Review

The Structural, Functional and Evolutionary Impact of Transposable Elements in Eukaryotes

Dareen Almojil [1], Yann Bourgeois [2], Marcin Falis [1], Imtiyaz Hariyani [1], Justin Wilcox [1,3] and Stéphane Boissinot [1,3,*]

1   New York University Abu Dhabi, Saadiyat Island, Abu Dhabi P.O. Box 129188, United Arab Emirates; da2451@nyu.edu (D.A.); mjf538@nyu.edu (M.F.); ieh211@nyu.edu (I.H.); jw5478@nyu.edu (J.W.)
2   School of Biological Sciences, University of Portsmouth, Portsmouth, UK; yann.bourgeois@port.ac.uk
3   Center for Genomics and Systems Biology, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi P.O. Box 129188, United Arab Emirates
*   Correspondence: sb5272@nyu.edu

Abstract: Transposable elements (TEs) are nearly ubiquitous in eukaryotes. The increase in genomic data, as well as progress in genome annotation and molecular biology techniques, have revealed the vast number of ways mobile elements have impacted the evolution of eukaryotes. In addition to being the main cause of difference in haploid genome size, TEs have affected the overall organization of genomes by accumulating preferentially in some genomic regions, by causing structural rearrangements or by modifying the recombination rate. Although the vast majority of insertions is neutral or deleterious, TEs have been an important source of evolutionary novelties and have played a determinant role in the evolution of fundamental biological processes. TEs have been recruited in the regulation of host genes and are implicated in the evolution of regulatory networks. They have also served as a source of protein-coding sequences or even entire genes. The impact of TEs on eukaryotic evolution is only now being fully appreciated and the role they may play in a number of biological processes, such as speciation and adaptation, remains to be deciphered.

Keywords: transposable elements; genome evolution; eukaryotes

1. Introduction

Since their early discovery by Barbara McClintock, the impact that transposable elements (TEs) may have on their hosts has been a topic of great controversy. Following McClintock's proposition that TEs may control gene action [1], a number of studies have confirmed an important role of TEs in gene regulation [2,3] and as a source of protein-coding sequences [4], both of which can benefit the host. In contrast, the observations that TEs segregate at low frequency in natural populations of many organisms [5–11] and that uncontrolled TE amplification can cause hybrid dysgenesis in *Drosophila* [12–15] suggest a negative impact of TEs on fitness. Another perspective on TEs came from the realization that differences in haploid genome size result from the differential amplification of TEs and are unrelated to the complexity of organisms. This led to the view that TEs may just be parasitic components of genomes and of little evolutionary significance (reviewed in [16]). Advances in genome sequencing, as well as progress in molecular biology methods, shed new light on the impact TEs have on eukaryotic genomes and the picture that has emerged is more nuanced and more complex than anticipated. Although the vast majority of TE insertions are, like other types of mutations, either neutral or deleterious, TEs can also serve as an extraordinary source of evolutionary novelties and have been involved in the evolution of a number of important biological processes across eukaryotes. Here we will review the impact TEs may have on their eukaryotic host in terms of genome size, structural variation, gene and genome regulation, protein-coding evolution, adaptation and

speciation, with the goal of contributing to a more holistic understanding of the interactions between TEs and their host.

2. TEs in Eukaryotic Genomes

2.1. What Are Transposable Elements?

The term "transposable element" covers a large diversity of sequences that have in common their ability to move from one genomic location to another via an active process mediated by an integrase-related enzyme. There are a number of excellent reviews describing in detail the classification of TEs (see for instance [17–20]) and we present here a brief overview of the diversity of TEs. TEs are typically classified based on their mode of transposition and sequence similarity. Elements that are mobilised via an RNA intermediate are called Class I TEs, and are generally referred to as retrotransposons, while those that propagate as DNA are Class II TEs and are called DNA transposons.

All autonomous Class I elements encode a reverse transcriptase suggesting a common origin [21] and mobilise by a copy-and-paste mechanism. They are further classified based on the presence or absence of Long Terminal Repeats (LTR). In LTR-containing retrotransposons, reverse transcription occurs in the cytoplasm (or possibly in the nucleus [22]) and the resulting cDNA is subsequently integrated into the genomic DNA. This group is widespread in eukaryotes and includes the LTR-retrotransposons sensu stricto (from which infectious and endogenised retroviruses are derived [21]) and the *DIRS* elements [23]. Non-LTR retrotransposons (which include the Long Interspersed Nuclear Elements, or LINEs, and the *Penelope* elements [24]) use a mechanism of transposition called Target Primed Reverse Transcription, in which reverse transcription occurs in the nucleus at the site of insertion [25–27]. Despite a strong *Cis* preference [28], the retrotransposition machinery of non-LTR retrotransposons can also act in *Trans* on other transcripts and mediate the transposition of non-autonomous elements, such as the Short INterspersed Elements, or SINEs, which can be derived from tRNA or other non-coding RNAs [29–31].

Class II elements include four main categories that are not evolutionarily related and do not have much in common, except that their mode of transposition does not include an RNA intermediate. The most diverse and widespread group of DNA transposons (DDE transposons or TIR transposons depending on the author) have a very simple structure consisting of a single open-reading frame (ORF) encoding for a DDE transposase [32], which recognises short flanking Terminal Inverted Repeats (TIR), cuts the element and inserts it elsewhere in the genome, hence the name cut-and-paste transposons [33]. A second group, called *Cryptons*, also use a cut and paste transposition mechanism, but in this case, the process is affected by a tyrosine recombinase [34]. Although the cut-and-paste mode of transposition involves the excision of a copy and its insertion at a different location, these elements can increase to extremely large numbers if transposition occurs during replication or if the excised site is repaired using homologous recombination with a chromosome still containing the insertion [33]. The *Helitrons* use a rolling-circle replicative mechanism mediated by a protein with an endonuclease and a helicase domain [35,36]. Finally, the *Polintons* or *Mavericks* are very large elements containing up to 20 ORFs [37,38]. They mobilise by a self-synthesizing mechanism involving direct synthesis of a DNA copy by a protein-primed polymerase B. DNA transposons also mediate the amplification of non-autonomous elements that are typically derived from autonomous copies that have suffered internal deletions [39–41]. The diversity and number of non-autonomous elements are often far greater than their autonomous relatives.

2.2. The Abundance and Diversity of TEs in Eukaryotic Genomes

Although TEs are nearly ubiquitous in eukaryotes, their abundance differs considerably among species and is largely responsible for differences in haploid genome size [42]. At one extreme, the 2.9 Mbp genome of *Encephalitozoon intestinalis* and the 23 Mbp genome of the malaria parasite *Plasmodium falciparum* are completely devoid of TEs [43,44]. At the other extreme, the 43,000 Mbp genome of the lungfish *Neoceratodus forsteri* is 90% repeti-

tive [45]. Large differences can also be found within eukaryotic lineages, as exemplified in plants, arthropods and vertebrates (Figure 1A,C,E). For instance in insects [46], the 99 Mbp genome of the Antarctic midge *Belgica antarctica* contains less than 1% TEs [47] while 65% of the 6500 Mbp genome of the migratory locust *Locusta migratoria* consists of TEs [48]. Substantial differences can even be found among closely related organisms. For instance, in the genus *Entamoeba*, the genome of *E. histolytica* contains ~20% TEs, while the proportion is ~10% in E. *dispar* [49]. Similarly, in the genus *Oryza*, the genome of the wild rice species *O. australiensis* has more than doubled in size over a 3 my period because of TE amplification resulting in a 965 Mbp genome, while the genome of some of its closest relatives is 375 Mbp [50]. In the fungus *Leptosphaeria maculans*, some strains have experienced a 30% increase in genome size due to TE amplification [51]. In contrast, some lineages exhibit only minor variation in TE content, which may reflect some constraints on genome size or, conversely, relatively high tolerance for genome expansion. Birds, for instance, have a relatively conserved genome size [52] possibly because of the metabolic cost associated with active flight [53], despite the fact that TEs are active in most bird lineages [54,55]. Salamanders, on the other hand, have experienced extreme and independent TE amplification resulting in gigantic genomes, ranging from 14,000 to 120,000 Mbp [56–58].
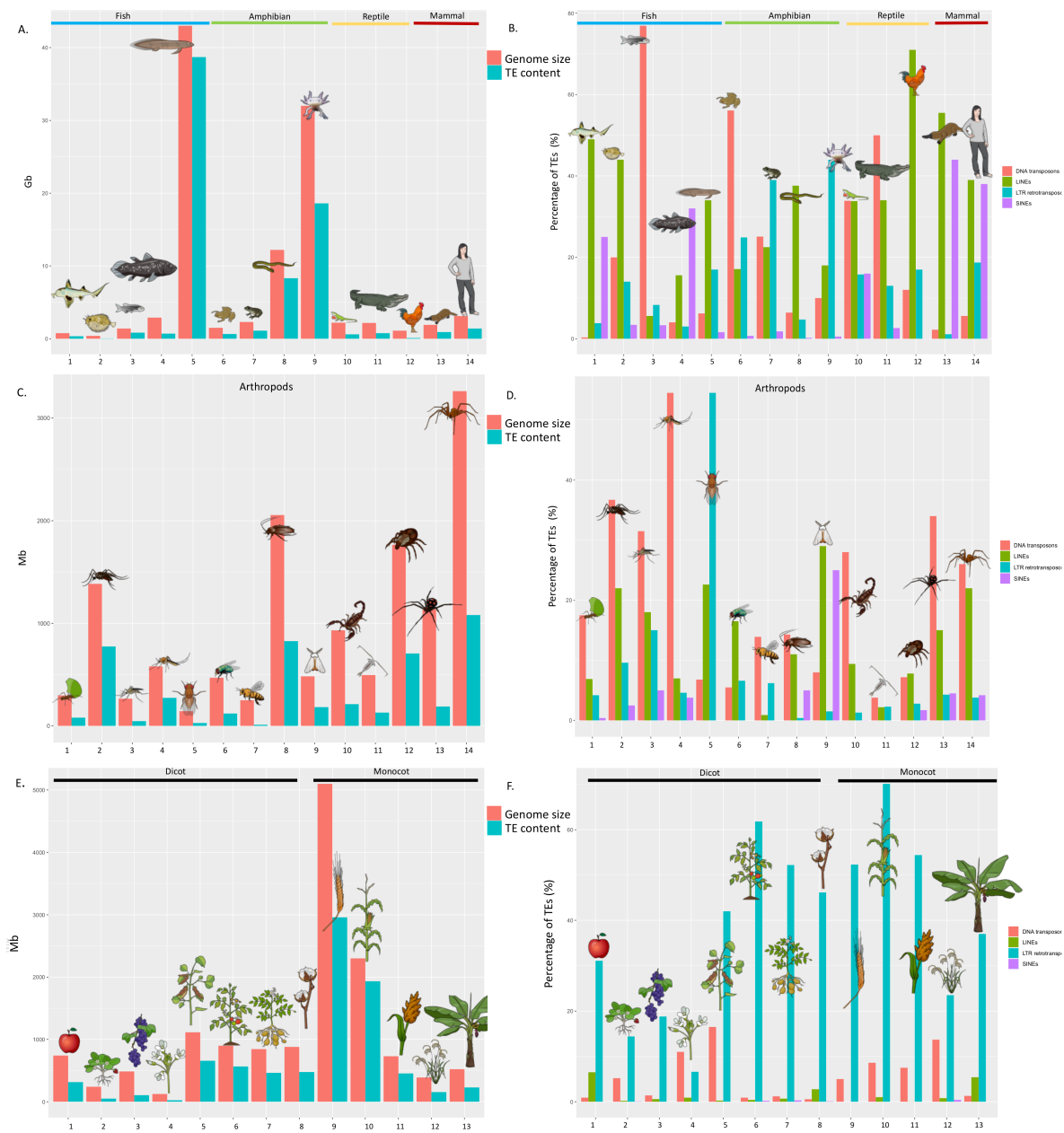
Figure 1. Genome size and its TE content across different organisms. (**A**) Genome size and its TE content in Gb across different vertebrates, from left to right: (1) *Callorhinchus milii*, (2) *Tetraodon nigroviridis*, (3) *Danio rerio*, (4) *Latimeria chalumnae*, (5) *N. forsteri*, (6) *Xenopus tropicalis*, (7) *Nanorana parkeri*, (8) *Ichthyophis bannanicus*, (9) *Ambystoma mexicanum*, (10) *Anolis carolinensis*, (11) *Alligator mississipiensis*, (12) *Gallus gallus*, (13) *Ornithorhynchus anatinus*, and (14) *Homo sapiens*. (**B**) Percentages of TEs types (DNA transposons, LINEs, LTR retrotransposons, and SINEs) of the vertebrates presented in (**A**). (**C**) Genome size and its TE content in Mb across different arthropods, from left to right: (1) *Acromyrmex echinatior*, (2) *Aedes aegypti*, (3) *Anopheles gambiae*, (4) *Culex quinquefasciatus*, (5) *Drosophila melanogaster*, (6) *Lucilia cuprina*, (7) *Apis mellifera*, (8) *Blattella germanica*, (9) *Bombyx mori*, (10) *Centruroides exilicauda*, (11) *Eurytemora affinis*, (12) *Ixodes scapularis*, (13) *Latrodectus Hesperus* and 14) *Loxosceles reclusa*. (**D**) Percentages of TEs types of the arthropods presented in (**C**). (**E**) Genome size and its TE content in Mb across different plants, from left to right: (1) *Malus domestica* (2) *Fragaria vesca*, (3) *Vitus vinifera*, (4) *Arabidopsis thaliana*, (5) *Glycine max*, (6) *Solanum lycopersicum*, (7) *Solanum tuberosum*, (8) *Gossypium raimondii*, (9) *Hordeum vulgare* (10) *Zea mays*, (11) *Sorghum bicolor*, (12) *Oryza sativa* and (13) *Musa acuminate*. (**F**) percentages of TEs types of the plants presented in (**E**). Data from [45,46,56,59–64].

TE diversity also differs considerably among organisms (Figure 1B,D,F). Within vertebrates, fish, amphibian and reptilian genomes typically harbour a large diversity of TEs, with both class I and class II elements [59,65,66], each being represented by a myriad of families, as exemplified by the genome of the green anole *A. carolinensis* [40,67]. In contrast, the genome of most birds and mammals is dominated by a single type of non-LTR retrotransposon *CR1* in bird [54], *L1* in placental mammals [17,66] and *L2* in monotremes [68]. In plants, TE diversity is also variable among species but differences in genome size result mostly from the differential amplification of LTR retrotransposons. Differences in diversity can also be observed at a finer scale and the relative proportion of different types of TEs can evolve rapidly. For instance, in the *Heliconius* butterfly radiation, the relative representation of different categories of TEs has changed markedly [69]. Interestingly, the relationship between TE diversity and abundance is not straightforward. Genome size and TE diversity are positively correlated up to a genome size of 500 Mbp [70], although some very compact genomes harbour very few TE families, such as the baker's yeast [71]. Genomes that are larger than 500 Mbp tend to have a lower TE diversity due to the amplification of a specific type of element, which accounts for most of the genome mass.

