Published in final edited form as:
Cell. 2011 March 18; 144(6): 970–985. doi:10.1016/j.cell.2011.02.017.

Evolution of Gene Regulatory Networks that Control Embryonic Development of the Body Plan

Isabelle S. Peter* and **Eric H. Davidson***
Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125, USA

SUMMARY

Alteration of the functional organization of the gene regulatory networks (GRNs) that control development of the body plan causes evolutionary change in animal morphology. A major mechanism of evolutionary change in GRN structure is alteration of *cis*-regulatory modules that determine regulatory gene expression. Both evolutionary conservation and evolutionary innovation must be considered in terms of GRN structure. Here we consider the causes and consequences of GRN evolution, both from an a priori point of view, and in light of extensive recent research on developmental regulatory alterations occurring at different levels of GRN hierarchy. Some GRN subcircuits are of great antiquity while other aspects are highly flexible and thus in any given genome more recent. Both evolutionary conservation and evolutionary innovation occur at the level of whole GRN subcircuits. This mosaic view of the evolution of GRN structure explains major aspects of evolutionary process, such as hierarchical phylogeny and discontinuities of paleontological change and stasis.

INTRODUCTION

In each generation of each animal species, the body plan is formed by the execution of an inherited genomic regulatory program for embryonic development. The basic control task is to determine transcriptional activity throughout embryonic time and space, and here ultimately lies causality in the developmental process. The genomic control apparatus for any given developmental episode has a physical reality: it consists of the specifically expressed genes which encode the transcription factors required to direct the events of that episode, most importantly including the *cis*-regulatory control regions of these genes. The *cis*-regulatory sequences combinatorially determine which regulatory inputs will affect the expression of each gene and what other genes it will affect; that is, they hardwire the functional linkages among the regulatory genes, forming network subcircuits. The subcircuits perform biologically meaningful jobs, for example, acting as logic gates, interpreting signals, stabilizing given regulatory states, or establishing specific regulatory states in given cell lineages; here the term “regulatory state” means the total of active transcription factors in any given cell at any given time. In turn the subcircuits are “wired” together to constitute the gene regulatory network (GRN), the overall genomically encoded developmental control system.

*Correspondence: ipeter@caltech.edu and davidson@caltech.edu.

Developmental GRN Structure

As with any operational control system, the particular structure of a developmental GRN determines its particular functions. GRN structure has a unique character, and in this review we return repeatedly to the way these structural characteristics affect the processes by which evolution of the animal body plan occurs. GRNs are inherently hierarchical: the networks controlling each phase of development are assemblages of subcircuits, the subcircuits are assemblages of specific regulatory linkages amongst specific genes, and the linkages are individually determined by assemblages of *cis*-regulatory A's, C's, G's, and T's. But at the highest level of its organization, the developmental GRN is hierarchical in an additional, and as we see below, very important sense. Development progresses from phase to phase and this fundamental phenomenon reflects the underlying sequential hierarchy of the GRN control system. In the earliest embryonic phases the function of the developmental GRN is establishment of specific regulatory states in the spatial domains of the developing organism. Thereby the design of the future body plan is mapped out in regional regulatory landscapes, which differentially endow the potentialities of the future parts. Lower down, GRN apparatus continues regional regulatory specification on finer scales. Ultimately, precisely confined regulatory states determine how the differentiation and morphogenetic gene batteries at the terminal periphery of the GRN will be deployed.

Since developmental GRN structure determines GRN function, and since derived evolutionary change in animal body plans must occur because of change in the genomic apparatus controlling development, evolution of the body plan must be effected by alterations in the structure of developmental GRNs. Most changes in GRN structure are rooted in *cis*-regulatory alterations, both in principle and in fact, a fundamental theme of the following review. The result of relevant change in GRN structure is derived change in GRN operation, compared to the immediately ancestral GRNs. This will cause changes in developmental process, and ultimately in the product of that process, the body plan (Britten and Davidson, 1971; Davidson and Erwin, 2006; Erwin and Davidson, 2009). As analysis of developmental GRNs accelerates, the rules of GRN structure/function relations emerge, and this in turn provides novel pathways into evolutionary mechanism which could never have been anticipated in advance. Although a variety of ways of thinking about evolution have been proposed, the evolution of the body plan is fundamentally a system-level problem to which GRN structure/function provides the most compelling direct access.

