

Meeting report

Genomic parasites and genome evolution

Zoltán Ivics

Address: Max Delbrück Center for Molecular Medicine, Robert Rössle Str. 10, D-13092 Berlin, Germany. Email: zivics@mdc-berlin.de

Published: 15 April 2009

Genome Biology 2009, **10**:306 (doi:10.1186/gb-2009-10-4-306)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2009/10/3/306>

© 2009 BioMed Central Ltd

A report of the Second International Conference/Workshop on the Genomic Impact of Eukaryotic Transposable Elements, Pacific Grove, USA, 6-10 February 2009.

Transposable elements (TEs) are genetic elements with the unique ability to move in the genome. TEs are major components of the repetitive fraction of genomes; for example, TE-derived sequences make up about 45% of the human genome. The most abundant transposons in mammals are non-long terminal repeat (non-LTR) retrotransposons represented by the long interspersed nuclear elements (LINEs) and the short interspersed nuclear elements (SINEs). DNA 'cut-and-paste' transposons are less abundant in mammals, and typically encode a transposase protein in their simple genome. Transposition can be exploited to harness these elements as gene vectors for diverse genome manipulations (see the review series in a special issue of *Genome Biology* [<http://genomebiology.com/supplements/8/S1>]).

Beyond their present-day use as research tools, TEs have been shaping genome structure and function for millions of years, and the impact of transposons on eukaryotic genomes was the central theme of a conference held recently at Asilomar. Nearly 40 years ago, Roy Britten (who spoke at the meeting) and Eric Davidson proposed that the spread of repetitive elements in the genome may play a key role in the evolution of gene regulatory networks. Today, TEs are no longer viewed as 'junk DNA'; they can undergo 'exaptation' (a term frequently used at the meeting), an evolutionary process in which a characteristic that evolved under natural selection for a particular function is placed under selection for a different function. For example, the feathers of birds were first used to retain heat and only later used for flight. There are now numerous examples of exaptation of TE-derived sequences described in the literature, and several were presented at the meeting. Here I cover a few of the highlights.

Transposon exaptation

David Haussler (University of California Santa Cruz, USA) presented data on TE sequences undergoing natural selection to control nearby genes. TEs are perfect genomic vehicles for distributing repetitive genetic material over the genome where, as Haussler pointed out, they might then act as binding sites for 'master regulators' represented by transcription factors (Figure 1). For example, binding sites for the tumor suppressor protein p53 are highly enriched in the LTRs of some human endogenous retroviruses (ERVs), and these sites represent more than 30% of the p53-binding sites in the genome. Expression of many genes that are linked to these LTRs are thus under the transcriptional control of p53. It appears, therefore, that even though many ERV insertions close to genes were selected against (probably because their effect on gene expression reduced fitness), a significant fraction became exapted to expand the p53 transcriptional network.

The thought-provoking hypothesis that multiple retrotransposon insertions made our brain mammalian was put forward by Norihiro Okada (Tokyo Institute of Technology, Japan). His group has characterized a SINE family called AmnSINE1 that constitutes a conserved noncoding element in mammalian genomes, suggesting that these sequences have acquired some function useful to the host. Okada used an *in vivo* enhancer assay in mice to show that a SINE locus closely linked to the *FGF8* (fibroblast growth factor 8) gene acts as a tissue-specific enhancer that drives *FGF8* expression in the developing forebrain. Moreover, another SINE locus linked to the gene *SATB2* appears to control tissue-specific expression of this gene in the lateral telencephalon. Okada suggested that particular SINE insertions might have been involved in the evolution of a neuronal gene regulatory network, leading to the exaptation of these elements for these functions in an ancestral mammalian species.

