

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

## Human and mouse transcriptome databases

ArticleInfo		
ArticleID	:	4361
ArticleDOI	:	10.1186/gb-2002-3-9-reports0047
ArticleCitationID	:	reports0047
ArticleSequenceNumber	:	21
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2002-7-1 Received : 2002-7-1 OnlineDate : 2002-9-2
ArticleCopyright	:	BioMed Central Ltd2002
ArticleGrants	:	

Chris Berrie

## Summary

The specificity of tissue transcriptional activities may be understood via high-throughput gene-expression profiling of a significant fraction of the human and mouse transcriptomes

## Significance and context

High-throughput gene-expression profiling is now used to define the normal transcriptome, to identify further potential markers for disease, and to augment annotation of genes with no known physiological function. Data generated by such methods are being included in a number of reference databases that can be accessed over the worldwide web. The study by Su *et al.* continues this trend, providing a free, publicly accessible and searchable website that represents a significant fraction of the human and mouse transcriptomes across diverse tissue types. The expansion in the number of gene-expression databases is just beginning, but will allow us to make combined analyses of the whole transcriptome. In turn, these analyses will enable us to define new networks of interactions, while also helping us to make informed decisions relating to specific gene(s) or protein(s) of interest.

## Key results

The authors have analysed 25 different human and 45 different mouse tissues using the Affymetrix human (U95A) and mouse (U74A) high-density oligonucleotide arrays and the GENECHIP 3.2 software. The principal result of this study is the generation of their human and mouse transcriptome databases. The paper includes a brief outline of how the databases can be used, although these suggestions are secondary to the information itself and are essentially based on specific ways of representing or grouping the data; for example, tissue specificity of gene expression within and between human and mouse, potential definition of gene function, and potential markers for human disease.

## Links

Supplementary data to *Proc Natl Acad Sci USA* 2002, **99**:4465-4470 are available as [Table 1, Tissue-specific gene expression in human](#) and [Table 2, Tissue-specific gene expression in mouse](#). The Genomics Institute of the Novartis Research Foundation provides a [Gene Expression Atlas](#). In addition various tools are available for analyzing such data, including cluster analysis (CLUSTER) and visualization (TREEVIEW), which can be downloaded from the website of [Michael Eisen](#), finding promoter region conserved motifs: ([AlignACE](#)) and ([SCANACE](#)); and finding potential human/mouse orthologs - [LocusLink](#).

## Reporter's comments

The main benefit of this study to biomedical research is the availability of the searchable database. This has allowed the massive amount of data obtained to be available to others, particularly as the actual retrieval of the data does not in itself provide anything conclusive. Indeed, the large amount of data of unknown significance that are generated by a single-throughput analysis can in itself be defeating. The authors have also expanded their study in an attempt to rationalize and confirm their data through their own examples of the types of analyses to which the data can be subjected. This is combined with a number of specific examples in each class of analysis, although some of this information could also be said to be potentially misleading. To take but one example: as detailed by the authors, the enriched expression of the orphan G-protein-coupled receptors (GPCRs) GPR31 and GPR9 in the pancreas could indicate they have a role in digestion or hormone secretion; however, the arginine vasopressin AVPR2 GPCR (which is known to be expressed predominantly in the kidney tubule) shows a lack of kidney-specific expression here. Thus, the data need careful consideration and continual expansion in order for researchers to trust that tissue specificity does indeed give potential functional information.

The authors also consider one of the potentially biggest areas of benefit that such analyses will provide in the future: the identification of markers of human carcinomas. They analyzed a cross-section of normal tissues, normal prostate and prostate cancer samples with a view to providing an analysis for potential markers for prostate cancer. In this case, at least, the authors provided a wider variety of prostate tumor samples (24), although this could also be self-defeating because of the potential variety of tumor types and progression stages seen even within a single tissue type. Furthermore, if one considers the top eight gene candidates as individual 'markers', a cross-check through the authors' database indicates that only one (Hs.301947) showed a high level of expression in their prostate cancer sample, with most appearing to be more prostate-specific than prostate-carcinoma-specific. Even with the wider inclusion of prostate cancer samples, therefore, any data obtained at this level need to be taken as merely indicative. At the same time, it is unfortunate that normal human mammary tissue was not included in the analysis, as this is one area of gene-expression profiling that would be further used by others already involved in the search for mammary tumor markers. Despite these criticisms of the 'supporting data' provided, if the study is accepted as comprising the collecting of the data then the authors have succeeded in providing a potentially very useful and usable analysis.

## Table of links

[Proceedings of the National Academy of Sciences of the United States of America](#)

[Table 1, Tissue-specific gene expression in human](#)

[Table 2, Tissue-specific gene expression in mouse](#)

[Gene Expression Atlas](#)

[Michael Eisen](#)

[AlignACE](#)

[SCANACE](#)

[LocusLink](#)

## References

1. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, et al: Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA. 2002, 99: 4465-4470.