New insight into the human genome through the lens of evolution

July 11 2013

By comparing the human genome to the genomes of 34 other mammals, Australian scientists have described an unexpectedly high proportion of functional elements conserved through evolution.

Less than 1.5% of the human genome is devoted to conventional genes, that is, encodes for proteins. The rest has been considered to be largely junk. However, while other studies have shown that around 5-8% of the genome is conserved at the level of DNA sequence, indicating that it is functional, the new study shows that in addition much more, possibly up to 30%, is also conserved at the level of RNA structure.

DNA is a biological blueprint that must be copied into another form before it can be actualised. Through a process known as 'transcription', DNA is copied into RNA, some of which 'encodes' the proteins that carry out the biological tasks within our cells. Most RNA molecules do not code for protein, but instead perform regulatory functions, such as determining the ways in which genes are expressed.

Like infinitesimally small Lego blocks, the nucleic acids that make up RNA connect to each other in very specific ways, which force RNA molecules to twist and loop into a variety of complicated 3D structures.

Dr Martin Smith and Professor John Mattick, from Sydney's Garvan Institute of Medical Research, devised a method for predicting these complex RNA structures – more accurate than those used in the past – and applied it to the genomes of 35 different mammals, including bats,

mice, pigs, cows, dolphins and humans. At the same time, they matched mutations found in the genomes with consistent RNA structures, inferring conserved function. Their findings are published in *Nucleic Acids Research*, now online.

"Genomes accumulate mutations over time, some of which don't change the structure of associated RNAs. If the sequence changes during evolution, yet the RNA structure stays the same, then the principles of natural selection suggest that the structure is functional and is required for the organism," explained Dr Martin Smith.

"Our hypothesis is that structures conserved in RNA are like a common template for regulating gene expression in mammals – and that this could even be extrapolated to vertebrates and less complex organisms."

"We believe that RNA structures probably operate in a similar way to proteins, which are composed of structural domains that assemble together to give the protein a function."

"We suspect that many RNA structures recruit specific molecules, such as proteins or other RNAs, helping these recruited elements to bond with each other. That's the general hypothesis at the moment – that non-coding RNAs serve as scaffolds, tethering various complexes together, especially those that control genome organization and expression during development."

"We know that many RNA transcripts are associated with diseases and developmental conditions, and that they are differentially expressed in distinct cells."

"Our structural predictions can serve as an annotative tool to help researchers understand the function of these RNA transcripts."

"That is the first step – the next is to describe the structures in more detail, figure out exactly what they do in the cell, then work out how they relate to our normal development and to disease."

Provided by Garvan Institute of Medical Research