GENOME EVOLUTION

A relaxed approach to errors

The expansion of early genomes presents something of a paradox. For a genome to expand, enzymes that prevent accumulation of errors are required, but a large genome is needed to encode these enzymes. A recent study indicates that this might not have been a serious problem after all — it predicts that primordial genomes were able to tolerate more errors than previously appreciated.

Kun and colleagues investigated the maximum genome size that could have been maintained before accurate replication evolved. This is determined by the rate of replication errors that can be tolerated before fitness is severely affected, as a larger genome will undergo more mutation, which can lead to loss of function.

The authors estimated the error threshold for ribozymes — RNA-based enzymes — as representatives of early RNA genomes. Ribozyme activity is not only affected by nucleotide sequence, but also by secondary structure. Combining this knowledge with results from previous mutagenesis experiments, the authors calculated the activity of many possible ribozyme sequences. Based on this, the maximum error rate that could be tolerated without loss of fitness was significantly higher than previous estimates that took only nucleotide sequence into account.

The error rate of the ribozymal replicators that copied the first genomes is unknown, so the authors used rates from RNA virus replicases as a rough estimate. Using the threshold for error tolerance calculated for ribozymes, an RNA-based protocell would support a genome of ~7,000 nucleotides. This is enough to contain 100 genes — many more than was previously thought.

By helping to resolve a long-standing paradox Kun and colleagues have revealed how primordial genomes could have increased in complexity — a key step towards understanding how life as we know it evolved.

Louisa Flintoft

References and links

ORIGINAL RESEARCH PAPER Kun, A., Santos, M. & Szathmáry, E. Real ribozymes suggest a relaxed error threshold. *Nature Genet.* 28 August 2005 (doi:10.1038/ng1621)

WEB SITE

Eörs Szathmáry's web site: http://www.colbud.hu/fellows/szathmary.shtml

TRANSCRIPTOMICS

The no-longer uncharted territory

Results of a genome-wide, collaborative effort to characterize the mouse transcriptome have been published. They reveal important information about the extent and the complexities of transcription.

The two recently published reports describe data from FANTOM3 (FANTOM stands for Functional Annotation of the Mouse), which is an international collaboration led by RIKEN, Japan. The group started by combining several approaches, including CAGE (new cap analysis gene expression) and two ditag technologies named as GIS/GSC (GIS, gene identification signature; GSC, gene signature cloning), to identify transcriptional start sites and termination sites. Corresponding pairs of these sites were identified for 181,047 independent transcripts — this number is an order of magnitude greater than the estimated number of genes in the mouse genome. The discrepancy can be attributed, at least in part, to the fact that alternative promoters and polyadenylation sites are associated with most transcriptional units. And at least 65% of transcriptional units contain several splice variants.

More than a third of the cDNAs in the FANTOM3 data set represent non-coding RNAs. The authors find that although non-coding RNAs are less conserved than 5′ and 3′ UTRs, their promoters are more conserved than those of protein-coding RNAs.

Mouse–human comparison revealed that human transcriptome is comparably complex. The authors point out that the abundance of transcriptional units and the overlap between them has implications for interpreting results from microarray and genome manipulation studies.

The FANTOM3 data also indicate that antisense transcription is more widespread than previously thought: 72% of transcriptional units overlap, at least partially, with a unit on the opposite strand. Co-expressed sense–antisense pairs show complex and tissue-specific regulation. Intriguingly, the authors find that expression of these pairs frequently co-varies — this is the opposite of what would be expected if the antisense transcript negatively regulated sense transcription. The authors suggest that "[i]f antisense transcripts do reflect the transcriptional activity of enhancers, the act of transcription from the antisense promoter may generate the regulatory interaction."

Although mouse transcriptome characterization will continue, it is already clear that the transcriptome landscape holds many surprises in store. We are only beginning to get our bearings.

Magdalena Skipper

References and links

ORIGINAL RESEARCH PAPERS The FANTOM Consortium and RIKEN Genome Exploration Research Group and Genome Science Group. The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005) | The FANTOM Consortium and RIKEN Genome Exploration Research Group and Genome Science Group. Antisense transcription in the mammalian genome. *Science* **309**, 1564–1566 (2005)