

FUNCTIONAL GENOMICS

Structure provides clues

Most genomic functions are encoded in the ~98% of the genome that does not encode proteins, but the identification of non-coding functional elements has proved tricky. A new study describes a structure-based comparative method for identifying functional elements and shows that DNA structure is a better predictor of function and evolutionary constraint than primary nucleotide sequence.

Sequence conservation in non-coding regions provides evidence that these regions may be functionally important, but previous sequence-based algorithms have identified few non-coding functional elements. Instead, Parker and colleagues looked at whether conserved structural features of DNA regions could be used to predict functional non-coding elements.

They used a previously developed method, based on the hydroxyl

radical cleavage pattern of DNA, to predict the shape of the DNA backbone and of the grooves of DNA at single nucleotide resolution, yielding structural profiles for DNA regions. The authors found that the relationship between DNA structure and the underlying sequence is not always clear cut. Very different nucleotide sequences can have similar local structures; moreover, even the alteration of a single base in an 11-mer sequence can have a significant effect on structure.

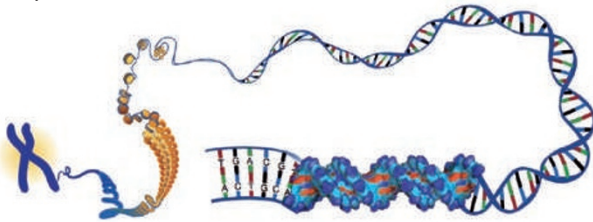
Next, the authors developed a computer program, Chai, which incorporates structural information to identify evolutionarily constrained regions. Using comparative sequence data from 36 species, Chai identified more known evolutionarily constrained bases than a previously developed sequence-based algorithm (binCons). Regions detected by Chai also overlapped with a higher proportion of DNase I hypersensitive sites and previously identified enhancers than the regions identified using binCons. Confirming that Chai can successfully identify functional enhancers, 8 out of 12 predicted enhancers that

contained elements identified by Chai drove upregulated expression in a luciferase reporter assay.

The results suggest that variants that affect DNA structure might have phenotypic consequences. Parker *et al.* explored this possibility for over 700 non-coding single nucleotide variants in the human genome that are associated with a phenotype. They quantified differences in the structural profiles between the mutant and non-mutant sequences, and compared this variation with the distribution of variation in structural profiles among ~17,000 neutrally evolving SNPs. The phenotype-associated distribution showed larger changes in structure compared with the baseline distribution, suggesting that non-coding nucleotide substitutions that affect DNA structure also affect phenotypes. The authors have constructed a database of changes in the structural profiles of all known SNPs in the human genome to be used as a resource to help identify functional SNPs on the basis of structure.

Meera Swami

DNA molecular topography. Image courtesy of D. Leja, National Human Genome Research Institute, Bethesda, Maryland, USA.



ORIGINAL RESEARCH PAPER Parker S. C. J. *et al.* Local DNA topography correlates with functional noncoding regions of the human genome. *Science* 12 Mar 2009 (doi:10.1126/science.1169050)

FURTHER READING Margulies, E. H. & Birney, E. Approaches to comparative sequence analysis: towards a functional view of vertebrate genomes. *Nature Rev. Genet.* 9, 303–313 (2008)