

## REVIEW

# Use with caution: Developmental systems divergence and potential pitfalls of animal models

Vincent J. Lynch

*Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut*

Transgenic animal models have played an important role in elucidating gene functions and the molecular basis development, physiology, behavior, and pathogenesis. Transgenic models have been so successful that they have become a standard tool in molecular genetics and biomedical studies and are being used to fulfill one of the main goals of the post-genomic era: to assign functions to each gene in the genome. However, the assumption that gene functions and genetic systems are conserved between models and humans is taken for granted, often in spite of evidence that gene functions and networks diverge during evolution. In this review, I discuss some mechanisms that generate functional divergence and highlight recent examples demonstrating that gene functions and regulatory networks diverge through time. These examples suggest that annotation of gene functions based solely on mutant phenotypes in animal models, as well as assumptions of conserved functions between species, can be wrong. Therefore, animal models of gene function and human disease may not provide appropriate information, particularly for rapidly evolving genes and systems.

Biomedical studies of animal models have played an important role in elucidating the molecular basis and pathogenesis of human disease, as well as in developing and testing novel therapies [1,2]. As early as the 1930s, animal models such as the fruit fly (*Drosophila melanogaster*), zebrafish (*Danio rerio*), and mouse (*Mus musculus*) were being used to study the embryology

and development of tissue and organ systems [1,2]. These early studies significantly advanced our understanding of embryology and development by taking advantage of forward-genetic approaches to describe and characterize mutants, but it was not until the development of target transgenesis that it became possible to generate specific animal models of human disease.

---

To whom all correspondence should be addressed: Vincent J. Lynch, Department of Ecology and Evolutionary Biology, Yale University, PO Box 208106, New Haven, CT 06520-8106; E-mail: vincent.j.lynch@yale.edu.

†Abbreviations: Dfd, *Drosophila* homolog Deformed; PRL, prolactin; PLCZ1, phospholipase C zeta 1; NLS, nuclear localization signal; MHC, major histocompatibility; PGR, progesterone receptor; tin, tinman; DSD, developmental system drift; R-fng, Radical-fringe; TSPY, testis-specific protein, Y-encoded; DAF, Derived Allele Frequency; ACR, Acrosin; DAZ, Deleted in Azoospermia; DAZL, DAZ-like; PLCZ1, phospholipase C zeta 1; NLS, nuclear localization signal; KOMP, Knockout Mouse Project.

The development of transgenic animals in the early 1980s was a major breakthrough in molecular genetics and marked the beginning of modern biomedical studies. Gene targeting in model organisms enabled the deletion (“knockout”) or replacement (“knockin”) of any gene in the genome, allowing investigation of that gene’s function in development, morphology, physiology, and even behavior [1-3]. Gene targeting has been so successful in elucidating the molecular function of genes that it has become a standard tool in molecular genetics and biomedical studies and now is being used to resolve one of the major goals of the post-genomic era: assigning a function to each gene in the genome [4]. Despite the great success of animal models in studies of gene function and disease, those studies are based on the critical underlying assumption that gene functions and developmental systems are conserved between models and humans — which may not always be the case.

In this review, I outline some of the mechanisms that lead to evolutionary divergence in protein functions and gene regulatory networks and discuss how divergence between species can complicate studies of model organisms. I begin with a brief summary of the biological justifications for using animal models to study development and disease and discuss how principles emerging from evolutionary developmental biology challenge the assumption of conserved genetic architecture. Finally, I provide some examples of divergence in protein function and gene regulation between mouse — a common model organism for studying gene function and disease — and humans. While model organisms are a vitally important part of biomedical research, the assumption that functions are conserved between models and other organisms needs to be critically examined on a case-by-case basis.

## EVOLUTIONARY AND DEVELOPMENTAL JUSTIFICATIONS FOR ANIMAL MODELS

Comparative anatomists have long recognized the similarity of tissue and organ systems between diverse animals, particu-

larly within mammals. Indeed, these shared similarities led early comparative embryologists to formalize the concept of “homology,” or similarity between characteristics because of shared ancestry, as a distinct biological phenomenon [5]. The numerous similarities in nervous, cardiovascular, endocrine, immune, musculoskeletal, and reproductive system development and function between model organisms such as the mouse, fish, fly, and humans provide a strong evolutionary foundation that the developmental processes giving rise to homologous structures is likely to be similar between different species.

Perhaps the strongest evidence that homologous genes control the development of homologous structures is the finding that proteins between remarkably divergent organisms can be functionally equivalent, often despite great divergence in both protein similarity and time. For example, the first use of transgenics to demonstrate that homologous genes functioned in homologous ways was between the human Hox transcription factor *HoxB-4* and its *Drosophila* homolog *Deformed* (*Dfd*). McGinnis and colleagues (1990) tested whether the human *HoxB-4* gene could substitute for the regulatory functions of *Dfd* in *Drosophila* embryos by inserting the *HoxB-4* gene into the *Drosophila* genome. Amazingly, the human *HoxB-4* gene can properly regulate endogenous *Dfd* expression in developing embryonic and larval cells through its autoregulatory element [6]. Thus, the autoregulatory function of the human and *Drosophila* homologs has been conserved after more than 800 million years of divergence.

Probably the best-known example of homologous proteins initiating homologous developmental programs is the transcription factor *Pax6*. Loss of function mutations in *Pax6* result in similar eye malformations in mice, humans, and *Drosophila* [7-9], indicating that *Pax6* has a conserved function in eye formation, despite the vast differences between vertebrate and invertebrate eyes. Ectopic expression of *eyeless*, the *Drosophila* homolog of the vertebrate *Pax6* gene, in the

antennae, legs, and wings of flies induces eye development on these body parts, indicating *eyeless* functions a “master regulator” that switches on the developmental program for eye formation [10]. Remarkably, ectopic expression of the mouse *Pax6* gene in *Drosophila* embryos induces the formation of well developed compound eyes [10], while expression of the *Drosophila eyeless* gene in *Xenopus* embryos leads to the induction of vertebrate eyes [11]. These data strongly indicate the function of *Pax* genes as high-level regulatory switches that activate the gene regulatory network leading to eye development is conserved between animals.

Numerous other examples of functional equivalence have been identified between orthologous genes from very different organisms [12-19] and between paralogous genes from ancient gene and genome duplications [20-27]. The broad consensus resulting from these studies is that developmental systems do not change much through evolution, and, therefore, studies of gene function in one species are directly applicable to another. But is this conclusion justified?