The Roots of Change in GRN Architecture: Evolution at *Cis*-Regulatory Nodes

Because GRN topology is encoded directly in *cis*-regulatory sequence at its nodes, evolutionary changes in this sequence have great potency to alter developmental GRN structure and function. But there are many kinds of *cis*-regulatory change that affect function in different ways, ranging from loss of function, to quantitative change in function, to qualitative gain of function resulting in redeployment of gene expression.

Types of *cis*-regulatory mutation and their diverse possible consequences—

Two general kinds of genomic change affecting *cis*-regulatory modules are internal changes affecting sequence within *cis*-regulatory modules, and contextual sequence changes which alter the physical disposition of entire *cis*-regulatory modules. Table 1 provides a list of both kinds of change and their possible functional connotations. Of the internal changes in Table 1, note that loss or gain of a given target site might each cause either loss or gain of function, depending on whether the factor binding the site is an activator or repressor.

Many internal *cis*-regulatory sequence changes will produce quantitative effects only, so long as the qualitatively complete set of interactions is assured by the identity of the target sites. The arrangement, spacing, and number of these sites is relatively insignificant (Balhoff

and Wray, 2005; Cameron and Davidson, 2009; Dermitzakis et al., 2003; Liberman and Stathopoulos, 2009; Ludwig et al., 2000; Walters et al., 2008). In a convincing recent example Hare et al (2008) showed that >70% of specific *Drosophila melanogaster eve* stripe 2 sites are not conserved in some other Drosophilidae, even though these modules produce identical output patterns. They all, however, respond to the same qualitative inputs. Furthermore, four different *eve cis*-regulatory modules (three pair rule stripe modules and a heart expression module) isolated from flies perhaps 100 my removed from their last common ancestor with *Drosophila*, were shown to function identically when introduced into *D. melanogaster* despite extremely different site order, number and spacing. Another example is found in a comparison of orthologous *otx cis*-regulatory modules in distantly related ascidians, which again revealed extremely different module organization despite identical spatial regulatory function (Oda-Ishii et al., 2005). These results indicate great freedom of *cis*-regulatory design, given only the constraint on input identity and of course the requirement that all the relevant sites lie within functional interaction range, in practice usually the several hundred base pairs of the module sequence.

There is however one notable exception, *viz*. the high conservation of arrangement of sites found very closely apposed, presumably because the proteins bound to them interact directly with each other or with third parties, e.g., Dorsal and Twist sites in multiple *Drosophila* neurogenic ectoderm genes (Hong et al., 2008); or Otx and Gatae sites in orthologous *cis*-regulatory modules of echinoderm *otx* genes (Hinman and Davidson, 2007). Also many vertebrate *cis*-regulatory modules are known where the order of closely packed target sites has been conserved, resulting in high levels of sequence identity across *cis*-regulatory modules that have been evolving separately for 350–450 my years (Elgar and Vavouri, 2008; Pennacchio et al., 2006; Rastegar et al., 2008; Siepel et al., 2005; Vavouri et al., 2007; Wang et al., 2009). Because of this exception to the general rule of relaxed *cis*-regulatory design, in Table 1 site spacing is considered as a possible cause of input gain or loss. As Table1 indicates, only those intra-modular *cis*-regulatory sequence changes which produce qualitative gain or loss of target sites can result in the co-option of the respective network node to a new temporal/spatial expression domain and thus in the alteration of functional GRN topology.

