EMBO *reports*

meeting report

Diamonds and rust: how transposable elements influence mammalian genomes

Conference on Mobile Elements in Mammalian Genomes

John V. Moran[1+] *& Harmit S. Malik*[2++]

[1]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan, USA, and [2]Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

The summer research conference on Mobile Elements in Mammalian Genomes took place between 5 and 10 July 2009 in Snowmass Village, Colorado, USA, and was organized by S. Martin, G.G. Schumann and P. Deininger. Photo credit: Deborah Bourc'his.

Keywords: transposable element; retrotransposon; long interspersed element 1; short interspersed element; genome; host defence

EMBO *reports* (2009) **10,** 1306–1310. doi:10.1038/embor.2009.248

See Glossary for abbreviations used in this article.

[1]Howard Hughes Medical Institute, Departments of Human Genetics and Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA
[2]Howard Hughes Medical Institute, Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA
[+]Corresponding author. Tel: +1 734 615 0456; Fax: +1 734 763 3784;
E-mail: moranj@umich.edu
[++]Corresponding author. Tel: +1 206 667 5204; Fax: +1 206 667 6522;
E-mail: hsmalik@fhcrc.org

Submitted 10 September 2009; accepted 20 October 2009; published online 13 November 2009

Introduction

Mammalian genomes are replete with both active and dead transposable elements (TEs), which can be grouped into two categories on the basis of their mobility intermediate. Classical transposons generally move (transpose) through a DNA intermediate by a non-replicative 'cut and paste' mechanism, whereas retrotransposons move through an RNA intermediate by a replicative 'copy and paste' mechanism. There is now incontrovertible evidence that TEs—once dismissed as 'junk' DNA—have had a crucial role in forging mammalian genome structure and function. This conference on Mobile Elements in Mammalian Genomes was a forum for researchers to discuss mechanistic aspects of TE mobility, the consequences of TE insertions, host factors that regulate TE activity and the use of TEs for genome engineering (Fig 1). Although previous editions of this conference have focused on mammalian non-long terminal repeat (LTR) retrotransposons, this year several presentations discussed advances in DNA transposon and LTR-retrotransposon biology in mammalian and non-mammalian systems. Here, we discuss some of the highlights of this exciting conference.

Non-LTR retrotransposition: when, where, how?

T. Eickbush (Rochester, NY, USA) gave the keynote address on the mobility mechanism of R2, a site-specific non-LTR retrotransposon that resides in the ribosomal DNA loci of many animals. R2 has a single open reading frame (ORF) that encodes reverse transcriptase and endonuclease activities required for retrotransposition by target-site primed reverse transcription (TPRT; Luan *et al*, 1993). His lab previously discovered that the R2 protein binds to a specific sequence in the R2 RNA 3′UTR and initiates TPRT by making a site-specific nick in the ribosomal DNA locus. The R2 reverse transcriptase uses the liberated 3′ hydroxyl as a primer to copy the R2 RNA into genomic DNA. Eickbush discussed recent *in vitro* data indicating that the R2 protein can bind to the 5′ region of R2 RNA, synthesize cDNA from a DNA template and displace RNA from an RNA/DNA heteroduplex. These data indicate that the R2 protein is needed for first-strand and second-strand R2 cDNA synthesis, and provide new insight into the TPRT mechanisms used by other non-LTR retrotransposons.

