RNA WORLD

The expanding universe of tiny RNAs

The finding that RNA can regulate gene expression is among the most exciting discoveries made in recent years. As well as RNAi, PTGS in plants and quelling in *Neurospora*, there are the stRNAs that regulate developmental timing. Discovered in *Caenorhabditis elegans*, the stRNAs *lin-4* and *let-7* have been the defining members of a group of small RNAs that are now referred to as microRNAs (miRNAs). Typical miRNAs are ~22 nt long and are cleaved from larger (~70 nt) precursors that form a characteristic stem-loop structure. Families of miRNA genes are present in both plant and animal genomes. Now, Bartel and colleagues take a computational genomics approach to identify miRNAs that are conserved across vertebrates. Their computational procedure (MiRscan) predicts that vertebrate genomes contain 200–255 miRNA genes, representing nearly 1% of the predicted genes in the human.

MiRscan — the details of which are being published elsewhere — identifies the evolutionarily conserved stem-loop precursors. Within each potential precursor, it scans 21 nt at a time to find the closest match to the original worm miRNAs. The authors compared the human, mouse and *Fugu rubripes* genomes and identified ~15,000 stem-loop segments in the human. All of these fell outside of protein coding regions and were at least partially conserved in mouse and *Fugu*. MiRscan narrowed this number down to 188, but the sensitivity of the scoring indicated that this number might represent 74% of all miRNA genes, setting the maximum number at 255.

Given that some miRNA loci were already known, and that MiRscan identified 107 new candidates, Lim *et al.* point out that no more than ~40 new miRNA loci remain to be discovered in the human. This estimate depends on the accuracy of the MiRscan prediction, so the authors set out to verify their candidates. Although some were closley related to previously cloned miRNAs and others could be detected in a zebrafish cDNA library that had been constructed specifically to contain miRNAs and siRNAs, Lim *et al.* were left with 55 candidates that could not be verified. So, the authors calculated the minimum specificity value and, taking into account the sensitivity of the zebrafish experiment and the incompleteness of the genome, proposed 200 as the lower limit for the total number of human miRNA genes.

Although MiRscan was 'trained' on worm miRNAs, it was able to identify most of the vertebrate counterparts, indicating that although most miRNA sequences have not been conserved, some of the generic features of miRNAs and their precursors have been. The authors also provide a parallel between protein coding and miRNA gene families: miRNA genes represent nearly 1% of the predicted human genes, a proportion that is similar for other families of regulatory genes. Because miRNA genes are absent from yeast, Lim *et al.* speculate that they might have evolved to regulate cell differentiation and developmental patterning. This is certainly true for some of the miRNAs that are already known; undoubtedly, functions will soon be assigned to the newly identified miRNAs.

Magdalena Skipper

References and links

ORIGINAL RESEARCH PAPER Lim, L. P. *et al.* Vertebrate microRNA genes. *Science* **299**, 1540 (2003)

FURTHER READING Eddy, S. Non-coding RNA genes and the modern RNA world. *Nature Rev. Genet.* **2**, 919–929 (2001)

WEB SITES

Dave Bartel's lab: http://web.wi.mit.edu/bartel/pub

Chris Burge's lab: http://genes.mit.edu/burgelab

IN BRIEF

EVOLUTION

Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes.

Wittkopp, P. J. *et al. Proc. Natl Acad. Sci. USA* **100**, 1808–1813 (2003)

In this quantitative trait analysis, at least four loci were identified by marker association that contribute to the different pigmentation patterns that are observed in *Drosophila novamexicana* and *Drosophila americana*. Although the pigmentation in these species is similar to that seen in other *Drosophila* species that have been studied, the genetic basis of the convergent phenotypes is different. Of the four loci found, only one (*ebony*) had been previously associated with interspecific variation in pigmentation, showing that convergent phenotypes can result from divergent genotypes.

GENE REGULATION

Genome-wide identification of *in vivo Drosophila* Engrailed-binding DNA fragments and related target genes.

Solano, P. J. *et al. Development* **130**, 1243–1254 (2003)

After ultraviolet crosslinking of DNA–protein interactions, the authors used a chromatin immunoprecipitation protocol to find potential targets of Engrailed in the *Drosophila* genome. There were 203 Engrailed-binding fragments situated in intergenic or intronic regions, and the putative target genes that were located near these binding sites were found to be involved in a wide range of developmental processes. Engrailed regulation was confirmed for 12 of the 14 genes, including *frizzled2*, by examining their expression in flies that ectopically expressed *engrailed*.

TECHNOLOGY

A discrete self-assembled metal array in artificial DNA.

Tanaka, F. *et al. Science* **299**, 1212–1213 (2003)

DNA molecule provides a computing machine with both data and fuel.

Benenson, F. *et al. Proc. Natl Acad. Sci. USA* 4 March 2003 (10.1073/pnas.0535624100)

Fifty years after the structure of DNA was determined, its unique chemical properties are increasingly being put to good use. Two new studies take advantage of DNA's highly selective base pairing to use it as a building block for supramolecular ensembles. By replacing hydrogen-bonded base pairs in the double helix with metal-bonded base pairs, Tanaka *et al.* assemble an array of five Cu^{2+} in the middle of the DNA. Uniquely, this method allows metal ions to be arrayed in solution and opens up the possibility of DNA-based nanodevices such as molecular magnets and wires. Benenson *et al.* focus on the potential of DNA for molecular computing. For the first time, they show that the energy generated by hydrolysis of the DNA backbone can drive a molecular computation. The authors suggest that the ability of DNA to act as an energetically efficient information-processing device helps to explain its selection as the mechanism for genetic information transfer.