Molecular Cell, Vol. 10, 961–971, November, 2002, Copyright ©2002 by Cell Press

Constructing the Vertebrate Genome: Evidence from Eels that LINEs Mobilize SINEs

It has been speculated that long interspersed nuclear elements (LINEs), mobile elements present in vertebrate genomes, could mobilize other types of sequences in *trans*. Now we have a direct demonstration that LINEs can mobilize short interspersed nuclearelements (SINEs).

Much of the genetic substrate for natural selection is created by the action of mobile DNA elements (Bushman, 2001). Of course other mechanisms that generate genetic variation, such as accumulation of point mutations over time, are also important, but the more extreme changes brought about by DNA rearrangements allow for more radical genetic steps. In humans, for example, the formation of some 49 genes can be attributed to incorporation of fragments of mobile element sequences (Lander et al., 2001). Mobile elements also potentiate genomic rearrangements by providing portable regions of homology that act as targets for cellular homologous recombination enzymes. In another example, mobile element-encoded recombination enzymes can act directly on the host cell genome, allowing diverse rearrangements of cellular sequences.

In vertebrates, much of the mass of mobile element sequences is contributed by LINEs (long interspersed nuclear elements) and SINEs (short interspersed nuclear elements) (Lander et al., 2001; Venter, 2001). In humans, LINEs contribute about 20% of the genome, and SINEs such as Alu and MIR elements contribute 13%. The isolation of LINE elements from murine and human cells that are clearly capable of transposition has established that these elements are active transposons (Dombroski et al., 1991; Naas et al., 1998). However, there is also evidence that SINEs can move about in human cells as well, but unlike LINEs, SINEs do not encode enzymes that could explain their mobility. This has led to the hypothesis that SINEs may move by hijacking machinery encoded by LINE elements. Though plausible, this view has resisted experimental confirmation until now, when in a recent issue of *Cell*, Kajikawa and Okada (2002) demonstrate that LINEs can indeed mobilize SINEs in *trans* using, of all things, eels as a model system.

Kajikawa and Okada began their study by surveying the LINEs and SINEs present in a variety of organisms, including eels. A major complication in studying mobile elements in general is the high frequency of defective copies littering genomic sequences. Kajikawa and Okada sequenced many copies of the eel UnaL2 LINE, which allowed them to compile a consensus sequence likely to be close to that of the primordial active element. (The Una part of the name derives from *unagi*, Japanese for eel, a term familar to sushi enthusiasts). One copy of UnaL2 was quite close to the consensus, and subsequent studies showed that it was indeed an active element.

The near-consensus UnaL2 element was marked to allow detection of retrotransposition. A DNA segment was inserted into the element that contained a neomycin resistance gene in backward orientation relative to the UnaL2 sequence. Interrupting the *neo* gene was an intron in the forward orientation. Thus the construct could not confer neomycin resistance, because the "antisense" intron could not be removed by splicing. However, a round of UnaL2 transcription, splicing to remove the artificial intron, reverse transcription, and integration restores the *neo* reading frame, allowing positive selection of cells that have hosted retrotransposition events.

Using this assay, Kajikawa and Okada were able to show that the UnaL2 near-consensus copy was indeed active, and that the expected coding regions (reverse transcriptase and endonuclease) were required for retrotransposition. Up to this point, results with the UnaL2 LINE paralleled previous studies on mammalian LINEs (Feng et al., 1996; Kazazian, 2000). The authors further noticed that sequences in the 3′ untranslated region had homologs in the UnaSINE1 family of eel SINE elements. Sequences in this region are expected to be close to the probable start point of reverse transcription based on current models for LINE element replication. A potential RNA stem loop and a set of short (5 base) repeated sequences were conserved between UnaL2 and UnaSINE1, and assays using the marked UnaL2 element confirmed that mutations in these sequences inhibited retrotransposition.

Kajikawa and Okada then created a SINE element similarly marked with the interrupted *neo* gene and showed that it could also retrotranspose, but only in the presence of proteins supplied by UnaL2. Together with appropriate controls, this established that the UnaSINE1 can indeed be mobilized by UnaL2, providing proof of the principle that LINEs can indeed mobilize SINEs.

Marking the repeated sequences in the element 3′ end by mutation led to another finding with wide implications. Mutating some but not all copies of the short repeated sequence revealed that after a round of retrotransposition, the new element had a new arrangement of repeats. Evidently the reverse transcriptase "stutters" in copying these sequences—forming one new DNA copy of the RNA repeat, dissociating, rebinding to a new repeat, copying that, dissociating again, etc. This is an important clue to the mechanism of initiating reverse transcription, but in addition this provides surprising parallels with function of telomerase enzymes.

Telomerase enzymes qualify as reverse transcrip tases because they use an RNA template as the internal guide sequence to direct polymerization at chromosome ends to make a series of short repeats. Telomerase needs to carry out exactly the "stuttering" reaction seen with UnaL2 reverse transcriptase in forming the new sequence repeats at the ends of chromosomes. In fact, in *Drosophila*, a LINE-like element has been recruited to substitute for telomerase. The new data from UnaL2 emphasize the deep parallels between LINE reverse transcriptases and telomerases and reinforce speculations about their common evolutionary origin.