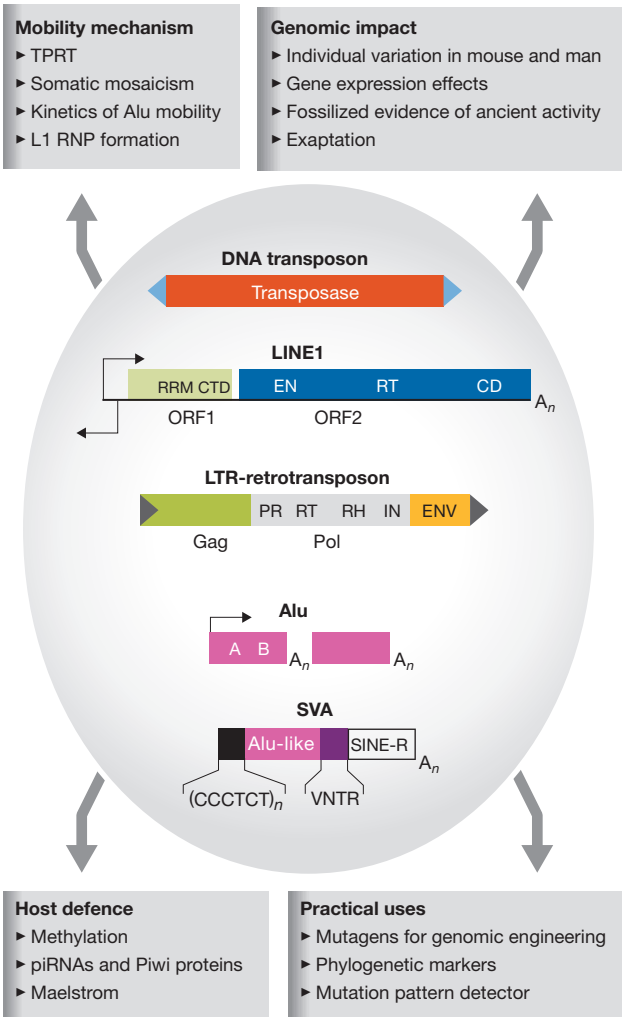

Fig 1 | Types of TE found in mammalian genomes and broad themes of TE biology discussed at this meeting. Autonomous TEs (such as some DNA transposons, LINE1 elements and LTR-retrotransposons) can encode proteins to promote their mobility. Almost all DNA transposons and LTR-retrotransposons in mammalian genomes are dead, although some LINE1 elements remain active. Non-autonomous TEs (such as Alu and SVA elements) rely on the proteins encoded by autonomous TEs to promote their mobility. Functional domains in the respective TEs (not drawn to scale) and the main themes that emerged from the meeting are highlighted. The A and B boxes represent *cis*-acting sequences in Alu DNA that are needed for transcription by RNA polymerase III. A_n, poly(A) tail; CD, cysteine-rich domain; CTD, carboxy-terminal domain; EN, endonuclease; ENV, envelope gene; IN, integrase; LINE1, long interspersed element 1; LTR, long terminal repeat; ORF, open reading frame; piRNA, Piwi-associated RNA; PR, protease domain; RH, RNase H; RNP, ribonucleoprotein; RRM, RNA recognition motif; RT, reverse transcriptase, SINE-R, short interspersed element repeat; SVA, SINE-R/VNTR/Alu; TE, transposable element; TPRT, target-site primed reverse transcription; VNTR, variable number tandem repeat.

TPRT-like mechanisms also are used by other retroelements; for example, for group II intron mobility and the extension of chromosomal ends by telomerase. Eickbush and colleagues previously used reverse transcriptase phylogenetic analyses to infer evolutionary relationships between retroelements. During his talk, Eickbush warned that phylogenetic artefacts—such as long branch attraction—might make it difficult to draw conclusions about distant evolutionary relationships among retroelements and urged caution when interpreting relationships on the basis of phylogenetic inference alone.

The keynote address provided a logical segue into an important focus of the meeting: understanding the LINE1 retrotransposition mechanism. LINE1s are ubiquitous mammalian TEs that comprise around 20% of genomic DNA. Active LINE1s encode two proteins that are required for retrotransposition: ORF1 and ORF2. Previous biochemical analyses by M. Singer (Bethesda, MD, USA) and S. Martin (Aurora, CO, USA) revealed that ORF1 contains an amino-terminal coiled-coil domain that mediates ORF1 trimer formation, a carboxy-terminal domain crucial for RNA binding, and a nucleic acid chaperone activity that might facilitate the initial steps of TPRT. However, the mechanism by which ORF1 binds LINE1 RNA was unknown. Recent structural analyses by O. Weichenrieder (Tübingen, Germany) demonstrated that human ORF1 contains a non-canonical RNA recognition motif in its central region (Khazina & Weichenrieder, 2009), and that a functional interplay between this motif and the C-terminal domain allows ORF1 to efficiently bind single-stranded nucleic acids. Because the RNA recognition motif is present in other non-LTR retrotransposons, these studies provide a foundation for understanding how ORF1 functions in retrotransposition.

