

RESEARCH HIGHLIGHT

Transposable elements: not as quiet as a mouse

Rita Rebollo^{1,2}, Ying Zhang^{1,2} and Dixie L Mager^{1,2*}

See research article: <http://genomebiology.com/2012/13/6/R45>

Abstract

In this issue of *Genome Biology*, Nellåker *et al.* show massive purging of deleterious transposable element variants, through negative selection, in 18 mouse strains.

Keywords Transposable elements, variants, mouse strains, next-generation sequencing, evolution, natural selection, comparative genetics

Genomes have been bombarded by insertions of transposable elements (TEs), which have contributed nearly half of our genetic material [1]. The major TE families in the human and mouse genomes are LINEs and SINEs (long/short interspersed nuclear elements) along with endogenous retroviruses (ERVs). Although they still retain a parasitic life cycle and can still cause harmful mutations, TEs can also participate in gene regulation and donate protein functions to the host, and they may increase the evolvability of species [2,3]. Owing to their multitude of potential effects, there is increasing interest in identifying TE insertional variants (TEVs) and assessing their role in genetic and phenotypic variation within a species or between related species. Methods for detecting such TEVs have evolved from the early days of Southern blotting to high-throughput sequencing technologies (Figure 1). Efforts to document TEVs among mouse strains have been ongoing for years and include recent studies that catalogued ERV TEVs in eight inbred mouse strains and LINE TEVs in four inbred strains [4-6].

In this issue of *Genome Biology*, Nellåker *et al.* [7] have massively increased the catalog of mouse TEVs by identifying 103,798 such variants of all the major TE classes (LINEs, SINEs and ERVs) in a total of 18 mouse genomes

(14 laboratory strains and 4 wild-derived mouse species) using whole-genome next-generation sequencing. Such an impressive catalog will undoubtedly be of great value for mouse geneticists. However, as pointed out by the authors [7], the technology used does have drawbacks. Owing to the short sequence reads, fine classification of TEs into subfamilies was not possible and the false-positive and -negative rates are somewhat higher than those obtained with other methods [4,6]. In contrast, PCR-based sequencing techniques, such as the transposon junction assay designed by Li *et al.* [6], allow specific TEV subfamily analysis in different mouse strains and localization of the exact insertion site, but are, nevertheless, limited in the number of TE subfamilies studied (Figure 1). Hence, whole-genome sequencing using short reads to find new TEVs is a tradeoff between the number of new TEVs found and the information that each TEV read carries. With continued advances in next-generation sequencing technologies producing longer and higher quality reads, it is expected that error rates and fine resolution of TE families will improve.