These comparisons of TE content across the eukaryotic tree of life raise the question: why do genomes differ so much in their TE composition and consequently in their genome size? It is easier to understand why some genomes remained small despite the long-term persistence of TE activity. In a compact genome, such as the one of *D. melanogaster*, new euchromatic insertions are more likely to be deleterious and are therefore eliminated by purifying selection, rarely reaching high population frequency or fixation [10]. Because selection against TEs is dependent on the length and copy number of the insertions [72], the more likely cause for the deleterious effect of new insertions results from their ability to mediate ectopic recombination events [10]. It is, however, more difficult to explain the evolutionary increase in genome size, which can result in genome gigantism in some groups. Genetic drift is a process that will counteract the effect of selection and may favour the fixation of slightly deleterious TE insertions. This is supported by the fact that organisms with large effective population size tend to have smaller genomes than organisms with small population size, in which drift is stronger [73]. However, the resulting accumulation of TEs will pose additional problems because the increase in the number of copies increases the chance of ectopic recombination. In addition, if a genome contains a low diversity of TEs but if these TEs are represented by many copies the chance of ectopic recombination will be higher than in a genome with a large diversity of TEs represented by a small number of copies [74]. A possible solution to this problem is the evolution of a lower rate of ectopic recombination in genomes that contain a large number of similar copies, possibly via increased surveillance by the mismatch repair system [74,75]. Since ectopic recombination events can be drivers of DNA loss [76–78], a higher rate of DNA loss, via large deletions, is expected in species with a diverse repertoire of TEs represented by small copy numbers (as in most non-mammalian vertebrates) than in genomes containing many similar copies (such as mammals and salamanders). This hypothesis is supported by a faster decay of non-LTR retrotransposon insertions caused by large deletions in reptiles and fish than in mammals [67,79]. Consistent with the hypothesis above, the gigantic genome of some salamanders and caecilian is characterised by an extremely reduced rate of DNA loss and ectopic recombination [56,80,81]. Interestingly, the rate of transposition and the rate of DNA loss can counteract each other perfectly, resulting in the long-term stability in genome size observed in birds and mammals, a process termed the "accordion" model of genome evolution [82].

2.3. The Population Dynamics of TEs

The abundance, diversity and genomic distribution of TEs depend on several factors, which include the number of copies produced (the transposition rate) and the number of copies that accumulate in the genome, with the latter depending on population processes such as selection and drift. The impact of TEs on their host's fitness is a key parameter

to understand their population dynamics. TEs are more likely to rise to high frequency in populations and accumulate in the genome if they have little to no deleterious effect. However, the impact of TEs on host fitness is often neutral or negative and depends on (1) whether TEs insert in functional regions and display regulatory features that affect nearby genes, (2) their abundance, which may trigger aberrant meiosis through ectopic recombination, and (3) whether TEs are involved in the build-up of genetic conflicts and reproductive isolation.

The first aspect, insertional mutagenesis, is supported by many examples, such as disease-causing insertions in several organisms, including humans [83]. In general, there are less TEs inserting in coding regions than expected if insertions occurred at random [10,84]. TEs such as retrotransposons also contain promoter sequences that may alter the expression of genes at their vicinity (see Section 3.3). TEs can also trigger epigenetic silencing extending to genes nearby (e.g., [85]), or produce RNAs or proteins that induce damages. For example, competition of TEs with host genes for transcription factors [86] may lead to a loss in fitness. Another mechanism by which TEs may have a negative impact on their host's fitness is ectopic recombination between non-allelic copies, which can lead to deleterious chromosomal rearrangements. Support for this scenario comes from the absence of full-length elements from highly recombining regions of genomes [84,87,88], and their accumulation in non-recombining regions [89–91]. Large families of TEs are more likely to be deleterious than smaller ones, since more aberrant pairings may occur during meiosis [10,72,92]. Heterozygous TEs are expected to be involved in ectopic recombination more often because of the lack of an allelic copy on the other chromosome [93]. Overall, this implies that the effects of ectopic recombination may be maximal for long, recently active TEs with many copies found at low frequency in populations. Finally, TEs may have an impact on the fitness of hybrids coming from divergent populations. The recent invasion of the *P*-element in *D. melanogaster* [12,13] and in *D. simulans* [9,94,95], which results in hybrid dysgenesis between individuals carrying the element or not is a classic example. Such mechanisms generate genomic conflicts and may promote speciation by reducing the fitness of hybrids.

This is not to say that TEs only have deleterious effects. TEs can also be recruited by positive selection, for example, because of their regulatory effects on neighbouring genes. This is exemplified by one of the most iconic cases of natural selection, the industrial melanism in peppered moths (*Biston betularia*), which is caused by a TE insertion at the *cortex* locus [96]. Cases of selective advantage have also been detected in *Drosophila*, where TE insertions are involved in resistance to insecticides [97] or response to oxidative stress [98], and in plants, where TE insertions are associated with adaptation to a broad range of environments [99–101]. However, it remains unclear how common adaptive insertions are. In a scan for adaptive TE in *D. melanogaster*, hundreds of putative adaptive insertions were identified [102] while similar population genetics approaches in *A. thaliana* [103] *D. suzukii* [104], human [105] and the lizard *A. carolinensis* [106] revealed much smaller numbers (less than 20) of positively selected insertions.

The likelihood for a TE to have an impact on fitness strongly depends on its intrinsic properties, such as insertion preference or transposition dynamics. These properties of TEs can be seen as the equivalent of life-history traits for species adapted to a specific ecological niche [107]. These niches often differ in their biochemical properties, such as DNA conformation, chromatin state or CpG content [108,109]. Some TE families [110] and endoviruses [108] are found more frequently near recombination hotspots, which may ultimately make them more deleterious if they engage in ectopic recombination. This might be due to a shared preference of recombination and transposition machineries for open chromatin [111]. In *Caenorhabditis elegans*, the cut-and-paste mechanism of transposition of DNA transposons may benefit from the double-stranded breaks that initiate recombination events, explaining their abundance near recombination hotspots [112]. Some TEs target specific DNA motifs, such as the ones found in *Drosophila* telomeres [113].

The abundance of TEs in a genome depends on the rate of transposition of progenitors, and the rate at which their copies become unable to transpose. This inactivation can be caused by the host's control mechanisms or be due to incomplete duplication and truncation. These factors vary strongly across families and hosts. Moreover, the transposition rate varies over time, with waves of transposition associated with high activity followed by quieter periods, possibly due to the activation of the host's control mechanisms [55,114–117]. These bursts produce cohorts of TEs having approximately the same age. Recent bursts of transposition may explain the excess of rare TEs observed in *Drosophila*, in association with purifying selection [118]. On the other hand, TEs produced by ancient bursts of transposition are generally at a much higher frequency or even fixed. This means that non-constant transposition may produce a signature in the TE frequency spectrum that mimic expectations under scenarios of purifying selection (recent burst) or balancing selection (ancient burst).

All the factors detailed above impact the host's fitness, and the selective coefficient associated with each TE insertion, s. However, the efficiency of selection also depends on genetic drift, which becomes stronger as the effective population size N_e decreases: if the product $N_e s$ is lower than 1, the insertion will behave as if effectively neutral. The effective population size is highly variable across species and strongly depends on their demographic history [8,119,120]. Recent population bottlenecks reduce N_e and may lower the efficiency of selection against TEs. In addition, demographic changes will shape the allele frequency spectrum of TEs [6]: reductions in N_e should lead to a genome-wide excess of alleles at intermediate frequencies, while population expansions may lead to an excess of rare insertions. A complicating factor is that estimators of N_e vary along the genome, due to the effects of selection at linked sites. Regions near a positively or negatively selected site display lower allelic diversity, and a higher rate of fixation of neutral and deleterious variants. This effect becomes stronger in regions of low recombination and is analogous to a local reduction in N_e [121]. This leads to an accumulation of TEs at both ends of the frequency spectrum: more TEs will reach fixation in the region of low recombination than in regions of high recombination due to Hill-Robertson interference [122], but the remaining polymorphic TEs will also be at a lower frequency on average. In addition, since selection is less effective in regions of low recombination, they may accumulate TEs that are more deleterious, on average, than the ones inserting in regions of high recombination.

The interaction between selection, linked selection, transposition and preferential insertion has a strong influence on where TEs accumulate and reach high frequencies in the genome. It may explain the broad diversity of correlations between TE density or frequency and various genomic features. For example, in the green anole, simulations combining linked selection and preferential insertion in regions of high recombination reproduced the correlations observed between TE density, frequency, and recombination for near-neutral TEs [106]. While linked selection led to lower average TE frequency in regions of low recombination, preferential insertion led to a higher TE density in regions of high recombination, contrary to expectations under linked selection alone.

3. The Impact of TEs on Eukaryotic Genes and Genomes

3.1. The Impact of TEs on Structural Variation

TEs vary in their mechanisms and molecular processes of transposition, but they are unified by their capacity to produce structural variation. While definitions of structural variation have varied over time, contemporary definitions consider structural variation as insertions, deletions, inversions, or translocations, equal to or greater than 50 base pairs [123]. As the smallest active DNA transposons (MITEs) and retrotransposons (i.e., SINEs) are approximately 100 and 300 bp, respectively [33,124], TEs are structural variants. Depending on the mechanism of replication, they may produce additional structural variation at the site of insertion or the site of origin, in the case of cut-and-paste TEs. Most, if not all, TEs can also replicate host sequences [125], so structural variation arising from TEs can be diverse in origin, and due to the myriad of replicative mechanisms

among TEs, diverse in form. While typified by insertion-deletion events, TE-associated structural variation commonly includes inversions and can be a source of copy number variations, large deletions, duplications, and translocations [126–128]. Beyond these basics, TE-associated structural variation is best understood in the context of specific TE types and their lifecycles.

Non-LTR retrotransposons can transduce their flanking DNA. Transduction of the 5′ flanking sequence can occur if the upstream flanking sequence contains an active promoter, which drives transcription and is eventually reverse-transcribed together with the element [129]. These cases are very rare. More common is the transduction of the 3′ flanking sequence due to the weak polyadenylation signal contained at the 3′ end of elements. In this case, transcription of the element continues until the next polyadenylation signal, taking with it the downstream sequence, which is reverse transcribed with the element [130]. The transduction of the 3′ flanking sequence, up to 1.6 Kbp long, is found associated with 20% of human *L1* elements [131,132]. LTR-retrotransposons can also mobilise non-TE DNA, but this typically occurs by transport of DNA (up to ~4 kb) captured between terminal repeats [125,133]. In addition, non-LTR retrotransposons can dramatically alter their site of insertion, resulting in structural variation. The insertion of an *L1* element is typically associated with the formation of a target site duplication, which, in some rare cases, can reach a few 100 bp in length [134]. The insertion of *L1* and *Alu* elements can also produce large deletions, up to a few Kbp, duplications or inter-chromosomal translocations [134–138]. Finally, non-LTR retrotransposons constitute a powerful source of microsatellites derived from the 3′ terminus of elements [139]. This phenomenon is particularly pronounced in some snakes whose genomes contain the highest density of microsatellites known in vertebrates [65]. These microsatellites are believed to be derived from the repetitive 3′ end of *CR1* LINE elements.

DNA transposons can give rise to a wide diversity of structural variation [33]. The classic "cut-and-paste" subclass of DNA transposons causes insertions at their target sites and can cause deletions or other anomalies through excision at the site of origin [140]— although many likely regenerate at the donor site by double-stranded-break-induced homologous recombination and concordant gene conversion to eliminate the deletion site at the origin [141,142]. "Cut-and-paste" transposons can also capture host DNA sequences between their terminal inverted repeats during transposition, and can generally exceed retroelements in terms of maximum transported bases, with some types being capable of relocating 10–20 kb of additional host DNA [125]. In rice, it was shown that a large number of genic fragments were mobilised and amplified by the *Pack-MULE* DNA transposon [143]. *Helitrons* use a rolling-circle replication [144], contingent upon a single-stranded DNA break and excision at the origin site. This would seem to exclude deletion variants at the origin site—as the single strand is typically replaced—but creates a seemingly unrivalled capacity for replication of host DNA sequences downstream from the origin site, although this is typically limited to a maximum of a few Kbp [125,145]. The *Mavericks/Polintons* [37,38] subclass are very large DNA transposons that are typically around 20 kb and can reach 40 kb. The capacity of their independent DNA polymerases to perform long-range amplification of up to approximately 45 kb allows them to potentially move very large fragments of host DNA along with them [125].

While the movement and replication of TEs is an obvious source of structural variation, TEs can cause structural variation through several indirect processes, particularly those related to DNA repair and recombination [75,126,146,147]. These processes can have a profound impact on genomic stability but can also be an important source of evolutionary novelty, for instance by mediating segmental duplications [148]. TEs can drive ectopic recombination events when similar TE sequences align in non-homologous locations of the genome. When this occurs on the same chromosome it can cause major chromosomal rearrangements, such as large inversions [149,150], and when it occurs between different chromosomes it can cause translocations. Ectopic recombination events between closely located TEs can also cause any number of large localised structural variants [126]. For

example, if two nearby TEs of the same orientation align, this can cause a reciprocal deletion on one chromatid and duplication on the other. If the repeats are in inverted orientation this can cause an inversion of the flanked sequences. Non-reciprocal crossover between non-homologous repeats can likewise result in large uncoupled deletions or duplications. Similar processes can occur when the end-joining of broken DNA occurs between non-homologous regions through single-stranded annealing or microhomology-mediated end-joining of shared repetitive sequences [147]. TEs also directly cause strand breaks during insertion [75] and excision, and can indirectly cause strand breaks during DNA replication due to secondary structures formed during DNA replication [127]; these breaks can be associated with recombination and homologous end-joining associated duplication and deletion events [146].

3.2. The Impact of TEs on the Genomic Landscape

As a consequence of their effect on fitness, TEs are not uniformly distributed in genomes and their accumulation in some regions and their absence from others has a defining impact on genomic landscapes. Not surprisingly, there is a general tendency for TEs to accumulate in gene-poor and low recombining regions because their negative effect is reduced there (see Section 2.3) [60,89,151–154]. As the negative effect of TEs is often length-dependent, long elements, in particular, will accumulate in low or non-recombining regions [89,91]. The long term accumulation of TEs in gene-poor regions will eventually increase the length of intergenic regions and overall physical distances within genomes. This process is exemplified in human where long stretches of intergenic sequences are composed almost exclusively of TE sequences, mostly from the *L1* family, while other regions, such as the HoxD gene cluster and recombination hotspots are almost completely devoid of TEs [60,88]. Interestingly, *Alu* elements, which are SINEs relying on *L1* for their mobility, have a different genomic distribution than *L1* and tend to be distributed in genic regions [154], possibly because of an insertion bias or selective processes. The clustering of TE-derived sequences in a specific genomic location can result in the evolution of satellites, as exemplified by the accumulation of sequences derived from the *Bari* DNA transposon in *D. melanogaster* [155].

The accumulation of TEs can potentially affect the regional base composition of genomes since TEs often harbour base compositions that are substantially different from the genome average [156–158]. In lizard and fish, for instance, the GC content of non-LTR retrotransposons ranges from 33% to 55% while the genome average is ~41% [158]. Even larger differences were reported when comparing class I and class II elements in fish [159]. An accumulation of AT-rich TEs, such as the mammalian *L1*, in a specific region of the genome, could have a substantial effect on the heterogeneity in base composition and could contribute to the isochoric organisation of some vertebrate genomes, although the main determinant of base composition heterogeneity is biased gene conversion [160]. In fungi of the genus *Leptosphaeria*, the differential amplification of AT-rich TEs resulted in substantial differences in the overall genomic base composition among strains, with ~45% GC in strains that have experienced TE amplification and 51% in strains that have not [51].

