

RESEARCH HIGHLIGHT

The similarity of gene expression between human and mouse tissues

Robin D Dowell*

See research article: <http://genomebiology.com/2010/11/12/R124>

Abstract

Meta-analysis of human and mouse microarray data reveals conservation of patterns of gene expression that will help to better characterize the evolution of gene expression.

The phenotype of an animal cell originates from a combination of its gene content and the regulation of those genes. In general, vertebrates have roughly equivalent numbers of genes, so differences in gene regulation are postulated to explain the vast amount of phenotypic variation within this group [1]. Most vertebrates also share a common body plan in which many organs behave essentially identically across diverse species and conditions. The large potential for variation in gene expression to influence phenotype is evident from the broad range of cell types present in an organism, all originating from a single genome. Understanding how evolutionary changes in gene expression contribute to phenotypic divergence is an important and open question. Gene-expression profiles across species can be compared to determine the conservation and divergence of transcription. The transcript abundance for each gene is measured and the collection of these measurements for all genes examined is the expression profile. In a recent issue of *Genome Biology*, Zheng-Bradley *et al.* [2] present a meta-analysis of recent work comparing mouse and human transcriptomes that confirms a greater degree of conservation of gene expression than previously thought.

Most studies of gene-expression profiling have used DNA microarrays, but cross-species microarray comparisons are unfortunately complicated by noise, assignment of homology, probe quality, platform variations, laboratory effects, genetic background, dynamic environments,

and organism status (such as age and sex). These technical difficulties effectively stack the deck against finding conservation, and often lead to overestimation of differences in gene expression between species. Effective cross-species comparison, therefore, requires these technical challenges to be addressed [3].

Mouse models are crucial in biomedical research, so understanding the differences and similarities between mouse and human is of fundamental importance. Consequently, tremendous attention has been paid to the comparison of human and mouse expression profiles. Early comparisons indicated that the expression profiles of orthologous genes differ substantively between human and mouse [4]. Subsequent analysis found that the divergence had been overestimated, mainly because of the large variation in sensitivity among the probe sets [5]. Building on improved computational approaches that help to correct for such variation [3], a series of human-mouse transcriptome comparisons, including the study by Zheng-Bradley *et al.*, has recently been published [2,6,7].

Meta-analysis of gene expression

Zheng-Bradley *et al.* [2] have gathered a large, quality-controlled collection of human and mouse microarray datasets, in which they can examine the clustering of expressed genes by tissue type in both species (rather than directly comparing the expression of individual genes between the species). They focus on a single Affymetrix platform for each species, selected on the basis of the platform's popularity in ArrayExpress. An earlier paper from the same laboratory [8] showed that a large set of curated human microarrays could be separated by principal component analysis (PCA) into three distinct classes: hematopoietic, neurological and malignancy. In the current study, Zheng-Bradley *et al.* [2] first run PCA analysis on a set of mouse experiments, finding a partitioning of the mouse expression datasets into remarkably similar classes: liver (hematopoietic), nervous (neurological), muscle and other cell types. Their focus then shifts to the analysis of a normalized merged dataset of both human and mouse microarrays, further filtered to include only high-quality orthologous probe

*Correspondence: robin.dowell@colorado.edu
Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309, USA

sets. PCA on the merged dataset is consistent with the patterns observed in the individual species, as half the variance is explained by the three dominant principal components (nervous, muscle, and liver).

The conserved expression signature found by Zheng-Bradley *et al.* [2] for brain and neural tissue is consistent with other recent meta-analyses. Chan *et al.* [6] compared multiple tissue-expression datasets across five vertebrate species: human, mouse, chicken, frog and pufferfish. They found tissue-specific expression patterns, identifying brain as the tissue with the most conserved transcription pattern across the five species. Miller *et al.* [7] undertook a brain-specific comparison of human and mouse transcription profiles. They used a meta-analysis strategy that groups genes on the basis of their co-expression relationships, and in agreement with the studies of Chan *et al.* and Zheng-Bradley *et al.*, they found that both gene expression and the summation of gene coexpression relationships are generally well conserved [7]. By eliminating modules with consistent transcription patterns across the species, Miller *et al.* were able to identify key between-species differences that may explain the inability to construct a satisfactory mouse model of Alzheimer's disease.

Both Zheng-Bradley *et al.* [2] and Chan *et al.* [6] reach the conclusion that related tissues within different species are more similar than unrelated tissues within the same species. Interestingly, this finding appears true even over broad evolutionary distances, as the Chan *et al.* study indicates that a major component of tissue gene expression has remained intact since the last common ancestor of vertebrates. Conservation of expression in tissues across diverse species is consistent with the notion that functionally important biological processes should be conserved. In fact, the recent 'tissue-driven hypothesis' postulates that a gene's tissue-expression pattern might constrain the permissible variation in its expression [9].

Are orthologous genes expressed in the same tissues across species?

