples clearly illustrate that TEs are able to directly influence genomic and morphological evolution [13,98-100].

5. TE Research in “Under-Represented Clades from the Metazoan Tree of Life”

As the number of publicly available non-model genomes increases, so too does the diversity of bioinformatic tools designed to analyse them. Molecular biologists interested in identifying TEs in a novel genome can choose from a large number of bioinformatic methods designed to screen for these elements in large datasets [19,43,44,101–108]. Because the fundamental approaches employed by these algorithms can vary dramatically, carrying out a comparison of the results they can generate appears to be a suitable way of acquiring reliable results [19,44]. For example, a combined screening approach with PILER [102] and RepeatScout [105] was used recently to identify TEs in the first reptilian genome to be sequenced, the green anole lizard *Anolis carolinensis* [37]. TEs total 30% of this reptilian genome, and display a much wider variety of TE families than was previously recognized in the genomes of birds and mammals [109] with LINEs and SINEs being the most abundant TEs [37,110]. While many of these TEs are still active in the lizard (which probably reflects the ancestral condition of the amniote ancestor), 96 of these were identified as having been exapted in the human genome. For example, Alföldi *et al.* [37] identified a protein-coding exon that is highly conserved across 29 mammals, and was exapted from a LINE2 sequence that is now part of the MIER1 (mesoderm induction early response 1) protein in mammals. TEs are also known to be involved in the evolution of venom toxin genes in reptiles [111]. For instance, introns of the PIII-SVMP gene in the highly poisonous snake *Echis ocellatus* contain a number of different LINEs, Sauria SINEs, and a hAT transposon which may have contributed to the functional recruitment and duplication of this gene in the venom gland [112].

Further examples of the influence of TEs on the evolution of non-model genomes can be found in the loci of four *Hox* genes (the products of which control many highly conserved aspects of embryonic development across the Metazoa) in the green anole lizard. These loci are known to have massively accumulated TEs [113], mostly PLEs [114] and Sauria SINEs [35]. This is especially interesting because *Hox* genes are thought to lack TE sequences in other vertebrates [115]. Considering this

unique situation in vertebrates, and because TEs are also present in other developmental gene regions in squamate reptiles (lizards and snakes), a correlation between TE activity and the morphological diversity of squamate species has been postulated [110,113]. In a related example, the ParaHox cluster in the genome of the cephalochordate *Branchiostoma floridae* is known to be a hotspot for TE insertion [116] illustrating that not all genomic regions are equally receptive to TE invasion.

Phylogenetically broad searches for TEs based on recent genome releases have identified highly conserved metazoan retroposons. For example, de la Chaux and Wagner [117] identified BEL/Pao retroposons in 53 metazoan genomes, including the sponge genome (*Amphimedon queenslandica*) [118], and concluded that these elements evolved during early metazoan evolution. In contrast, the genome of the phylogenetically enigmatic *Trichoplax* [119] was recently shown to be relatively devoid of conserved TEs [120]. In our own efforts to identify highly conserved and ancient TEs we characterized SINE sequences with a deeply conserved domain (the Nin-domain) in the genomes of cnidarians, molluscs, annelids, and arthropods. The Nin-domain can be traced back to the origins of the Eumetazoa > 600 million years ago, making this SINE domain the most phylogenetically widespread, vertically transferred SINE sequence currently known [121]. Going further back in time, DIRS1-like and PLE retroposons are thought to have emerged during the radiation of the eukaryotes [122,123]. However, DIRS1-like TEs remain undetected in streptophytes and mammals, while PLEs have not been detected in mammals.

6. Conclusions

It is a testament to their diversity and pervasiveness that after 60 years of their discovery TEs remain scientifically topical. Even in genomic datasets that have long been available, novel TEs with high intra-genomic copy numbers and signatures of deep evolutionary conservation are still being identified. This is in part due to the relatively recent appreciation that TEs in fact have the capacity to directly influence functional genomic output. This paradigm shift away from the notion of TEs being parasitic “junk-DNA”, coupled with the growth of the fields of evo-devo, comparative genomics and concomitant technical advances in DNA sequencing technologies, has yielded great insights into the mechanisms of how complex genomes and morphological novelty can evolve. As whole genome datasets from ‘exotic’ metazoans continue to accumulate, we can expect the rate and breadth of TE discovery to accelerate and deepen. What is difficult to predict is the new ways in which TEs will be shown to interact with and influence their host genomes.