The inability to detect LINE1 ORF2—which encodes endonuclease and reverse transcriptase activities—in cultured cells has hampered the study of LINE1 mobilization *in vivo*. Studies by N. Gilbert (Montpellier, France) and J. Goodier (Philadelphia, PA, USA) have overcome this technical hurdle. Gilbert used epitope-tagged engineered human LINE1s to visualize ORF1, ORF2 and LINE1 RNA in discrete cytoplasmic structures. He found that missense mutations in the ORF1 RNA binding domain could perturb cytoplasmic ORF1 localization. By using a complementary approach, Goodier also reported the presence of ORF1, ORF2 and LINE1 RNA in cytoplasmic foci. Interestingly, both groups found that a subset of LINE1 cytoplasmic foci are associated with stress granules, which confirmed and advanced previous studies presented by Goodier at the 2007 meeting. Future experiments should clarify the consequence of this association on LINE1 retrotransposition.

LINE1 ORF2 can also act *in trans* to mobilize SINE RNAs, such as Alu elements and potentially SVA elements, certain non-coding RNAs and some mRNAs to form processed pseudogenes. Together, these sequences comprise more than 11% of human DNA. Previous studies by T. Heidmann's laboratory (Villejuif, France) revealed that Alu RNA efficiently hijacks LINE1 ORF2 and retrotransposes by TPRT. Here, A. Roy-Engel (New Orleans, LA, USA) reported that Alu retrotransposition is kinetically faster than LINE1 retrotransposition in cultured cells, which could partly explain the evolutionary success of Alu (Krouter *et al*, 2009). Similarly, P. Deininger (New Orleans, LA, USA) extended previous findings from V. Belancio (New Orleans, LA, USA) and described how inactive LINE1s or alternative LINE1 transcripts could provide a source of ORF2 to promote Alu retrotransposition *in vitro*.

Glossary

Dnmt3L	*de novo* methyl transferase 3L
ERCC1/XPF	excision repair cross-complementing 1/xeroderma pigmentosa complementation group F. This heterodimer protein complex functions in nucleotide excision DNA repair
LINE	long interspersed element
mRNA	messenger RNA
piRNA	Piwi-associated RNA
R2	a retrotransposon found in the ribosomal DNA of various animals
SINE	short interspersed element
SVA	SINE-R/VNTR/Alu elements, a class of SINEs in primate DNA
VNTR	variable number tandem repeat
UTR	untranslated region

Thus, Alu seems to have used several strategies to parasitize LINE1 ORF2 to ensure its propagation.

As non-LTR retrotransposons rely predominantly on a vertical mode of inheritance, it was assumed that LINE1 retrotransposition would be restricted mainly to the germ line. Contrary to this expectation, H. Kazazian (Philadelphia, PA, USA) found that LINE1 retrotransposition frequently occurs during early development. By using transgenic mice and rats that contained an engineered human or mouse LINE1, Kazazian found that LINE1 RNA transcribed in germ cells can be deposited into the zygote and can undergo retrotransposition during the early stages of development (Kano *et al*, 2009). Remarkably, most LINE1 retrotransposition events were not heritable and resulted in animals showing LINE1 somatic mosaicism. A. Muotri (San Diego, CA, USA) also touched on the theme of LINE1 retrotransposition-mediated somatic mosaicism, and reported that retrotransposition of an engineered human LINE1 can be regulated by methyl-binding proteins during neuronal differentiation. These findings complement his earlier work, which indicates that LINE1 retrotransposition leads to somatic mosaicism in the brain.

Transposable elements in the genome: history and impact

Whole-genome sequencing has ushered in a golden age for TE analyses and has profoundly increased our knowledge of the history and impact of both ancient and modern TEs on their hosts. A. Smit (Seattle, WA, USA) discussed how the landscape of TE retention differs dramatically between mammalian and bird genomes and noted that birds have many LINE lineages whereas mammalian genomes are dominated by one LINE1 lineage. He speculated that SINE elements, which are abundant in mammals but not in birds, could explain this difference. Indeed, the single mammalian LINE lineage might have emerged as a result of surviving the selective cost imposed by SINEs.

