

 **IMPRINTING**

Human genome gets full marks



By training a computer to tell the difference between imprinted and non-imprinted genes, a group of researchers at Duke University (Durham, USA) has, for the first time, produced a catalogue of potential imprinted loci in the human genome. The list contains four times as many genes as were previously known, and two of those that might be involved in human disease have been experimentally validated as being imprinted.

Before the publication of this paper we knew of only 40 imprinted genes — defined as those with monoallelic, parent-of-origin-dependent expression in one or more tissues. Because the imprinted status of a gene might apply to just one isoform or tissue, identifying such loci experimentally is not easy. The authors therefore turned to computational means, relying on the knowledge that imprinted and non-imprinted genes are surrounded by distinct DNA-sequence patterns.

The approach involved teaching a computer to recognize the sequence features of imprinted genes by training

it on 600 genes of two types: in one class were known imprinted genes, whereas the other class comprised probable non-imprinted genes.

The resulting algorithm correctly identified all imprinted loci (100% sensitivity) and gave virtually no false positives (99% specific) in a cross-validation study. When the authors turned this method on the human genome, 156 autosomal genes showed up as being imprinted with high confidence, many residing in regions that are associated with important diseases such as cancer and diabetes. A separate algorithm performed well when asked to predict the parental preference of the expressed allele — that is, whether a locus was expressed from the maternal (56%) or paternal (44%) allele.

Two of the newly identified genes were on chromosome 8, on which no imprinted locus had previously been reported, and these were experimentally validated as being imprinted.

The potassium channel gene *KCNK9* is active primarily in the brain and is linked to cancer, epilepsy and bipolar disorder; the membrane-associated

guanylate kinase gene *DLGAP2* is frequently deleted in bladder cancers and is therefore a candidate tumour-suppressor gene.

A similar computational approach had previously been taken to identify imprinted genes in mice but, surprisingly, the overlap was small. Therefore, the authors suggest that mice might not be a good model system for studying imprinting. More of these predicted imprinted human genes remain to be validated. However, this is an exceptionally good starting point, not only for the investigation of individual genes but for detecting patterns in imprinted loci as a whole — these can be used to formulate models for the evolution and genomic distribution of this gene category and its role in development.

Tanita Casci

ORIGINAL RESEARCH PAPER Luedi, P. P. et al. Computational and experimental identification of novel human imprinted genes. *Genome Res.* 30 November 2007 (doi:10.1101/gr.6584707)

WEB SITES

Randy Jirtle's laboratory:

<http://www.geneimprint.com/lab>

Alexander Hartemink's web site:

www.cs.duke.edu/~amink