An important implication of Table 1 is that contextual (external) *cis*-regulatory changes of several kinds may be a major source of evolutionary GRN re-design. Co-optive redeployment of *cis*-regulatory modules can be due to translocation by mobile elements; spatial repression functions can disappear by deletion of whole modules; *cis*-regulatory recruitment can be altered by functions that tether them to different promoters. In some branches of evolution duplication of regulatory genes followed by sub-functionalization has been a major source of evolutionary novelty (Jimenez-Delgado et al., 2009;Ohno, 1970). But while it is possible to estimate computationally the rate of single target site sequence appearance and disappearance, or for specific cases observe it, we have virtually no fix on the rates of processes that move *cis*-regulatory modules into new genomic contexts. Because *cis*-regulatory modules may be carried around by transposing mobile elements, and because the transposition of mobile elements is the most rapid type of large scale genomic sequence change in animal genomes, this is likely to be a major mechanism of GRN evolution. In human, mouse and *Drosophila*, estimates suggest insertion rates for certain types of mobile elements on the order of 10^{-1} per genome-generation (Garza et al., 1991;Ostertag and Kazazian, 2001), and it is clear that there have been great bursts of mobile element insertion in the evolutionary history of many animal lineages including our own (e.g., Ohshima et al., 2003;Ostertag and Kazazian, 2001). DNA transposons, LTR-containing retrotransposons, and non-LTR containing retrotransposons, both autonomous and non-autonomous (the latter meaning that enzymatic machinery from another retrotransposon is required for mobility) are all capable of altering genomic sequence. Their various excision, copy and integration

mechanisms lie beyond the scope of this paper (for reviews see Gogvadze and Buzdin, 2009;Kazazian, 2004); suffice it to say that the diverse types of rearrangements they cause may directly affect transcriptional processes, positively or negatively. The LTRs of retrotransposons have intrinsic *cis*-regulatory activity, and when transposed into the vicinity of a gene may cause its transcription (Gogvadze and Buzdin, 2009). In mammals non-LTR retrotransposons (such as L1 in humans) have the ability to mobilize non-autonomous mobile elements (such as Alu repeats in humans), and these frequently carry with them adjacent sequence elements. Thus Alu repeats have apparently picked up *cis*-regulatory apparatus during their non-autonomous transpositions, and moved them to the locations of new genes, and in addition their own sequence may mutate to produce *cis*-active transcription factor target sites, as shown in a number of specific examples (for review, Britten, 1997). A very important aspect of this mode of *cis*-regulatory target site insertion has recently been emphasized with the observation that, on a genome wide basis, many such sites are species- (or genus- or order-) specific (e.g., Odom et al., 2007). An excellent case in point is a recent study of sites recognizing the neural repressor REST (Johnson et al., 2009) where it is clear from comparison among mammalian genomes that primate-specific sites have been inserted in recent evolutionary time all over the genome by Alu and L1 transposition, though most of the primate-specific sites are (as yet) probably functionless. In another case a non-LTR retrotransposon has inserted an auto- and cross-regulatory site into a duplicate copy of the *dmrt* sex control gene in Medaka within the last 10 my, which generates a functional species-specific control circuit determining developmental interplay between these two genes (Herpin et al., 2010). In summary, as previously speculated (Britten and Davidson, 1971), mobile elements could have provided a major mechanism of GRN evolution. They have the potential to produce exactly the kinds of genomic *cis*-regulatory change which a priori might be the most potent mechanisms for GRN change, i.e., gain of function co-options of regulatory gene expression (Table 1).

Evolutionary consequences of *cis*-regulatory gain of function changes— Evolutionary change in GRN structure may follow directly from qualitative gain of *cis*-regulatory linkages among regulatory and/or signaling genes. If the phenotypic functionality of this type of evolutionary process were to require the homozygosity of the underlying DNA alteration, as in classic microevolutionary theory, GRN evolution would be essentially inconceivable. But in fact phenotypic functionality of a co-optive change in regulatory gene expression will not depend on homozygosity. As initially pointed out by Ruvkun et al. (1991) and further discussed by Davidson and Erwin (2010), gain of function *cis*-regulatory co-options that produce regulatory gene expression in new domains act dominantly, and this has fundamental consequences for evolutionary process. Thousands of routine lab experiments in which regulatory systems are systematically redesigned to produce ectopic expression show that for most regulatory genes, particularly in early development, a single copy of the gain of function allele produces the regulatory effect.