Kajikawa and Okada point out that several vertebrate

LINE-SINE pairs in addition to UnaL2-UnaSINE1 have conserved sequences at their 3′ ends, including elements from salmon and cichlids. Human LINEs and SINEs, in contrast, do not show such conservation, leaving open the question of what is going on in us. Do human LINEs mobilize human SINEs in a straightforward way, but without involvement of obvious conserved sequences between the two? Could there be more involved? The Kajikawa and Okada study of eel elements is sure to invigorate a new round of studies. In any event, their work clearly establishes that the LINE element machinery can mobilize SINEs, with profound consequences for vertebrate genomes.

Frederic Bushman
Infectious Disease Laboratory
Salk Institute
10010 North Torrey Pines Road
La Jolla, California 92037

Selected Reading

Bushman, F.D. (2001). Lateral DNA Transfer: Mechanisms and Consequences (New York, NY: Cold Spring Harbor Laboratory Press).

Dombroski, B., Mathias, S., Nanthakumar, E., Scott, A., and Kazazian, H.H., Jr. (1991). Science *254*, 1805–1808.

Feng, Q., Moran, J.V., Kazazian, H.H., and Boeke, J.D. (1996). Cell *87*, 905–916.

Kajikawa, M., and Okada, N. (2002). Cell *111*, 433–444.

Kazazian, H.H. (2000). Science *289*, 1152–1153.

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. (2001). Nature *409*, 860–921.

Naas, T.P., DeBerardinis, R.J., Moran, J.V., Ostertag, E.M., Kingsmore, S.F., Seldin, M.F., Hayashizaki, Y., Martin, S.L., and Kazazian, H.H., Jr. (1998). EMBO J. *17*, 590–597.

Venter, J.C. (2001). Science *291*, 1304–1351.

Enzymatic Cytosine Deamination: Friend AND Foe

AID and Apobec proteins have several, normally beneficial cellular functions. It has now been discovered that they act as DNA mutators in *E. coli* by deaminating cytosine in DNA. Thus, they may produce genome instability in mammals if not controlled.

In the time it takes to read this preview, the genome in every cell in your body will be damaged many times by the normal cellular environment. Processes such as deamination, depurination, oxidation, alkylation, lipid peroxidation, and strand breakage produce a wide variety of different lesions in DNA (Lindahl, 1993). Fortunately, these lesions are efficiently removed by several different DNA repair processes (Lindahl and Wood, 1999). However, repair processes are not perfect and they can be inactivated, leaving damage in the genome that can be converted to mutations. The fact that mutations arising from unrepaired DNA damage often have adverse biological effects (e.g., cancer) has been one of the major motivations for half a century of investigations of the sources and repair of DNA damage. It is therefore ironic that, while investigating the origin of mutations that are actually highly beneficial to human health, Neuberger and colleagues (Harris et al., 2002; Petersen-Mahrt et al., 2002) have now found a previously unappreciated source of mutagenic and potentially carcinogenic DNA damage, enzymatic deamination of cytosine in DNA by cytidine deaminase family proteins.

A fully competent human immune response against foreign antigens requires antibodies whose high-affinity binding capacity depends on somatic hypermutation of immunoglobulin genes (Diaz and Storb, 2002). These "beneficial" base substitutions are introduced into the variable regions of expressed immunoglobulin genes in a transcription-dependent manner at an extraordinary rate of about one per 1000 base pairs per generation of B cells in germinal centers. Somatic hypermutation requires the Activation-Induced Cytidine Deaminase (AID) protein (Muramatsu et al., 2000). AID belongs to a family of proteins (Gerber and Keller, 2001) that also includes the apolipoprotein B mRNA editing catalytic subunit 1 (APOBEC1). In gastrointestinal tissues, APOBEC1 deaminates cytosine 6666 in *apoB* mRNA to uracil to create a nonsense codon. In this issue of *Molecular Cell,* Harris et al. (2002) report that, as recently shown for AID (Petersen-Mahrt et al., 2002), APOBEC1 and two related Apobec proteins are mutagenic when ectopically expressed in *E. coli*. AID and the Apobec proteins are all related to cytidine deaminase, an enzyme already known to deaminate cytidine monomers (Gerber and Keller, 2001). The mutations induced by AID and Apobec in *E. coli* are C•G to T•A transitions whose frequencies are substantially higher in cells lacking uracil DNA glycosylase. These facts strongly suggest that AID and the Apobec proteins enzymatically deaminate cytosine residues in DNA to produce uracil. The mutagenic potential of uracil is particularly high because, unlike many other lesions, it can be efficiently replicated like normal thymine to yield C to T base substitution mutations. Thus, AID and Apobec proteins, enzymes with obviously beneficial cellular functions, including targeted somatic hypermutation, RNA editing, and possibly resistance to infectious nucleic acid (Harris et al., 2002), may produce widespread genome instability if not controlled properly.

It has been estimated that perhaps 400 uracil residues could arise per day in a human genome as a result of spontaneous, nonenzymatic cytosine deamination occurring at physiological pH and 37° (Lindahl, 1993). This implies that uracil is one of the more common potential sources of spontaneous mutagenesis. That enzymatic cytosine deamination also contributes to C to T transition mutations was previously suggested for DNA (cytosine-5)-methyltransferases (Shen et al., 1992). Just how much this challenge to genome stability might be further increased by AID and Apobec proteins awaits a direct biochemical demonstration that they actually do deami-