References

1. McClintock, B. A Cytological and Genetical Study of Triploid Maize. *Genetics* **1929**, *14*, 180–222.
2. McClintock, B. The order of the genes *c*, *sh* and *wx* in *zea mays* with reference to a cytologically known point in the chromosome. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 485–491.
3. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* **1950**, *36*, 344–355.
4. McClintock, B. Induction of Instability at Selected Loci in Maize. *Genetics* **1953**, *38*, 579–599.

5. Feschotte, C.; Jiang, N.; Wessler, S.R. Plant transposable elements: Where genetics meets genomics. *Nat. Rev. Gen.* **2002**, *3*, 329–341.
6. Doolittle, W.F.; Sapienza, C. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **1980**, *284*, 601–603.
7. Orgel, L.E.; Crick, F.H. Selfish DNA: The ultimate parasite. *Nature* **1980**, *284*, 604–607.
8. Gould, S.J.; Vrba, E.S. Exaptation—a missing term in the science of form. *Paleobiology* **1982**, *8*, 4–15.
9. Brosius, J. Retroposons—Seeds of evolution. *Science* **1991**, *251*, 753.
10. Miller, W.J.; McDonald, J.F.; Nouaud, D.; Anxolabehere, D. Molecular domestication—More than a sporadic episode in evolution. *Genetica* **1999**, *107*, 197–207.
11. Kidwell, M.G.; Lisch, D.R. Perspective: Transposable elements, parasitic DNA, and genome evolution. *Evolution* **2001**, *55*, 1–24.
12. Bejerano, G.; Lowe, C.; Ahituv, N.; King, B.; Siepel, A.; Salama, S.; Rubin, E.; Kent, W.; Haussler, D. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **2006**, *441*, 87–90.
13. Volff, J.N. Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays* **2006**, *28*, 913–922.
14. Feschotte, C.; Pritham, E. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* **2007**, *41*, 331–368.
15. Sasaki, T.; Nishihara, H.; Hirakawa, M.; Fujimura, K.; Tanaka, M.; Kokubo, N.; Kimura-Yoshida, C.; Matsuo, I.; Sumiyama, K.; Saitou, N.; *et al.* Possible involvement of SINEs in mammalian brain formation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4220–4225.
16. Tashiro, K.; Teissier, A.; Kobayashi, N.; Nakanishi, A.; Sasaki, T.; Yan, K.; Tarabykin, V.; Vigier, L.; Sumiyama, K.; Hirakawa, M.; *et al.* A mammalian conserved element derived from SINE displays enhancer properties recapitulating *Satb2* expression in early-born callosal projection neurons. *PLoS One* **2011**, *6*, e28497.
17. Pace, J.K.; Gilbert, C.; Clark, M.S.; Feschotte, C. Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc. Natl. Acad. Sci. USA* **2012**, *105*, 17023–17028.
18. Ainscough, R.; Bardill, S.; Barlow, K.; Basham, V.; Baynes, C.; Beard, L.; Beasley, A.; Berks, M.; Bonfield, J.; Brown, J.; *et al.* Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **1998**, *282*, 2012–2018.
19. Bergman, C.M.; Quesneville, H. Discovering and detecting transposable elements in genome sequences. *Brief. Bioinform.* **2007**, *8*, 382–392.
20. Lerat, E. Identifying repeats and transposable elements in sequenced genomes: How to find your way through the dense forest of programs. *Heredity* **2010**, *104*, 520–533.
21. Chaparro, C.; Sabot, F. Methods and software in NGS for TE analysis. *Methods Mol. Biol.* **2012**, *859*, 105–114.
22. Wicker, T.; Sabot, F.; Hua-van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; *et al.* A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982.
23. Kapitonov, V.V.; Jurka, J. A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat. Rev. Genet.* **2008**, *9*, 411–412.