### CHALLENGES TO THE ASSUMPTION OF CONSERVED GENETIC ARCHITECTURE

I have briefly reviewed some of the evidence suggesting that protein functions and the developmental systems to which they contribute are so conserved during evolution that biomedical studies in model organisms are directly applicable to human development, physiology, and disease. However, numerous studies have found that homologous proteins do not necessarily remain functionally equivalent during evolution [28-55]. For example, the transcription factor HoxA-11 has evolved a novel ability to activate *prolactin* (*PRL*) expression in endometrial stromal cells in placental mammals, while HoxA-11 from non-placental mammals like opossum, platypus, and chicken lack activation functions and can only repress *PRL* expression. Similarly, mouse phospholipase C zeta 1 (PLCZ1) is actively transported into the nucleus shortly after fertilization, but

PLCZ1 from other mammals lack the nuclear localization signal (NLS) found in mouse PLCZ1 and do not enter the nucleus. Furthermore, evolutionary modification of gene regulatory networks and developmental systems are mediated through changes in how genes are regulated and function [56-60], thus the assumption that homologous genes function in homologous ways between models and humans needs to be critically examined. In the following section, I review some of the ways that developmental systems can diverge and discuss how these factors can bias biomedical studies that assume conservation of regulatory systems between models and humans.

### *Positive selection promotes functional divergence*

Many powerful methods for detecting adaptive molecular evolution have been developed over the last decade. These methods compare the rate of synonymous (non-amino acid changing) and non-synonymous (amino acid changing) nucleotide substitutions to infer the strength and direction of selection acting on protein coding genes [61-63]. Since synonymous nucleotide substitutions do not lead to amino acid changes, they are not exposed to natural selection acting on the protein structure/function and accumulate in the genome at a more or less clock-like, or neutral, rate. On the other hand, non-synonymous nucleotide changes generate amino acid changes and are, therefore, exposed to the action of selection. If an amino acid change is deleterious for the structure/function of a protein, the new alleles' frequency in the population will remain low and eventually disappear. If, however, the amino acid change is beneficial because it generates a novel function or contributes other beneficial effects, the frequency of that allele in the population will increase until it completely replaces the ancestral allele. Thus, when amino acid changes are adaptive, the rate of fixation in the population will increase above the background rate of synonymous substitutions [61-63].

Numerous cases of molecular adaptation have been identified in various systems

from viruses to humans with the common theme that positive selection is very often associated with an escalating arms race, such as between immune genes and parasites, or the emergence of novel functions. For example, one of the first demonstrations of molecular adaptation was in the antigen recognition site of *major histocompatibility (MHC)* genes. It was shown that amino acid substitutions occurred more rapidly at this locus than either synonymous substitutions or amino acid changes in other parts of the protein [64,65]. Positive selection can also act to modify ancestral functions or generate novel ones. For instance, the *progesterone receptor (PGR)* evolved extremely rapidly in humans and chimpanzees with the majority of amino acid substitutions occurring in parts of the protein important for transcriptional activity [66]. Remarkably, the episodes of rapid *PGR* evolution are coincident with changes in the mechanism of parturition in higher apes, suggesting that using rodents as a model organism to study *PGR* actions in humans and other primates may not produce meaningful results.

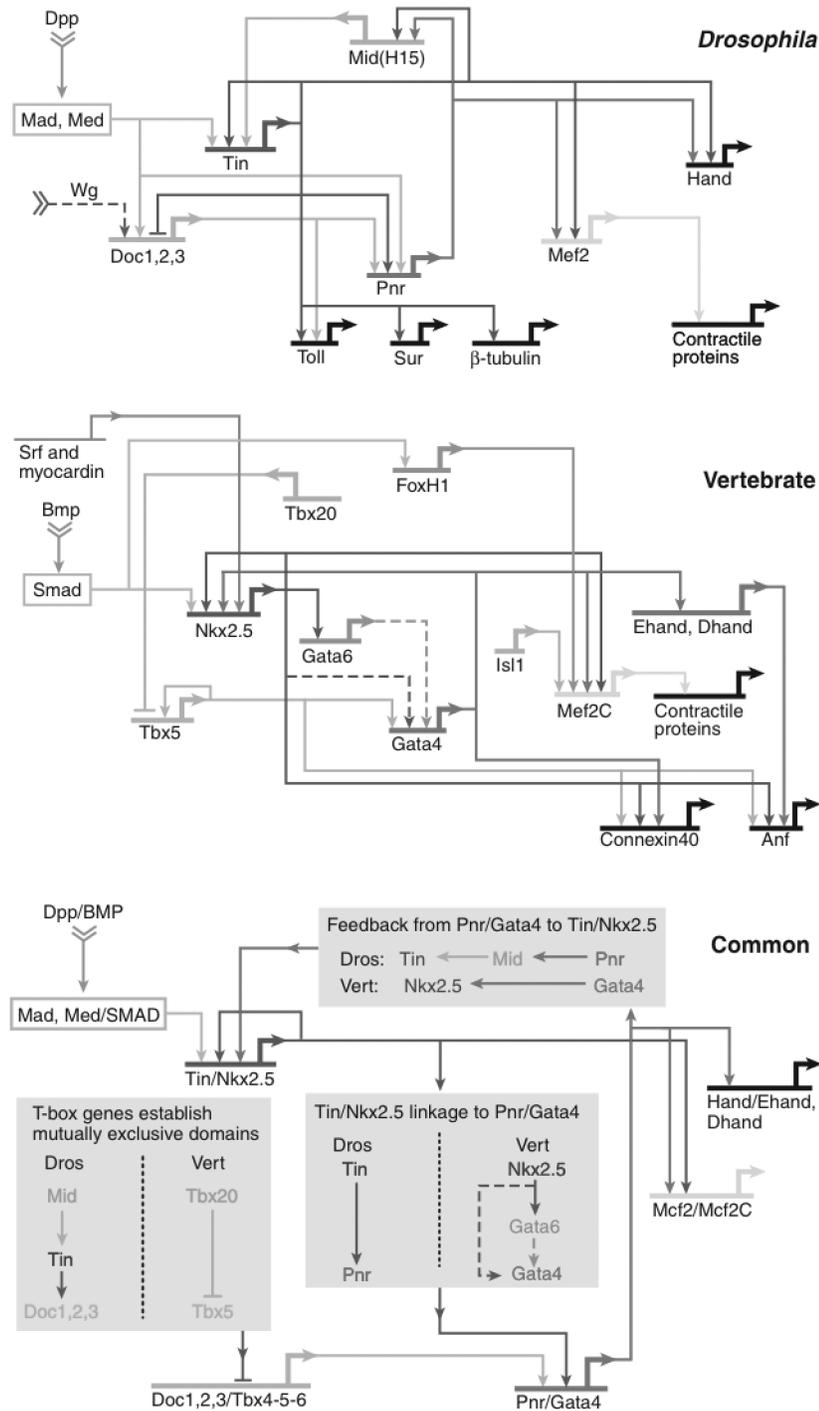
One of the major findings from recent genome-wide scans for adaptively evolving genes is that positive selection may be more common than previously thought [67-69]. These analyses suggest that positive selection in humans is strongest for X-linked genes and genes related to the immune system, reproduction, and sensory perception [69]. Unlike early studies of positive selection in humans, which focused on a few genes of interest (*PRM1*, *PRM2*, *SRY*, *FOXP2*, *G6PD*, and *MC1R*), many of the genes identified from genome-wide scans are novel [69]. They include several genes with testis-specific expression (*USP26*, *C15orf2*, and *HYAL3*), immune regulation (*CD58*, *APOBEC3F*, and *CD72*), tumor antigens (*SAGE1* and *MAGEC2*), and others with unknown functions (*FLJ46156I*, *ABHD1*, and *LOC389458*) [67-69]. Many of the genes that were positively selected in humans are involved in resistance to malaria (*HBB*, *CD40L*, *FY*, and *CD36*), HIV (*CCR5-Δ32*) or other infectious diseases (*MHC*). Other genes under selection are involved in

