Is somatic retrotransposition a parasitic or symbiotic phenomenon?

Citation for published version:

Digital Object Identifier (DOI):
10.4161/mge.18422

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Mobile Genetic Elements

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Is somatic retrotransposition a parasitic or symbiotic phenomenon?

Kyle R. Upton,* J. Kenneth Baillie and Geoffrey J. Faulkner*
Division of Genetics and Genomics; The Roslin Institute and Royal (Dick) School of Veterinary Studies; University of Edinburgh; Easter Bush, UK

Keywords: LINE-1, Alu, SVA, retrotransposon, transposable element, parasitism, symbiotism, somatic retrotransposition

Abbreviations: TE, transposable element; LINE-1 or L1, long interspersed nuclear element; ORF, open reading frame

Submitted: 09/14/11
Revised: 10/13/11
Accepted: 10/14/11

The extraordinary evolutionary success of transposable elements (TEs) invites us to question the nature of the co-evolutionary dynamics between TE and host. Although sometimes assumed to be wholly parasitic, TEs have penetrated and spread throughout eukaryotic genomes at a rate unparalleled by other parasites. This near-ubiquity, occurring despite the potentially deleterious effects of insertional mutagenesis, raises the possibility that a counterbalancing benefit exists for the host. Such a benefit may act at the population level to generate genomic diversity within a species and hence greater adaptability under new selective pressures, or at the level of primary gain for the individual. Recent studies have highlighted the occurrence of retrotransposition events in the germline and discovered a surprisingly high rate of mobilization in somatic cells. Here we examine the available evidence for somatic retrotransposition and discuss how this phenomenon may confer a selective advantage upon an individual or species.

Transposable elements are a prominent feature of our genetic heritage. In addition to providing nearly half of the human genome, TEs have generated numerous sequences that distinguish our DNA from that of other primates and more distant relatives. Whether these differences are a cause or effect of evolution, and whether TEs are parasitic or symbiotic mobile genetic elements, is the subject of long-term debate. Three retrotransposon families remain mobile in the human genome: L1, Alu and SVA. Of these, L1 is considered the main driver of retrotransposition (Fig. 1A). Proteins translated from its two open reading frames mobilize L1 RNAs in cis as well as Alu, SVA and other RNAs incorporating a polyA tail in trans (Fig. 1D). Approximately 3,000 retrotransposons (~100 L1, ~3,000 Alu, < 100 SVA) are transposition-competent per individual, in contrast to the millions of immobile sequences produced by ancestral TEs.

Other than a common pattern of near-exclusion from exons, the genomic distributions of L1, Alu and SVA are markedly different. L1 sequences are depleted in introns and very recent L1 insertions are more likely to be excluded from protein-coding genes than older insertions, suggesting that these events are strongly selected against. By contrast, recent Alu insertions are almost randomly distributed in the genome and SVA insertions are enriched in protein-coding genes. As noted above, the L1 machinery mediates L1, Alu and SVA mobilization, implying that each family is inserted in a similar genomic pattern and then redacted from the genome by natural selection depending on their impact. It is also possible that insertion site preference is modulated by unknown host factor interactions specific to each family.

An obvious consequence of insertional mutagenesis is genetic disease; TEs are associated with more than 75 human disorders. Likewise there are numerous documented cases of alternative transcripts and chimeric genes produced by TE insertions, often leading to expression of a host gene in a new spatiotemporal context (Fig. 1E). Several L1 sequence features, including a long polyA tail and strong internal 5' and 3' promoters, can also dramatically alter the expression of a host gene in cases of intronic integration.
while the epigenetic marks associated with L1 and other retrotransposons can modify chromatin state at integration sites and thereby drive rapid shifts in gene expression (Fig. 1E).

Given the multiple routes by which TEs can deleteriously alter the functional landscape of a genome, it is perhaps surprising that the global human population presents such a large number of dimorphic insertions. Recent studies using high-throughput sequencing (for reviews, see refs. 13, 24 and 25) have yielded a wealth of new insertion sites in healthy and diseased individuals, suggesting the full catalog of dimorphic and private insertions has been vastly underestimated and that roughly 1/20 live births harbor de novo retrotransposition events. Most of these new insertions are thought to be neutral and are ultimately lost or fixed through genetic drift. The overall impact of the remaining insertions is likely to be overwhelmingly deleterious, raising the question of why retrotransposition is allowed to continue at an apparently high rate. More effective TE suppression would prevent harmful mutations, both in the germline and during somatic development. A model of successful parasitism would suggest that we have simply failed; that somehow despite a clear selective advantage to the host in silencing retrotransposons, L1 has managed to evade all attempts to prevent its activity. However, suppression of L1 has been effective during our recent evolution: less than 0.002% of human L1 copies are transposition-competent, and even fewer are frequently active or “hot” (Fig. 1F). While the current state of L1 activity is a snapshot of a dynamic system, this could nonetheless suggest that it is evolutionarily advantageous to limit retrotransposition but not to totally eradicate it. For example, the South American rat genus Oryzomys has won this evolutionary arms race, apparently achieving L1 quiescence but, interestingly, this outcome coincides with a notable increase in karyotypic instability.