24. Kapitonov, V.V.; Jurka, J. Helitrons on a roll: Eukaryotic rolling-circle transposons. *Trends Genet.* **2007**, *23*, 521–529.
25. Kapitonov, V.V.; Jurka, J. Self-synthesizing DNA transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4540–4545.
26. Ohshima, K.; Hamada, M.; Terai, Y.; Okada, N. The 3' ends of tRNA-derived short interspersed repetitive elements are derived from the 3' ends of long interspersed repetitive elements. *Mol. Cell. Biol.* **1996**, *16*, 3756–3764.
27. Okada, N.; Hamada, M.; Ogiwara, I.; Ohshima, K. SINEs and LINEs share common 3' sequences: A review. *Gene* **1997**, *205*, 229–243.
28. Kajikawa, M.; Okada, N. LINEs mobilize SINEs in the eel through a shared 3' sequence. *Cell* **2002**, *111*, 433–444.
29. Lovsin, N.; Gubensek, F.; Kordis, D. Evolutionary dynamics in a novel L2 clade of non-LTR retrotransposons in Deuterostomia. *Mol. Biol. Evol.* **2001**, *18*, 2213–2224.
30. Kordis, D.; Lovsin, N.; Gubensek, F. Phylogenomic analysis of the L1 retrotransposons in Deuterostomia. *Syst. Biol.* **2006**, *55*, 886–901.
31. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.
32. Kazazian, H.H., Jr. Mobile elements: Drivers of genome evolution. *Science* **2004**, *303*, 1626–1632.
33. Mikkelsen, T.S.; Wakefield, M.J.; Aken, B.; Amemiya, C.T.; Chang, J.L.; Duke, S.; Garber, M.; Gentles, A.J.; Goodstadt, L.; Heger, A.; *et al.* Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* **2007**, *447*, 167–177.
34. Nilsson, M.A.; Churakov, G.; Sommer, M.; Tran, N.V.; Zemmann, A.; Brosius, J.; Schmitz, J. Tracking marsupial evolution using archaic genomic retroposon insertions. *PLoS Biol.* **2010**, *8*, e1000436.
35. Piskurek, O.; Austin, C.C.; Okada, N. Sauria SINEs: Novel short interspersed transposable elements that are widespread in reptile genomes. *J. Mol. Evol.* **2006**, *62*, 630–644.
36. Kordis, D. Transposable elements in reptilian and avian (sauropsida) genomes. *Cytogenet. Genome Res.* **2009**, *127*, 94–111.
37. Alföldi, J.; Di Palma, F.; Grabherr, M.; Williams, C.; Kong, L.; Mauceli, E.; Russell, P.; Lowe, C.B.; Glor, R.E.; Jaffe, J.D.; *et al.* The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* **2011**, *477*, 587–591.
38. Hua-Van, A.; Le Rouzic, A.; Boutin, T.S.; Filé, J.; Capy, P. The struggle for life of the genome's selfish architects. *Biol. Direct.* **2011**, *17*, 6–19.
39. Dunn, C.W.D.; Hejnol, A.; Matus, D.Q.; Pang, K.; Browne, W.E.; Smith, S.A.; Seaver, E.C.S.; Rouse, G.W.; Obst, M.; Edgecombe, G.D.; *et al.* Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **2008**, *452*, 745–749.
40. Philippe, H.P.; Derelle, R.; Lopez, P.; Pick, K.; Borchellini, C.; Boury-Esnault, N.; Vacelet, J.; Renard, E.; Houliston, E.; Qu énnec, E.; *et al.* Phylogenomics revives traditional views on deep animal relationships. *Curr. Biol.* **2009**, *19*, 706–712.