Although the recombination rate is one of the main determinants of TE distribution in genomes, the interplay between TEs and recombination is quite intricate [161]. If the local recombination rate determines the probability of fixation of a TE, TEs in turn can modify the recombination rate. TE activity tends to be deleterious to the host and is thus repressed epigenetically. TE repression can cause chromatin compaction, which in turn represses recombination since recombination initiation more often occurs in an open chromatin state [160]. This model is supported by the correlation between the methylation status of chromatin and recombination rate [161–163], by the experimental disruption of methylation that results in increased recombination [164,165] and by the repression of recombination triggered by specific insertions [166,167]. The reduced recombination rate caused by TEs can eventually affect the dynamics of linked alleles in the genome and a recent study proposed that polymorphic TEs increase linkage disequilibrium and consequently the size

of chromosomal blocks subject to selection [168]. In some other cases, it was shown that TEs can increase the recombination rate locally, as was observed in the vicinity of *Alu* elements in human [169], possibly because *Alu* sequences contain recombinogenic motifs.

3.3. The Impact of TEs on Gene Regulation

TEs have been shown to influence patterns of gene expression in multiple ways and scientists are just beginning to appreciate the considerable impact TE-derived sequences have had on the regulation of host genes [2,3,170–173]. Although most studies in this field come from model organisms, TEs likely constitute the main source of regulatory novelty in some lineages such as primates [174,175]. To maintain their ability to move around the genome, TEs encode intrinsic regulatory sequences necessary for their own transcription. These sequences can, in turn, affect the expression of host genes. In addition, regulatory elements may evolve by exaptation, i.e., the evolution of a new function out of non-adaptive material [176]. This can occur when TE-derived sequences accumulate mutations, which can eventually evolve into a regulatory element. A number of studies in the past two decades have demonstrated that TEs constitute an important source of *cis*-regulatory motifs in the form of transcription factor binding sites (TFBS), distal enhancer, suppressor or insulator. For instance, a substantial fraction of TFBS in mammalian genomes was shown to be derived from TEs [177–179]. These include binding sites for TFs implicated in immune response such as *STAT1* [180] or pregnancy such as *cAMP* [181]. In addition to a direct effect on genes, some TFBS generated by TEs can affect the three-dimensional organization of chromatin and thus participate in the regulation of genes that are not in close proximity. This process is exemplified in mammals where many of the *CTCF* TFBS, which play an important role in the definition of topologically associated domains (TADs), are derived from TEs [182,183]. A similar effect of TEs on the spatial organization of the genome is found in plants where miniature inverted-repeat transposable elements (MITEs) are associated in rice and sorghum with the presence of matrix attachment regions, which serve as anchors for loops [184].

The spread of TFBS tends to be lineage-specific and thus promotes regulatory divergence among lineages [177], but the invasion of different eukaryotic lineages by the same horizontally-transferred TE can also result in the independent spread of the same TFBS among distantly related organisms [185]. In some cases, TEs can contribute directly to the core promoter sequence of genes and replace pre-existing promoters [186–190]. Bioinformatic and experimental evidence have now demonstrated a direct effect of TE-derived regulatory sequences on a vast number of genes, such as *POMC*, an important brain-expressed gene which controls food intake in human [191]; *Agouti*, a colour-coat gene in mice [192]; *ZmCCT9*, a transcription factor regulating flowering time in *Z. mays* [99] and *Zmr1*, a gene that controls melanisation in *Zymoseptoria tritici*, a trait that affects the survival of this fungus under stressful conditions [193]. It should be noted however that, although substantial, the recruitment of TEs in gene regulation remains a rare occurrence and there is a tendency for TEs to be under-represented upstream of genes [194]. The impact of TEs on gene regulation also differs considerably among TE categories. For instance, SINEs and LTR retrotransposons seem to have been recruited disproportionately as gene regulators in mammals, compared to non-LTR retrotransposons [2,195,196].

The effect of TEs on gene regulation may be particularly significant in the evolution of complex functions that require the concerted regulation of multiple genes interconnected in a regulatory network. A gene regulatory network is a framework of interactions between molecular regulators (e.g., transcription factors) and substrates (e.g., transcription factors binding sites) that orchestrate the expression of genes involved in complex biological processes [197]. It has been suggested that the evolution of gene regulatory networks can be induced by a wave of TEs invading and amplifying within a eukaryotic lineage [3,198]. This suggestion is based on the observed associations between expansions of specific TE families and the dispersal of regulatory elements, which coincides with major evolutionary events, such as the emergence of mammalian pregnancy [181], and the evolution of the mammalian

neocortex [199]. It remains possible that insertion bias or selection favours the retention of specific TEs in genes that share similar profiles of expression but that the TEs have a minimal impact on the expression of those genes. However, there is a growing number of studies demonstrating a direct role of TEs in the evolution of gene regulatory networks [3,178] which are involved in important biological processes such as pluripotency [200], innate immunity [180], placentation [181,201], early mammalian embryonic development [202], and dosage compensation on sex chromosomes [203]. In *A. thaliana* and in *Z. mays*, the activation of stress-sensitive TEs affects the expression of nearby genes and may be involved in the evolution of stress-response regulatory networks [204,205]. Not only are TEs involved in the 'wiring' of connections between genes but they can also participate in the 're-wiring, and thus the evolution, of these regulatory networks [178,206,207]. One of the mechanisms that enable TEs to rewire regulatory networks could be through non-allelic gene conversion between TE copies [208,209], which could aid the spread of beneficial mutations, and allows selection to improve the functionality of the network [208,210]. As a mechanism, non-allelic gene conversion between copies of TEs is suggested to allow a faster share of beneficial mutations across different TE copies than the whole rewiring process by transposition alone [208].

There are many other ways TE insertions can affect the expression of genes, independently from their ability to provide specific regulatory motifs. TEs inserted in introns can reduce the efficacy of transcription by inserting premature polyadenylation sites or because the unusual base composition of the TE sequence reduces the amount of transcript being produced, as demonstrated for the A–rich L1 element of mammals [211]. TEs can also introduce splicing sites that can result in the production of incomplete or chimeric transcripts [212]. These effects are reflected in the biased orientation of intronic TEs which tend to be in the orientation that has the least negative effect on the transcription of the gene they are inserted in [154,213]. TEs inserted in the untranslated regions (UTRs) of genes can also have a profound impact on their regulation. For instance, a *POGON1* DNA transposon insertion in the 3′UTR of the *CG11699* gene in *D. melanogaster* produces a shorter 3′UTR which is adaptive and confers resistance to xenobiotic stress [214]. Another example is found in the flowering plant *Capsella rubella* where a *Helitron* insertion in the 3′UTR of the *FLC* gene affects mRNA stability and result in a reduced expression level of *FLC* [215]. In mouse and human, the presence of SINEs in the 3′UTR of a gene can result in the formation of a STAU1 binding site and target the mRNA to *Staufen*-mediated decay, thus downregulating expression [216,217]. Finally, TEs can play a role in gene regulation in *Trans* by participating in the evolution of regulatory non-coding RNAs (such as miRNA, piRNA or long non-coding RNAs) [2,218]. A number of miRNAs are derived from TEs (~20% in the human genome [199]) and TE-derived sequences confer the functionality of different types of small and long non-coding RNAs [2].

An indirect way TEs can affect the regulation of genes results from their epigenetic repression by the host. The activity of TEs is deleterious in general and a number of mechanisms have evolved to silence TEs in eukaryotes, and one of the known strategies for silencing TEs is performed at the epigenetic level [219–222]. Epigenetic silencing of TEs can be performed through different mechanisms, including DNA methylation, which is arguably the most widely used strategy by higher eukaryotes for repressing TEs [223,224], the KRAB-zinc finger proteins (KZFPs) [225], and RNA interference by PIWI-interacting RNAs (piRNAs) or small interfering RNAs (siRNAs) [226]. These different mechanisms can repress TEs but can also affect host genes. This effect is more pronounced if TEs are at a close distance from host genes, but epigenetic repressive marks on TEs can extend beyond 20 kb from a repressed TE insertion [227]. In *A. thaliana* the repression of TEs by DNA methylation extends 500 base pairs on each side of the element [228], and in many cases can affect neighbouring genes [229–232]. In mice and *Drosophila*, enrichment level with the repressive H3K9me2 mark decreases with increased distance from polymorphic TE insertions [227,233]. Since the spreading of repressive marks from the selected TE insertions may influence the expression of nearby genes, it is likely to influence the

fitness of the host [234,235]. It is thus expected that TEs with repressive marks acting on functional genes will be deleterious, and therefore would be eliminated by purifying selection. Interestingly, genome-wide population studies of TEs in plants show consistency with this prediction, as unmethylated TEs are usually closer to genes than methylated TEs [234,236]. Moreover, studies that examined population frequencies of TEs reported that low-frequency TEs are more likely to be TEs with higher enrichment of repressive mark at neighbouring genes [227,229], suggesting that these TEs are more strongly selected against. The unintended consequence of the epigenetic repression of TEs led to the intriguing idea that epigenetic repression is a "double-edged sword", which is beneficial to the host as a defence mechanism but also has a cost since the repression of TEs will affect gene expression [221].

3.4. The Impact of TEs on Protein-Coding Sequence

TE-derived sequences can also significantly impact the protein-coding capabilities of eukaryotic genomes, either by providing motifs that can become embedded in the protein-coding sequence of host genes, by becoming a protein-coding gene through a process called domestication or by mediating the duplication of host genes. TE activity may lead to the development of evolutionary novelties through gene fusion events—phenomena, in which a single transcriptional unit, known as a chimeric gene, originates from different source genes [237,238]. Several TE-mediated fusion mechanisms have been proposed, including gene capture by LTR-retrotransposons or DNA transposons (*Pac-MULEs*, *Helitrons*), retroposition, and reverse end transposition [237,239]. These mechanisms can create novel sequence combinations through the juxtaposition of exons from different genes (i.e., exon shuffling) or insertion of retrocopies within introns of existing genes. Numerous examples of chimeric genes resulting from the activity of TEs have been identified in a variety of eukaryotes. In a recent study, 94 cases of fusion genes containing a transposase motif were discovered in vertebrates [240] demonstrating that this mechanism of gene evolution is relatively common. Interesting examples of fusions include *TRIM5-CypA*, an antiviral gene that evolved independently in New and Old World monkeys [241–243], the primate *SETMAR* gene that plays a role in histone methylation and DNA repair [244,245] and the *KRABINER* transcription factor in bats [240]. The *KRABINER* example is particularly telling because it illustrates the features of the transposases that make it particularly prone to recruitment in a function that necessitates recognition and binding of specific DNA motifs [240].

Under certain circumstances, the host may harness parts or whole sequences of TEs and repurpose them for its own benefits in a process known as molecular domestication. Genome-wide studies documented numerous examples of functionally important non-coding [218,246] and coding sequences that had been co-opted from a variety of TEs [247,248], including transposases, the critical enzymes used by DNA transposons for their mobility. One of the most iconic examples of transposase domestication is the case of the recombination-activating genes (RAGs), which play a pivotal role in V(D)J somatic recombination, a fundamental element of the adaptive immune response of jawed vertebrates [249]. Guided by the localization of recombination signal sequences (RSSs), *RAG1* and *RAG2* facilitate the separation, rearrangement, and rejoining of DNA segments in numerous permutations, creating a diverse reservoir of T- and B-cell antibodies. While the origin of *RAG1* has been traced to transposases encoded by the *Transib* family of DNA transposons [250,251], the origin of *RAG2* remained uncertain until the recent discovery of an active transposase in the genome of the lancelet, which contains both *RAG1-* and *RAG2*-like genes in the correct orientation flanked by terminal inverted repeats (TIRs) that resemble the RSSs [251,252]. Another famous example of transposase domestication comes from plants with the *FAR1* and *PHY3* genes, which are transcription factors that modulate light signalling [253]. Interestingly, independent domestication of transposases derived from the same TEs can result in the convergent evolution of genes with similar function. This is the case of centromere-binding proteins in fission yeast and mammals that have

convergently evolved from a transposase encoded by a *pogo* element [254]. Domestication is not limited to DNA transposons and numerous co-options of retrotransposon-derived sequences have been documented. The domestication of the *L1* reverse transcriptase has been documented as exemplified by the *TERT* gene, which plays a role in the maintenance of telomere length [255]. Numerous genes have been classified as derivatives of proteases, integrases, *pol*, *gag* and *env* derived from LTR retrotransposons [247]. Several of these products have been linked to the formation and function of the placenta, a transient organ facilitating interactions between the mother and the developing foetus. *Env*-derived *syncytins* have been independently co-opted numerous times across the mammalian tree of life, with homologs described in both eutherians [256–258] and marsupials [259]. In rodents, certain *syncytins* play a pivotal role in the formation of the placenta, mediating the fusion of maternal and fetal trophoblasts [260]. Furthermore, some of them are known to possess immunosuppressive properties, which are hypothesised to play a role in defending the foetus from the maternal immune system [261].

The examples of molecular domestication cited above resulted from the modification of TE-derived protein-coding sequences from TEs for a new purpose. However, in some rare cases, the transposition activity itself has been domesticated such as in the case of *Drosophila*, where TEs are involved in facilitating the elongation of telomeres at chromosome ends. Since chromosome ends may lose nucleotides with every DNA replication cycle (also known as the "end replication problem"), organisms have evolved various mechanisms to prevent the shortening of chromosomes and retain their telomeres. Several eukaryotic organisms use the telomerase enzyme, which counteracts chromosome erosion by promoting the synthesis of repetitive sequences to the ends of chromosomes [262]. In contrast, some *Drosophila* species like *D. melanogaster*, have recruited specialised transposable elements from the *Jockey* subfamily (*HeT-A*, *TART* and *TAHRE* families) to protect their telomeres by transposing to the ends of the chromosomes [263–266]. While these retrotransposons appear to have evolved in a symbiotic manner, another hypothesis emphasises the selfish nature of these TEs, suggesting that they selfishly over-replicate in these genomes [267]. This over-replication may lead to a diversification of the specialised TEs and an eventual disappearance in species such as *Drosophila biarmipes* [266]. Another example of domestication with a direct functional impact on chromosomal evolution is the role of TEs in maintaining the stability of eukaryotic centromeres and the silencing of heterochromatin [267]. For instance, the genome of the social amoeba *Dictyostelium discoideum*, whose centromere is rich in TEs (86%, 171–361 bp), is dominated by LTR retrotransposons [268]. The centromere and the surrounding region containing the tRNA genes appear to harbour several copies of these TEs, which have a strong insertional preference for the region as a result of low mutagenesis and the presence of relatively safe integration sites [269]. This hypothesised domestication of TEs is thought to be involved in a complex control network that comprises LTR retrotransposon families (*Skipper* and *DGLT-A*), the RNAi machinery, and centromeric histones that regulate retrotransposition and kinetochore assembly [269]. Thus, TEs in the centromere of *D. discoideum* may have co-evolved with its host to provide the necessary substrate for kinetochore complex formation in return for the ability to accumulate in the genome near "low-danger" tRNA genes [270].