This leads us to consider the position that regulated germline retrotransposition confers a benefit upon a host population. L1 provides clues for how this system may have co-evolved with the governing transcriptional programs of the host. Paradoxically, the canonical L1 promoter has retained motifs necessary for its transcriptional suppression in the germline and throughout development (e.g., SOX2 binding sites) while new, usually 5' truncated, L1 insertions are rapidly inactivated despite breaking free of the suppression inherent to the canonical L1 promoter (Fig. 1B and C). Thus, L1 maintains its own suppression but is not entirely silenced, leading to a tolerable rate of insertional mutagenesis while maintaining increased genomic malleability and genetic diversity that may be selected on when a population is strongly...

Figure 1. L1 is the main driver of retrotransposition in human cells. (A) L1 structure. ORF1 encodes an RNA-packaging protein and ORF2 encodes a protein (ORF2p) with endonuclease and reverse transcriptase domains. (B) Expression of L1 is limited by transcriptional repression and (C) post transcriptional regulation. (D) L1 ORF1p and ORF2p form an RNP with a marked cis preference, but ORF2p can also mobilize other RNAs with a polyA tail in trans. (E) Diverse effects of L1 insertional mutagenesis on gene expression. (F) L1 is known to be highly active during embryogenesis, and in neural cells (G), resulting in somatic mosaicism. Somatic retrotransposition in other adult tissues may also occur.
pressed (e.g., in the cases of pandemic or famine). For example, an L1-mediated TRIM5-CypA gene fusion following the divergence of Old and New World primates provides owl monkeys with HIV resistance not seen in other New World monkeys.

Nonetheless, a model founded exclusively upon observations of germline retrotransposition may be critically incomplete. We propose that L1 activity during ontogenesis is

(4) which transcription factors govern L1 activation in somatic cells other than neurons and (5) whether the same rules that apply to germ cells (e.g., a limited number of "hot" donor elements and families) also apply to somatic cells.

Moreover, as somatic events are by definition non-heritable, it is the propensity for L1 mobilization, rather than its consequences, on which natural selection may apply. If true, this may suggest that the brain is enriched for somatic mobilization as an innocent bystander in an evolutionary arms race occurring primarily in the germline. A large percentage of genes expressed in the brain are also expressed in the testis (the "brains and balls" phenomenon), meaning that L1 transcription may be activated in somatic cells as an accident of evolution. The mutagenic effects of these insertions may then be simply tolerated by somatic cells; in addition to a reduction in impact due to heterozygosity, each mutation is expected to affect only a small sub-population of mature cells.

Another, more striking, possibility is that somatic retrotransposition confers some primary gain upon the individual host. As noted by others, Barbara McClintock’s celebrated discovery of transposition-derived kernel variegation in maize was also the first description of somatic mosaicism caused by a transposable element. Singer et al. more recently provided a compelling case for the potential action of L1 in producing somatic mosaicism in neural cells, resulting in greater genetic diversity and thus a greater variety of behavioral phenotypes in isogenic animals. As at the population level, genetic diversity may be beneficial at the cellular level. One classic example, driven by RAG proteins domesticated from an ancient transposon, is V(D)J recombination, where somatic rearrangements in immunoglobins and T-cell receptors provide genetic diversification crucial for the adaptive immune system.

The contribution of TEs to the fitness and success of species may not be limited to their well-documented effects on the genome mediated through germline retrotransposition. Their potential role in driving genetic diversity both within and between individuals adds yet another layer to the complex relationship between TEs and their hosts. Characterization of the regulation and functional impact of somatic retrotransposition is now feasible, and may soon settle debate on whether TEs are merely globally successful parasites, or diverse genomic symbiotes.

Acknowledgments

G.J.F. receives the support of a C.J. Martin Overseas Based Biomedical Fellowship from the Australian NHMRC (575585). J.K.B. is supported by a Welcome Trust Clinical Fellowship (090385/Z/09/Z) through the Edinburgh Clinical Academic Track (ECAT). G.J.F. is funded by an Institute Strategic Programme Grant and a New Investigator Award from the British BBSRC (BB/H005935/1) and an EU FP7 Collaborative Project research grant (FP7-HEALTH-2010-259743).

References

3. McClintock B. The mélanisme of the maize. Science 1969; 165:349-57; PMID:575585). J.K.B. is supported by a Welcome Trust Clinical Fellowship (090385/Z/09/Z) through the Edinburgh Clinical Academic Track (ECAT). G.J.F. is funded by an Institute Strategic Programme Grant and a New Investigator Award from the British BBSRC (BB/H005935/1) and an EU FP7 Collaborative Project research grant (FP7-HEALTH-2010-259743).