As well as noncoding regulatory sequences, DNA transposons encode potentially useful and elaborate

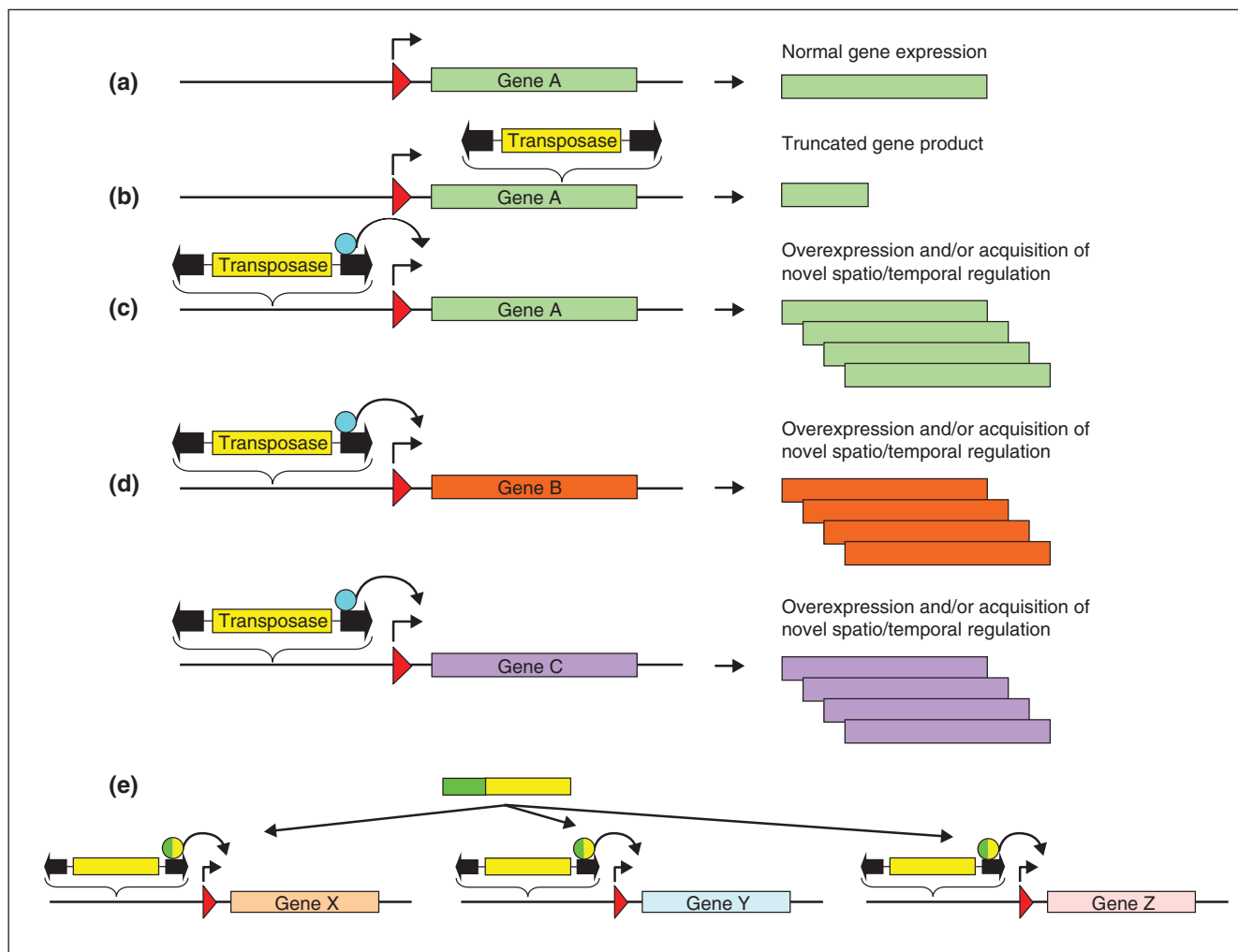


Figure 1
Possible consequences of transposon integration in or close to a transcription unit. **(a)** A hypothetical host genomic transcription unit with a promoter (red arrow) driving expression of Gene A. **(b)** Insertion of a transposon into the coding region results in a truncated gene product. This example shows a DNA transposon, but retroelement insertion can have similar consequences. The black arrows represent terminal inverted repeats flanking a transposase coding region (yellow box). **(c)** Transposon insertion into the 5' transcriptional regulatory region of the gene might introduce a binding site for a transcription factor (blue sphere), resulting in ectopic and/or overexpression of Gene A. **(d)** Transposition into multiple genes brings Genes A, B and C into a regulatory network under the control of a master transcriptional regulator. **(e)** The transposase coding region gets fused to a transcriptional regulatory domain, but can still bind to the inverted repeats of transposons dispersed in the genome. The transposase fusion protein might thereby become a master regulator of genes that have a transposon insertion.