Another way that TEs can contribute to the emergence of new genes is by mediating the duplication of host genes. The 'birth' of a gene is associated with a duplication event, in which a gene copy emerges as a result of replication (DNA-mediated) or reverse transcription (RNA-mediated) [237,271]. In the latter process, known as retroposition, the enzymatic machinery of active retrotransposons or retroviruses orchestrates the reverse transcription of an mRNA transcript and its integration into the genome, leading to the emergence of a retrocopy (see [271,272] for mechanisms). Retrocopies are intron-less, may contain polyA tails at the 3' end, and, in the case of LTR-mediated retroposition, flanking LTR regions [239,271]. Due to the initial bias in sequencing primarily mammalian genomes, early studies of retroposition focused on understanding the role of non-LTR retrotransposons in mediating the process. While the contribution of some members of

this group, particularly *L1* elements, is well-understood [273], there is growing evidence of the occurrence of LTR-mediated retroposition, with evidence from plants [274] and animals [275]. Importantly, not all non-LTR and LTR retrotransposons are capable of effective retroposition, as evident by the example of *CR1*s, the predominant retroelements in bird genomes, which are generally unable to generate retrocopies [276]. There are several evolutionary pathways that can be followed by a retroposed gene copy. The vast majority of retrocopies exist in the form of retropseudogenes [212], non-functional relics of the parental gene that had acquired frameshift mutations and premature stop codons [237]. While pseudogenization is commonly associated with loss of function, it does not necessarily represent a "dead-end" for a gene. There is growing evidence that neutrally-decaying retropseudogenes and non-functional retrocopies may serve as components of functionally important non-coding RNAs, such as miRNAs [277], siRNAs [278] and lncRNAs [218]. However, some retrocopies can maintain their protein-coding abilities and are referred to as retrogenes. This process has led to the increase in the copy number of ribosomal protein genes in fission yeast, which may improve glucose consumption rate through enhanced ribosome biogenesis [279]. Another interesting example of retrocopy evolution is represented by the cell-cycle gene *CDC14B*, which encodes a protein that stabilises microtubules, and its ape-specific retrocopy dubbed *CDC14Bretro* [280]. In the African ape lineage (gorilla-chimpanzee-human), the protein product of the retrocopy has been found to occur in association with the endoplasmic reticulum, indicating acquisition of a novel function, while in the Asian ape lineage (orangutan and gibbon), the retrogene was found to associate with microtubules, similar to the original copy of *CDC14B* [280]. This retention of the parental gene function represents another scenario of retrogene evolution, which occurs if the increased expression of the encoded product provides a fitness advantage or the functionality of the parental gene is lost due to pseudogenization or deletion. The latter case leads to the emergence of the so-called "orphan" retrogenes, which have been identified in several studies [281,282].

4. The Role of TEs in the Diversification of Life

The previous sections of this review have described the numerous mechanisms by which TEs may impact their host genomes. Either through exaptation or positive selection on new insertions, TEs have been an important source of evolutionary novelties and have been instrumental in the evolution of biological processes that have affected the diversification of eukaryotes [283]. For instance, many of the traits that are responsible for the evolutionary success of mammals have been profoundly impacted by TEs, from placenta formation [182,202] to the evolution of a complex brain [199,284]. TEs could also affect the diversification of life by directly driving the formation of species [285,286]. The role of TEs in speciation remains controversial and it is unclear to what extent TEs actually contribute to the processes that lead to reproductive isolation. A role of TEs in speciation is supported by the correlation between speciation and bursts of TE amplification. In a number of taxa, including rodents [117,287], bats [288], and woodpeckers [55], rapid speciation correlates with the amplification of TEs, leading to the question of whether one causes the other. Because TEs amplify in bursts and could thus trigger species formation, as well as the evolution of novel phenotypes, over a short period of time, TE amplification could provide an explanation for the observation that evolution often proceeds in jumps (at the time of a burst of TE amplification) followed by a period of stasis (when TE activity is low), a model called the "punctuated equilibrium" theory of evolution [289].

Although the correlation between the rate of speciation and transposition seems solid, at least in some lineages [290], it remains a correlation and it does not, in itself, provide a mechanism. In fact, this observation is consistent both with the idea that TEs are drivers of speciation [291,292] but also with the view that TE amplification is a consequence of the speciation process [293]. The "TE thrust" hypothesis [291] proposes that through active (transposition and integration) or passive (ectopic recombination) thrust, TEs can cause significant structural changes in the genome, allowing the development of phenotypic

diversity, which enables populations to adapt and occupy new ecological niches. Consider the event of a bottleneck in the presence of abiotic differences, where effective population size is low, and there is a loss of genetic variation due to the founder effect. In line with the genomic shock model proposed by McClintock [294], TEs may mobilise in response to environmental stress, possibly because of disruption of epigenetic repression [285,292], and allow an increase in genetic diversity, with neutral or slightly deleterious TEs actively transposing and becoming fixed in the genome. Abiotic differences may drive ecological speciation and the diversification of lineages, with TE thrust playing a key role in the realization of adaptive potential [291]. The TE thrust hypothesis has been invoked to explain the difference in diversity of gymnosperms and angiosperms [63]. Angiosperms have younger and more active TE families in comparison with extant gymnosperms and this could explain the lack of evolutionary innovations and speciation in gymnosperms. The TE Thrust hypothesis, as well as the epi-transposon hypothesis [292] present TEs as drivers of speciation, suggesting a causative role of TEs in achieving speciation.

However, the carrier subpopulation (CASP) hypothesis argues from a neutralist perspective [293]. It proposes that TE amplification may be a byproduct of speciation when populations subdivide, leading to the random assortment, fixation and amplification of TEs due to genetic drift. Thus, neutral or slightly deleterious active TE families may have a large impact on genome evolution if they amplify into larger families, thereby resulting in TE-mediated diversification, which may further lead to speciation by increasing the chances of reproductive isolation. This hypothesis implies that the TE landscape of a species may be indicative of structural changes in populations over time and may reflect recent or ongoing subdivisions in the population that may be associated with speciation events. We caution here that these hypotheses are not mutually exclusive—TEs may be largely drifters accompanying demes that result in new species, but at the same time, some TEs may have properties that are adaptive and may drive ecological speciation by allowing populations to evolve independently.

Another indication that TEs may play a role in speciation comes from their potential ability to promote the evolution of reproductive isolation. TEs can mediate genomic divergence between nascent species, for instance by causing chromosomal rearrangements [295], which could result in hybrid inviability or sterility [296,297]. Another possibility is an increased activity of TEs in hybrids [298–301]. Since the early discovery of hybrid dysgenesis in *Drosophila* [13–15], the uncontrolled amplification of TEs in hybrids between populations or species that differ in their ability to regulate specific transposable elements provides an appealing mechanism for the evolution of post-zygotic reproductive isolation. The disruption of the epigenetic repression of TEs observed in some crosses could certainly result in reduced viability or fertility in the hybrid [285], yet the number of documented cases remains very low and several known examples involve artificial crosses between relatively distant species [302,303]. In fact, recent hybrids between incipient yeast species [304] or between well-differentiated species [304,305] have failed to detect any evidence of TE de-repression. Thus the role of TEs in mediating post-zygotic isolation remains unclear and will require further studies [306].

5. Conclusions

The goal of this review was to emphasise the considerable number of ways in which TEs can impact the genome of their host. With the exponential increase of genomic data and improvement of annotation methods, researchers will likely discover new ways by which eukaryotic genomes have been affected by TEs. There are however a number of questions that remained open and that will benefit from the use of novel analytical and experimental approaches. Although TEs are subject to the same evolutionary forces as any type of mutation, the range of effects produced by TEs likely differ and will require further investigation. Overall, further work is needed to quantify the distribution of fitness effects of TEs relative to other types of variation. In particular, the proportion of TEs that are under the effect of selection may vary substantially, depending on their

intrinsic properties and their host's dynamics. Future work may gain from the study of the distribution of fitness effect of coding variation [307], which will provide a theoretical framework upon which models incorporating TE specificities could be built. Such models including important aspects of TE biology will certainly benefit from the development of fast and efficient simulators (e.g., [308]), but also functional studies in laboratory models. Genomics analyses have generated many hypotheses about the functional impact of TEs that are still in need of experimental validation. Approaches that link the distribution of TEs with functional analyses are particularly promising [309] and the increasing use of the CRISPR-Cas9 technology should provide experimental validation on the impact of specific TE insertions [180,310,311].

References

1. McClintock, B. Controlling Elements and the Gene. *Cold Spring Harb. Symp. Quant. Biol.* **1956**, *21*, 197–216. [CrossRef] [PubMed]
2. Ali, A.; Han, K.; Liang, P. Role of Transposable Elements in Gene Regulation in the Human Genome. *Life* **2021**, *11*, 118. [CrossRef] [PubMed]
3. Chuong, E.B.; Elde, N.C.; Feschotte, C. Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **2017**, *18*, 71–86. [CrossRef] [PubMed]
4. Sinzelle, L.; Izsvak, Z.; Ivics, Z. Molecular domestication of transposable elements: From detrimental parasites to useful host geness. *Cell. Mol. Life Sci.* **2009**, *66*, 1073–1093. [CrossRef] [PubMed]
5. Ruggiero, R.P.; Bourgeois, Y.; Boissinot, S. LINE insertion polymorphisms are abundant but at low frequencies across populations of Anolis carolinensis. *Front. Genet.* **2017**, *8*, 44. [CrossRef]
6. Xue, A.T.; Ruggiero, R.P.; Hickerson, M.J.; Boissinot, S. Differential effect of selection against LINE retrotransposons among vertebrates inferred from whole-genome data and demographic modeling. *Genome Biol. Evol.* **2018**, *10*, 1265–1281. [CrossRef]
7. Boissinot, S.; Davis, J.; Entezam, A.; Petrov, D.; Furano, A.V. Fitness cost of LINE-1 (L1) activity in humans. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9590–9594. [CrossRef]
8. Lockton, S.; Ross-Ibarra, J.; Gaut, B.S. Demography and weak selection drive patterns of transposable element diversity in natural populations of Arabidopsis lyrata. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13965–13970. [CrossRef]
9. Kofler, R.; Hill, T.; Nolte, V.; Betancourt, A.J.; Schlötterer, C. The recent invasion of natural Drosophila simulans populations by the P-element. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6659–6663. [CrossRef]
10. 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. [CrossRef] [PubMed]
11. Stritt, C.; Gordon, S.P.; Wicker, T.; Vogel, J.P.; Roulin, A.C. Recent Activity in Expanding Populations and Purifying Selection Have Shaped Transposable Element Landscapes across Natural Accessions of the Mediterranean Grass Brachypodium distachyon. *Genome Biol. Evol.* **2018**, *10*, 304–318. [CrossRef] [PubMed]
12. Daniels, S.B.; Peterson, K.R.; Strausbaugh, L.D.; Kidwell, M.G.; Chovnick, A. Evidence for horizontal transmission of the P transposable element between Drosophila species. *Genetics* **1990**, *124*, 339–355. [CrossRef] [PubMed]
13. Kidwell, M.G. Hybrid dysgenesis in Drosophila melanogaster: The relationship between the P–M and I–R interaction systems. *Genet. Res.* **1979**, *33*, 205–217. [CrossRef]
14. Kidwell, M.G.; Kidwell, J.F.; Sved, J.A. Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. *Genetics* **1977**, *86*, 813–833. [CrossRef]
15. Picard, G. Non-mendelian female sterility in Drosophila melanogaster: Hereditary transmission of I factor. *Genetics* **1976**, *83*, 107–123. [CrossRef]

16. Gregory, T.R. Coincidence, coevolution, or causation? DNA content, cellsize, and the C-value enigma. *Biol. Rev.* **2001**, *76*, 65–101. [CrossRef] [PubMed]
17. Tollis, M.; Boissinot, S. The evolutionary dynamics of transposable elements in eukaryote genomes. *Repetitive DNA* **2012**, *7*, 68–91.
18. Wells, J.N.; Feschotte, C. A Field Guide to Eukaryotic Transposable Elements. *Annu. Rev. Genet.* **2020**, *54*, 539–561. [CrossRef]
19. 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. [CrossRef] [PubMed]
20. Kojima, K.K. Structural and sequence diversity of eukaryotic transposable elements. *Genes Genet. Syst.* **2018**, *94*, 233–252. [CrossRef] [PubMed]
21. Eickbush, T.H.; Malik, H.S. Origins and evolution of retrotransposons. In *Mobile DNA II*; American Society of Microbiology: Washington, DC, USA, 2002; pp. 1111–1144.
22. Selyutina, A.; Persaud, M.; Lee, K.; KewalRamani, V.; Diaz-Griffero, F. Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. *Cell Rep.* **2020**, *32*, 108201. [CrossRef]
23. Goodwin, T.J.; Poulter, R.T. The DIRS1 group of retrotransposons. *Mol. Biol. Evol.* **2001**, *18*, 2067–2082. [CrossRef]
24. Arkhipova, I.R. Distribution and phylogeny of Penelope-like elements in eukaryotes. *Syst. Biol.* **2006**, *55*, 875–885. [CrossRef] [PubMed]
25. Luan, D.D.; Korman, M.H.; Jakubczak, J.L.; Eickbush, T.H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. *Cell* **1993**, *72*, 595–605. [CrossRef]
26. Cost, G.J.; Feng, Q.; Jacquier, A.; Boeke, J.D. Human L1 element target-primed reverse transcription in vitro. *EMBO J.* **2002**, *21*, 5899–5910. [CrossRef] [PubMed]
27. Ichiyanagi, K.; Okada, N. Mobility pathways for vertebrate L1, L2, CR1, and RTE clade retrotransposons. *Mol. Biol. Evol.* **2008**, *25*, 1148–1157. [CrossRef] [PubMed]
28. Wei, W.; Gilbert, N.; Ooi, S.L.; Lawler, J.F.; Ostertag, E.M.; Kazazian, H.H.; Boeke, J.D.; Moran, J.V. Human L1 retrotransposition: Cispreference versus trans complementation. *Mol. Cell. Biol.* **2001**, *21*, 1429–1439. [CrossRef]
29. Dewannieux, M.; Esnault, C.; Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.* **2003**, *35*, 41–48. [CrossRef]
30. Dewannieux, M.; Heidmann, T. LINEs, SINEs and processed pseudogenes: Parasitic strategies for genome modeling. *Cytogenet. Genome Res.* **2005**, *110*, 35–48. [CrossRef]
31. Ohshima, K.; Okada, N. SINEs and LINEs: Symbionts of eukaryotic genomes with a common tail. *Cytogenet. Genome Res.* **2005**, *110*, 475–490. [CrossRef]
32. Doak, T.G.; Doerder, F.P.; Jahn, C.L.; Herrick, G. A proposed superfamily of transposase genes: Transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 942–946. [CrossRef] [PubMed]
33. Feschotte, C.; Pritham, E.J. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* **2007**, *41*, 331–368. [CrossRef] [PubMed]
34. Kojima, K.K.; Jurka, J. Crypton transposons: Identification of new diverse families and ancient domestication events. *Mob. DNA* **2011**, *2*, 1–17. [CrossRef] [PubMed]
35. Kapitonov, V.V.; Jurka, J. Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8714–8719. [CrossRef] [PubMed]
36. Kapitonov, V.V.; Jurka, J. Helitrons on a roll: Eukaryotic rolling-circle transposons. *Trends Genet.* **2007**, *23*, 521–529. [CrossRef]
37. Kapitonov, V.V.; Jurka, J. Self-synthesizing DNA transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4540–4545. [CrossRef] [PubMed]
38. 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**, *390*, 3–17. [CrossRef] [PubMed]
39. Hartl, D.; Lozovskaya, E.; Lawrence, J. Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica* **1992**, *86*, 47–53. [CrossRef]
40. Novick, P.A.; Smith, J.D.; Floumanhaft, M.; Ray, D.A.; Boissinot, S. The evolution and diversity of DNA transposons in the genome of the lizard Anolis carolinensis. *Genome Biol. Evol.* **2011**, *3*, 1–14. [CrossRef]
41. Yang, G.; Nagel, D.H.; Feschotte, C.; Hancock, C.N.; Wessler, S.R. Tuned for transposition: Molecular determinants underlying the hyperactivity of a Stowaway MITE. *Science* **2009**, *325*, 1391–1394. [CrossRef]
42. Kidwell, M.G. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **2002**, *115*, 49–63. [CrossRef]
43. Gardner, M.J.; Hall, N.; Fung, E.; White, O.; Berriman, M.; Hyman, R.W.; Carlton, J.M.; Pain, A.; Nelson, K.E.; Bowman, S.; et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **2002**, *419*, 498–511. [CrossRef]
44. Corradi, N.; Pombert, J.F.; Farinelli, L.; Didier, E.S.; Keeling, P.J. The complete sequence of the smallest known nuclear genome from the microsporidian Encephalitozoon intestinalis. *Nat. Commun.* **2010**, *1*, 77. [CrossRef]
45. Meyer, A.; Schloissnig, S.; Franchini, P.; Du, K.; Woltering, J.M.; Irisarri, I.; Wong, W.Y.; Nowoshilow, S.; Kneitz, S.; Kawaguchi, A.; et al. Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature* **2021**, *590*, 284–289. [CrossRef]
46. Petersen, M.; Armisen, D.; Gibbs, R.A.; Hering, L.; Khila, A.; Mayer, G.; Richards, S.; Niehuis, O.; Misof, B. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evol. Biol.* **2019**, *19*, 11. [CrossRef]