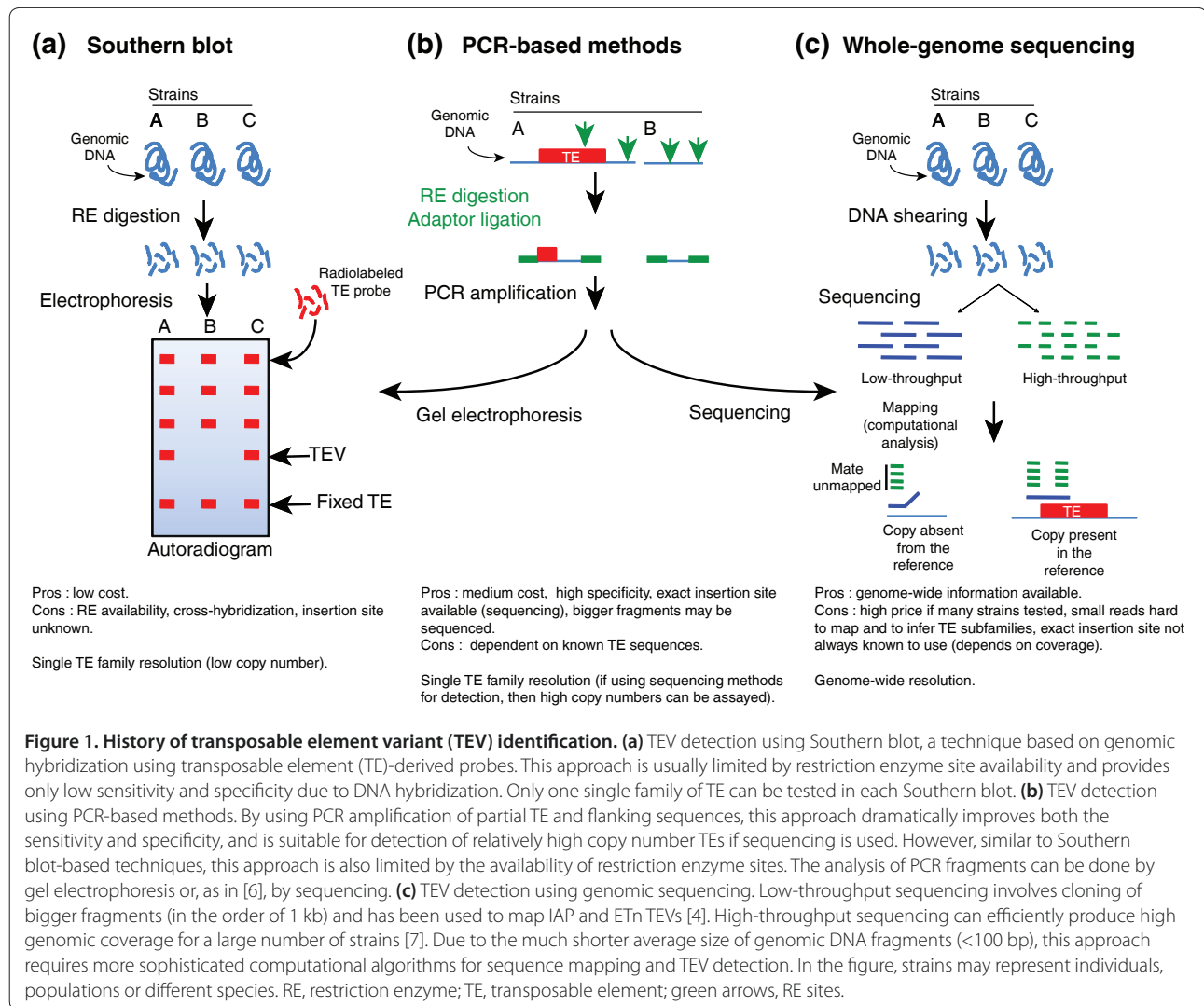
ERV activity abounds in mice

It has been known for many years that ERVs have been and continue to be highly active in mouse [4,6,8] and indeed the Nellåker *et al.* study [7] confirms this fact. It is interesting to compare the types of ERVs most prevalent among the TEVs with the lists of ERVs known to cause new germline mutations in inbred strains. Such a list was published in 2006 [8] and revealed that, of 63 cases, 32 were due to intracisternal A-type particle (IAP) insertions, 23 to early transposon (ETn) insertions and 8 to other ERV types. The predominance of IAPs among both the new, mutation-causing insertions and the old ERV TEVs [7] indicates that this family has maintained high activity, at least in some strains, to the present day. The majority of new IAP-induced mutations occur in the C3H background [8] and, as expected, IAPs account for 80% of all TEVs unique to this strain (Figure S1 in Nellåker *et al.* [7]). In contrast, the low fraction of ERV TEVs due to ETn insertions does not correspond to the prominent role of this ERV family in contributing to recent germline insertions. This finding suggests that, again at least in some strains, ETn elements have become

*Correspondence: dmager@bccrc.ca

¹Terry Fox Laboratory, British Columbia Cancer Agency, 675 West 10th Avenue, Vancouver, BC, V5Z1L3 Canada

²Department of Medical Genetics, University of British Columbia, Vancouver, BC, V6H 3N1 Canada



more active or are less likely to be fixed, even within a strain. The opposite situation occurs with mammalian long terminal repeat transposon (MaLR) elements, which contribute a relatively large fraction of ERV TEVs and of strain-specific TEVs but which have only one reported case causing a germline mutation [8]. These observations suggest either very recent extinction of MaLR activity or a lower likelihood of causing detrimental effects on insertion.

Blame the father!

Compared with the chromosomal distribution patterns for fixed TEs, Nellåker *et al.* [7] found that TEVs have a more uneven distribution, with depletion of all three TEV classes (LINEs, SINEs and ERVs) on the X chromosome. As discussed by the authors [7], this X depletion suggests that the vast majority of TE insertions occur in the male germline. Several differences between the

female and male mouse germline could account for this male bias for TE insertions, including differential epigenetic reprogramming and different mechanisms to silence TEs [9]. Moreover, the male germline undergoes many more divisions than the female one, providing increased opportunity for retrotranspositions. In addition, the authors [7] confirmed other studies showing positive selection for LINEs on the X, but they also noted evidence for retention of ERVs and SINE TEVs on this chromosome. Such retention could be due to several factors, including reduced opportunity for recombinational loss compared with the autosomes.