The potency of a *cis*-regulatory gain of function co-option for altering GRN structure is easily imagined, as in the cartoon of Fig. 1. Here we see how the co-option could have occurred by addition of sites to a pre-existing *cis*-regulatory module, or by insertion of a new *cis*-regulatory module; and then how this co-option could alter function downstream of the GRN. In the examples of Fig. 1 the regulatory gene newly incorporated in the GRN might control the deployment of a signal system, or of a differentiation gene battery, but of course it could have many different effects.

There is an intrinsically high possibility of evolutionary re-organization of GRN structure by *cis*-regulatory gain of function co-options, given the general rapidity of *cis*-regulatory evolution, and the haplo-dominance of gain of function changes in regulatory gene expression. Any organism in which such a change had occurred in either the maternal or

paternal germ line would, if viable, become a clonal founder; (Davidson and Erwin, 2010). A *cis*-regulatory gain of function event of any of the kinds listed in Table 1 could have an immediate operational effect on a GRN, if a newly incident addition to a regulatory state caused additional GRN subcircuits to be deployed (as in Fig.1B). Or it could perhaps result in a regulatory gene expression that is for the nonce functionless, though harmless, but which could later become functional when additional co-optive events add to the regulatory state other factors with which the first can cooperate combinatorially, or when additional *cis*-regulatory changes provide new functional targets. An almost revolutionary revision emerges from the realization that GRN function can change in creative ways by mechanisms that are likely rather than unlikely to occur; that will be dominant and haplo-sufficient when they do occur; and that may be driven by a plethora of diverse processes at the *cis*-regulatory DNA level, some of which continuously or stochastically alter genomes with relatively high frequency. Periods of rapid evolutionary change may be thought of in these terms, but this also raises the obverse question: we now need an explanation for the paleontological demonstration of very long periods of evolutionary stasis in the basic body plans of many animal lineages.

The Hierarchical Organization of Developmental GRNS and its Impact on Evolution

Knowing that the basic events causing GRN evolution are *cis*-regulatory alterations, particularly those resulting in qualitative additions to or subtractions from developmental regulatory state, we can sharpen the question we are asking: how do the structural properties of GRNs affect the developmental consequences of such *cis*-regulatory alterations?

The consequences of hierarchical GRN structure—As discussed above, the GRNs controlling embryonic development of the body plan are intrinsically hierarchical, essentially because of the number of successive spatial regulatory states that must be installed in the course of pattern formation, cell type specification and differentiation. This property of GRNs fundamentally affects the way we need to consider the question just put. The consequences of any given *cis*-regulatory mutation will depend entirely on where in the GRN hierarchy the affected *cis*-regulatory node lies. As Fig. 2 shows, changes that occur in the *cis*-regulatory control apparatus of a given differentiation gene could cause redeployment of that gene; changes in the *cis*-regulatory system determining expression of a controller of the battery could cause redeployment of the whole battery; changes upstream of that could effect redeployment of whole regulatory states, or of many other features. The circuitry drawn in Fig.2 is of course arbitrary but its import is general. So in order to understand predictively the effect of a given *cis*-regulatory change, the GRN architecture and the position of the mutation therein must be known. This may seem a demanding requirement, but from the point of view of understanding evolution mechanistically, it places a powerful lever in our hands. First, it should enable a rational interpretation of evolutionary differences in development between related animals in terms of GRN structure (we consider examples below); second, in principle it could enable predicted effects to be tested experimentally by inserting the *cis*-regulatory change into a related form expressing the pleisiomorphic GRN (“synthetic experimental evolution”;(Erwin and Davidson, 2009)).

Another direct evolutionary consequence of GRN hierarchy has also been discussed (Davidson and Erwin, 2006, 2009), and this is the phenomenon of canalization. In developmental terms the establishment of a spatial regulatory state constrains subsequent processes: like a decrease in entropy, the number of possible regulatory states downstream is now decreased. If the regulatory state defines a progenitor field for a given organ, then all the subsequent stages in the development of that organ must take place within that domain. As in development so in evolution, and thus a co-optive mutation leading to qualitative evolutionary reorganization at *cis*-regulatory nodes of an upper level GRN subcircuit is

much more likely to entail numerous deleterious problems downstream, than if the change were to occur further down in the hierarchy. Therefore upper levels of GRN hierarchy are much less likely to change once a hierarchical GRN has evolved than are more peripheral levels, and this is the empirical mark of the classical canalization phenomenon.