41. Pick, K.; Hervé P.; Fabian, S.; Erpenbeck, D.; Jackson, D.J.; Wrede, P.; Matthias, W.; Alie, A.; Burkhard, M.; Manuel, M.; *et al.* Improved phylogenomic taxon sampling noticeably affects non-bilaterian relationships. *Mol. Biol. Evol.* **2010**, *27*, 1983–1987.
42. Ewing, A.D.; Kazazian, H.H., Jr. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res.* **2010**, *20*, 1262–1270.
43. Fiston-Lavier, A.S.; Carrigan, M.; Petrov, D.A.; González, J. T-lex: A program for fast and accurate assessment of transposable element presence using next-generation sequencing data. *Nucleic Acids Res.* **2011**, *39*, e36.
44. Flutre, T.; Duprat, E.; Feuillet, C.; Quesneville, H. Considering transposable element diversification in de novo annotation approaches. *PLoS One* **2011**, *6*, e16526.
45. Touchon, M.; Rocha, E.P.C. Causes of insertion sequences abundance in prokaryotic genomes. *Mol. Biol. Evol.* **2007**, *24*, 969–981.
46. Gladyshev, E.A.; Meselson, M.; Arkhipova, I.R. Massive horizontal gene transfer in bdelloid rotifers. *Science* **2008**, *320*, 1210–1213.
47. Gladyshev, E.A.; Arkhipova, I.R. A single-copy IS5-like transposon in the genome of a bdelloid rotifer. *Mol. Biol. Evol.* **2009**, *26*, 1921–1929.
48. Daniels, S.B.; Petterson, K.R.; Strausbaugh, L.D.; Kidwell, M.G.; Chovnick, A.C. Evidence for horizontal transmission of the P transposable elements between *Drosophila* species. *Genetics* **1990**, *124*, 339–355.
49. Robertson, H.M. The mariner transposable element is widely distributed in insects. *Nature* **1993**, *362*, 241–245.
50. Gorinsek, B.; Gubensek, F.; Kordis, D. Evolutionary genomics of chromoviruses in eukaryotes. *Mol. Biol. Evol.* **2004**, *21*, 781–798.
51. Silva, J.C.; Loreto, E.L.; Clark, J.B. Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol.* **2004**, *6*, 57–72.
52. Loreto, E.L.; Carareto, C.M.; Capy, P. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity* **2008**, *100*, 545–554.
53. Bartolomé C.; Bello, X.; Maside, X. Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. *Genome Biol.* **2009**, *10*, R22.
54. Schaack, S.; Gilbert, C.; Feschotte, C. Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **2010**, *25*, 537–546.
55. Kordis, D.; Gubensek, F. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10704–10709.
56. Zupunski, V.; Gubensek, F.; Kordis, D. Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons. *Mol. Biol. Evol.* **2001**, *18*, 1849–1863.
57. Piskurek, O.; Okada, N. Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12046–12051.
58. Houck, M.A.; Clark, J.B.; Peterson, K.R.; Kidwell, M.G. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* **1991**, *253*, 1125–1128.
59. Gilbert, C.; Schaack, S.; Pace, J.K., 2nd.; Brindley, P.J.; Feschotte, C. A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* **2010**, *464*, 1347–1350.