skin pigmentation (*OCA2*, *MYO5A*, *DTNBP1*, *TYRP1*, and *SLC24A5*), diet (*LCT* and *ADH*), brain development (*MCPHI*, *ASPM*, *CDK5RAP2*, *CENPJ*, *GABRA4*, *PSEN1*, *SYT1*, *SLC6A4*, and *SNTG1*) and bone morphogenesis (*BMP3*, *BMP2*, and *BMP5*), among many others [67-69]. Analysis of Affymetrix array data indicated that positively selected genes are more tissue-specific than other genes and more frequently expressed in spleen, testes, liver, and breast than other tissues [67]. Remarkably, the number of positively selected genes is substantially smaller in humans than in chimpanzees [70]. These results suggest that knockouts of mouse genes to identify their functions in humans may not identify human-specific functions, particularly for quickly evolving genes under positive selection. Indeed, a recent study (discussed in more detail below) found that genes with differential effects when mutated in humans vs. mice are commonly associated with increased rates of evolution and positive selection [96].

#### *Orthologous proteins, divergent networks*

One of the best studied gene regulatory networks specifies the development of the heart [71,72]. The fundamental functional components of hearts are the cardiac muscle cells that express an array of contractile proteins, such as actin, myosin, troponin, and tropomyosin. The first heart-like organs evolved more than 500 million years ago in the ancestor of extant bilaterians and probably resembled the simple tube-like hearts of tunicates, amphioxus, and *Drosophila* [71]. During evolution, this simple, single-layered tube, which pumped via peristaltic contractions, evolved into a more efficient pump, beating via synchronous contractions with thick muscular chambers functionally specialized for receiving (atria) and sending (ventricles) blood [71]. Interestingly, muscle cells are developmentally and evolutionarily derived from mesoderm, a phylogenetically older tissue type [71].

While the structure of invertebrate and vertebrate hearts are dramatically different,



**Figure 1.** Gene regulatory networks for heart development in *Drosophila* (top) and vertebrates (middle); shared linkages are shown at the bottom. A core set of regulatory genes are used in common between insect and vertebrates and are linked in a similar way in conserved subcircuits of the gene network. Gray boxes highlight different ways that the same two nodes of the network are linked in *Drosophila* and vertebrates. Orthologs of many regulatory genes in vertebrate heart formation are not known in *Drosophila*. Image used with permission from [72].

the core gene regulatory network for heart formation is highly conserved and governed by homologous transcription factors and signaling molecules in all animals (Figure 1) [71,72]. This conserved core of regulators controls cardiac cell fate, the expression of genes for contractile proteins, and the morphogenesis of the heart [71]. For example, initiation of both *Drosophila* and mouse hearts during development is dependent on the homologous transcription factors *tinman* (*tin*) and *Nkx2.5*, respectively. Detailed network analysis showed that both *tin* and *Nkx2.5* function by regulating many of the same target genes in their respective species. However, gene swap experiments that attempted to rescue *tin*-null flies with its mammalian homolog *Nkx2.5* uncovered that they were only partially equivalent. Specifically, the mouse *Nkx2.5* gene could completely rescue the expression of some target genes, like *FascIII*, and partially rescue others like *MEF2*, but the *Nkx2.5* gene could not rescue the expression of *eve* and *zfb-1* in the developing *Drosophila* heart [47,50].

Similar to *tin* and *Nkx2.5*, human *Otx1/2* genes can rescue their *Drosophila* ortholog, *otd*, in nervous system development [46,73]. The reverse, however, is only partially true. *Otd* can replace most *Otx1/2* functions in mouse nervous system development, but it is unable to rescue the development of the mesencephalon, cerebellar foliation, or the lateral semicircular canals of the inner ear [74]. These differences in function between *otd* and *Otx1/2* likely result from divergence in their ability to regulate downstream target genes. Montalta-He et al. (2002) tested this possibility using whole genome microarrays in *Drosophila*, overexpressing *otd* and *Otx2*. Amazingly, of the 287 and 682 genes that responded to *otd* and *Otx2* overexpression, respectively, only 90 were shared target genes. Thus, even transcription factors that are functionally equivalent with respect to the development of a particular structure can be nonequivalent with respect to regulation of gene networks at the systems level.

The functional equivalence of *tin/Nkx2.5* with respect to *FascIII* and *MEF2*

expression, but not *eve* and *zfb-1*, in the developing heart and the divergence of *otd/Otx2* target genes and functions in specific parts of nervous system development is particularly revealing. *FascIII* is a cell surface antigen expressed in derivatives of visceral mesoderm, and *MEF2* is a transcription factor that is a central component of gene regulatory networks in all muscle types [47,50]. Thus, they are likely ancient targets of *tin/Nkx2.5*, whose expression in muscle cells predates the origin of the heart. However, the ability of *tin*, but not *Nkx2.5*, to regulate *eve* and *zfb-1* suggests these genes are either derived, clade-specific targets of *tin* in *Drosophila* or *Nkx2.5* has lost the ability to regulate them during the evolution of mammals. Similarly, the function of *otd/Otx2* in nervous system development is ancient, and their ability to regulate partially overlapping sets of target genes reflects the underlying homologies in nervous system development in bilaterians. Nevertheless, clade-specific functions have evolved for both genes, suggesting that while the basic developmental plan of nervous system development is similar between bilaterians, *otd/Otx2* have functionally diverged with respect to their target gene repertoires and the ability to direct the development of species-specific neural structures.

Like the divergence in protein functions, gene regulatory networks can diverge over time. Thus, homologous proteins may not have equivalent functions in regulatory networks, or may not even be part of the same network between different species. This suggests that functional genomic and biomedical studies should be careful when determining gene function and network position in models and applying it to other species, particularly when the phylogenetic distance between the species grows.

#### *Developmental and physiological systems drift*

It is generally a reasonable assumption — and fundamental to biomedical studies using animal models — that homologous developmental processes govern the development and function of homologous cells,

tissues, and organs. Indeed, animal models have provided invaluable information on development, disease, and gene function. However, a growing number of studies have uncovered surprising divergence in developmental pathways, often without overt changes in phenotype [75]. This phenomenon, the development of homologous systems via divergent processes, has been termed developmental system drift (DSD).