Currently, no GRN is analyzed to a degree that we know its linkages and functions from its upstream to downstream peripheries, that is, from the beginning of the developmental process to the terminal differentiated state. We do know however that the GRN output is observable as individual gene expression patterns, and ultimately as the developmental process. We can use these outputs to infer a framework within which to position individual regulatory subcircuits or evolutionary changes within the hierarchical GRN. To facilitate the discussion on GRN evolution we now define GRN parts according to the developmental functions they control and then go on to consider abstractly the impact of evolutionary changes occurring in each of these parts.

As shown in Fig. 3, we can distinguish four causally connected developmental functions which are encoded by sections of the GRN represented by Boxes I–IV. The most upstream part of the GRN indicated in Box I, controls postgastrular pattern formation. It is animated by pregastrular spatial and signaling inputs (maternal anisotropies, maternal factors, early interblastomere signals, all used as directional cues, and then by the outputs of the initial zygotic GRNs). The functions of the GRNs set up in this phase of development, including their signaling interactions, are to establish broad domains which section the organism with respect to the major body axes. The immediate output of the GRNs of Box I is to set up regulatory state domains within spatially defined areas of the organism. These domains, such as the neuraxis, or mesodermal layers, constrain the position of future body parts and also now provide initial regulatory inputs that will be utilized in subsequent derivatives of their territories. The fate patterns they produce are often broadly conserved within clades (the early postembryonic “phylotype”).

In Box II progenitor fields for specific body parts (for example the heart progenitor field or the limb bud) are defined within these early domains. These are sets of cells each expressing the specific GRNs indicated at the level of Box II. The progenitor field then must be subdivided into regions that give rise to the future constituent pieces of the body part, each of which is foreshadowed by a new GRN (for example the aorta or ventricle of the heart or the autopod of the limb). Within Box III thus lie the GRNs which control both the identity and the spatial boundaries of these subparts. This patterning GRN thus implements a coordinate system within the progenitor domain which is crucial for morphology and function of the body part. Both patterning GRNs (e.g. Box I and Box III) are oriented along the same axes and the downstream body-part specific patterning GRN therefore depends at least indirectly on the upper level postgastrular patterning GRNs. Depending on the complexity of the body part, multiple rounds of spatial regulatory state subdivision and installation of further regional GRNs may be required. Thus, the progression from Box II- and Box III-type GRNs may be reiterated (backwards arrow in Fig. 3). Only following these patterning processes, the terminal cell fate specification GRNs (Box IV) become activated in spatially restricted domains within the body part progenitor field. At the lower periphery of developmental GRNs are the differentiation gene batteries, that is, the protein coding effector genes plus their immediate transcriptional regulatory drivers.

What kinds of subcircuit topologies are found at these different levels of GRN hierarchy? So far, a number of GRNs have been elaborated which indicate the recurrent use of subcircuits in given developmental contexts (Peter and Davidson, 2009). One such subcircuit, the positive feedback subcircuit, links two or more regulatory genes by multiple activating regulatory interactions and acts to stabilize regulatory states. This is necessary in upper level

body-part specific GRNs (Box II) or cell fate GRNs (Box IV), since pattern formation processes usually occur only in a limited temporal window. Recurrent activating linkages keep the genes expressed even when the initial activating regulatory input fades. A positive intercellular feedback subcircuit can result in a 'community effect' (Bolouri and Davidson, 2010), the stabilizing activation of similar regulatory states within a field of cells. Here a gene encoding an intercellular signaling ligand is expressed under the control of the same signal transduction system it activates. The pattern forming GRNs of Box I and Box III in Fig. 3, in contrast, operate largely by means of transient signal inputs as well as repressive exclusion functions that control spatial subdivision. Patterning processes are not concerned with stabilization or homogenization of regulatory states and they contain few positive feedback loops. The biological function of individual subcircuit topologies predicts the probability of its occurrence at specific positions within the GRN hierarchy.