C. Feschotte (Arlington, TX, USA) exploited the 'fossil' record of TEs in several vertebrate genomes to reveal that 'space invader' transposons, which are extinct in most mammals, underwent remarkable amplification 40 million years ago and were transferred horizontally to several orders of tetrapods (Gilbert *et al*, 2009a). Feschotte is now pursuing possible vectors to account for this unprecedented burst of horizontal transmission. Similarly, C. Gilbert (Arlington, TX, USA) discussed the endogenization of lentiviruses in two lemur species. These findings force us to rethink the age of lentiviruses and provide insight into how they shaped primate evolution (Gilbert *et al*, 2009b). Finally, G. Schumann (Langen, Germany) used computational and molecular approaches to detail the role of external promoters in the transcriptional regulation of SVA elements, which are recently uncovered SINEs responsible for disease-producing insertions in humans (Damert *et al*, 2009).

Retrospective views are a powerful method to glean information about TE biology, but genome sequences provide an incomplete snapshot of TE diversity and new tools are needed to identify and annotate them. D. Pollock (Denver, CO, USA) presented a P-cloud computational algorithm to identify previously unannotated TEs in human DNA. Remarkably, he found that the TE content of the human genome might exceed greatly the 50% ascribed by existing methodologies. Several presentations then focused on using computational analyses, modern genomic technologies and second-generation DNA sequencing to gauge the impact of TEs on genome variation. L. Jorde (Salt Lake City, UT, USA) reported on the identification of mobile-element-associated structural variations in a recently completed human diploid genome sequence (Xing *et al*, 2009). Jorde's group identified approximately 700 novel mobile element insertions and about 140 novel mobile element deletions that result from intramolecular recombination events between TEs. Thus, mobile genetic elements are responsible for around 10% of structural variations in human DNA. Similarly, C. Beck (Ann Arbor, MI, USA), A. Ewing (Philadelphia, PA, USA) and J. Boeke (Baltimore, MD, USA) reported on the use of modern genomic tools and next-generation DNA sequencing technologies to identify polymorphic LINE1 insertions in humans. D. Symer (Columbus, OH, USA) used a similar methodology and advanced his previous findings to identify polymorphic retrotransposon insertions in various mouse lineages. The central theme in each presentation was that TE insertions provide a rich source of inter-individual genetic variation and that improved methods are needed to gauge the extent of TE diversity, not only in one genome but also in populations.

Various speakers highlighted how TEs could interfere with transcription or alter the post-transcriptional processing of mRNAs. M. Speek (Tallinn, Estonia) used *in vitro* assays to demonstrate that transcription from either the native LINE1 promoter or an antisense promoter located in its 5′UTR could interfere with the transcription of cellular mRNAs. Similarly, Kazazian reported that splice acceptor sites in SVA elements could capture exons from cellular genes and thereby alter their expression (Hancks *et al*, 2009), and Symer described retrotransposon polymorphisms that are linked to transcriptional and phenotypic differences between mouse strains. G. Carmichael (Farmington, CT, USA) reported that the presence of inverted Alu repeats in the 3′UTRs of some mRNAs made them targets for adenosine-to-inosine editing in intranuclear paraspeckle-associated complexes, leading to their retention in the nuclei of certain differentiated cells. Carmichael also reported that some of these mRNAs are not retained in the nuclei of human embryonic stem cells because these cells do not express an abundant 3.7 kb non-coding RNA (hNEAT1) required for intranuclear paraspeckle formation (Chen & Carmichael, 2009). These data indicate that repetitive elements can alter the transcription profile of genes and highlight the fact that human embryonic stem cells differ from differentiated cell types in the way in which they regulate the export of mRNAs containing inverted Alu elements in their 3′UTRs.

Not all consequences of TE activity are negative. Once they become established in the genome, TEs can be exapted by the host for useful purposes. A. Huda (Atlanta, GA, USA) reported on studies examining TE histone modification profiles to identify TEs that might have been exapted for host functions. N. Okada (Tokyo, Japan) reported that sequences in a small number of SINEs can function as specific enhancers of gene expression, which might have facilitated the development of features in the mammalian brain (Sasaki *et al*, 2008). Similarly, J. Jurka (Mountain View, CA, USA) extended his previous studies and reported the computational identification of novel evolutionarily conserved TE sequences that might function as *cis*-acting transcriptional control sequences. Finally, J. Chow (Paris, France) suggested that LINE1s might have a role in the nucleation of heterochromatin formation by establishing a transcriptionally silent nuclear compartment. In contrast to previous models, she reported that some 'young' LINE1s are expressed from the inactive X chromosome, and suggested that their transcription could influence the local spread of X-inactivation. Thus, the incremental accumulation of LINE1s on the X chromosome during the course of evolution might improve the efficiency of dosage compensation, supporting Mary Lyon's original hypothesis for LINE1's role in X-inactivation (reviewed in Lyon, 2006).