47. Kelley, J.L.; Peyton, J.T.; Fiston-Lavier, A.S.; Teets, N.M.; Yee, M.C.; Johnston, J.S.; Bustamante, C.D.; Lee, R.E.; Denlinger, D.L. Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nat. Commun.* **2014**, *5*, 4611. [CrossRef] [PubMed]
48. Wang, X.; Fang, X.; Yang, P.; Jiang, X.; Jiang, F.; Zhao, D.; Li, B.; Cui, F.; Wei, J.; Ma, C.; et al. The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* **2014**, *5*, 2957. [CrossRef] [PubMed]
49. Lorenzi, H.; Thiagarajan, M.; Haas, B.; Wortman, J.; Hall, N.; Caler, E. Genome wide survey, discovery and evolution of repetitive elements in three Entamoeba species. *BMC Genom.* **2008**, *9*, 595. [CrossRef]
50. Piegu, B.; Guyot, R.; Picault, N.; Roulin, A.; Sanyal, A.; Kim, H.; Collura, K.; Brar, D.S.; Jackson, S.; Wing, R.A.; et al. Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in Oryza australiensis, a wild relative of rice. *Genome Res.* **2006**, *16*, 1262–1269. [CrossRef]
51. Grandaubert, J.; Lowe, R.G.; Soyer, J.L.; Schoch, C.L.; Van de Wouw, A.P.; Fudal, I.; Robbertse, B.; Lapalu, N.; Links, M.G.; Ollivier, B. Transposable element-assisted evolution and adaptation to host plant within the Leptosphaeria maculans-Leptosphaeria biglobosa species complex of fungal pathogens. *BMC Genom.* **2014**, *15*, 891. [CrossRef]
52. Zhang, G.; Li, B.; Li, C.; Gilbert, M.T.; Jarvis, E.D.; Wang, J.; Avian Genome, C. Comparative genomic data of the Avian Phylogenomics Project. *Gigascience* **2014**, *3*, 26. [CrossRef] [PubMed]
53. Wright, N.A.; Gregory, T.R.; Witt, C.C. Metabolic 'engines' of flight drive genome size reduction in birds. *Proc. Biol. Sci.* **2014**, *281*, 20132780. [CrossRef] [PubMed]
54. Kapusta, A.; Suh, A. Evolution of bird genomes-a transposon's-eye view. *Ann. N. Y. Acad. Sci.* **2017**, *1389*, 164–185. [CrossRef]
55. Manthey, J.D.; Moyle, R.G.; Boissinot, S. Multiple and independent phases of transposable element amplification in the genomes of piciformes (woodpeckers and allies). *Genome Biol. Evol.* **2018**, *10*, 1445–1456. [CrossRef]
56. Wang, J.; Itgen, M.W.; Wang, H.; Gong, Y.; Jiang, J.; Li, J.; Sun, C.; Sessions, S.K.; Lockridge Mueller, R. Gigantic Genomes Provide Empirical Tests of Transposable Element Dynamics Models. *Genom. Proteom. Bioinform.* 2021; in press. [CrossRef]
57. Sun, C.; Shepard, D.B.; Chong, R.A.; Lopez Arriaza, J.; Hall, K.; Castoe, T.A.; Feschotte, C.; Pollock, D.D.; Mueller, R.L. LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome Biol. Evol.* **2012**, *4*, 168–183. [CrossRef]
58. Nowoshilow, S.; Schloissnig, S.; Fei, J.F.; Dahl, A.; Pang, A.W.C.; Pippel, M.; Winkler, S.; Hastie, A.R.; Young, G.; Roscito, J.G.; et al. The axolotl genome and the evolution of key tissue formation regulators. *Nature* **2018**, *554*, 50–55. [CrossRef]
59. Chalopin, D.; Naville, M.; Plard, F.; Galiana, D.; Volff, J.-N. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. *Genome Biol. Evol.* **2015**, *7*, 567–580. [CrossRef] [PubMed]
60. 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.
61. Shao, F.; Han, M.; Peng, Z. Evolution and diversity of transposable elements in fish genomes. *Sci. Rep.* **2019**, *9*, 15399. [CrossRef] [PubMed]
62. Sun, Y.B.; Xiong, Z.J.; Xiang, X.Y.; Liu, S.P.; Zhou, W.W.; Tu, X.L.; Zhong, L.; Wang, L.; Wu, D.D.; Zhang, B.L.; et al. Whole-genome sequence of the Tibetan frog Nanorana parkeri and the comparative evolution of tetrapod genomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1257–E1262. [CrossRef]
63. Oliver, K.R.; McComb, J.A.; Greene, W.K. Transposable elements: Powerful contributors to angiosperm evolution and diversity. *Genome Biol. Evol.* **2013**, *5*, 1886–1901. [CrossRef] [PubMed]
64. Wang, K.; Wang, Z.; Li, F.; Ye, W.; Wang, J.; Song, G.; Yue, Z.; Cong, L.; Shang, H.; Zhu, S.; et al. The draft genome of a diploid cotton Gossypium raimondii. *Nat. Genet.* **2012**, *44*, 1098–1103. [CrossRef]
65. Pasquesi, G.I.M.; Adams, R.H.; Card, D.C.; Schield, D.R.; Corbin, A.B.; Perry, B.W.; Reyes-Velasco, J.; Ruggiero, R.P.; Vandewege, M.W.; Shortt, J.A.; et al. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. *Nat. Commun.* **2018**, *9*, 2774. [CrossRef] [PubMed]
66. Sotero-Caio, C.G.; Platt, R.N., 2nd; Suh, A.; Ray, D.A. Evolution and Diversity of Transposable Elements in Vertebrate Genomes. *Genome Biol. Evol.* **2017**, *9*, 161–177. [CrossRef] [PubMed]
67. Novick, P.A.; Basta, H.; Floumanhaft, M.; McClure, M.A.; Boissinot, S. The evolutionary dynamics of autonomous non-LTR retrotransposons in the lizard Anolis carolinensis shows more similarity to fish than mammals. *Mol. Biol. Evol.* **2009**, *26*, 1811–1822. [CrossRef] [PubMed]
68. Warren, W.C.; Hillier, L.W.; Marshall Graves, J.A.; Birney, E.; Ponting, C.P.; Grutzner, F.; Belov, K.; Miller, W.; Clarke, L.; Chinwalla, A.T.; et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **2008**, *453*, 175–183. [CrossRef] [PubMed]
69. Ray, D.A.; Grimshaw, J.R.; Halsey, M.K.; Korstian, J.M.; Osmanski, A.B.; Sullivan, K.A.M.; Wolf, K.A.; Reddy, H.; Foley, N.; Stevens, R.D.; et al. Simultaneous TE Analysis of 19 Heliconiine Butterflies Yields Novel Insights into Rapid TE-Based Genome Diversification and Multiple SINE Births and Deaths. *Genome Biol. Evol.* **2019**, *11*, 2162–2177. [CrossRef] [PubMed]
70. Elliott, T.A.; Gregory, T.R. Do larger genomes contain more diverse transposable elements? *BMC Evol. Biol.* **2015**, *15*, 69. [CrossRef]
71. Bleykasten-Grosshans, C.; Friedrich, A.; Schacherer, J. Genome-wide analysis of intraspecific transposon diversity in yeast. *BMC Genom.* **2013**, *14*, 399. [CrossRef]
72. Petrov, D.A.; Aminetzach, Y.T.; Davis, J.C.; Bensasson, D.; Hirsh, A.E. Size matters: Non-LTR retrotransposable elements and ectopic recombination in Drosophila. *Mol. Biol. Evol.* **2003**, *20*, 880–892. [CrossRef]
73. Lynch, M.; Conery, J.S. The origins of genome complexity. *Science* **2003**, *302*, 1401–1404. [CrossRef]

74. Furano, A.V.; Duvernell, D.D.; Boissinot, S. L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. *Trends Genet.* **2004**, *20*, 9–14. [CrossRef]

75. Hedges, D.J.; Deininger, P.L. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. *Mutat. Res.* **2007**, *616*, 46–59. [CrossRef]

76. Ma, J.; Devos, K.M.; Bennetzen, J.L. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res.* **2004**, *14*, 860–869. [CrossRef]

77. Jedlicka, P.; Lexa, M.; Kejnovsky, E. What Can Long Terminal Repeats Tell Us About the Age of LTR Retrotransposons, Gene Conversion and Ectopic Recombination? *Front. Plant Sci.* **2020**, *11*, 644. [CrossRef]

78. Sen, S.K.; Han, K.; Wang, J.; Lee, J.; Wang, H.; Callinan, P.A.; Dyer, M.; Cordaux, R.; Liang, P.; Batzer, M.A. Human genomic deletions mediated by recombination between Alu elements. *Am. J. Hum. Genet.* **2006**, *79*, 41–53. [CrossRef]

79. Blass, E.; Bell, M.; Boissinot, S. Accumulation and rapid decay of non-LTR retrotransposons in the genome of the three-spine stickleback. *Genome Biol. Evol.* **2012**, *4*, 687–702. [CrossRef]

80. Sun, C.; Lopez Arriaza, J.R.; Mueller, R.L. Slow DNA loss in the gigantic genomes of salamanders. *Genome Biol. Evol.* **2012**, *4*, 1340–1348. [CrossRef]

81. Frahry, M.B.; Sun, C.; Chong, R.A.; Mueller, R.L. Low levels of LTR retrotransposon deletion by ectopic recombination in the gigantic genomes of salamanders. *J. Mol. Evol.* **2015**, *80*, 120–129. [CrossRef]

82. Kapusta, A.; Suh, A.; Feschotte, C. Dynamics of genome size evolution in birds and mammals. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1460–E1469. [CrossRef]

83. Hancks, D.C.; Kazazian, H.H. Roles for retrotransposon insertions in human disease. *Mobile DNA* **2016**, *7*, 1–28. [CrossRef] [PubMed]

84. Kofler, R.; Betancourt, A.J.; Schlötterer, C. Sequencing of pooled DNA samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in Drosophila melanogaster. *PLoS Genet.* **2012**, *8*, e1002487. [CrossRef]

85. Morgan, H.D.; Sutherland, H.G.; Martin, D.I.; Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **1999**, *23*, 314–318. [CrossRef]

86. Conte, C.; Dastugue, B.; Vaury, C. Promoter competition as a mechanism of transcriptional interference mediated by retrotransposons. *EMBO J.* **2002**, *21*, 3908–3916. [CrossRef] [PubMed]

87. Nam, K.; Ellegren, H. Recombination drives vertebrate genome contraction. *PLoS Genet.* **2012**, *8*, e1002680. [CrossRef] [PubMed]

88. Myers, S.; Bottolo, L.; Freeman, C.; McVean, G.; Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **2005**, *310*, 321–324. [CrossRef] [PubMed]

89. Boissinot, S.; Entezam, A.; Furano, A.V. Selection against deleterious LINE-1-containing loci in the human lineage. *Mol. Biol. Evol.* **2001**, *18*, 926–935. [CrossRef] [PubMed]

90. Charlesworth, B. The organization and evolution of the human Y chromosome. *Genome Biol.* **2003**, *4*, 226. [CrossRef]

91. Song, M.; Boissinot, S. Selection against LINE-1 retrotransposons results principally from their ability to mediate ectopic recombination. *Gene* **2007**, *390*, 206–213. [CrossRef]

92. Montgomery, E.; Charlesworth, B.; Langley, C.H. A test for the role of natural selection in the stabilization of transposable element copy number in a population of Drosophila melanogaster. *Genet. Res.* **1987**, *49*, 31–41. [CrossRef] [PubMed]

93. Montgomery, E.; Huang, S.; Langley, C.; Judd, B. Chromosome rearrangement by ectopic recombination in Drosophila melanogaster: Genome structure and evolution. *Genetics* **1991**, *129*, 1085–1098. [CrossRef] [PubMed]

94. Hill, T.; Schlötterer, C.; Betancourt, A.J. Hybrid dysgenesis in Drosophila simulans associated with a rapid invasion of the P-element. *PLoS Genet.* **2016**, *12*, e1005920.

95. Yoshitake, Y.; Inomata, N.; Sano, M.; Kato, Y.; Itoh, M. The P element invaded rapidly and caused hybrid dysgenesis in natural populations of Drosophila simulans in Japan. *Ecol. Evol.* **2018**, *8*, 9590–9599. [CrossRef] [PubMed]

96. Van't Hof, A.E.; Campagne, P.; Rigden, D.J.; Yung, C.J.; Lingley, J.; Quail, M.A.; Hall, N.; Darby, A.C.; Saccheri, I.J. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **2016**, *534*, 102–105. [CrossRef] [PubMed]

97. Daborn, P.; Yen, J.; Bogwitz, M.; Le Goff, G.; Feil, E.; Jeffers, S.; Tijet, N.; Perry, T.; Heckel, D.; Batterham, P. A single P450 allele associated with insecticide resistance in Drosophila. *Science* **2002**, *297*, 2253–2256. [CrossRef] [PubMed]

98. Guio, L.; Barrón, M.G.; González, J. The transposable element Bari-Jheh mediates oxidative stress response in Drosophila. *Mol. Ecol.* **2014**, *23*, 2020–2030. [CrossRef]

99. Huang, C.; Sun, H.; Xu, D.; Chen, Q.; Liang, Y.; Wang, X.; Xu, G.; Tian, J.; Wang, C.; Li, D. ZmCCT9 enhances maize adaptation to higher latitudes. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E334–E341. [CrossRef]

100. Kanazawa, A.; Liu, B.; Kong, F.; Arase, S.; Abe, J. Adaptive evolution involving gene duplication and insertion of a novel Ty1/copia-like retrotransposon in soybean. *J. Mol. Evol.* **2009**, *69*, 164–175. [CrossRef]

101. Chu, C.; Tan, C.; Yu, G.; Zhong, S.; Xu, S.; Yan, L. A novel retrotransposon inserted in the dominant Vrn-B1 allele confers spring growth habit in tetraploid wheat (*Triticum turgidum* L.). *G3 Genes Genomes Genet.* **2011**, *1*, 637–645. [CrossRef]

102. Rech, G.E.; Bogaerts-Marquez, M.; Barrón, M.G.; Merenciano, M.; Villanueva-Cañas, J.L.; Horváth, V.; Fiston-Lavier, A.-S.; Luyten, I.; Venkataram, S.; Quesneville, H. Stress response, behavior, and development are shaped by transposable element-induced mutations in Drosophila. *PLoS Genet.* **2019**, *15*, e1007900. [CrossRef]

103. Li, Z.-W.; Hou, X.-H.; Chen, J.-F.; Xu, Y.-C.; Wu, Q.; González, J.; Guo, Y.-L. Transposable elements contribute to the adaptation of Arabidopsis thaliana. *Genome Biol. Evol.* **2018**, *10*, 2140–2150. [CrossRef]