DSD is emerging as a general process that arises after species diverge. Indeed, the divergence in target genes between insect and mammalian *tin/Nxk2.5* and *otd/Otx2* likely are special cases of DSD. While the divergence between insect and mammalian developmental systems may have occurred over 100 million years, DSD can arise over much shorter time frames. For example, thoracic bristles are missing in hybrids between *Drosophila melanogaster* and *Drosophila simulans*, even though the pattern of thoracic bristles is identical between *D. melanogaster* and *D. simulans* [76]. More dramatically, the second (T2) and third (T3) thoracic segments of hybrids between *D. subobscura* and *D. maderensis* are partially transformed into the first segment (T1) [77,78].

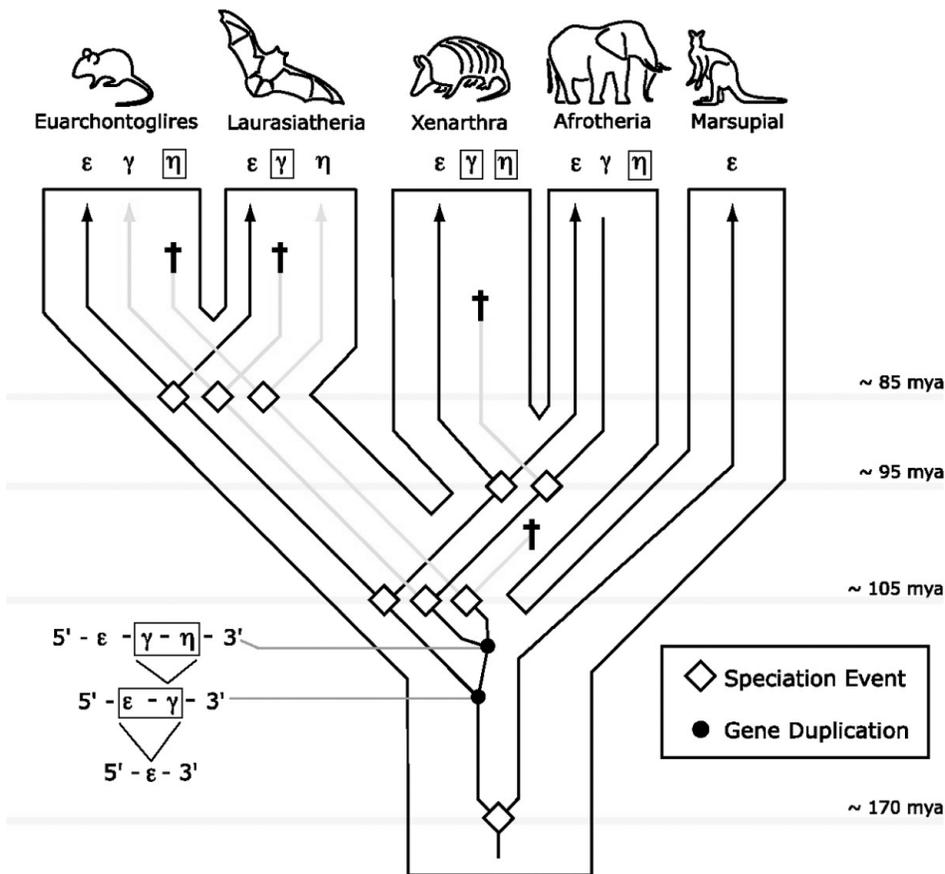
Hints of DSD also have been discovered in limb development and spermatogenesis. *Radical-fringe (R-fng)*, an extracellular protein that binds Notch during cell signaling, is expressed at the dorsal-ventral boundary of the limb bud in mouse and chicken, where it promotes the formation of the apical epidermal ridge (AER; Johnston et al. 1997). Remarkably, while *R-fng* is sufficient to induce AER formation in chickens [79], mouse knockouts of *R-fng* have no defects in limb development [80]. A particularly interesting example of DSD is the *testis-specific protein, Y-encoded (TSPY)* gene. One or more copies of the *TSPY* gene are found on the Y chromosome of all placental mammals, and, as its name implies, its expression is restricted to the testis [81]. Remarkably, although *TSPY* is functional in rat, *Mus musculus* has a single-copy *TSPY* pseudogene that does not generate a functional transcript; knockin of multiple human *TSPY*

does not alter spermatogenesis [81]. Obviously, this natural knockout does not affect fertility in mice and restoration of functional copies of *TSPY* does not alter fertility, suggesting that the spermatogenesis has drifted away from *TSPY*-dependence in mice.

Recent studies suggest that the molecular basis of DSD likely is transcriptional regulation. Various elements of transcriptional regulation play central roles in development, homeostasis, and evolution. Initial studies comparing genome-wide binding sites for two stem cell-specific transcription factors found large differences in binding site locations between human and mouse [82]. In a more comprehensive analysis, Odom et al. (2007) systematically compared the binding of four tissue-specific transcription factors (*FOXA2*, *HNF1A*, *HNF4A*, and *HNF6*) using ChIP-chip on hepatocytes purified from human and mouse livers. Despite the conserved function of these transcription factors, between 41 percent and 89 percent of their binding sites were species-specific [83]. Moreover, genome-wide analysis of polymorphisms in human populations using the Derived Allele Frequency (DAF) test of natural selection found that many human- and primate-specific transcription factor binding sites in cis-regulatory elements are evolving under positive selection [68]. These results suggest that divergence in gene regulation is adaptively driven and can accumulate particularly fast without leading to obvious differences in phenotype [54,84]. Thus, even morphologically similar organs can develop from divergent gene regulatory networks.

### *Pseudoorthology*

The differential gain and loss of genes is an important source of functional variation between species. Differences in gene content between species arise from lineage-specific gains via duplication and losses via deletion or inactivation. A consequence of gene duplication is often functional differentiation because one copy is no longer constrained to perform the ancestral function and is free to acquire novel functions. This process of neofunctionalization has been



**Figure 2.**  $\beta$ -globin gene family evolution in eutherian mammals. Two successive duplications of a proto- $\epsilon$  gene gave rise to the  $\gamma$  and  $\eta$  genes in the stem-lineage of eutherian mammals after their divergence from marsupials. Consequently, the full complement of embryonic globin genes (hemoglobin  $\epsilon$ ,  $\gamma$  and  $\eta$ ) was present in the common ancestor of the eutherian mammals. The  $\eta$  gene was lost in the common ancestor of Xenarthrans and Afrotherians, and the  $\gamma$  gene was lost in Xenarthrans after divergence from the Afrotherians. The  $\eta$  and  $\gamma$  genes were independently lost in the Euarchontoglires and Laurasiatherians, respectively. Crosses indicate lineage-specific gene losses. Image used with permission from [87].

proposed to be the main source of new protein functions during evolution [85].

While gene duplication followed by neofunctionalization is a major source of molecular novelty, it can also mask true orthology — particularly if there is differential loss of duplicate genes in two species. For example, if gene A duplicates before a speciation event to produce paralogs A and A', differential loss of duplicates after a subsequent speciation can lead to one species maintaining A while the other maintains A'. If the initial duplication event was associated with neofunctionalization, then even

though both species apparently have a copy of the A gene, the gene's functions will not be the same and can be very different.

