Molecular Systems Biology 7; 552 Article number ; doi:10.1038/msb.2011.86
Citation: *Molecular Systems Biology* 7: 552

www.molecularsystemsbiology.com

molecular systems biology

NEWS AND VIEWS

A panorama of mammalian gene expression evolution

Jianzhi Zhang*

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA
*Corresponding author. Department of Ecology and Evolutionary Biology, University of Michigan, 1075 Natural Science Building, 830 North University Avenue, Ann Arbor, MI 48109, USA. Tel.: +1 734 763 0527; Fax: +1 734 763 0544; E-mail: jianzhi@umich.edu

Molecular Systems Biology **7**: 552; published online 22 November 2011; doi:10.1038/msb.2011.86

Thirty-six years ago, King and Wilson (1975) noticed in their comparison of humans and chimps that the two species, despite having numerous phenotypic differences, are extremely similar in protein sequences. They speculated that changes in gene expression levels and patterns must have played a more important role in phenotypic evolution than changes in protein sequences and functions. However, testing King and Wilson's hypothesis has not been easy, in part because of the difficulty in comparing gene expressions between species, especially at the genomic scale. Since ~10 years ago, microarray has been used to quantify and compare genome-wide gene expressions across species, but this technology has serious limitations. First, if different probes are used to measure the expressions of orthologous genes in two species, the estimated expression levels are usually not directly comparable because of differential affinities of different probes to their respective targets (Liao and Zhang, 2006a). Second, if the same probes are used for different species, the probe-targeting sequences must be identical between the species to avoid creating spurious expression differences. For these reasons, it has not been possible to examine genome-wide evolutionary patterns of gene expression except in a few closely related species (Khaitovich *et al*, 2005). RNA sequencing (RNA-seq) quantifies messenger RNA concentrations by directly sequencing their complementary DNAs, and therefore is immune to the hybridization-caused non-comparability problem of microarray. RNA-seq data have been generated from a number of genetic model organisms such as yeast, *Arabidopsis*, and mouse, and have been used to address a few evolutionary questions (Qian *et al*, 2010). What is lacking, however, are RNA-seq data from multiple conditions or tissues of non-model organisms that fill the huge evolutionary gaps among the few model organisms. In a recent issue of *Nature*, Brawand *et al* (2011) reported a large RNA-seq data set from six organs (cerebral cortex or whole brain without cerebellum, cerebellum, heart, kidney, liver, and testis) of nine mammals (human, chimpanzee, bonobo, gorilla, orangutan, macaque, mouse, opossum, and platypus), and one bird (chicken), including both males and females. This data set, largest of its kind in the literature, promises to provide a grand view of gene expression evolution in mammals.

What have the authors found in this data set? First, they observed that expression differences are greater among different organs of the same species than among the same organs of different species. This observation is not surprising, because (i) the evolutionary origins of the six organs predated the evolutionary separations of the mammals and (ii) the same organs of different species are functionally more similar than different organs of the same species. However, it is notable that an earlier comparison of human and mouse microarray data yielded an opposite pattern (Yanai *et al*, 2004), although it was reversed after data normalization that mitigated the non-comparability problem (Liao and Zhang, 2006a).

Second, the authors found that, for most organs, expression-level differences among species roughly reflect the evolutionary distances among the species, suggesting that the evolutionary time separating two species is the primary determinant of their expression differences. However, a few intriguing exceptions exist. For example, although chimpanzee and bonobo are the closest human relatives, gorilla appears more similar to human in testis, brain, and heart gene expressions. Whether this phenomenon is simply due to incomplete lineage sorting that renders gene trees inconsistent with species trees or has deeper biological reasons is worth further scrutiny. Another unexpected finding is that the mouse transcriptomes of several organs have evolved more slowly than those of other mammals. This is in sharp contrast to the observation that protein sequences evolve faster in rodents than in many other mammals. It is unclear how these seemingly inconsistent patterns of rodent expression evolution and protein evolution are explainable by rodents' high per year mutation rate and large effective population size.