104. Mérel, V.; Gibert, P.; Buch, I.; Rada, V.R.; Estoup, A.; Gautier, M.; Fablet, M.; Boulesteix, M.; Vieira, C. The worldwide invasion of Drosophila suzukii is accompanied by a large increase of transposable element load and a small number of putatively adaptive insertions. *bioRxiv* **2020**. [CrossRef]
105. Rishishwar, L.; Wang, L.; Wang, J.; Yi, S.V.; Lachance, J.; Jordan, I.K. Evidence for positive selection on recent human transposable element insertions. *Gene* **2018**, *675*, 69–79. [CrossRef]
106. Bourgeois, Y.; Ruggiero, R.; Hariyani, I.; Boissinot, S. Disentangling the determinants of transposable elements dynamics in vertebrate genomes using empirical evidences and simulations. *PLoS Genet.* **2020**, *16*, e1009082. [CrossRef] [PubMed]
107. Venner, S.; Feschotte, C.; Biémont, C. Dynamics of transposable elements: Towards a community ecology of the genome. *Trends Genet.* **2009**, *25*, 317–323. [CrossRef]
108. Campos-Sánchez, R.; Cremona, M.A.; Pini, A.; Chiaromonte, F.; Makova, K.D. Integration and fixation preferences of human and mouse endogenous retroviruses uncovered with functional data analysis. *PLoS Comput. Biol.* **2016**, *12*, e1004956. [CrossRef] [PubMed]
109. Yoshida, J.; Akagi, K.; Misawa, R.; Kokubu, C.; Takeda, J.; Horie, K. Chromatin states shape insertion profiles of the piggyBac, Tol2 and Sleeping Beauty transposons and murine leukemia virus. *Sci. Rep.* **2017**, *7*, 43613. [CrossRef]
110. Myers, S.; Freeman, C.; Auton, A.; Donnelly, P.; McVean, G. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat. Genet.* **2008**, *40*, 1124. [CrossRef] [PubMed]
111. Kawakami, T.; Mugal, C.F.; Suh, A.; Nater, A.; Burri, R.; Smeds, L.; Ellegren, H. Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds. *Mol. Ecol.* **2017**, *26*, 4158–4172. [CrossRef]
112. Duret, L.; Marais, G.; Biémont, C. Transposons but not retrotransposons are located preferentially in regions of high recombination rate in Caenorhabditis elegans. *Genetics* **2000**, *156*, 1661–1669. [CrossRef] [PubMed]
113. Pardue, M.-L.; DeBaryshe, P. Retrotransposons that maintain chromosome ends. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20317–20324. [CrossRef]
114. de Boer, J.G.; Yazawa, R.; Davidson, W.S.; Koop, B.F. Bursts and horizontal evolution of DNA transposons in the speciation of pseudotetraploid salmonids. *BMC Genom.* **2007**, *8*, 1–10. [CrossRef]
115. Pace, J.K.; Feschotte, C. The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage. *Genome Res.* **2007**, *17*, 422–432. [CrossRef]
116. Vieira, C.; Lepetit, D.; Dumont, S.; Biémont, C. Wake up of transposable elements following Drosophila simulans worldwide colonization. *Mol. Biol. Evol.* **1999**, *16*, 1251–1255. [CrossRef]
117. Pascale, E.; Valle, E.; Furano, A.V. Amplification of an ancestral mammalian L1 family of long interspersed repeated DNA occurred just before the murine radiation. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9481–9485. [CrossRef] [PubMed]
118. Blumenstiel, J.P.; Chen, X.; He, M.; Bergman, C.M. An age-of-allele test of neutrality for transposable element insertions. *Genetics* **2014**, *196*, 523–538. [CrossRef]
119. Garcia Guerreiro, M.P.; Chavez-Sandoval, B.E.; Balanya, J.; Serra, L.; Fontdevila, A. Distribution of the transposable elements bilbo and gypsy in original and colonizing populations of Drosophila subobscura. *BMC Evol. Biol.* **2008**, *8*, 234. [CrossRef]
120. Gonzalez, J.; Macpherson, J.M.; Messer, P.W.; Petrov, D.A. Inferring the strength of selection in Drosophila under complex demographic models. *Mol. Biol. Evol.* **2009**, *26*, 513–526. [CrossRef]
121. Burri, R.; Nater, A.; Kawakami, T.; Mugal, C.F.; Olason, P.I.; Smeds, L.; Suh, A.; Dutoit, L.; Bureš, S.; Garamszegi, L.Z. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Res.* **2015**, *25*, 1656–1665. [CrossRef] [PubMed]
122. Dolgin, E.S.; Charlesworth, B. The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* **2008**, *178*, 2169–2177. [CrossRef] [PubMed]
123. Ho, S.S.; Urban, A.E.; Mills, R.E. Structural variation in the sequencing era. *Nat. Rev. Genet.* **2020**, *21*, 171–189. [CrossRef]
124. Finnegan, D.J. Retrotransposons. *Curr. Biol.* **2012**, *22*, R432–R437. [CrossRef]
125. Arkhipova, I.R.; Yushenova, I.A. Giant Transposons in Eukaryotes: Is Bigger Better? *Genome Biol. Evol.* **2019**, *11*, 906–918. [CrossRef] [PubMed]
126. Mieczkowski, P.A.; Lemoine, F.J.; Petes, T.D. Recombination between retrotransposons as a source of chromosome rearrangements in the yeast Saccharomyces cerevisiae. *DNA Repair* **2006**, *5*, 1010–1020. [CrossRef]
127. Polleys, E.J.; House, N.C.M.; Freudenreich, C.H. Role of recombination and replication fork restart in repeat instability. *DNA Repair* **2017**, *56*, 156–165. [CrossRef] [PubMed]
128. Bariah, I.; Keidar-Friedman, D.; Kashkush, K. Where the Wild Things Are: Transposable Elements as Drivers of Structural and Functional Variations in the Wheat Genome. *Front. Plant Sci.* **2020**, *11*, 585515. [CrossRef]
129. Chen, J.; Rattner, A.; Nathans, J. Effects of L1 retrotransposon insertion on transcript processing, localization and accumulation: Lessons from the retinal degeneration 7 mouse and implications for the genomic ecology of L1 elements. *Hum. Mol. Genet.* **2006**, *15*, 2146–2156. [CrossRef] [PubMed]
130. Moran, J.V.; DeBerardinis, R.J.; Kazazian, H.H., Jr. Exon shuffling by L1 retrotransposition. *Science* **1999**, *283*, 1530–1534. [CrossRef]
131. Goodier, J.L.; Ostertag, E.M.; Kazazian, H.H., Jr. Transduction of 3′-flanking sequences is common in L1 retrotransposition. *Hum. Mol. Genet.* **2000**, *9*, 653–657. [CrossRef]

132. Pickeral, O.K.; Makalowski, W.; Boguski, M.S.; Boeke, J.D. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* **2000**, *10*, 411–415. [CrossRef]

133. Kalendar, R.; Vicient, C.M.; Peleg, O.; Anamthawat-Jonsson, K.; Bolshoy, A.; Schulman, A.H. Large retrotransposon derivatives: Abundant, conserved but nonautonomous retroelements of barley and related genomes. *Genetics* **2004**, *166*, 1437–1450. [CrossRef]

134. Gilbert, N.; Lutz, S.; Morrish, T.A.; Moran, J.V. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol. Cell. Biol.* **2005**, *25*, 7780–7795. [CrossRef] [PubMed]

135. Symer, D.E.; Connelly, C.; Szak, S.T.; Caputo, E.M.; Cost, G.J.; Parmigiani, G.; Boeke, J.D. Human l1 retrotransposition is associated with genetic instability in vivo. *Cell* **2002**, *110*, 327–338. [CrossRef]

136. Gilbert, N.; Lutz-Prigge, S.; Moran, J.V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **2002**, *110*, 315–325. [CrossRef]

137. Callinan, P.A.; Wang, J.; Herke, S.W.; Garber, R.K.; Liang, P.; Batzer, M.A. Alu retrotransposition-mediated deletion. *J. Mol. Biol.* **2005**, *348*, 791–800. [CrossRef]

138. Han, K.; Sen, S.K.; Wang, J.; Callinan, P.A.; Lee, J.; Cordaux, R.; Liang, P.; Batzer, M.A. Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. *Nucleic Acids Res.* **2005**, *33*, 4040–4052. [CrossRef] [PubMed]

139. Grandi, F.C.; An, W. Non-LTR retrotransposons and microsatellites: Partners in genomic variation. *Mob. Genet. Elem.* **2013**, *3*, e25674. [CrossRef]

140. Hickman, A.B.; Dyda, F. DNA Transposition at Work. *Chem. Rev.* **2016**, *116*, 12758–12784. [CrossRef]

141. Engels, W.R.; Johnson-Schlitz, D.M.; Eggleston, W.B.; Sved, J. High-frequency P element loss in Drosophila is homolog dependent. *Cell* **1990**, *62*, 515–525. [CrossRef]

142. Hagemann, A.T.; Craig, N.L. Tn7 transposition creates a hotspot for homologous recombination at the transposon donor site. *Genetics* **1993**, *133*, 9–16. [CrossRef]

143. Jiang, N.; Bao, Z.; Zhang, X.; Eddy, S.R.; Wessler, S.R. Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **2004**, *431*, 569–573. [CrossRef]

144. Grabundzija, I.; Messing, S.A.; Thomas, J.; Cosby, R.L.; Bilic, I.; Miskey, C.; Gogol-Doring, A.; Kapitonov, V.; Diem, T.; Dalda, A.; et al. A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat. Commun.* **2016**, *7*, 10716. [CrossRef] [PubMed]

145. Thomas, J.; Pritham, E.J. Helitrons, the Eukaryotic Rolling-circle Transposable Elements. *Microbiol. Spectr.* **2015**, *3*, 893–926. [CrossRef]

146. Gusa, A.; Jinks-Robertson, S. Mitotic Recombination and Adaptive Genomic Changes in Human Pathogenic Fungi. *Genes* **2019**, *10*, 901. [CrossRef]

147. Wicker, T.; Buchmann, J.P.; Keller, B. Patching gaps in plant genomes results in gene movement and erosion of colinearity. *Genome Res.* **2010**, *20*, 1229–1237. [CrossRef]

148. Bailey, J.A.; Liu, G.; Eichler, E.E. An Alu transposition model for the origin and expansion of human segmental duplications. *Am. J. Hum. Genet.* **2003**, *73*, 823–834. [CrossRef]

149. Caceres, M.; Ranz, J.M.; Barbadilla, A.; Long, M.; Ruiz, A. Generation of a widespread Drosophila inversion by a transposable element. *Science* **1999**, *285*, 415–418. [CrossRef] [PubMed]

150. Delprat, A.; Negre, B.; Puig, M.; Ruiz, A. The transposon Galileo generates natural chromosomal inversions in Drosophila by ectopic recombination. *PLoS ONE* **2009**, *4*, e7883. [CrossRef]

151. Schwartz, A.; Chan, D.C.; Brown, L.G.; Alagappan, R.; Pettay, D.; Disteche, C.; McGillivray, B.; de la Chapelle, A.; Page, D.C. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum. Mol. Genet.* **1998**, *7*, 1–11. [CrossRef] [PubMed]

152. Bartolomé, C.; Maside, X.; Charlesworth, B. On the abundance and distribution of transposable elements in the genome of Drosophila melanogaster. *Mol. Biol. Evol.* **2002**, *19*, 926–937. [CrossRef] [PubMed]

153. Wright, S.I.; Agrawal, N.; Bureau, T.E. Effects of recombination rate and gene density on transposable element distributions in Arabidopsis thaliana. *Genome Res.* **2003**, *13*, 1897–1903. [PubMed]

154. Medstrand, P.; van de Lagemaat, L.N.; Mager, D.L. Retroelement distributions in the human genome: Variations associated with age and proximity to genes. *Genome Res.* **2002**, *12*, 1483–1495. [CrossRef]

155. Caizzi, R.; Caggese, C.; Pimpinelli, S. Bari-1, a new transposon-like family in Drosophila melanogaster with a unique heterochromatic organization. *Genetics* **1993**, *133*, 335–345. [CrossRef] [PubMed]

156. Lerat, E.; Capy, P.; Biemont, C. Codon usage by transposable elements and their host genes in five species. *J. Mol. Evol.* **2002**, *54*, 625–637. [CrossRef]

157. Ruggiero, R.P.; Boissinot, S. Variation in base composition underlies functional and evolutionary divergence in non-LTR retrotransposons. *Mob. DNA* **2020**, *11*, 1–18. [CrossRef]

158. Symonova, R.; Suh, A. Nucleotide composition of transposable elements likely contributes to AT/GC compositional homogeneity of teleost fish genomes. *Mob. DNA* **2019**, *10*, 49. [CrossRef]

159. Mugal, C.F.; Weber, C.C.; Ellegren, H. GC-biased gene conversion links the recombination landscape and demography to genomic base composition: GC-biased gene conversion drives genomic base composition across a wide range of species. *Bioessays* **2015**, *37*, 1317–1326. [CrossRef]