Unfortunately, there are few studies of how common differential gene loss is, thus, it is not clear how often it may bias functional genomic annotation of gene function or biomedical studies of gene functions [86]. However, a recent study of the mammalian beta-globin gene family suggests it can be common among some genes. Opazo et al. (2008) used a comparative genomic approach to investigate  $\beta$ -globin gene turnover in placental mammals. Interestingly, they

found that there was large-scale differential loss and retention of genes after an initial expansion of the non-adult portion of the  $\beta$ -globin gene cluster in the common ancestor of placental mammals (Figure 2). The differential sorting and species-specific duplication/neofunctionalization of  $\epsilon$ -,  $\gamma$ -, and  $\eta$ -globin gene lineages among groups of placental mammals has produced species differences in the functions of hemoglobin isoforms, thereby generating variation in the complement of globin genes among mammals [87].

### FUNCTIONAL DIVERGENCE BETWEEN MOUSE AND HUMAN GENES

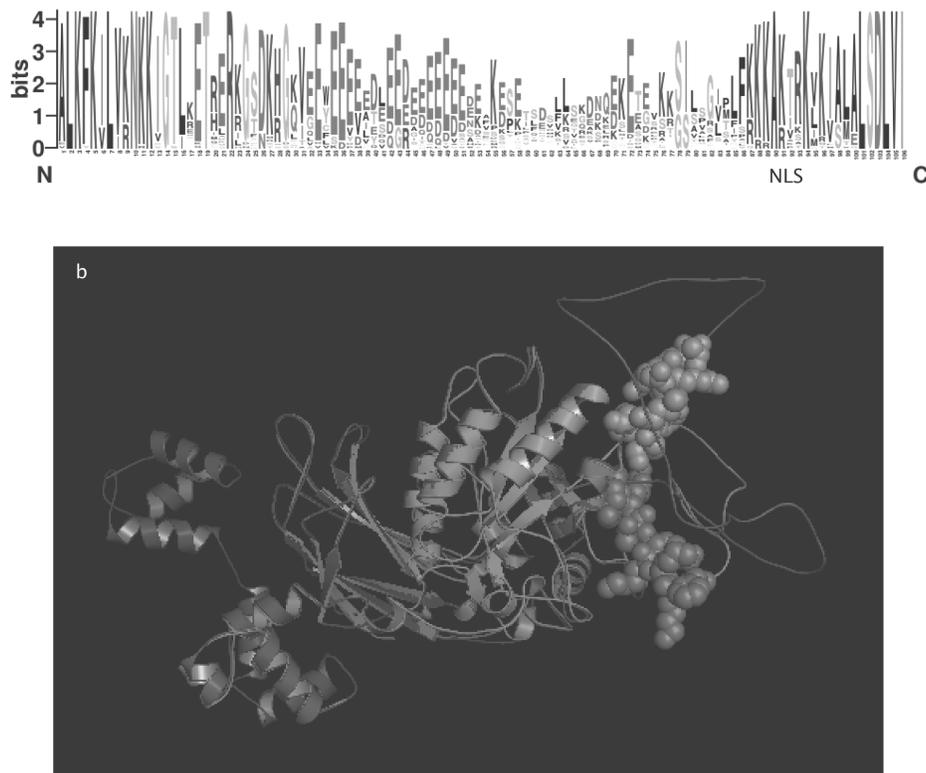
The examples discussed above indicate that functional divergence in genes and regulatory networks accumulate during evolution. This suggests that a fundamental assumption of comparative genomics and biomedical studies, i.e., that gene functions are conserved between model organisms and humans, may be too simplistic. Indeed, beyond case-by-case studies and anecdotal evidence, there have been few systematic examinations to verify the assumption of functional conservation between models and humans. Examples of functional divergence, however, have received little attention within the biomedical community, despite a growing number of studies finding that changes in protein function may be common during mammalian evolution.

Perhaps one of the most dramatic examples of an essential human gene that is non-essential in mouse is *Acrosin* (*ACR*). *ACR* enables sperm to penetrate the extracellular matrix and fertilize oocytes by proteolyzing the zona pellucida. Given the role of *ACR* in fertilization, it is not surprising that numerous studies have implicated *ACR* in human male infertility [88,89]. Unexpectedly, however, male *ACR*-null mice are fertile in spite of the complete absence of acrosin protease activity in the sperm [90]. Thus, *ACR* is not essential for sperm penetration of the zona pellucida or fertilization in mice, despite its necessity for human fertility. Interestingly,

mouse sperm have much lower levels of *ACR* than other rodents, such as rat or hamster, suggesting that other proteases may function in place of *ACR* during mouse fertilization [91]. In fact, other serine proteases in addition to *ACR* are present in mouse sperm, but not in rat or hamster, which may be an indication of developmental system drift in mammalian fertilization systems [91].

Like *ACR*, *Deleted in Azoospermia* (*DAZ*) and *DAZ-like* (*DAZL*) are involved in male germ cell differentiation and function. *DAZL* is highly conserved during evolution and has been isolated from species as diverse as mammals, frogs, fish, and worms; *DAZ* is only found in apes and other Old World monkeys [92]. Thus, the *DAZ* gene duplicated sometime after Old World and New World monkeys diverged but before the radiation of the Old World monkeys. In most mammals, the ancestral *DAZL* is sufficient to complete gametogenesis; however, in humans, deletions removing the Y-chromosomal *DAZ* gene are often associated with azoospermia or oligospermia [93], indicating *DAZ* is necessary for spermatogenesis. Interestingly, most amino acids in *DAZ* are under strong selective constraint, suggesting they are important for protein structure/function. Nevertheless, a few sites evolving under diversifying selection suggest they are involved in a molecular arms race or that these sites have different species-specific functions [94]. Lineage-specific analysis indicated that human members of this gene family were evolving by positive Darwinian selection, suggesting *DAZ* may have evolved human-specific functions [94].

Vogel et al. (2002) studied the functions of the human *DAZL* and *DAZ* genes in mouse *DAZL*-null mice and found that both *DAZL* and *DAZ* enabled production of prophase spermatocytes. However, both human genes failed to promote differentiation into mid- to late pachytenes. The human *DAZL* transgene led to more early germ cells compared with that observed in *DAZ*. Human *DAZL* and *DAZ* can only substitute for early functions of the mouse *DAZL*, sug-



**Figure 3.** Evolution of a novel nuclear localization signal (NLS) in PLCZ1. (a) Logo of the X-Y linker region from 31 species of amniotes. In this representation of a multiple sequence alignment the conservation of amino acid residues in the protein are indicated by the height of the column (Bit Score) while the conservation of specific amino acids in at that site are shown by the height of letters within columns. Note that conservation is generally poor. (b) Superimposed structural models of human and mouse PLCZ1 proteins. Note that while the structure of human and mouse PLCZ1 proteins are generally the same, the X-Y linker region is much longer in humans leading to a longer disordered loop than in mouse.

gesting gene- and species-specific function have evolved in *DAZL/DAZ* after human and mouse diverged and/or after *DAZL/DAZ* duplicated, particularly in their ability to drive mid- to late spermatogenesis. Thus, even though both human *DAZL/DAZ* and mouse *DAZ* are essential genes, the essentiality has diverged.