The host response to retrotransposition

Although TE insertions are occasionally useful to the host, transposition poses a significant fitness burden and must be tightly regulated to ensure host survival (Malone & Hannon, 2009). Several talks described the role of piRNAs in restricting germline TE activity. G. Hannon (Cold Spring Harbor, NY, USA) described how piRNAs might 'immunize' the *Drosophila* germ line against potentially sterilizing transposition events. He discussed how the *flamenco* locus encodes piRNAs against several *Drosophila* retroelements, and then presented data showing that the lack of maternally loaded piRNAs is responsible for the de-repression of TE activity in the *Drosophila* germ line. R. Ketting (Utrecht, The Netherlands) described his studies on the zebrafish Piwi proteins, Ziwi and Zili. These proteins are essential for piRNA biogenesis and a subset of the resultant piRNAs target and regulate TE activity by a 'ping-pong' model, previously described by Hannon in mice. Interestingly, Ketting, Hannon and A. Bortvin (Baltimore, MD, USA) also discussed how interactions between Piwi proteins and specific Tudor proteins seem to be crucial for piRNA production in zebrafish and mice. Bortvin also discussed his studies on the Maelstrom protein, which is implicated in piRNA function in mice and *Drosophila* (Soper *et al*, 2008). Previous studies indicated that unregulated TE activity—owing to either piRNA dysfunction or to the de-repression of TE expression—might lead to chromosomal asynapsis in male meiosis, resulting in sterility. However, Bortvin suggested that the controlled de-repression of TE activity during male meiosis might facilitate rather than impede chromosomal synapsis.

Other presentations also focused on germline defence against TEs. D. Bourc'his (Paris, France) discussed how the depletion of Piwi family genes or Dnmt3L affects TE expression and described how the interplay between DNA methylation and piRNA production regulates germline TE expression. T. Bestor (New York, NY, USA) described interesting parallels between the way in which embryonic stem cells defend themselves against retroviral integration events and how the mammalian germ line defends itself against LINE1s. Finally, Roy-Engel and P. Deininger (New Orleans, LA, USA) described how proteins involved in DNA repair, such as ERCC1/XPF, might restrict LINE1 retrotransposition in cultured mammalian cells, implicating the DNA repair machinery as a host-defence mechanism (Gasior *et al*, 2008).

The awesome power of TEs

Harnessing TEs for practical purposes was another important theme at the meeting. Y. Voziyanov (Ruston, LA, USA) described his efforts to re-engineer Flp recombinase to recognize novel sites for genome engineering. N. Craig (Baltimore, MD, USA) reported on her laboratory's efforts to characterize and re-engineer TEs from both the hAT and piggyBac superfamilies. Z. Izsvak (Berlin, Germany) and Z. Ivics (Berlin, Germany) discussed clever selection schemes to re-engineer the Sleeping Beauty transposon to achieve high transposition efficiencies in mouse and rat spermatogonial stem cells. They also discussed how these engineered transposons could be used as gene delivery vectors (Ivics *et al*, 2009). Finally, T. Xu (New Haven, CT, USA) demonstrated that piggyBac transposons are a remarkable resource for the creation of genome-wide mutations in mice and rats. Although some discussion focused on the advantages and disadvantages of specific TEs, overall it was clear that engineered DNA transposons, which are effectively dead in the human and mouse genomes, are unparalleled tools for mammalian genome manipulation.

As retrotransposons are homoplasy-free genetic markers that are identical by descent and have a known ancestral state—that is, the absence of an element from a given locus—they are a powerful resource for inferring phylogenetic relationships among organisms. M. Batzer (Baton Rouge, LA, USA) reported on the power of this approach to resolve phylogenetic relationships among species of macaques (Li *et al*, 2009). He also discussed potential pitfalls, including lineage sorting and multiple parallel insertions, that complicate such analyses.