160. Kent, T.V.; Uzunović, J.; Wright, S.I. Coevolution between transposable elements and recombination. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *372*, 20160458. [CrossRef] [PubMed]
161. Choi, K.; Zhao, X.; Kelly, K.A.; Venn, O.; Higgins, J.D.; Yelina, N.E.; Hardcastle, T.J.; Ziolkowski, P.A.; Copenhaver, G.P.; Franklin, F.C.; et al. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* **2013**, *45*, 1327–1336. [CrossRef]
162. Rodgers-Melnick, E.; Bradbury, P.J.; Elshire, R.J.; Glaubitz, J.C.; Acharya, C.B.; Mitchell, S.E.; Li, C.; Li, Y.; Buckler, E.S. Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3823–3828. [CrossRef]
163. Pan, J.; Sasaki, M.; Kniewel, R.; Murakami, H.; Blitzblau, H.G.; Tischfield, S.E.; Zhu, X.; Neale, M.J.; Jasin, M.; Socci, N.D.; et al. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **2011**, *144*, 719–731. [CrossRef] [PubMed]
164. Zamudio, N.; Barau, J.; Teissandier, A.; Walter, M.; Borsos, M.; Servant, N.; Bourc'his, D. DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes Dev.* **2015**, *29*, 1256–1270. [CrossRef]
165. Mirouze, M.; Lieberman-Lazarovich, M.; Aversano, R.; Bucher, E.; Nicolet, J.; Reinders, J.; Paszkowski, J. Loss of DNA methylation affects the recombination landscape in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5880–5885. [CrossRef] [PubMed]
166. Ben-Aroya, S.; Mieczkowski, P.A.; Petes, T.D.; Kupiec, M. The compact chromatin structure of a Ty repeated sequence suppresses recombination hotspot activity in Saccharomyces cerevisiae. *Mol. Cell* **2004**, *15*, 221–231. [CrossRef]
167. He, L.; Dooner, H.K. Haplotype structure strongly affects recombination in a maize genetic interval polymorphic for Helitron and retrotransposon insertions. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8410–8416. [CrossRef] [PubMed]
168. Choudhury, R.R.; Rogivue, A.; Gugerli, F.; Parisod, C. Impact of polymorphic transposable elements on linkage disequilibrium along chromosomes. *Mol. Ecol.* **2019**, *28*, 1550–1562. [CrossRef]
169. Witherspoon, D.J.; Watkins, W.S.; Zhang, Y.; Xing, J.; Tolpinrud, W.L.; Hedges, D.J.; Batzer, M.A.; Jorde, L.B. Alu repeats increase local recombination rates. *BMC Genom.* **2009**, *10*, 530. [CrossRef]
170. Sundaram, V.; Wysocka, J. Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2020**, *375*, 20190347. [CrossRef]
171. Drongitis, D.; Aniello, F.; Fucci, L.; Donizetti, A. Roles of Transposable Elements in the Different Layers of Gene Expression Regulation. *Int. J. Mol. Sci.* **2019**, *20*, 5755. [CrossRef]
172. Elbarbary, R.A.; Lucas, B.A.; Maquat, L.E. Retrotransposons as regulators of gene expression. *Science* **2016**, *351*, aac7247. [CrossRef]
173. Moschetti, R.; Palazzo, A.; Lorusso, P.; Viggiano, L.; Massimiliano Marsano, R. "What You Need, Baby, I Got It": Transposable Elements as Suppliers of Cis-Operating Sequences in Drosophila. *Biology* **2020**, *9*, 25. [CrossRef] [PubMed]
174. Trizzino, M.; Park, Y.; Holsbach-Beltrame, M.; Aracena, K.; Mika, K.; Caliskan, M.; Perry, G.H.; Lynch, V.J.; Brown, C.D. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Res.* **2017**, *27*, 1623–1633. [CrossRef] [PubMed]
175. Jacques, P.E.; Jeyakani, J.; Bourque, G. The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet.* **2013**, *9*, e1003504. [CrossRef]
176. Brosius, J. Exaptation at the molecular genetic level. *Sci. China Life Sci.* **2019**, *62*, 437–452. [CrossRef]
177. Sundaram, V.; Cheng, Y.; Ma, Z.; Li, D.; Xing, X.; Edge, P.; Snyder, M.P.; Wang, T. Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res.* **2014**, *24*, 1963–1976. [CrossRef]
178. Bourque, G.; Leong, B.; Vega, V.B.; Chen, X.; Lee, Y.L.; Srinivasan, K.G.; Chew, J.L.; Ruan, Y.; Wei, C.L.; Ng, H.H.; et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* **2008**, *18*, 1752–1762. [CrossRef] [PubMed]
179. Hermant, C.; Torres-Padilla, M.E. TFs for TEs: The transcription factor repertoire of mammalian transposable elements. *Genes Dev.* **2021**, *35*, 22–39. [CrossRef]
180. Chuong, E.B.; Elde, N.C.; Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **2016**, *351*, 1083–1087. [CrossRef] [PubMed]
181. 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. [CrossRef]
182. Choudhary, M.N.K.; Friedman, R.Z.; Wang, J.T.; Jang, H.S.; Zhuo, X.; Wang, T. Publisher Correction: Co-opted transposons help perpetuate conserved higher-order chromosomal structures. *Genome Biol.* **2020**, *21*, 28. [CrossRef]
183. Diehl, A.G.; Ouyang, N.; Boyle, A.P. Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. *Nat. Commun.* **2020**, *11*, 1796. [CrossRef]
184. Avramova, Z.; Tikhonov, A.; Chen, M.; Bennetzen, J.L. Matrix attachment regions and structural colinearity in the genomes of two grass species. *Nucleic Acids Res.* **1998**, *26*, 761–767. [CrossRef]
185. Yellan, I.; Yang, A.W.H.; Hughes, T.R. Diverse Eukaryotic CGG Binding Proteins Produced by Independent Domestications of hAT Transposons. *Mol. Biol. Evol.* **2021**. [CrossRef] [PubMed]
186. Wang, J.; Xie, G.; Singh, M.; Ghanbarian, A.T.; Rasko, T.; Szvetnik, A.; Cai, H.; Besser, D.; Prigione, A.; Fuchs, N.V.; et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **2014**, *516*, 405–409. [CrossRef]
187. Dunn, C.A.; van de Lagemaat, L.N.; Baillie, G.J.; Mager, D.L. Endogenous retrovirus long terminal repeats as ready-to-use mobile promoters: The case of primate beta3GAL-T5. *Gene* **2005**, *364*, 2–12. [CrossRef]

188. Conley, A.B.; Piriyapongsa, J.; Jordan, I.K. Retroviral promoters in the human genome. *Bioinformatics* **2008**, *24*, 1563–1567. [CrossRef] [PubMed]

189. Kellner, M.; Makalowski, W. Transposable elements significantly contributed to the core promoters in the human genome. *Sci. China Life Sci.* **2019**, *62*, 489–497. [CrossRef] [PubMed]

190. Samuelson, L.C.; Wiebauer, K.; Snow, C.M.; Meisler, M.H. Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Mol. Cell. Biol.* **1990**, *10*, 2513–2520. [CrossRef]

191. Franchini, L.F.; Lopez-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. [CrossRef]

192. Duhl, D.M.; Vrieling, H.; Miller, K.A.; Wolff, G.L.; Barsh, G.S. Neomorphic agouti mutations in obese yellow mice. *Nat. Genet.* **1994**, *8*, 59–65. [CrossRef] [PubMed]

193. Krishnan, P.; Meile, L.; Plissonneau, C.; Ma, X.; Hartmann, F.E.; Croll, D.; McDonald, B.A.; Sanchez-Vallet, A. Transposable element insertions shape gene regulation and melanin production in a fungal pathogen of wheat. *BMC Biol.* **2018**, *16*, 78. [CrossRef] [PubMed]

194. Simonti, C.N.; Pavlicev, M.; Capra, J.A. Transposable Element Exaptation into Regulatory Regions Is Rare, Influenced by Evolutionary Age, and Subject to Pleiotropic Constraints. *Mol. Biol. Evol.* **2017**, *34*, 2856–2869. [CrossRef]

195. Zeng, L.; Pederson, S.M.; Cao, D.; Qu, Z.; Hu, Z.; Adelson, D.L.; Wei, C. Genome-Wide Analysis of the Association of Transposable Elements with Gene Regulation Suggests that Alu Elements Have the Largest Overall Regulatory Impact. *J. Comput. Biol.* **2018**, *25*, 551–562. [CrossRef]

196. Nikitin, D.; Penzar, D.; Garazha, A.; Sorokin, M.; Tkachev, V.; Borisov, N.; Poltorak, A.; Prassolov, V.; Buzdin, A.A. Profiling of Human Molecular Pathways Affected by Retrotransposons at the Level of Regulation by Transcription Factor Proteins. *Front. Immunol.* **2018**, *9*, 30. [CrossRef]

197. Davidson, E.; Levin, M. Gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4935. [CrossRef]

198. Britten, R.J.; Davidson, E.H. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q. Rev. Biol.* **1971**, *46*, 111–138. [CrossRef]

199. Notwell, J.H.; Chung, T.; Heavner, W.; Bejerano, G. A family of transposable elements co-opted into developmental enhancers in the mouse neocortex. *Nat. Commun.* **2015**, *6*, 6644. [CrossRef]

200. Kunarso, G.; Chia, N.Y.; Jeyakani, J.; Hwang, C.; Lu, X.; Chan, Y.S.; Ng, H.H.; Bourque, G. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **2010**, *42*, 631–634. [CrossRef] [PubMed]

201. Chuong, E.B.; Rumi, M.A.; Soares, M.J.; Baker, J.C. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat. Genet.* **2013**, *45*, 325–329. [CrossRef] [PubMed]

202. Gerdes, P.; Richardson, S.R.; Mager, D.L.; Faulkner, G.J. Transposable elements in the mammalian embryo: Pioneers surviving through stealth and service. *Genome Biol.* **2016**, *17*, 100. [CrossRef]

203. Ellison, C.E.; Bachtrog, D. Dosage compensation via transposable element mediated rewiring of a regulatory network. *Science* **2013**, *342*, 846–850. [CrossRef]

204. Ito, H.; Gaubert, H.; Bucher, E.; Mirouze, M.; Vaillant, I.; Paszkowski, J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **2011**, *472*, 115–119. [CrossRef] [PubMed]

205. Makarevitch, I.; Waters, A.J.; West, P.T.; Stitzer, M.; Hirsch, C.N.; Ross-Ibarra, J.; Springer, N.M. Transposable elements contribute to activation of maize genes in response to abiotic stress. *PLoS Genet.* **2015**, *11*, e1004915.

206. Schmidt, D.; Schwalie, P.C.; Wilson, M.D.; Ballester, B.; Goncalves, 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. [CrossRef] [PubMed]

207. Ellison, C.; Bachtrog, D. Contingency in the convergent evolution of a regulatory network: Dosage compensation in Drosophila. *PLoS Biol.* **2019**, *17*, e3000094. [CrossRef] [PubMed]

208. Fawcett, J.A.; Innan, H. The Role of Gene Conversion between Transposable Elements in Rewiring Regulatory Networks. *Genome Biol. Evol.* **2019**, *11*, 1723–1729. [CrossRef]

209. Sun, X.; Wang, X.; Tang, Z.; Grivainis, M.; Kahler, D.; Yun, C.; Mita, P.; Fenyo, D.; Boeke, J.D. Transcription factor profiling reveals molecular choreography and key regulators of human retrotransposon expression. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5526–E5535. [CrossRef] [PubMed]

210. Chen, J.M.; Cooper, D.N.; Chuzhanova, N.; Ferec, C.; Patrinos, G.P. Gene conversion: Mechanisms, evolution and human disease. *Nat. Rev. Genet.* **2007**, *8*, 762–775. [CrossRef]

211. Han, J.S.; Szak, S.T.; Boeke, J.D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **2004**, *429*, 268–274. [CrossRef]

212. Zhang, Y.; Romanish, M.T.; Mager, D.L. Distributions of transposable elements reveal hazardous zones in mammalian introns. *PLoS Comput. Biol.* **2011**, *7*, e1002046. [CrossRef]

213. van de Lagemaat, L.N.; Medstrand, P.; Mager, D.L. Multiple effects govern endogenous retrovirus survival patterns in human gene introns. *Genome Biol.* **2006**, *7*, R86. [CrossRef]

214. Mateo, L.; Ullastres, A.; Gonzalez, J. A transposable element insertion confers xenobiotic resistance in Drosophila. *PLoS Genet.* **2014**, *10*, e1004560. [CrossRef]

215. Niu, X.M.; Xu, Y.C.; Li, Z.W.; Bian, Y.T.; Hou, X.H.; Chen, J.F.; Zou, Y.P.; Jiang, J.; Wu, Q.; Ge, S.; et al. Transposable elements drive rapid phenotypic variation in Capsella rubella. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6908–6913. [CrossRef]

216. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288. [CrossRef] [PubMed]

217. Lucas, B.A.; Lavi, E.; Shiue, L.; Cho, H.; Katzman, S.; Miyoshi, K.; Siomi, M.C.; Carmel, L.; Ares, M., Jr.; Maquat, L.E. Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 968–973. [CrossRef]

218. Kapusta, A.; Kronenberg, Z.; Lynch, V.J.; Zhuo, X.; Ramsay, L.; Bourque, G.; Yandell, M.; Feschotte, C. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.* **2013**, *9*, e1003470. [CrossRef]

219. Qin, S.; Jin, P.; Zhou, X.; Chen, L.; Ma, F. The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human. *PLoS ONE* **2015**, *10*, e0131365. [CrossRef]

220. Slotkin, R.K.; Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **2007**, *8*, 272–285. [CrossRef] [PubMed]

221. Choi, J.Y.; Lee, Y.C.G. Double-edged sword: The evolutionary consequences of the epigenetic silencing of transposable elements. *PLoS Genet.* **2020**, *16*, e1008872. [CrossRef]

222. Kelleher, E.S.; Barbash, D.A.; Blumenstiel, J.P. Taming the Turmoil Within: New Insights on the Containment of Transposable Elements. *Trends Genet.* **2020**, *36*, 474–489. [CrossRef] [PubMed]

223. Deniz, O.; Frost, J.M.; Branco, M.R. Regulation of transposable elements by DNA modifications. *Nat. Rev. Genet.* **2019**, *20*, 417–431. [CrossRef] [PubMed]

224. Yoder, J.A.; Walsh, C.P.; Bestor, T.H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **1997**, *13*, 335–340. [CrossRef]

225. Wolf, G.; de Iaco, A.; Sun, M.A.; Bruno, M.; Tinkham, M.; Hoang, D.; Mitra, A.; Ralls, S.; Trono, D.; Macfarlan, T.S. KRAB-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *eLife* **2020**, *9*, e56337. [CrossRef] [PubMed]

226. Czech, B.; Hannon, G.J. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem. Sci.* **2016**, *41*, 324–337. [CrossRef] [PubMed]

227. Lee, Y.C.G.; Karpen, G.H. Pervasive epigenetic effects of Drosophila euchromatic transposable elements impact their evolution. *eLife* **2017**, *6*, e25762. [CrossRef]

228. Ahmed, I.; Sarazin, A.; Bowler, C.; Colot, V.; Quesneville, H. Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in Arabidopsis. *Nucleic Acids Res.* **2011**, *39*, 6919–6931. [CrossRef]

229. Hollister, J.D.; Gaut, B.S. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **2009**, *19*, 1419–1428. [CrossRef]

230. Dubin, M.J.; Zhang, P.; Meng, D.; Remigereau, M.S.; Osborne, E.J.; Paolo Casale, F.; Drewe, P.; Kahles, A.; Jean, G.; Vilhjalmsson, B.; et al. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* **2015**, *4*, e05255. [CrossRef]

231. Quadrana, L.; Bortolini Silveira, A.; Mayhew, G.F.; LeBlanc, C.; Martienssen, R.A.; Jeddeloh, J.A.; Colot, V. The Arabidopsis thaliana mobilome and its impact at the species level. *eLife* **2016**, *5*, e15716. [CrossRef]

232. Stuart, T.; Eichten, S.R.; Cahn, J.; Karpievitch, Y.V.; Borevitz, J.O.; Lister, R. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* **2016**, *5*, e20777. [CrossRef]

233. Rebollo, R.; Karimi, M.M.; Bilenky, M.; Gagnier, L.; Miceli-Royer, K.; Zhang, Y.; Goyal, P.; Keane, T.M.; Jones, S.; Hirst, M.; et al. Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet.* **2011**, *7*, e1002301. [CrossRef]

234. Kinoshita, Y.; Saze, H.; Kinoshita, T.; Miura, A.; Soppe, W.J.; Koornneef, M.; Kakutani, T. Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. *Plant J.* **2007**, *49*, 38–45. [CrossRef]

235. Sigman, M.J.; Slotkin, R.K. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *Plant. Cell* **2016**, *28*, 304–313. [CrossRef] [PubMed]

236. Parinov, S.; Sundaresan, V. Functional genomics in Arabidopsis: Large-scale insertional mutagenesis complements the genome sequencing project. *Curr. Opin. Biotechnol.* **2000**, *11*, 157–161. [CrossRef]

237. Kaessmann, H.; Vinckenbosch, N.; Long, M. RNA-based gene duplication: Mechanistic and evolutionary insights. *Nat. Rev. Genet.* **2009**, *10*, 19–31. [CrossRef]

238. Long, M.; Betran, E.; Thornton, K.; Wang, W. The origin of new genes: Glimpses from the young and old. *Nat. Rev. Genet.* **2003**, *4*, 865–875. [CrossRef] [PubMed]

239. Krasileva, K.V. The role of transposable elements and DNA damage repair mechanisms in gene duplications and gene fusions in plant genomes. *Curr. Opin. Plant Biol.* **2019**, *48*, 18–25. [CrossRef]

240. Cosby, R.L.; Judd, J.; Zhang, R.; Zhong, A.; Garry, N.; Pritham, E.J.; Feschotte, C. Recurrent evolution of vertebrate transcription factors by transposase capture. *Science* **2021**, *371*, eabc6405. [CrossRef]

