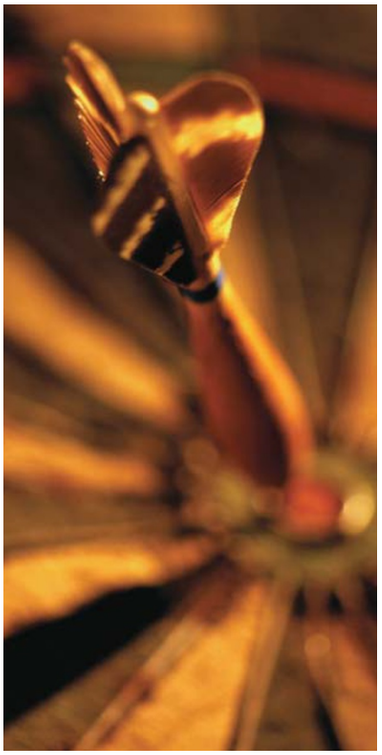

COMPARATIVE GENOMICS

It's a bullseye!

Sequencing genomes is still an expensive and time-consuming business; so, the question is, how can we reap the full benefits of comparative genomics while waiting for many more genome sequences to become available? The first data from the National Institutes of Health Intramural Sequencing Centre (NISC) Comparative Sequencing Program show that a strategy of targeting key genomic regions in a range of vertebrates 'hits the bullseye'.

The authors of the new *Nature* paper generated more than 12 Mb from 12 selected vertebrate species ranging from chimpanzee to zebrafish. They targeted the region that is orthologous to a 1.8 Mb section of human chromosome 7, which contains ten genes including the cystic fibrosis gene.

Comparative analyses confirmed that rodents and primates are sister groups to the exclusion of carnivores (such as cats and dogs) and artiodactyls (such as cows and pigs) because they share multiple transposon insertions that are not present in the other species studied. This conclusive result would not have been possible without data from an orthologous section of the genome in a range of mammals.

The analyses also identified sequences that are conserved across multiple species — 'multi-species conserved sequences' (MCSs) — which are strong candidates for being functionally significant. Approximately one-third of the 1,194 MCSs examined (with an average length of 58 bp) overlapped with coding sequences or UTRs. Surprisingly, 950 MCSs were either intronic (648) or intergenic (302), most of which (98%) are not known regulatory elements.

By assessing how well different sets of species could detect a subset of these MCSs, the authors showed that most would have gone unnoticed if just the completed mammalian genome sequences, mouse and human, had been used. Similarly, almost all possible subsets of the species studied are much worse at detecting MCSs than the complete set.

The diversity of species examined in this study allowed the authors to get a real handle on the dynamics that underlie differences in vertebrate genome sizes. Large indels accounted for the greatest fraction of sequence differences among mammals, which was particularly evident for primates. These differences mostly reflect the insertion of repeats and the deletion of ancestral

EVOLUTION

Problems with premature termination?

Premature termination codons (PTCs) would spell disaster were it not for the quality control system of the cell, known as nonsense-mediated decay (NMD), which rapidly eliminates mRNAs with PTCs. Now, a report in the *EMBO Journal* by Izaurralde and colleagues shows that the NMD pathway in *Drosophila* seems to have elements of both the yeast and mammalian pathways.

A crucial step in NMD is the recognition of mRNAs that contain PTCs; so, how does a cell determine whether a stop codon is premature or legitimate? In mammalian cells, the presence of a downstream exon-junction complex (EJC) during translation termination provides warning of a premature stop codon. Components of the EJC interact with the UPF complex of proteins, which mediate NMD, triggering the rapid digestion (or degradation) of the offending transcript. In yeast, where there is usually no EJC, a *cis*-acting downstream sequence element (DSE) is the NMD trigger.

Previous studies have shown that three genes that are essential for NMD (*UPF1*, *UPF2* and *UPF3*) are conserved in humans and yeast. However, other human genes (*SMG1* and *SMG5–7*) have no yeast orthologues.

To further investigate the molecular mechanism of NMD, Izaurralde and colleagues created a smart assay to determine whether or not specific genes were vital for NMD, based on *Drosophila* cell lines that constitutively express alcohol dehydrogenase (*adh*) mRNAs, into which PTCs had been introduced. If a gene was essential to NMD, then silencing using RNAi would increase levels of these reporter mRNAs. Using this technique, Gatfield *et al.* showed that *Drosophila* orthologues of *UPF1–3*, *SMG1*, *SMG5* and *SMG6* are essential for NMD. Intriguingly, no *SMG7* orthologue was discovered.

Given the presence of components of the EJC in *Drosophila,* it was previously assumed that, as for mammalian cells, PTC detection would involve the EJC and exon–exon boundaries. However, Izaurralde and colleagues' reporter assay clearly showed that neither a downstream exon–exon boundary nor the EJC proteins that they tested were essential for NMD. So, against all expectations, *Drosophila* would seem to use an alternative — as yet unknown — method of PTC identification.

These studies provide us with a tantalizing glimpse into the evolution of the NMD mechanism, as *Drosophila* shares with yeast the ability to identify PTCs independently of exon boundaries, while at the same time requiring the products of *SMG1*, *SMG5* and *SMG6*, in common with higher eukaryotes.

Ruth Kirby,
Nature Publishing Group

References and links

ORIGINAL RESEARCH PAPER Gatfield, D. *et al.* Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. *EMBO J.* **22**, 3960–3970 (2003)

WEB SITE

The Izaurralde group: http://www-db.embl-heidelberg.de/jss/emblGroups/g_127.html

FURTHER READING Cartegni, L. *et al.* Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Rev. Genet.* **3**, 285–298 (2002) | Wilkinson, M. F. & Shyu, A. B. RNA surveillance by nuclear scanning? *Nature Cell Biol.* **4**, 144–147 (2002)