The vast number of extinct orthologous TE insertions also present a means through which to query actual—that is, neutral—mutation rates. A. Furano (Bethesda, MD, USA) reported analyses from his lab that determine CpG and non-CpG mutation rates by comparing the sequence divergence of thousands of orthologous LINE1 sequences in chimpanzees and humans (Walser *et al*, 2008). He reported a steady decline in CpG content with evolutionary time, as might be expected given the higher mutability of CpG sites. Much less intuitive was his finding that the non-CpG mutation rate, as well as the ratio of transition to transversion, was dependent on CpG content. The results indicate that the CpG content might even influence the mutation rates of neighbouring sites; a remarkable insight that might not have been revealed without rummaging through primate 'junk' DNA.

Concluding remarks

Once considered evolutionary oddities, TEs have been established as one of the most important forces driving the evolution of mammalian genomes. The 2009 conference on Mobile Elements in Mammalian Genomes presented a unique opportunity for scientists with research interests in TE mobility mechanisms, TE evolution and the genomic impact of TEs to exchange ideas and form inter-disciplinary collaborations to tackle some of the cutting-edge questions in this field. We thank S. Martin, G.G. Schumann, and P. Deininger for making the effort to organize this wonderful meeting and eagerly await the 2011 conference.

ACKNOWLEDGEMENTS
J.V.M. was supported by grants from the National Institutes of Health (GM060518 and GM082970) and the Howard Hughes Medical Institute (HHMI). H.S.M. was supported by a National Science Foundation CAREER grant and is an Early Career Scientist of the HHMI. 'Diamonds and Rust' is the title of a song written by Joan Baez. We apologize to colleagues whose work is not covered in this meeting report owing to space constraints.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

REFERENCES
Chen LL, Carmichael GG (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear non-coding RNA. *Mol Cell* **35:** 467–478
Damert A, Raiz J, Horn AV, Löwer J, Wang H, Xing J, Batzer MA, Löwer R, Schumann GG (2009) 5′-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res* [published online] 29 Sep
Gasior SL, Roy-Engel AM, Deininger PL (2008) ERCC1/XPF limits L1 retrotransposition. *DNA Repair (Amst)* **7:** 983–989
Gilbert C, Pace JK, Feschotte C (2009a) Horizontal SPINning of transposons. *Commun Integr Biol* **2:** 117–119
Gilbert C, Maxfield DG, Goodman SM, Feschotte C (2009b) Parallel germline infiltration of a lentivirus in two Malagasy lemurs. *PLoS Genet* **5:** e1000425
Hancks DC, Ewing AD, Chen JE, Tokunaga K, Kazazian HH Jr (2009) Exon-trapping mediated by the human retrotransposon SVA. *Genome Res* [published online] 29 Sep
Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, Izsvak Z (2009) Transposon-mediated genome manipulation in vertebrates. *Nat Methods* **6:** 415–422
Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH Jr (2009) L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev* **23:** 1303–1312
Khazina E, Weichenrieder O (2009) Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci USA* **106:** 731–736
Kroutter EN, Belancio VP, Wagstaff BJ, Roy-Engel AM (2009) The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. *PLoS Genet* **5:** e1000458
Li J, Han K, Xing J, Kim HS, Rogers J, Ryder OA, Disotell T, Yue B, Batzer MA (2009) Phylogeny of the macaques (Cercopithecidae: Macaca) based on Alu elements. *Gene* [published online] 2 Jun
Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72:** 595–605
Lyon MF (2006) Do LINEs have a role in X-chromosome inactivation? *J Biomed Biotechnol* **1:** 59746
Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. *Cell* **136:** 656–668
Sasaki T *et al* (2008) Possible involvement of SINEs in mammalian-specific brain formation. *Proc Natl Acad Sci USA* **105:** 4220–4225
Soper SF, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A (2008) Mouse Maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev Cell* **15:** 285–297
Walser JC, Ponger L, Furano AV (2008) CpG dinucleotides and the mutation rate of non-CpG DNA. *Genome Res* **18:** 1403–1414
Xing J *et al* (2009) Mobile elements create structural variation: analysis of a complete human genome. *Genome Res* **9:** 1516–1526

John V. Moran

Harmit S. Malik