Zheng-Bradley *et al.* [2] then addressed the question of whether orthologous gene pairs have similar patterns of gene expression across the datasets. When comparing orthologous genes on species-specific arrays, the different hybridization properties of the probe sets used on each platform can profoundly bias estimates of expression levels [5]. To minimize this effect, Zheng-Bradley *et al.* [2] use the correlation of correlation coefficient, also known as the integrative correlation coefficient, to explore the conservation of expression between orthologous genes. This metric assumes that whereas the raw expression values may vary from study to study, the intergene correlations will be more invariant. In this way,

the measure borrows statistical strength across many genes and experiments to estimate the strength of conservation between orthologs. Using this metric and considering all tissue types, they found that a number of orthologous genes were expressed in a correlated fashion between mouse and human [2]. Within a single tissue type, Zheng-Bradley *et al.* found that a rank-based comparison of expression variance indicates that 42% of the most variable genes in one species have an orthologous counterpart in the most variable gene set identified in the other species.

Both Zheng-Bradley *et al.* [2] and Chan *et al.* [6] found that, in general, genes expressed in a highly tissue-specific manner show the greatest similarity of expression pattern between species. Perhaps not surprisingly, Zheng-Bradley *et al.* also observe substantial overlap between species in the expression patterns of genes involved in basic cellular processes such as transcription and protein phosphorylation. They also find a small, but statistically significant, overlap with Chan *et al.* with respect to the genes identified as having evolutionarily conserved expression patterns in different tissues [2].

Implications of conserved tissue expression

Our understanding of the evolution of gene expression is rudimentary compared to our understanding of protein sequence evolution. The fundamental paradigm of comparative genomics is that conservation is related to functional importance. It is therefore reassuring that the studies by Zheng-Bradley *et al.* [2] and others show that similar tissues share significant expression patterns across evolutionary time. It is, however, also somewhat surprising, as this conservation of expression appears despite rapid divergence of promoter sequence [6]. It is possible that the capacity of regulatory sequence to diverge rapidly while maintaining roughly equivalent functional outcomes may be necessary for maintaining both robustness and plasticity in regulation [10]. In general, we have much to learn about the rules governing the evolution of transcription.

As our understanding of the evolution of gene expression improves, it should become possible to use expression patterns to improve inference about gene function. Expression is a dynamic and continuous variable that changes with developmental and physiological states. Yet, a gene's transcriptional response provides important clues to its function. Expression profiles are therefore an important piece of information that should be considered, in addition to structure and sequence information, when annotating genes, particular in closely related gene families. Detectable expression differences between species or individuals can logically be divided into two sets, those that are selectively neutral (or nearly neutral) differences and those underlying observable phenotypic

difference. Only by accurately inferring expression conservation is it then possible to consider the impact of expression variation on phenotype.

Acknowledgements

This work was supported by funds from the Boettcher Foundation's Webb-Waring Biomedical Research Program.

Published: 17 January 2011

References

1. King M, Wilson A: Evolution at two levels in human and chimpanzees. *Science* 1975, **188**:107-116.
2. Zheng-Bradley X, Rung J, Parkinson H, Brazma A: Large scale comparison of global gene expression patterns in human and mouse. *Genome Biol* 2010, **11**:R124.
3. Lu Y, Huggins P, Bar-Joseph Z: Cross species analysis of microarray expression data. *Bioinformatics* 2009, **25**:1476-1483.
4. Yanai I, Graur D, Ophir R: Incongruent expression profiles between human and mouse orthologous genes suggests widespread neutral evolution of transcription control. *Omic*s 2004, **8**:15-24.
5. Liao BY, Zhang J: Evolutionary conservation of expression profiles between human and mouse orthologous genes. *Mol Biol Evol* 2006, **23**:530-540.
6. Chan ET, Quon GT, Chua G, Babak T, Trochesset M, Zirngibl RA, Aubin J, Ratcliffe MJ, Wilde A, Brudno M, Morris QD, Hughes TR: Conservation of core gene expression in vertebrate tissues. *J Biol* 2009, **8**:33.
7. Miller JA, Horvath S, Geschwind DH: Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathway. *Proc Natl Acad Sci USA* 2010, **107**:12698-12703.
8. Lukk M, Kapushesky M, Nikkila J, Parkinson H, Goncalves A, Huber W, Ukkonen E, Brazma A: A global map of human gene expression. *Nat Biotechnol* 2010, **28**:322-324.
9. Gu X, Su Z: Tissue-driven hypothesis of genomic evolution and sequence-expression correlations. *Proc Natl Acad Sci USA* 2007, **104**:2779-2784.
10. Weirauch MT, Hughes TR: Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends Genet* 2010, **26**:66-74.

doi:10.1186/gb-2011-12-1-101

Cite this article as: Dowell RD: The similarity of gene expression between human and mouse tissues. *Genome Biology* 2011, **12**:101.