During fertilization, sperm trigger a series of oscillations in the concentration of intracellular calcium, which initiates egg activation. The major protein delivered by the sperm that induces the calcium flux is an isoform of phospholipase C termed *phospholipase C zeta 1 (PLCZ1)*. Remarkably, mouse *PLCZ1* is transported into the nucleus after fertilization, but *PLCZ1* from

other species, such as human, rat, and fish, are not. Ito et al. (2008) traced this difference in nuclear localization ability to a novel nuclear localization signal (NLS) that evolved in mouse within a preexisting lysine/arginine rich region of the protein (Figure 3). All *PLCZ1* genes tested by Ito et al. (2008) could induce  $Ca^{2+}$  oscillations when expressed in mouse eggs, but the activity was highly variable. For example,  $Ca^{2+}$  oscillations induced by human *PLCZ1* continued far beyond the time of pronuclear formation, whereas oscillations induced by rat *PLCZ1* stopped far before pronuclear formation. Ito and colleagues found that *PLCZ1* sequestration into the pronucleus participates in termination of

Ca<sup>2+</sup> oscillations of mouse embryos, but not other mammals, indicating the mechanism that terminates Ca<sup>2+</sup> oscillations after fertilization has diverged significantly, even between the closely related mouse and rat [95].

The examples of *ACR*, *DAZ/DAZL*, and *PLCZ1* show that homologous proteins are not necessarily functionally equivalent, but these case studies only hint that a larger pattern of functional divergence may exist. Recently, Liao and Zhang (2008) systematically compared the phenotype of null mutations for 120 essential human genes, i.e., genes associated with death before puberty or infertility, which have been knocked out in mice. Remarkably, ~22 percent of human essential genes were non-essential in mouse. Changes in essentiality between species were associated with adaptive evolution of the proteins, consistent with previous findings that changes in protein function are often associated with rapid evolution. By examining the evolution of genes that have changed essentiality between humans and mice in a broader context, they found that positive selection was much more common in the human lineage than in the rodent, suggesting that the differences in essentiality arose during human evolution [96].

## CONCLUSIONS AND OUTLOOK

Biomedical studies of animal models have played incredibly important roles in elucidating the molecular basis development, physiology, behavior, and pathogenesis of human disease. Transgenic models are so successful in elucidating the molecular function of genes that they are now a standard tool in molecular genetics and biomedical studies and are being used to assign functions to each gene in the genome through initiatives such as the Knockout Mouse Project (KOMP). Yet the assumption that gene functions and developmental systems are conserved between models and humans is taken for granted, in spite of the evidence that gene functions and gene networks can diverge through evolution. The mechanisms and examples of functional di-

vergence discussed above indicate that gene studies in model organisms need to be examined carefully for several potential biases. Annotation of gene functions based solely on mutant phenotypes in animal models, as well as assumptions of conserved functions between species, may be misleading. Perhaps more importantly, animal models of human disease may not provide appropriate information on gene function, particularly for rapidly evolving genes and systems. It is vitally important to demonstrate equivalence between the model and human gene with respect to the particular function under study to avoid spurious conclusions.

## REFERENCES

1. Lieschke GJ, Currie PD. Animal models of human disease: zebrafish swims into view. *Nature Reviews Genetics*. 2007;8:250-67.
2. Rosenthal N, Brown S. The mouse ascending: perspectives for human-disease models. *Nat Cell Biol*. 2007;9(9):993-9.
3. Hau J. *Animal models for human disease*. Berlin, Germany: Humana Press Inc.; 2008.
4. Koonin EV. Orthologs, paralogs and evolutionary genomics. *Annu Rev Genet*. 2005;39:309-38.
5. Amundson R. *The changing role of the embryo in evolutionary thought: roots of Evo-Devo*. Cambridge: Cambridge University Press; 2005.
6. McGinnis N, Kuziora MA, McGinnis W. Human Hox-4.2 and Drosophila Deformed encode similar regulatory specificities in Drosophila embryos and larvae. *Cell*. 1990;63(5):969-76.
7. Niimi T, Clements J, Gehring WJ, et al. Dominant-negative form of the Pax6 homolog eyeless for tissue-specific loss-of-function studies in the developing eye and brain in drosophila. *Genesis*. 2002;34(1-2):74-5.
8. Hill RE, Favor J, Hogan BLM, et al. Mouse Small eye results from mutations in a paired-like homeobox-containing gene. *Nature*. 1991;354:522-5.
9. Jordan T, Hanson I, Zaletayev D, et al. The human PAX6 gene is mutated in two patients with aniridia. *Nat Genet*. 1992;1:238-332.
10. Halder G, Callaerts P, Gehring WJ. Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science*. 1995;267(5205):1788-92.
11. Onuma Y, Takahashi S, Asashima M, et al. Conservation of Pax 6 function and upstream activation by Notch signaling in eye development of frogs and flies. *Proc Acad Nat Sci USA*. 2002;99(4):2020-5.