60. Laha, T.; Loukas, A.; Wattanasatitarp, S.; Somprakhon, J.; Kewgrai, N.; Sithithaworn, P.; Kaewkes, S.; Mitreva, M.; Brindley, P.J. The bandit, a new DNA transposon from a hookworm-possible horizontal genetic transfer between host and parasite. *PLoS Negl. Trop. Dis.* **2007**, *1*, e35.
61. Castillo, D.M.; Mell, J.C.; Box, K.S.; Blumenstiel, J.P. Molecular evolution under increasing transposable element burden in *Drosophila*: A speed limit on the evolutionary arms race. *BMC Evol. Biol.* **2011**, *11*, 258.
62. Shedlock, A.M.; Takahashi, K.; Okada, N. SINEs of speciation: Tracking lineages with retroposons. *Trends Ecol. Evol.* **2004**, *19*, 545–553.
63. Shedlock, A.M.; Okada, N. SINE insertions: Powerful tools for molecular systematics. *Bioessays* **2000**, *22*, 148–160.
64. Murata, S.; Takasaki, N.; Saitoh, M.; Okada, N. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6995–6999.
65. Nikaido, M.; Rooney, A.P.; Okada, N. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: Hippopotamuses are the closest extant relatives of whales. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10261–10266.
66. Nishihara, H.; Smit, A.F.; Okada, N. Functional noncoding sequences derived from SINEs in the mammalian genome. *Genome Res.* **2006**, *16*, 864–874.
67. Sasaki, T.; Yasukawa, Y.; Takahashi, K.; Miura, S.; Shedlock, A.M.; Okada, N. Extensive morphological convergence and rapid radiation in the evolutionary history of the family Geoemydidae (old world pond turtles) revealed by SINE insertion analysis. *Syst. Biol.* **2006**, *55*, 912–927.
68. Nikaido, M.; Piskurek, O.; Okada, N. Toothed whale monophyly reassessed by SINE insertion analysis: The absence of lineage sorting effects suggests a small population of a common ancestral species. *Mol. Phylogen. Evol.* **2007**, *43*, 216–224.
69. Nishihara, H.; Maruyama, S.; Okada, N. Retroposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5235–5240.
70. Schmitz, J.; Ohme, M.; Zischler, H. SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. *Genetics* **2001**, *157*, 777–784.
71. Kriegs, J.O.; Matzke, A.; Churakov, G.; Kuritzin, A.; Mayr, G.; Brosius, J.; Schmitz, J. Waves of genomic hitchhikers shed light on the evolution of gamebirds (Aves: Galliformes). *BMC Evol. Biol.* **2007**, *7*, 190.
72. Churakov, G.; Kriegs, J.O.; Baertsch, R.; Zemmann, A.; Brosius, J.; Schmitz, J. Mosaic retroposon insertion patterns in placental mammals. *Genome Res.* **2009**, *19*, 868–875.
73. Suh, A.; Paus, M.; Kiefmann, M.; Churakov, G.; Franke, F.A.; Brosius, J.; Kriegs, J.O.; Schmitz, J. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nat. Commun.* **2011**, *2*, 443.
74. Matzke, A.; Churakov, G.; Berkes, P.; Arms, E.M.; Kelsey, D.; Brosius, J.; Kriegs, J.O.; Schmitz, J. Retroposon insertion patterns of neoavian birds: Strong evidence for an extensive incomplete lineage sorting era. *Mol. Biol. Evol.* **2012**, *29*, 1497–1501.

75. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A.; *et al.* The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* **2009**, *326*, 1112–1115.
76. Chapman, J.A.; Kirkness, E.F.; Simakov, O.; Hampson, S.E.; Mitros, T.; Weinmaier, T.; Rattei, T.; Balasubramanian, P.G.; Borman, J.; Busam, D.; *et al.* The dynamic genome of *Hydra*. *Nature* **2010**, *464*, 592–596.
77. De Konin, A.P.; Gu, W.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* **2011**, *7*, e1002384.
78. Deininger, P.L.; Moran, J.V.; Batzer, M.A.; Kazazian, H.H., Jr. Mobile elements and mammalian genome evolution. *Curr. Opin. Genet. Dev.* **2003**, *13*, 651–658.
79. Callinan, P.A.; Batzer, M.A. Retrotransposable elements and human disease. *Genome Dyn.* **2006**, *1*, 104–115.
80. Petrov, D.A.; Fiston-Lavier, A.S.; Lipatov, M.; Lenkov, K.; González, J. Population genomics of transposable elements in *Drosophila melanogaster*. *Mol. Biol. Evol.* **2011**, *28*, 1633–1644.
81. Jordan, I.K.; Rogozin, I.B.; Glazko, G.V.; Koonin, E.V. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* **2003**, *19*, 68–72.
82. Nekrutenko, A.; Li, W.H. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* **2001**, *17*, 619–621.
83. Lunyak, V.V.; Prefontaine, G.G.; Nunez, E.; Cramer, T.; Ju, B.G.; Ohgi, K.A.; Hutt, K.; Roy, R.; Garcia-Diaz, A.; Zhu, X.; *et al.* Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* **2007**, *317*, 248–251.
84. Singer, S.S.; Mannel, D.N.; Hehlhans, T.; Brosius, J.; Schmitz, J. From “junk” to gene: Curriculum vitae of a primate receptor isoform gene. *J. Mol. Biol.* **2004**, *341*, 883–886.
85. Krull, M.; Petrusma, M.; Makalowski, W.; Brosius, J.; Schmitz, J. Functional persistence of exonized mammalian-wide interspersed repeat elements (MIRs). *Genome Res.* **2007**, *17*, 1139–1145.
86. Möller-Krull, M.; Zemmann, A.; Roos, C.; Brosius, J.; Schmitz, J. Beyond DNA: RNA editing and steps toward Alu exonization in primates. *J. Mol. Biol.* **2008**, *382*, 601–609.
87. Schmitz, J.; Brosius, J. Exonization of transposed elements: A challenge and opportunity for evolution. *Biochimie* **2011**, *93*, 1928–1934.
88. Kordis, D. Extensive intron gain in the ancestor of placental mammals. *Biol. Direct.* **2011**, *6*, 59.
89. Brosius, J.; Gould, S.J. On “genomenclature”: A comprehensive (and respectful) taxonomy for pseudogenes and other “junk DNA”. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10706–10710.
90. Bejerano, G.; Pheasant, M.; Makunin, I.; Stephen, S.; Kent, W.J.; Mattick, J.S.; Haussler, D. Ultraconserved elements in the human genome. *Science* **2004**, *304*, 1321–1325.
91. Gilbert, N.; Labuda, D. CORE-SINES: Eukaryotic short interspersed retroposing elements with common sequence motifs. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2869–2874.
92. Ogiwara, I.; Miya, M.; Ohshima, K.; Okada, N. V-SINES: A new superfamily of vertebrate SINES that are widespread in vertebrate genomes and retain a strongly conserved segment within each repetitive unit. *Genome Res.* **2002**, *12*, 316–324.
93. Akasaki, T.; Nikaido, M.; Nishihara, H.; Tsuchiya, K.; Segawa, S.; Okada, N. Characterization of a novel SINE superfamily from invertebrates: “Ceph-SINES” from the genomes of squids and cuttlefish. *Gene* **2010**, *454*, 8–19.

