The mammalian transcriptome

ArticleInfo		
ArticleID	:	4651
ArticleDOI	:	10.1186/gb-spotlight-20021205-01
ArticleCitationID	:	spotlight-20021205-01
ArticleSequenceNumber	:	317
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	3
ArticleHistory	:	RegistrationDate : 2002–12–5 OnlineDate : 2002–12–5
ArticleCopyright	:	BioMed Central Ltd2002
ArticleGrants	:	
ArticleContext	:	130593311

Jonathan B Weitzman
Email: jonathanweitzman@hotmail.com

In the December 5 Nature the FANTOM Consortium and the RIKEN Genome Exploration Groups report a comprehensive characterization of the mouse transcriptome (*Nature* 2002, **420:**563-573).

Only a small proportion of the mouse genome is transcribed into functional readouts - protein-coding and non-coding mRNAs. The Japanese initiative is a large-scale project to sequence full-length mouse cDNAs for functional annotation. The FANTOM2 clone set contains 60,770 cDNAs of which around 21,000 clones were previously reported. The term 'transcriptional unit' (TU) is used to avoid complications arising from alternative splicing that generates proteins with different sequences. Most clones come from cDNA libraries from C57BL/6J mice (the strain used by the Mouse Genome Sequencing Consortium - see spotlight).

Computational analysis reduced the clones to 33,409 representative transcript clusters, around of half of which are predicted to code for proteins. Around 41% of transcripts showed evidence of alternative splicing. Splicing was also found in non-coding transcripts. The functional significance of non-coding TUs, many of which are expressed at low levels, remains to be elucidated. There are 2,431 pairs of overlapping sense-antisense transcripts, suggesting that antisense mRNA regulation may be quite widespread. About 92% of predicted proteins contain a predicted protein domain. The FANTOM2 collection includes mouse homolog of human genes involved in disease.

This study has generated the largest set of independent full-length clones for any species to date and will provide a useful tool for future functional genomics approaches.

References

1. *Nature*, [http://www.nature.com]

2. RIKEN, [http://www.riken.go.jp]

3. FANTOM, [http://fantom2.gsc.riken.go.jp]

4. Functional annotation of a full-length mouse cDNA collection.

5. The mouse genome, [http://genomebiology.com/researchnews/default.asp?arx_id=gb-spotlight-20021205-02]

6. MouSDB (Database of splice variants in the mouse transcriptome), [http://genomes.rockefeller.edu/MouSDB]

7. Disease gene matrix, [http://fantom2.gsc.riken.go.jp/supplement/disease_genes/]