Third, among the six organs examined, Brawand *et al* found that the rate of expression evolution is the lowest in the brain and highest in the testis. This observation is consistent with previous microarray results from apes and mice (Khaitovich *et al*, 2005). Interestingly, the evolution of protein sequences also tends to be fast when the proteins are testis specific while slow when they are brain specific (Khaitovich *et al*, 2005). Apparently, genes functioning in the mammalian brain are under exceptionally strong evolutionary constraints in both expression levels and protein functions.

Fourth, the authors found many expression changes in X-linked genes shortly after the origin of the mammalian X chromosome, which occurred in the common ancestor of eutherians and marsupials since its divergence from monotremes. This is probably caused by sex-related selective pressures on X and/or dosage reduction of X-linked genes in males upon Y degeneration.

In addition to these general findings, the authors identified groups of genes that exhibit coordinated expression changes

during mammalian evolution, which may be phenotypically relevant and can serve as candidate genes for future studies. They also found numerous individual genes that have experienced significant expression changes across mammals, which, when augmented with other data, may help uncover the molecular basis of phenotypic evolution.

Despite the numerous findings by the authors, I believe that the value of this RNA-seq data set is well beyond the discoveries so far made. I predict that this data set, much like an early mammalian microarray data set (Su *et al*, 2004), will be analyzed again and again by different scientists for different purposes. Several interesting analyses come to mind immediately. For instance, previous studies in microbes showed that the expression level of a gene is the dominant determinant of the rate of its protein sequence evolution. But this is not the case in mammals when microarray expression data are analyzed (Liao *et al*, 2010). It would be interesting to see if the use of the more accurate RNA-seq data alters the conclusion. Reliable measures of expression levels across multiple tissues and species also permit the study of general rules in expression evolution. For instance, previous microarray studies found the rate of expression-profile evolution to be negatively correlated with the absolute expression level and the tissue specificity of the gene (Liao and Zhang, 2006b). Whether these results remain and whether new rules can be found in RNA-seq data are interesting to explore. The availability of the RNA-seq data from the platypus and chicken also allows a direct test of the currently debated Ohno's hypothesis (Xiong *et al*, 2010), which asserts that the expressions of X-linked genes would be doubled upon Y degeneration to restore their previous expressions. These being said, testing King and Wilson's conjecture remains challenging.

Conflict of interest

The author declares that he has no conflict of interest.

References

Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, Albert FW, Zeller U, Khaitovich P, Grutzner F, Bergmann S, Nielsen R, Pääbo S, Kaessmann H (2011) The evolution of gene expression levels in mammalian organs. *Nature* **478:** 343–348

Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, Weiss G, Lachmann M, Paabo S (2005) Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309:** 1850–1854

King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* **188:** 107–116

Liao BY, Weng MP, Zhang J (2010) Impact of extracellularity on the evolutionary rate of mammalian proteins. *Genome Biol Evol* **2:** 39–43

Liao BY, Zhang J (2006a) Evolutionary conservation of expression profiles between human and mouse orthologous genes. *Mol Biol Evol* **23:** 530–540

Liao BY, Zhang J (2006b) Low rates of expression profile divergence in highly expressed genes and tissue-specific genes during mammalian evolution. *Mol Biol Evol* **23:** 1119–1128

Qian W, Liao BY, Chang AY, Zhang J (2010) Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet* **26:** 425–430

Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci USA* **101:** 6062–6067

Xiong Y, Chen X, Chen Z, Wang X, Shi S, Zhang J, He X (2010) RNA sequencing shows no dosage compensation of the active X-chromosome. *Nat Genet* **42:** 1043–1047

Yanai I, Graur D, Ophir R (2004) Incongruent expression profiles between human and mouse orthologous genes suggest widespread neutral evolution of transcription control. *Omics* **8:** 15–24