Not such a huge impact on gene expression

The million dollar question is of course whether TEVs are responsible for functional differences between strains and species. Nellåker *et al.* [7] approached this question by examining quantitative trait loci (QTLs) and, out of all

the QTLs associated with 100 traits studied, found only 12 loci that seem to be associated with TEVs. Furthermore, TEV classes are not associated with global changes in gene expression between the strains studied, suggesting that the impact of TEVs on gene expression is very subtle or non-existent. However, when the authors [7] filtered their panel of genes for those differentially expressed between strains, they observed significant association of those genes with TEVs. Therefore, TEVs do not seem to be responsible for global gene expression changes but require unknown segregating factors that might be related to, for instance, gene structure, gene regulation or TEV type. The authors suggest that TEs surviving immediate selection have a small or no impact on host gene expression. Indeed, fixed TEs and TEVs have very similar distribution patterns with respect to genes, including a strong orientation bias in introns, as has been reported before [4,6]. All these characteristics suggest that the vast majority of TEVs described by the authors [7] have already survived natural selection, implying a rapid and efficient purge of the mouse genome against deleterious TEs. The effects of inbreeding on TEVs may indeed contribute to fixation of neutral and slightly detrimental traits but also to exclusion of TEVs that have major consequences on the host genome. However, the youngest ERV TEVs (for example, those specific to just one strain), which were not extensively analyzed by Nellåker *et al.* [7], do show less orientation bias in gene introns, suggesting that some could indeed be deleterious but have not yet been lost by selection [4]. Hence, it is likely that strain-specific TE copies will have a larger role in gene expression differences than older TEVs. Moreover, as proposed by Burns and Boeke [1], it is possible that TEVs could act as soft modulators of gene expression, with subtle effects on expression levels or transcript structure that are not easy to detect.

What about humans?

Because mouse is the organism of choice for modeling human development and disease, it is interesting to compare the relative contributions of TEVs in these two species during the more recent stages of evolution. Although fixed TEs overall make up similar percentages of the mouse and human genomes, mouse comes out on top in terms of the number of TEVs. Nellåker *et al.* [7] found 62,800 TEVs that differentiate the classical *Mus musculus* strain C57Bl/6 from *Mus spretus*, representing 2 million years of divergence. In contrast, the reference human and chimpanzee genomes, representing 5 to 6 million years of divergence, have accumulated fewer than 16,000 TEVs, nearly all SINEs and LINEs [10]. These

lower numbers are due mainly to an overall decrease in activity of TEs, particularly ERVs, in primates [10]. It is not unexpected, therefore, that the fraction of disease-causing mutations due to new TE insertions is much higher in mouse than in human [1,8]. Nonetheless, intensive efforts to detect TE insertional variants among humans have catalogued over 7,000 TEVs, and evidence for cancer-specific TE insertions and normal cell-to-cell diversity in TE insertions in the human brain are beginning to be revealed (reviewed in [1]). Determining the significance of such TE-based variation in normal processes and in disease is an intriguing challenge.

Abbreviations

ERV, endogenous retrovirus; ETn, early transposon; IAP, intracisternal A-type particle; LINE, long interspersed nuclear element; MaLR, mammalian long terminal repeat transposon; QTL, quantitative trait locus; SINE, short interspersed nuclear element; TE, transposable element; TEV, transposable element insertional variant.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We apologize to our colleagues whose work was not cited due to space limitations. Work in our laboratory is supported by a grant from the Canadian Institutes of Health Research with core support provided by the British Columbia Cancer Agency.

Published: 20 June 2012

References

1. Burns KH, Boeke JD: Human transposon tectonics. *Cell* 2012, **149**:740-752.
2. Feschotte C: Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* 2008, **9**:397-405.
3. Fablet M, Vieira C: Evolvability, epigenetics and transposable elements. *BioMol Concepts* 2011, **2**:333-341.
4. Zhang Y, Maksakova IA, Gagnier L, van de Lagemaat LN, Mager DL: Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements. *PLoS Genet* 2008, **4**:e1000007.
5. Akagi K, Li J, Stephens RM, Volfovsky N, Symer DE: Extensive variation between inbred mouse strains due to endogenous L1 retrotransposition. *Genome Res* 2008, **18**:869-880.
6. Li J, Akagi K, Hu Y, Trivett AL, Hlynialuk CJ, Swing DA, Volfovsky N, Morgan TC, Golubeva Y, Stephens RM, Smith DE, Symer DE: Mouse endogenous retroviruses can trigger premature transcriptional termination at a distance. *Genome Res* 2012, **22**:870-884.
7. Nellåker C, Keane TM, Yalcin B, Wong K, Agam A, Belgard G, Flint J, Adams DJ, Frankel WN, Ponting CP: The genomic landscape shaped by selection on transposable elements across 18 mouse strains. *Genome Biol* 2012, **13**:R45.
8. Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN, Mager DL: Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet* 2006, **2**:e2.
9. Castaneda J, Genzor P, Bortvin A: piRNAs, transposon silencing, and germline genome integrity. *Mutat Res* 2011, **714**:95-104.
10. Chimpanzee sequencing consortium: Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 2005, **437**:69-87.

doi:10.1186/gb-2012-13-6-159

Cite this article as: Rebollo R, *et al.*: Transposable elements: not as quiet as a mouse. *Genome Biology* 2012, **13**:159.