

Minireview

## The genomic 'inner fish' and a regulatory enigma in the vertebrates

John Malone and Brian Oliver

Address: Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA.

Correspondence: John Malone. Email: malonej@nidk.nih.gov. Brian Oliver. Email: oliver@helix.nih.gov

Published: 16 April 2009

*Journal of Biology* 2009, **8**:32 (doi:10.1186/jbiol131)

The electronic version of this article is the complete one and can be found online at <http://jbiol.com/content/8/3/32>

© 2009 BioMed Central Ltd

### Abstract

Information on how genomic information from fish to human encodes the same tissues has until now emerged one gene at a time. The study published in this issue now provides lists of genes and their expression levels for 20 vertebrate tissues spanning 450 million years of vertebrate evolution. It reveals a core set of genes with similar tissue-expression patterns yet no common regulatory signatures - a gene-expression paradox.

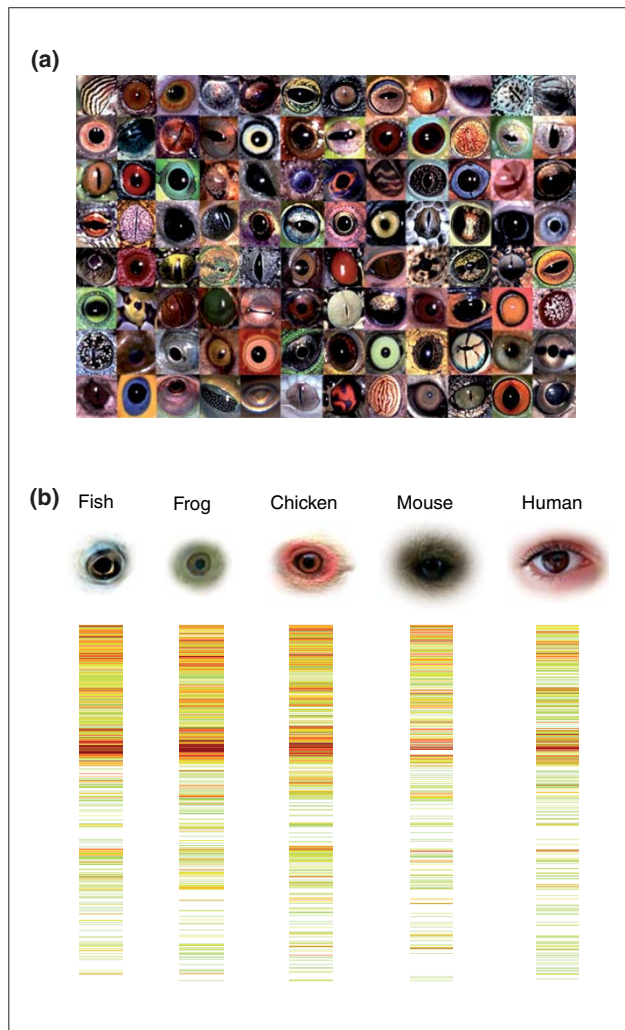
Even before the origin of species by descent from a common ancestor was posited, it was realized that groups of animals had related morphologies. Georges Cuvier, the father of comparative anatomy, viewed anatomical structures through the lens of form and function. Similar looking anatomical structures should have similar function, and anatomy could be used diagnostically to group organisms - a theory he termed "the correlation of parts" [1]. A famous story illustrates the idea. One of Cuvier's students dressed as the Devil with horns on his head and hoof-shaped shoes burst into Cuvier's bedroom when he was asleep and said, "I am the Devil. I have come to devour you!" Cuvier woke up and replied, "I doubt whether you can. You have horns and hooves. You eat only plants."

The relationship between form and function during evolution is a classic problem in biology. Yet to fully understand form and function at the level of anatomy and how those anatomical features change over time, it is important to probe the proximate mechanisms that create adult anatomy. It was therefore only natural that embryology became such an important tool. Karl Ernst von Baer showed that nearly all organs and tissues were derived from the same

embryonic layers in practically all animals. This similarity, an instructional 'inner fish', implies that core processes shared among all vertebrates shape much of ontogeny and ultimately adult anatomy [2].

DNA sequencing and the understanding that changes in sequence can also be used to deduce the relatedness of species marks another important landmark in the history of science. Conservative and non-conservative changes to codons have been a boon in understanding the influence of selection and chance on evolution. The form and function of vertebrate tissues and organ systems are a thought-provoking tour of the inner workings of organisms. Understanding what genes are deployed in a tissue- or organ-specific manner and across a variety of divergent species will be as valuable to our understanding of evolutionary processes and will inform us about what parts of the gene expression networks in an organism of particular interest, such as humans, have core functionality.