241. Brennan, G.; Kozyrev, Y.; Hu, S.L. TRIMCyp expression in Old World primates Macaca nemestrina and Macaca fascicularis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3569–3574. [CrossRef]

242. Virgen, C.A.; Kratovac, Z.; Bieniasz, P.D.; Hatziioannou, T. Independent genesis of chimeric TRIM5-cyclophilin proteins in two primate species. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3563–3568. [CrossRef] [PubMed]
243. Wilson, S.J.; Webb, B.L.; Ylinen, L.M.; Verschoor, E.; Heeney, J.L.; Towers, G.J. Independent evolution of an antiviral TRIMCyp in rhesus macaques. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3557–3562. [CrossRef]
244. Cordaux, R.; Udit, S.; Batzer, M.A.; Feschotte, C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8101–8106. [CrossRef]
245. Shaheen, M.; Williamson, E.; Nickoloff, J.; Lee, S.H.; Hromas, R. Metnase/SETMAR: A domesticated primate transposase that enhances DNA repair, replication, and decatenation. *Genetica* **2010**, *138*, 559–566. [CrossRef]
246. Bourque, G. Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Curr. Opin. Genet. Dev.* **2009**, *19*, 607–612. [CrossRef]
247. Naville, M.; Warren, I.A.; Haftek-Terreau, Z.; Chalopin, D.; Brunet, F.; Levin, P.; Galiana, D.; Volff, J.N. Not so bad after all: Retroviruses and long terminal repeat retrotransposons as a source of new genes in vertebrates. *Clin. Microbiol. Infect.* **2016**, *22*, 312–323. [CrossRef]
248. Volff, J.N. Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays* **2006**, *28*, 913–922. [CrossRef]
249. Jangam, D.; Feschotte, C.; Betrán, E. Transposable element domestication as an adaptation to evolutionary conflicts. *Trends Genet.* **2017**, *33*, 817–831. [CrossRef]
250. Kapitonov, V.V.; Jurka, J. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol.* **2005**, *3*, e181. [CrossRef] [PubMed]
251. Carmona, L.M.; Schatz, D.G. New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. *FEBS J.* **2017**, *284*, 1590–1605. [CrossRef] [PubMed]
252. Huang, S.; Tao, X.; Yuan, S.; Zhang, Y.; Li, P.; Beilinson, H.A.; Zhang, Y.; Yu, W.; Pontarotti, P.; Escriva, H.; et al. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. *Cell* **2016**, *166*, 102–114. [CrossRef] [PubMed]
253. Lin, R.; Ding, L.; Casola, C.; Ripoll, D.R.; Feschotte, C.; Wang, H. Transposase-derived transcription factors regulate light signaling in Arabidopsis. *Science* **2007**, *318*, 1302–1305. [CrossRef] [PubMed]
254. Casola, C.; Hucks, D.; Feschotte, C. Convergent domestication of pogo-like transposases into centromere-binding proteins in fission yeast and mammals. *Mol. Biol. Evol.* **2008**, *25*, 29–41. [CrossRef] [PubMed]
255. Curcio, M.J.; Belfort, M. The beginning of the end: Links between ancient retroelements and modern telomerases. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9107–9108. [CrossRef] [PubMed]
256. Heidmann, O.; Vernochet, C.; Dupressoir, A.; Heidmann, T. Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: A new "syncytin" in a third order of mammals. *Retrovirology* **2009**, *6*, 107. [CrossRef] [PubMed]
257. Vernochet, C.; Heidmann, O.; Dupressoir, A.; Cornelis, G.; Dessen, P.; Catzeflis, F.; Heidmann, T. A syncytin-like endogenous retrovirus envelope gene of the guinea pig specifically expressed in the placenta junctional zone and conserved in Caviomorpha. *Placenta* **2011**, *32*, 885–892. [CrossRef]
258. Cornelis, G.; Heidmann, O.; Bernard-Stoecklin, S.; Reynaud, K.; Veron, G.; Mulot, B.; Dupressoir, A.; Heidmann, T. Ancestral capture of syncytin-Car1, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E432–E441. [CrossRef]
259. Cornelis, G.; Vernochet, C.; Carradec, Q.; Souquere, S.; Mulot, B.; Catzeflis, F.; Nilsson, M.A.; Menzies, B.R.; Renfree, M.B.; Pierron, G.; et al. Retroviral envelope gene captures and syncytin exaptation for placentation in marsupials. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E487–E496. [CrossRef]
260. Dupressoir, A.; Marceau, G.; Vernochet, C.; Benit, L.; Kanellopoulos, C.; Sapin, V.; Heidmann, T. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 725–730. [CrossRef]
261. Mangeney, M.; Renard, M.; Schlecht-Louf, G.; Bouallaga, I.; Heidmann, O.; Letzelter, C.; Richaud, A.; Ducos, B.; Heidmann, T. Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20534–20539. [CrossRef]
262. Blackburn, E.H. Telomerases. *Annu. Rev. Biochem.* **1992**, *61*, 113–129. [CrossRef]
263. Biessmann, H.; Mason, J.M.; Ferry, K.; d'Hulst, M.; Valgeirsdottir, K.; Traverse, K.L.; Pardue, M.-L. Addition of telomere-associated HeT DNA sequences "heals" broken chromosome ends in Drosophila. *Cell* **1990**, *61*, 663–673. [CrossRef]
264. Abad, J.P.; De Pablos, B.; Osoegawa, K.; De Jong, P.J.; Martin-Gallardo, A.; Villasante, A. TAHRE, a novel telomeric retrotransposon from Drosophila melanogaster, reveals the origin of Drosophila telomeres. *Mol. Biol Evol.* **2004**, *21*, 1620–1624. [CrossRef] [PubMed]
265. Abad, J.P.; De Pablos, B.; Osoegawa, K.; De Jong, P.J.; Martin-Gallardo, A.; Villasante, A. Genomic analysis of Drosophila melanogaster telomeres: Full-length copies of HeT-A and TART elements at telomeres. *Mol. Biol Evol.* **2004**, *21*, 1613–1619. [CrossRef] [PubMed]
266. Saint-Leandre, B.; Nguyen, S.C.; Levine, M.T. Diversification and collapse of a telomere elongation mechanism. *Genome Res.* **2019**, *29*, 920–931. [CrossRef] [PubMed]

267. Gao, D.; Jiang, N.; Wing, R.A.; Jiang, J.; Jackson, S.A. Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front. Plant Sci.* **2015**, *6*, 216. [CrossRef]

268. Glockner, G.; Heidel, A.J. Centromere sequence and dynamics in Dictyostelium discoideum. *Nucleic Acids Res.* **2009**, *37*, 1809–1816. [CrossRef]

269. Malicki, M.; Iliopoulou, M.; Hammann, C. Retrotransposon Domestication and Control in Dictyostelium discoideum. *Front. Microbiol.* **2017**, *8*, 1869. [CrossRef]

270. Spaller, T.; Kling, E.; Glockner, G.; Hillmann, F.; Winckler, T. Convergent evolution of tRNA gene targeting preferences in compact genomes. *Mob. DNA* **2016**, *7*, 17. [CrossRef]

271. Casola, C.; Betran, E. The Genomic Impact of Gene Retrocopies: What Have We Learned from Comparative Genomics, Population Genomics, and Transcriptomic Analyses? *Genome Biol. Evol.* **2017**, *9*, 1351–1373. [CrossRef]

272. Kubiak, M.R.; Makalowska, I. Protein-Coding Genes' Retrocopies and Their Functions. *Viruses* **2017**, *9*, 80. [CrossRef]

273. Esnault, C.; Maestre, J.; Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nature Genet.* **2000**, *24*, 363–367. [CrossRef]

274. Wang, W.; Zheng, H.; Fan, C.; Li, J.; Shi, J.; Cai, Z.; Zhang, G.; Liu, D.; Zhang, J.; Vang, S.; et al. High rate of chimeric gene origination by retroposition in plant genomes. *Plant Cell* **2006**, *18*, 1791–1802. [CrossRef]

275. Tan, S.; Cardoso-Moreira, M.; Shi, W.; Zhang, D.; Huang, J.; Mao, Y.; Jia, H.; Zhang, Y.; Chen, C.; Shao, Y.; et al. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Res.* **2016**, *26*, 1663–1675. [CrossRef] [PubMed]

276. Suh, A. The Specific Requirements for CR1 Retrotransposition Explain the Scarcity of Retrogenes in Birds. *J. Mol. Evol.* **2015**, *81*, 18–20. [CrossRef] [PubMed]

277. Piriyapongsa, J.; Marino-Ramirez, L.; Jordan, I.K. Origin and evolution of human microRNAs from transposable elements. *Genetics* **2007**, *176*, 1323–1337. [CrossRef] [PubMed]

278. Tam, O.H.; Aravin, A.A.; Stein, P.; Girard, A.; Murchison, E.P.; Cheloufi, S.; Hodges, E.; Anger, M.; Sachidanandam, R.; Schultz, R.M.; et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **2008**, *453*, 534–538. [CrossRef]

279. Mullis, A.; Lu, Z.; Zhan, Y.; Wang, T.-Y.; Rodriguez, J.; Rajeh, A.; Chatrath, A.; Lin, Z. Parallel concerted evolution of ribosomal protein genes in Fungi and its adaptive significance. *Mol. Biol. Evol.* **2020**, *37*, 455–468. [CrossRef]

280. Rosso, L.; Marques, A.C.; Weier, M.; Lambert, N.; Lambot, M.-A.; Vanderhaeghen, P.; Kaessmann, H. Birth and rapid subcellular adaptation of a hominoid-specific CDC14 protein. *PLoS Biol.* **2008**, *6*, e140. [CrossRef]

281. Ciomborowska, J.; Rosikiewicz, W.; Szklarczyk, D.; Makałowski, W.; Makałowska, I. "Orphan" retrogenes in the human genome. *Mol. Biol. Evol.* **2012**, *30*, 384–396. [CrossRef]

282. Carelli, F.N.; Hayakawa, T.; Go, Y.; Imai, H.; Warnefors, M.; Kaessmann, H. The life history of retrocopies illuminates the evolution of new mammalian genes. *Genome Res.* **2016**, *26*, 301–314. [CrossRef]

283. Schrader, L.; Schmitz, J. The impact of transposable elements in adaptive evolution. *Mol. Ecol.* **2019**, *28*, 1537–1549. [CrossRef] [PubMed]

284. Etchegaray, E.; Naville, M.; Volff, J.N.; Haftek-Terreau, Z. Transposable element-derived sequences in vertebrate development. *Mob. DNA* **2021**, *12*, 1. [CrossRef]

285. Rebollo, R.; Horard, B.; Hubert, B.; Vieira, C. Jumping genes and epigenetics: Towards new species. *Gene* **2010**, *454*, 1–7. [CrossRef]

286. Serrato-Capuchina, A.; Matute, D.R. The Role of Transposable Elements in Speciation. *Genes* **2018**, *9*, 254. [CrossRef]

287. Verneau, O.; Catzeflis, F.; Furano, A.V. Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11284–11289. [CrossRef]

288. Ray, D.A.; Feschotte, C.; Pagan, H.J.; Smith, J.D.; Pritham, E.J.; Arensburger, P.; Atkinson, P.W.; Craig, N.L. Multiple waves of recent DNA transposon activity in the bat, Myotis lucifugus. *Genome Res.* **2008**, *18*, 717–728. [CrossRef]

289. Gould, S.J.; Eldredge, N. Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology* **1977**, *3*, 115–151. [CrossRef]

290. Ricci, M.; Peona, V.; Guichard, E.; Taccioli, C.; Boattini, A. Transposable Elements Activity is Positively Related to Rate of Speciation in Mammals. *J. Mol. Evol.* **2018**, *86*, 303–310. [CrossRef]

291. Oliver, K.R.; Greene, W.K. Transposable elements and viruses as factors in adaptation and evolution: An expansion and strengthening of the TE-Thrust hypothesis. *Ecol. Evol.* **2012**, *2*, 2912–2933. [CrossRef] [PubMed]

292. Zeh, D.W.; Zeh, J.A.; Ishida, Y. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays* **2009**, *31*, 715–726. [CrossRef] [PubMed]

293. Jurka, J.; Bao, W.; Kojima, K.K. Families of transposable elements, population structure and the origin of species. *Biol. Direct* **2011**, *6*, 44. [CrossRef] [PubMed]

294. Mc, C.B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* **1950**, *36*, 344–355. [CrossRef]

295. Langley, C.H.; Montgomery, E.; Hudson, R.; Kaplan, N.; Charlesworth, B. On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* **1988**, *52*, 223–235. [CrossRef] [PubMed]

296. Brown, J.D.; O'Neill, R.J. Chromosomes, conflict, and epigenetics: Chromosomal speciation revisited. *Annu. Rev. Genom. Hum. Genet.* **2010**, *11*, 291–316. [CrossRef]

297. Noor, M.A.; Grams, K.L.; Bertucci, L.A.; Reiland, J. Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12084–12088. [CrossRef]

298.　Labrador, M.; Farre, M.; Utzet, F.; Fontdevila, A. Interspecific hybridization increases transposition rates of Osvaldo. *Mol. Biol. Evol.* **1999**, *16*, 931–937. [CrossRef]

299.　Liu, B.; Wendel, J.F. Retrotransposon activation followed by rapid repression in introgressed rice plants. *Genome* **2000**, *43*, 874–880. [CrossRef]

300.　Dion-Cote, A.M.; Renaut, S.; Normandeau, E.; Bernatchez, L. RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Mol. Biol. Evol.* **2014**, *31*, 1188–1199. [CrossRef]

301.　Ungerer, M.C.; Strakosh, S.C.; Zhen, Y. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr. Biol.* **2006**, *16*, R872–R873. [CrossRef]

302.　Brown, J.D.; Golden, D.; O'Neill, R.J. Methylation perturbations in retroelements within the genome of a Mus interspecific hybrid correlate with double minute chromosome formation. *Genomics* **2008**, *91*, 267–273. [CrossRef]

303.　O'Neill, R.J.W.; O'Neill, M.J.; Graves, J.A.M. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **1998**, *393*, 68–72. [CrossRef]

304.　Henault, M.; Marsit, S.; Charron, G.; Landry, C.R. The effect of hybridization on transposable element accumulation in an undomesticated fungal species. *eLife* **2020**, *9*, e60474. [CrossRef] [PubMed]

305.　Smukowski Heil, C.; Patterson, K.; Hickey, A.S.; Alcantara, E.; Dunham, M.J. Transposable Element Mobilization in Interspecific Yeast Hybrids. *Genome Biol. Evol.* **2021**, *13*, evab033. [CrossRef] [PubMed]

306.　Henault, M. The challenges of predicting transposable element activity in hybrids. *Curr Genet.* **2021**, in press. [CrossRef]

307.　Eyre-Walker, A.; Keightley, P.D. The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **2007**, *8*, 610–618. [CrossRef]

308.　Haller, B.C.; Messer, P.W. SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Mol. Biol. Evol.* **2019**, *36*, 632–637. [CrossRef] [PubMed]

309.　Chen, D.; Cremona, M.A.; Qi, Z.; Mitra, R.D.; Chiaromonte, F.; Makova, K.D. Human L1 Transposition Dynamics Unraveled with Functional Data Analysis. *Mol. Biol. Evol.* **2020**, *37*, 3576–3600. [CrossRef]

310.　Saika, H.; Mori, A.; Endo, M.; Toki, S. Targeted deletion of rice retrotransposon Tos17 via CRISPR/Cas9. *Plant Cell Rep.* **2019**, *38*, 455–458. [CrossRef]

311.　Ding, Y.; Berrocal, A.; Morita, T.; Longden, K.D.; Stern, D.L. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* **2016**, *536*, 329–332. [CrossRef]