enzymatic machinery (Figure 1) that has been exapted by the host genome via an evolutionary process referred to as 'molecular domestication'. One recent example of the emergence of such a domesticated gene is the insertion of a *piggyBac* (PB) element into an intron of the human Cockayne syndrome Group B gene (*CSB*) that leads to alternative splicing and the generation of a *CSB-PB* transposase fusion protein, in which only the first five exons of *CSB* are retained. Alan Weiner (University of Washington, Seattle, USA) presented genetic evidence suggesting that this *CSB-PB* fusion protein is advantageous in the presence of the normal *CSB* gene product, but harmful in its absence in humans. Earlier work by others established that *CSB*

encodes a chromatin-remodeling protein required for repair of UV-induced DNA damage. The presence in the human genome of more than 600 non-autonomous transposons (*MER85* elements) derived from *piggyBac* by internal deletions has been reported previously, and it is believed that these non-autonomous *MER85* elements were mobilized *in trans* by the *piggyBac* transposase at least 37 million years ago in a primate ancestor. Intriguingly, as Weiner discussed, many of the *MER85*-associated genes are downregulated by UV irradiation and *CSB*, suggesting that the *CSB-PB* fusion protein and its binding sites embedded in the dispersed *MER85* elements might constitute a potential gene regulatory network.

Transposon mutagenesis and regulation

Transposon movement also leaves its mark in the genome by aberrant transposition events that induce genomic rearrangements, including deletions, translocations and duplications of chromosomal DNA. Gerald Schumann (Paul-Ehrlich-Institut, Langen, Germany) reported that the composite non-LTR retrotransposon SVA occasionally carries over 5'-flanking genomic sequences to new chromosomal locations. Schumann suggested that this is presumably due to the requirement for external promoters to drive transcription of the elements that produce transcripts containing the entire SVA element plus upstream sequences. These 5'-transduced SVA elements may give rise to entire subfamilies as a result of repeated rounds of retrotransposition. Thus, SVA elements might have contributed to human genome evolution by capturing and dispersing DNA with potential regulatory or coding functions.

Transposons are potentially mutagenic as their insertion can interfere with normal gene function (Figure 1), and a plethora of regulatory mechanisms exist to keep transposition under control. The LINE-1 human retrotransposon is regulated at various levels, including transcriptional control by DNA methylation and premature polyadenylation and aberrant splicing of the LINE-1 transcript. Prescott Deininger (Tulane Cancer Center, New Orleans, USA) described a further regulatory mechanism that operates on the level of cellular DNA repair factors recognizing and eliminating transpositional intermediates containing a flap structure that is heterologous to the target DNA. The ERCC1/XPF complex that is normally involved in nucleotide excision repair is highly efficient at removing a partially inserted LINE-1 cDNA from the genome.

Small interfering RNA (siRNA)-mediated gene silencing is believed to have evolved to control the activities of TEs in diverse organisms, especially in gametes that can transmit potentially mutagenic transposon insertions to the next generation. Keith Slotkin (Cold Spring Harbor Laboratory, Cold Spring Harbor, USA) described microarray transcriptional profiling experiments showing coordinate expression of diverse TEs in the pollen of *Arabidopsis* (also observed in maize and rice), suggesting loss of *trans*-acting factors that otherwise keep these elements silent. Marking of an LTR-retrotransposon with a gene trap insertion revealed that the site of retrotransposon expression in the pollen is the vegetative nucleus (VN), which controls the development of the pollen grain but does not contribute DNA to the next generation. Transcriptional derepression of TEs leads to transposition events in the pollen; however, these events are not passed onto the next generation, consistent with their occurring in the VN. Transposon activation in the VN is associated with loss of heterochromatic silencing modifications such as DNA methylation. The activation of

TEs in pollen results in the production of siRNAs that are enriched in the generative sperm cells, suggesting that epigenetic reprogramming in the VN leads to TE reactivation and to the genesis of small RNAs that mediate TE silencing in the sperm cells.

The discussions on the intriguing impact of TEs on genome evolution and function at the Asilomar meeting were a fine celebration of Darwin's 200th birthday by representatives of the transposon community. After all, as one of the speakers put it: "life is a total mess, and what brings order into this mess is natural selection".