In this issue, Chan *et al.* [3] dissect the inner workings of vertebrate tissues and organs with a genomic scalpel and show that gene-expression profiles of orthologs are correlated.



**Figure 1**

(a) "Eyes of the world" by Paul Freed. When isolated from other aspects of head morphology, the remarkable conservation of eye morphology among animals is striking. Reproduced with permission. (b) Photos of eyes of various vertebrates shown above their gene-expression profiles. Data taken from Chan *et al.* [3].

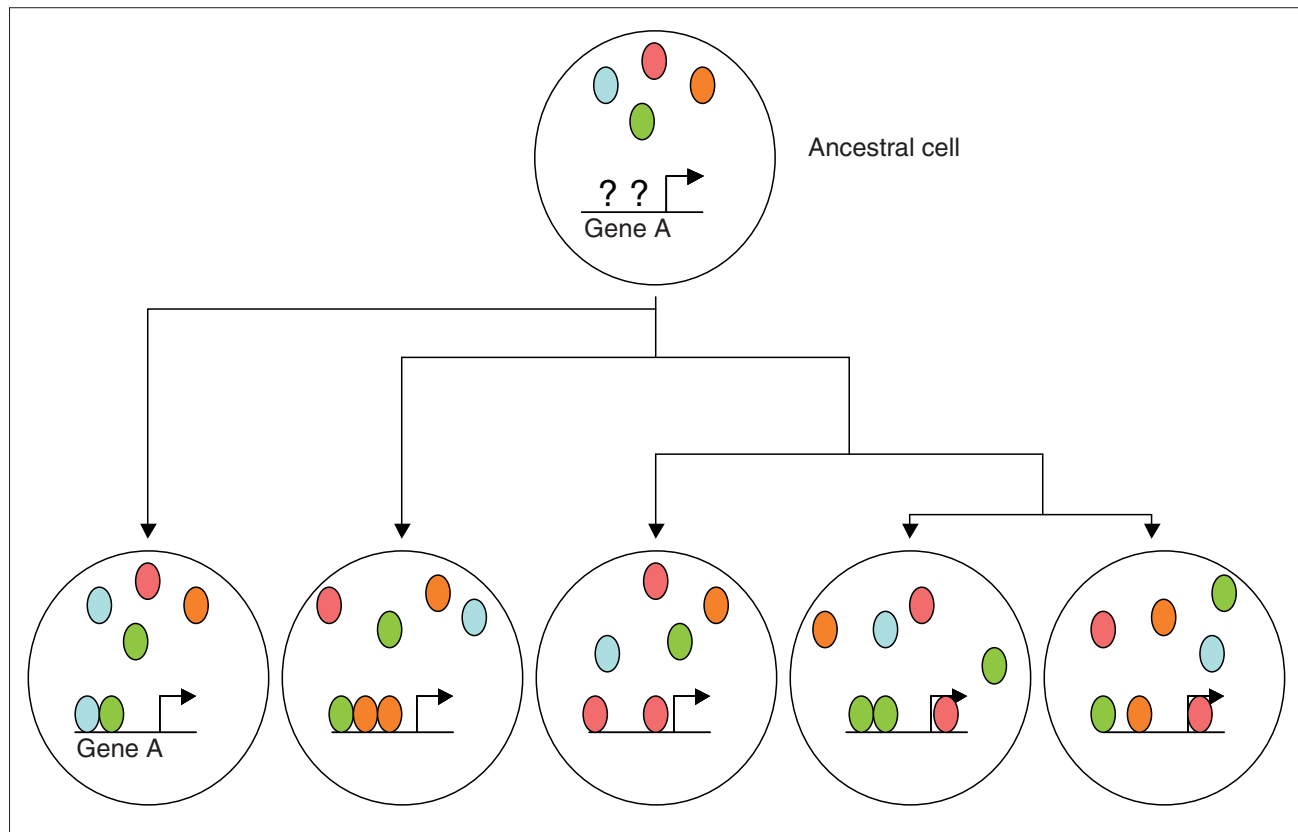
These data strongly suggest that both genes and gene-expression networks are derived from common ancestral genes and gene-expression networks. While maybe not surprising, this is important genomic confirmation of what comparative anatomists and embryologists have long described and believed. More important, these datasets are harbingers of the more quantitative and qualitative comprehensive description of morphology, a more theoretically grounded understanding of evolutionary processes that will follow. As pointed out by the authors, the contribution of selection and random drift in gene-expression profiles remains unclear. The theoretical framework for finding

meaning in comparative gene-expression and network topology data is still in its infancy [4]. These studies will provide the basis for a revolution in our understanding of evolution of gene-expression networks that is likely to thematically recapitulate the study of DNA sequence change in protein-coding regions.