94. Santangelo, A.M.; de Souza, F.S.; Franchini, L.F.; Bumashny, V.F.; Low, M.J.; Rubinstein, M. Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene. *PLoS Genet.* **2007**, *3*, 1813–1826.
95. Franchini, L.F.; López-Leal, R.; Nasif, S.; Beati, P.; Gelman, D.M.; Low, M.J.; de Souza, F.J.; Rubinstein, M. Convergent evolution of two mammalian neuronal enhancers by sequential exaptation of unrelated retroposons. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15270–15275.
96. Okada, N.; Sasaki, T.; Shimogori, T.; Nishihara, H. Emergence of mammals by emergency: Exaptation. *Genes Cells* **2010**, *15*, 801–812.
97. Lynch, V.J.; Leclerc, R.D.; May, G.; Wagner, G.P. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* **2011**, *43*, 1154–1159.
98. Batzer, M.A.; Deininger, P.L. Alu repeats and human genomic diversity. *Nat. Rev. Genet.* **2002**, *3*, 370–379.
99. Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* **2008**, *9*, 397–405.
100. Schmidt, D.; Schwalie, P.C.; Wilson, M.D.; Ballester, B.; Gonçalves, A.; Kutter, C.; Brown, G.D.; Marshall, A.; Flicek, P.; Odom, D.T. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **2012**, *148*, 335–348.
101. Bao, Z.; Eddy, S.R. Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Res.* **2002**, *12*, 1269–1276.
102. Edgar, R.C.; Myers, E.W. PILER: Identification and classification of genomic repeats. *Bioinformatics* **2005**, *21*, i152–i158.
103. Quesneville, H.; Bergman, C.M.; Andrieu, O.; Autard, D.; Nouaud, D.; Ashburner, M.; Anxolabehere, D. Combined evidence annotation of transposable elements in genome sequences. *PLoS Comput. Biol.* **2005**, *1*, 166–175.
104. Li, R.; Ye, J.; Li, S.; Wang, J.; Han, Y.; Ye, C.; Wang, J.; Yang, H.; Yu, J.; Wong, G.K.; *et al.* ReAS: Recovery of ancestral sequences for transposable elements from the unassembled reads of a whole genome shotgun. *PLoS Comput. Biol.* **2005**, *1*, e43.
105. Price, A.L.; Jones, N.C.; Pevzner, P.A. De novo identification of repeat families in large genomes. *Bioinformatics* **2005**, *21*, i351–i358.
106. Ichiyanagi, K.; Okada, N. Genomic alterations upon integration of zebrafish L1 elements revealed by the TANT method. *Gene* **2006**, *383*, 108–116.
107. Feschotte, C.; Keswani, U.; Ranganathan, N.; Guibotsy, M.L.; Levine, D. Exploring repetitive DNA landscapes using REPCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. *Genome Biol. Evol.* **2009**, *1*, 205–220.
108. Churakov, G.; Grundmann, N.; Kuritzin, A.; Brosius, J.; Makalowski, W.; Schmitz, J. A novel web-based TinT application and the chronology of the Primate Alu retroposon activity. *BMC Evol. Biol.* **2010**, *10*, 376.
109. Shedlock, A.M.; Botka, C.W.; Zhao, S.; Shetty, J.; Zhang, T.; Liu, J.S.; Deschavanne, P.J.; Edwards, S.V. Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2767–2772.