12. Shimeld SM. An Amphioxus Gli Gene Reveals Conservation of Midline Patterning and the Evolution of Hedgehog Signalling Diversity in Chordates. *PLoS ONE*. 2007;2(9):e864.
13. Yu H, Schurr MJ, Deretic V. Functional equivalence of *Escherichia coli* sigma E and *Pseudomonas aeruginosa* AlgU: *E. coli* rpoE restores mucoidy and reduces sensitivity to reactive oxygen intermediates in algU mutants of *P. aeruginosa*. *J Bacteriol*. 1995;177(11):3259-68.
14. Liu M, Su M, Lyons G, et al. Functional conservation of zinc-finger homeodomain gene *zfh1/SIP1* in *Drosophila* heart development. *Dev Genes Evol*. 2006;216(11):683-93.
15. Wang W, Grimmer JF, Van De Water TR, et al. *Hmx2* and *Hmx3* Homeobox Genes Direct Development of the Murine Inner Ear and Hypothalamus and Can Be Functionally Replaced by *Drosophila* *Hmx*. *Dev Cell*. 2004;7(3):439-53.
16. Bickar D, Lehninger A, Brunori M, et al. Functional equivalence of monomeric (shark) and dimeric (bovine) cytochrome c oxidase. *J Inorg Biochem*. 1985;23(3-4):365-72.
17. Zhang P, Tan HT, Pwee KH, et al. Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. *Plant J*. 2004;37(4):566-77.
18. McIntosh KB, Bonham-Smith PC. Establishment of *Arabidopsis thaliana* ribosomal protein RPL23A-1 as a functional homologue of *Saccharomyces cerevisiae* ribosomal protein L25. *Plant Mol Biol*. 2001;46(6):673-82.
19. Wang VY, Hassan BA, Bellen HJ, et al. *Drosophila* atonal Fully Rescues the Phenotype of *Math1* Null Mice: New Functions Evolve in New Cellular Contexts. *Curr Biol*. 2002;12(18):1611-6.
20. Hirth F, Loop T, Egger B, et al. Functional equivalence of Hox gene products in the specification of the tritocerebrum during embryonic brain development of *Drosophila*. *Development*. 2001;128(23):4781-8.
21. Coronado R, Ahern CA, Sheridan DC, et al. Functional equivalence of dihydropyridine receptor alpha1S and beta1a subunits in triggering excitation-contraction coupling in skeletal muscle. *Biol Res*. 2004;37(4):565-75.
22. Pan L, Yang Z, Feng L, et al. Functional equivalence of *Brn3* POU-domain transcription factors in mouse retinal neurogenesis. *Development*. 2005;132(4):703-12.
23. Roman BB, Wieringa B, Koretsky AP. Functional Equivalence of Creatine Kinase Isoforms in Mouse Skeletal Muscle. *J Biol Chem*. 1997;272(28):17790-4.
24. Chia IV, Costantini F. Mouse Axin and Axin2/Conductin Proteins Are Functionally Equivalent In Vivo. *Mol Cell Biol*. 2005;25(11):4371-6.
25. Bouchard M, Pfeffer P, Busslinger M. Functional equivalence of the transcription factors *Pax2* and *Pax5* in mouse development. *Development*. 2000;127(17):3703-13.
26. Fleischmann A, Hafezi F, Elliott C, et al. Fra-1 replaces c-Fos-dependent functions in mice. *Genes Dev*. 2000;14(21):2695-700.
27. Acampora D, Annino A, Puelles E, et al. OTX1 compensates for OTX2 requirement in regionalisation of anterior neuroectoderm. *Gene Expr Patterns*. 2003;3(4):497-501.
28. Adamska M, Matus DQ, Adamski M, et al. The evolutionary origin of hedgehog proteins. *Curr Biol*. 2007;17(19):R836-7.
29. Anan K, Yoshida N, Kataoka Y, et al. Morphological Change Caused by Loss of the Taxon-Specific Polyalanine Tract in *Hoxd-13*. *Mol Biol Evol*. 2007;24(1):281-7.
30. Averof M, Akam M. Hox genes and the diversification of insect and crustacean body plans. *Nature*. 1995;376(6539):420-3.
31. Branford WW, Benson GV, Ma L, et al. Characterization of *Hoxa-10/Hoxa-11* Transheterozygotes Reveals Functional Redundancy and Regulatory Interactions. *Dev Biol*. 2000;224(2):373-87.
32. Chen M, Abele R, Tampe R. Functional Non-equivalence of ATP-binding Cassette Signature Motifs in the Transporter Associated with Antigen Processing (TAP). *J Biol Chem*. 2004;279(44):46073-81.
33. Emili A, Greenblatt J, Ingles CJ. Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. *Mol Cell Biol*. 1994;14(3):1582-93.
34. Falciani F, Hausdorf B, Schroder R, et al. Class 3 Hox genes in insects and the origin of Zen. *Proc Natl Acad Sci USA*. 1996;93:8479-84.
35. Fondon JW III, Garner HR. Molecular origins of rapid and continuous morphological evolution. *Proc Natl Acad Sci USA*. 2004;101(52):18058-63.
36. Galant R, Carroll SB. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature*. 2002;415(6874):910-3.
37. Gerber HP, Seipel K, Georgiev O, et al. Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science*. 1994;263(5148):808-11.
38. Grenier JK, Carroll SB. Functional evolution of the Ultrabithorax protein. *Proc Natl Acad Sci USA*. 2000;97(2):704-9.
39. Hanks MC, Loomis CA, Harris E, et al. *Drosophila* engrailed can substitute for mouse *Engrailed1* function in mid-hindbrain, but not limb development. *Development*. 1998;125(22):4521-30.
40. Hsia CC, McGinnis W. Evolution of transcription factor function. *Current Opinion in Genetics & Development*. 2003;13(2):199-206.
41. Hyman CA, Bartholin L, Newfeld SJ, et al. *Drosophila* TGIF Proteins Are Transcriptional Activators. *Mol Cell Biol*. 2003;23(24):9262-74.
42. Kellerer S, Schreiner S, Stolt CC, et al. Replacement of the *Sox10* transcription factor by *Sox8* reveals incomplete functional equivalence. *Development*. 2006;133(15):2875-86.