Chan *et al.* [3] sampled a range of different tissues, which should probably become standard in comparative expression analyses, as not all organs tell the same story. For example, eyes must be very well adapted, as they show impressively conserved morphologies within the vertebrates [5]. Chan *et al.* show that they also express a set of core genes that have been highly conserved (Figure 1). As a counter example, Darwin suggested that sexual selection is an important driving force in evolution [6], and there are many studies showing that genes preferentially expressed in males, and in the testis in particular, are rapidly evolving in the vertebrates - in frogs [7], birds [8], rodents [9], and primates [10]. Chan *et al.* also find that testis gene expression is rapidly evolving in the vertebrates.

Curiously, they find that the kidney gene-expression profile may also be rapidly evolving. This could be related to the changes in water homeostasis in freshwater-dwelling organisms compared with those that inhabit drier or saltier environments, and the related excretion of urea or uric acid. Alternatively, it could be related to the closely linked embryonic origins and development of gonads and kidneys, both of which produce products that are passed from the body. The mesodermal urogenital ridge in the vertebrate embryo gives rise to both kidney and gonad, and the development of the kidney and the reproductive tract shows a remarkable development of functional embryonic nephritic tissue and an array of used, reused, and discarded plumbing arrangements (for example, Mullerian and Wolfian ducts) that connect the gonad and the kidney to the outside world [11]. It would be interesting to explore the idea that fast changes in testis expression profiles drive changes in the kidney as well.

Given that organs and organ expression profiles are derived from a common ancestor, one might expect that the regulation of gene expression should also be conserved. This is not really terribly different from the idea that similar organs should express similar genes. The logical idea that coexpression and co-regulation are linked was one of the early driving forces behind DNA microarray analysis, but links between coexpression of batteries of genes and co-regulation have not been as clearcut as initially expected. Indeed, Chan *et al.* [3] make the point that they fail to find conserved non-genic sequences that are expected to be driving the core organ-specific expression patterns.



**Figure 2**

The conservation of transcription factor production and target gene expression in a given cell type. This cartoon representation shows such conservation of gene expression driven by a lineage-specific arrangement of bound transcription factors (colored ovals). The arrow indicates gene expression. The factors expressed in the ancestral cell can be inferred, but the *cis*-regulatory arrangement cannot.

Not finding is a negative result, but if results like this continue to accumulate, it will be important to fully explore why. Although it is possible that we are not yet good enough at finding *cis*-regulatory sites, this negative result is becoming a common theme in array studies. Genomic features with highly conserved functions, such as core promoters and enhancers, clearly can be swapped between genes and species (where they function as expected as judged by the endogenous patterns) but show remarkable diversity in nucleotide sequence. For example, vertebrate transgenes expressed in hepatocytes of different species show similar expression even though the transcription factor occupancy profiles differ, and divergent enhancers from different species of *Drosophila* drive the same patterns of expression in *Drosophila melanogaster* embryos [12,13]. Remarkably, there have been natural wholesale exchanges of regulatory sequences to drive the expression of highly conserved ribosomal protein encoding genes in yeasts [14], suggesting that different transcription factors can coregulate large groups of genes in different species. It is beginning to

look as if there is a more profound explanation than technical limitations for our inability to find conserved *cis*-regulatory patterns among orthologs with similar expression patterns. Maybe conservation in *cis*-regulatory regions is difficult to find because they are highly malleable and transient.

It is perhaps worthwhile to step back and think about the unit of selection. For an animal to reproduce and pass on its genome, it needs to develop and use organs and organ systems. We know that early errors in organ development are catastrophic for adult viability and reproductive success. So, intuitively, the initial steps in a genetic pathway or the first committed step in a series of enzymatic reactions must be critical, but this is only true if there are few ways to generate a pattern or product. If there are multiple mechanisms, how an organism bootstraps to an acceptable outcome is less important. In a 'Christmas tree' model of evolution, this represents changing the branches on which ornaments are placed while maintaining the same

decorative appearance [15]. Sexual reproduction is a good example of this model. Sex results in remarkably similar gametes in a wide range of species, but the genetic pathways that govern the early steps of sex differentiation show astounding differences in theme and gene [16]. The malleability of sexual mode can be seen in the nematodes, where hermaphrodite and separate sexes have evolved multiple times using different underlying mechanisms, and within *Caenorhabditis elegans* the prime sex-determination signal can be experimentally switched from chromosomal to temperature [17].

