Transposable elements reveal a stem cell specific class of long noncoding RNAs

November 26 2012

Over a decade after sequencing the human genome, it has now become clear that the genome is not mostly 'junk' as previously thought. In fact, the ENCODE project consortium of dozens of labs and petabytes of data have determined that these 'noncoding' regions house everything from disease trait loci to important regulatory signals, all the way through to new types of RNA-based genes.

Yet over 70 years ago, it was first proclaimed that all this junk wasn't so junky. Barbara McClintock discovered the first utility of all of this junk DNA: jumping genes, also known as transposable elements. These genes serve only one purpose, which is to replicate themselves and reinsert randomly in the genome, or do they? Ironically, at the same time two other scientists (Roy Britten and Eric Davidson) proposed that jumping genes may be involved in regulating cell specificity. Indeed, in an exciting new study published in *Genome Biology*, John Rinn and David Kelley based at Harvard University and the Broad Institute in Boston, USA, provide genome-wide evidence that jumping genes may shape when a gene is turned on or off in stem cells.

"We set out to investigate how jumping genes have invaded the genome to potentially give rise to new genes in the 'junk regions'" says Rinn, the senior author of the study. "It has become very clear that there are thousands of long intergenic noncoding RNA genes (lincRNAs) that may herald a new paradigm for human health and disease." Yet how these genes have evolved from such a desert wasteland has remained a burning question. A new clue has emerged from the jumping genes that compose

nearly 50% of the human genome.

"I like to think of it as 'on the origins of lincRNAs'" says Rinn. "It doesn't take more than a brief survey of McClintock, Britten and Davidson's work in the 50s and 60s to realize that transposable elements were a great first place to look. The human genome is in a constant battle with transposable elements with them randomly hopping into new locations, for good or for bad." Kelley adds that "In my Ph.D. work assembling genomes from sequence fragments, these repetitive hopping genes were a major nuisance, which got me thinking about what they were doing in the genome". The study published by Rinn and Kelley finds a striking affinity for a class of hopping genes known as endogenous retroviruses, or ERVs, to land in lincRNAs. The study finds that ERVs are not only enriched in lincRNAs, but also often sit at the start of the gene in an orientation to promote transcription. Perhaps more intriguingly, lincRNAs containing an ERV family known as HERVH correlated with expression in stem cells relative to dozens of other tested tissues and cells. According to Rinn, "This strongly suggests that ERV transposition in the genome may have given rise to stem cell-specific lincRNAs. The observation that HERVHs landed at the start of dozens of lincRNAs was almost chilling; that this appears to impart a stem cell-specific expression pattern was simply stunning!"

These results also raise the tantalizing question of why transposable elements, derived from viruses, regulate stem cell-specific expression in mammals. Rinn hypothesizes that "transposable elements may not be limited to giving rise to new lincRNA genes, but may also provide an engine for the evolution of RNA-encoding genes. I like to think of it as the 'genome getting dirty': in the same way that kids that play in the dirt develop better immune systems, the genome may be 'getting dirty' with transposable elements, and once in a while, this has an advantageous effect of producing a new lincRNA gene."

What is clear is that transposable elements may control the tissue-specific expression of lincRNAs, thereby affecting the evolution and function of lincRNAs with important regulatory roles. Following on from these results, it will be interesting to determine other ways hopping genes may have shaped lincRNA evolution. Kelley notes that "This study merely scratches the surface of the possible roles of transposable elements influencing lincRNA function".

More information: Transposable elements reveal a stem cell specific class of long noncoding RNAs, David Kelley and John Rinn, *Genome Biology* (in press)

Provided by BioMed Central