43. Li X, Liang J, Yu H, et al. Functional consequences of new exon acquisition in mammalian chromodomain Y-like (CDYL) genes. *Trends Genet.* 2007;23(9):427-31.
44. Lohr U, Yussa M, Pick L. *Drosophila fushi tarazu*: a gene on the border of homeotic function. *Curr Biol.* 2001;11(18):1403-12.
45. Montalita-He H, Leemans R, Loop T, et al. Evolutionary conservation of otd/Otx2 transcription factor action: a genome-wide microarray analysis in *Drosophila*. *Genome Biol.* 2002;3(4):research0015.1-0015.15.
46. Nagao T, Leuzinger S, Acampora D, et al. Developmental rescue of *Drosophila* cephalic defects by the human Otx genes. *Proc Acad Natl Sci USA.* 1998;95(7):3737-42.
47. Park M, Lewis C, Turbay D, et al. Differential rescue of visceral and cardiac defects in *Drosophila* by vertebrate tinman-related genes. *Proc Acad Natl Sci USA.* 1998;95(16):9366-71.
48. Prager EM, Wilson AC. Slow evolutionary loss of potential for interspecific hybridization in birds. *Proc Acad Natl Sci USA.* 1975;72:200-4.
49. Punzo C, Plaza S, Seimiya M, et al. Functional divergence between eyeless and twin of eyeless in *Drosophila melanogaster*. *Development.* 2004;131(16):3943-53.
50. Ranganayakulu G, Elliott DA, Harvey RP, et al. Divergent roles for NK-2 class homeobox genes in cardiogenesis in flies and mice. *Development.* 1998;125(16):3037-48.
51. Ronshaugen M, McGinnis N, McGinnis W. Hox protein mutation and macroevolution of the insect body plan. *Nature.* 2002;415(6874):914-7.
52. Stolt CC, Lommes P, Friedrich RP, et al. Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. *Development.* 2004;131(10):2349-58.
53. Tanay A, Regev A, Shamir R. Conservation and evolvability in regulatory networks: The evolution of ribosomal regulation in yeast. *Proc Acad Natl Sci USA.* 2005;102(20):7203-8.
54. Tuch BB, Galgoczy DJ, Hernday AD, et al. The Evolution of Combinatorial Gene Regulation in Fungi. *PLoS Biology.* 2008;6(2):e38.
55. Lynch VJ, Tanzer A, Wang Y, et al. Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. *Proc Acad Natl Sci USA.* 2008;105(39):14928-33.
56. Carroll SB. Evolution at Two Levels: On Genes and Form. *PLoS Biology.* 2005;3(7):e245.
57. Carroll SB. *Endless Forms Most Beautiful.* New York: W.W. Norton & Company, Inc.; 2005.
58. Wagner GP, Lynch VJ. Molecular evolution of evolutionary novelties: the vagina and uterus of therian mammals. *J Exp Zool Part B Mol Dev Evol.* 2005;304B(6):580-92.
59. Lynch VJ, Wagner GP. Resurrecting the role of transcription factor change in developmental evolution. *Evolution.* 2008;62(9):2131-54.
60. Wagner GP, Lynch VJ. The gene regulatory logic of transcription factor evolution. *Trends in Ecology and Evolution.* 2008;23(7):377-85.
61. Kimura M. Evolutionary rate at the molecular level. *Science.* 1968;217:624-6.
62. Kimura M. *The Neutral Theory of molecular evolution.* Cambridge: Cambridge University Press; 1983.
63. King JL, Jukes TH. Non-Darwinian Evolution. *Science.* 1969;164(3881):788-98.
64. Hughes AL, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature.* 1988;335:167-70.
65. Chen FC, Chen CJ, Ho JY, et al. Identification and evolutionary analysis of novel exons and alternative splicing events using cross-species EST-to-genome comparisons in human, mouse and rat. *BMC Bioinformatics.* 2006;7(1):136.
66. Chen C, Opazo JC, Erez O, et al. The human progesterone receptor shows evidence of adaptive evolution associated with its ability to act as a transcription factor. *Molecular Phylogenetics and Evolution.* 2008;47(2):637-49.
67. Kosiol C, Vinar T, da Fonseca RR, et al. Patterns of Positive Selection in Six Mammalian Genomes. *PLoS Genet.* 2008;4(8):e1000144.
68. Voight BF, Kudravalli S, Wen X, et al. A Map of Recent Positive Selection in the Human Genome. *PLoS Biology.* 2006;4(3):e72.
69. Sabeti PC, Schaffner SF, Fry B, et al. Positive Natural Selection in the Human Lineage. *Science.* 2006;312(5780):1614-20.
70. Bakewell MA, Shi P, Zhang J. More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci USA.* 2007;104(18):7489-94.
71. Olsen ER. Gene regulatory networks in the evolution and development of the heart. *Science.* 2006;313:1922-7.
72. Davidson EH, Erwin DH. Gene regulatory networks and the evolution and animal body plans. *Science.* 2006;331:796-800.
73. Leuzinger S, Hirth F, Gerlich D, et al. Equivalence of the fly orthodenticle gene and the human OTX genes in embryonic brain development of *Drosophila*. *Development.* 1998;125(9):1703-10.
74. Acampora D, Avantaggiato V, Tuorto F, et al. Murine Otx1 and *Drosophila* otd genes share conserved genetic functions required in invertebrate and vertebrate brain development. *Development.* 1998;125(9):1691-702.
75. True JR, Haag ES. Developmental system drift and flexibility in evolutionary trajectories. *Evolution and Development.* 2001;3(2):109-19.
76. Takano T. Loss of notum macrochaetae as an interspecific hybrid anomaly between *Drosophila melanogaster* and *D. simulans*. *Genetics.* 1998;129:1435-50.
77. Khadem M, Krimbas CB. Studies of the species barrier between *Drosophila madeiren-*

- sis and *Drosophila subobscura*. *Hereditas*. 1991;114:189-95.
78. Johnston SH, Rauskolb C, Wilson R, et al. A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development*. 1997;124(11):2245-54.
  79. Laufer E, Dahn R, Orozco OE, et al. Expression of Radical fringe in limb-bud ectoderm regulates apical ectodermal ridge formation. *Nature*. 1997;386:366-73.
  80. Moran JL, Levorse JM, Vogt TF. Limbs move beyond the radical fringe. *Nature*. 1999;399:742-3.
  81. Schubert S, Skawran B, Dechend F, et al. Generation and Characterization of a Transgenic Mouse with a Functional Human TSPY. *Biol Reprod*. 2003;69(3):968-75.
  82. Loh YH, Wu Q, Chew JL, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet*. 2006;38(4):431-40.
  83. Odom DT, Dowell RD, Jacobsen ES, et al. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet*. 2007;39(6):730-2.
  84. Tuch BB, Li H, Johnson AD. Evolution of Eukaryotic Transcription Circuits. *Science*. 2008;319(5871):1797-9.
  85. Ohno S. *Evolution by gene duplication*. Berlin, Germany: Springer-Verlag; 1970.
  86. Studer RA, Robinson-Rechavi M. How confident can we be that orthologs are similar, but paralogs differ? *Trends Genet*. In Press, Corrected Proof.
  87. Opazo JC, Hoffmann FG, Storz JF. Differential loss of embryonic globin genes during the radiation of placental mammals. *Proc Natl Acad Sci USA*. 2008;105(35):12950-5.
  88. Koel C, Tuhin D, Baidyanath C, et al. Acrosin activity as a potential marker for sperm membrane characteristics in unexplained male infertility. *Fertility and Sterility*. 2005;83(1):104-9.
  89. Shimizu Y, Kodama H, Fukuda J, et al. Evidence of proacrosin molecule abnormality as a possible cause of low acrosin activity and unexplained failure of fertilization in vitro. *J Androl*. 1997;18(3):281-8.
  90. Baba T, Azuma S, Kashiwabara S, et al. Sperm from mice carrying a target mutation of the acrosin gene can penetrate the oocyte zona pellucida and effect fertilization. *J Biol Chem*. 1994;269(50):31845-9.
  91. Kazuo Y, Honda A, Kashiwabara S, et al. Difference of acrosomal serine protease system between mouse and other rodent sperm. *Developmental Genetics*. 1999;25(2):115-22.
  92. Vogel T, Speed RM, Ross A, et al. Partial rescue of the *Dazl* knockout mouse by the human *DAZL* gene. *Mol Hum Reprod*. 2002;8(9):797-804.
  93. Vogt P, Edelmann A, Kirsch S, et al. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet*. 1996;5(7):933-43.
  94. Bielawski JP, Yang Z. Positive and Negative Selection in the *DAZ* Gene Family. *Mol Biol Evol*. 2001;18(4):523-9.
  95. Ito M, Shikano T, Oda S, et al. Difference in  $Ca^{2+}$  Oscillation-Inducing Activity and Nuclear Translocation Ability of PLCZ1, an Egg-Activating Sperm Factor Candidate, Between Mouse, Rat, Human, and Medaka Fish. *Biol Reprod*. 2008;78(6):1081-90.
  96. Liao BY, Zhang J. Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proc Natl Acad Sci USA*. 2008;105(19):6987-92.