If the females and males of the same species can be built using such different basal genetic hierarchies while maintaining the expression of critical well-adapted 'terminal' functions like sperm and eggs, then maybe organ gene-expression patterns can also be maintained with different underlying sets of transcription factors. This really boils down to asking how many solutions exist for a given expression pattern problem. If more than 10% of genes in a genome encode transcription factors and some substantial fraction of those genes are expressed in a given cell type, then there may be many ways to achieve the same transcriptional output. In these circumstances, a rather fluid exchange of regulatory mechanisms might be expected during the evolution of the vertebrates (Figure 2). *De novo* evolution of transcription factor binding sites should be relatively simple as these are short (usually less than 10 base pairs) and degenerate. The combination of conserved factor function and site turnover might result in multiple functionally equivalent *cis*-regulatory elements. Indeed, the exchange of one *cis*-regulatory sequence for another can occasionally be spotted [18]. If this proves to be generally true, then the implications for evolution and for using phylogeny to discover *cis*-regulatory regions are significant.

Since Cuvier, careful cataloging of anatomy in the context of phylogeny and development has had a major impact on our understanding of how living organisms evolve. While there are occasional examples of convergent evolution that has, for example, resulted in wings and thermal homeostasis in both mammals and birds, the vast bulk of comparative anatomy data reveals the deep roots of tissues and organ systems. Morphology indicates that the basic sensory, digestive, reproductive and excretory functions in animals are conserved. Although we do not have a rigorous understanding of the role of selection and drift in the evolution of gene expression, form and function has probably required the conservation of much of the core organ-specific expression network in the vertebrates. The lack of a relationship between coexpression and co-regulation at evolutionary timescales indicates that either we still do not understand how to find *cis*-regulatory

modules, or time has erased the vestiges of the intermediates in the vertebrates sampled.

## References

1. Coleman W: *Georges Cuvier, Zoologist: A Study in the History of Evolution Theory*. Cambridge: Harvard University Press; 1964.
2. Shubin N: *Your Inner Fish*. New York: Pantheon Books; 2008.
3. Chan ET, Quon GT, Chua G, Babak T, Trochesset M, Zirngibl R, Aubin J, Ratcliffe M, Wilde W, Brudno M, Morris QD, Hughes TR: **Conservation of gene expression in vertebrate tissues.** *J Biol* 2009, **8**:33.
4. Lynch M: **The evolution of genetic networks by non-adaptive processes.** *Nat Rev Genet* 2007, **8**:803-813.
5. Lamb TD, Collin SP, Pugh EN Jr: **Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup.** *Nat Rev Neurosci* 2007, **8**:960-976.
6. Darwin C: *The Descent of Man, and Selection in Relation to Sex*. London: John Murray; 1871.
7. Malone JH, Hawkins DL Jr, Michalak P: **Sex-biased gene expression in a ZW sex determination system.** *J Mol Evol* 2006, **63**:427-436.
8. Mank JE, Hultin-Rosenberg L, Webster MT, Ellegren H: **The unique genomic properties of sex-biased genes: insights from avian microarray data.** *BMC Genomics* 2008, **9**:148.
9. Turner LM, Chuong EB, Hoekstra HE: **Comparative analysis of testis protein evolution in rodents.** *Genetics* 2008, **179**:2075-2089.
10. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, Weiss G, Lachmann M, Paabo S: **Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees.** *Science* 2005, **309**:1850-1854.
11. Wagner GP, Lynch VJ: **Molecular evolution of evolutionary novelties: the vagina and uterus of therian mammals.** *J Exp Zool B Mol Dev Evol* 2005, **304**:580-592.
12. Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS: **Conservation of RET regulatory function from human to zebrafish without sequence similarity.** *Science* 2006, **312**:276-279.
13. Ludwig MZ, Patel NH, Kreitman M: **Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change.** *Development* 1998, **125**:949-958.
14. Tanay A, Regev A, Shamir R: **Conservation and evolvability in regulatory networks: the evolution of ribosomal regulation in yeast.** *Proc Natl Acad Sci USA* 2005, **102**:7203-7208.
15. Wagner GP, Lynch VJ: **The gene regulatory logic of transcription factor evolution.** *Trends Ecol Evol* 2008, **23**:377-385.
16. Wilkins AS: **Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway.** *BioEssays* 1995, **17**:71-77.
17. Haag ES: **The evolution of nematode sex determination: C. elegans as a reference point for comparative biology.** *WormBook* 2005, 1-14.
18. Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, Barkai N: **Rewiring of the yeast transcriptional network through the evolution of motif usage.** *Science* 2005, **309**:938-940.