110. Piskurek, O.; Nishihara, H.; Okada, N. The evolution of two partner LINE/SINE families and a full-length chromodomain-containing Ty3/Gypsy LTR element in the first reptilian whole-genome of *Anolis carolinensis*. *Gene* **2009**, *15*, 111–118.
111. Kordis, D.; Gubensek, F. Bov-B long interspersed repeated DNA (LINE) sequences are present in *Vipera ammodytes* phospholipase A2 genes and in genomes of Viperidae snakes. *Eur. J. Biochem.* **1997**, *246*, 772–779.
112. Sanz, L.; Harrison, R.A.; Calvete, J.J. First draft of the genomic organization of a PIII-SVMP gene. *Toxicon* **2012**, in press.
113. Di Poi, N.; Montoya-Burgos, J.I.; Duboule, D. Atypical relaxation of structural constraints in *Hox* gene clusters of the green anole lizard. *Genome Res.* **2009**, *19*, 602–610.
114. Evgen'ev, M.B.; Arkhipova, I.R. Penelope-like elements—A new class of retroelements: Distribution, function and possible evolutionary significance. *Cytogenet. Genome Res.* **2005**, *110*, 510–521.
115. Simons, C.; Makunin, I.V.; Pheasant, M.; Mattick, J.S. Maintenance of transposon-free regions throughout vertebrate evolution. *BMC Genomics* **2007**, *8*, 470.
116. Cañestro, C.; Albalat, R. Transposon diversity is higher in amphioxus than in vertebrates: Functional and evolutionary inferences. *Brief. Funct. Genomics* **2012**, *11*, 131–141.
117. De la Chaux, N.; Wagner, A. BEL/Pao retrotransposons in metazoan genomes. *BMC Evol. Biol.* **2011**, *11*, 154.
118. Srivastava, M.; Simakov, O.; Chapman, J.; Fahey, B.; Gauthier, M.E.; Mitros, T.; Richards, G.S.; Conaco, C.; Dacre, M.; Hellsten, U.; *et al.* The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* **2010**, *466*, 720–726.
119. Srivastava, M.; Begovic, E.; Chapman, J.; Putnam, N.H.; Hellsten, U.; Kawashima, T.; Kuo, A.; Mitros, T.; Salamov, A.; Carpenter, M.L.; *et al.* The *Trichoplax* genome and the nature of placozoans. *Nature* **2008**, *454*, 955–960.
120. Wang, S.; Zhang, L.; Meyer, E.; Bao, Z. Genome-wide analysis of transposable elements and tandem repeats in the compact placozoan genome. *Biol. Direct* **2010**, *5*, 18.
121. Piskurek, O.; Jackson, D.J. Tracking the ancestry of a deeply conserved eumetazoan SINE domain. *Mol. Biol. Evol.* **2011**, *28*, 2727–2730.
122. Arkhipova, I.R. Distribution and phylogeny of Penelope-like elements in eukaryotes. *Syst. Biol.* **2006**, *55*, 875–885.
123. Piednoël, M.; Gonçalves, I.R.; Higuier, D.; Bonnivard, E. Eukaryote DIRS1-like retrotransposons: An overview. *BMC Genomics* **2011**, *12*, 621.