If one had to predict the GRN parts most likely modified in the evolution of body plans, a place to begin would be to define where in the developmental process and therefore in the GRN hierarchy differences occur. Morphological differences between species of different Phyla affect the basic body plan, the overall organization of the organism. During development, the body plan is established mainly by the upstream embryonic patterning mechanisms and the individual body part specification programs which they activate in given positions. Phylum-level morphological differences are therefore expected to occur in the GRNs underlying Boxes I and II. Among Classes within the same phylum, the position with respect to the body axes or the internal structures of individual body parts may differ. Differences in the position of body parts relative to each other, which basically insert body parts of different identities at given locations could occur even when embryonic patterning GRN and body part specification GRNs are conserved, simply by rewiring the connections between these functions (e.g., the linkages connecting Box I and Box II; c.f. discussion of *hox* gene functions below). Morphological differences within body parts are more likely to be caused by differences in the spatial assignment of cell fate domains determined by the body part patterning GRNs of Box III. Based on these arguments one would expect that mutations in regulatory linkages within the patterning functions are more likely to be the cause of morphological changes, whereas specification GRNs active within given cell types or body part progenitor fields are more likely to be conserved.

Given the predicted prevalence of specific network topologies for given biological functions, there might be a direct correlation between regional network topology and rate of evolutionary change. Regulatory linkages used for patterning embryos or body parts frequently rely on inductive signals which connect GRNs underlying specification in different domains, and ensure orchestrated progression of development. In organisms of different spatial geometry, inductive signaling relationships will differ, and thus, inductive signaling interactions are likely to show a higher rate of evolutionary change. Indeed they do, as discussed elsewhere (Davidson and Erwin, 2006; Erwin and Davidson, 2009). The high level of conservation of positive feedback subcircuits has been previously proposed in the Kernel theory of Davidson and Erwin (2006). These Kernels consist of a few regulatory genes linked by recursive positive regulatory interactions, and they are usually used upstream in specification GRNs and are conserved at large evolutionary distances.

In summary, evolution of GRNs to produce new developmental outcomes must involve new subcircuit deployments. This places a premium on co-optive change at the switches, signals, and inter-subcircuit inputs that encode subcircuit deployment. Evolution of novel developmental GRN features must thus proceed to some extent as a process in which diverse subcircuits are combined, recombined, activated and inactivated in given spatial domains of the embryo.

Evolutionary Consequences of Regulatory Changes in Single Genes Operating at Different Levels of GRN Hierarchy

Though the jobs of development require the outputs of multigene subcircuits of given topologies, we see from the above that there are points of "flexibility" in developmental GRNs, where co-optive gain of function, or loss of function, regulatory changes may have large effects. By focusing on naturally occurring variations between closely related animals where visible evolutionary change has occurred recently, the most evolutionarily flexible aspects of the regulatory system are revealed. In the examples that follow, in which single genes are responsible for the changes observed, it has furthermore been possible to obtain experimental evidence for the evolutionary mechanism underlying the phenotypic variation in form.

Genomic basis of rapid evolutionary trait loss—A canonical example, recently elaborated at the sequence level, and causally confirmed by experiment, is reduction of pelvic spines in stickleback fish. Following the end of the last Ice Age, marine stickleback fish were marooned in multiple lakes formed as the glaciers melted, and during the last 10,000–20,000 years independent populations of two different genera of these fish have repeatedly lost external pelvic spines. The exact selective advantages of pelvic reduction and spine loss are not defined, but as it has happened many times independently there clearly are some (Shapiro et al., 2006, and references therein). Genetic complementation tests show that diverse isolates bear the same or overlapping genetic lesions, and this is so even in crosses of species from different genera displaying the same spine reduction phenotype. The underlying genomic event turns out to be deletion of a *cis*-regulatory module which controls expression of the *pitx1* regulatory gene in the pelvic buds during larval development (Chan et al., 2010). Most significantly, when this *cis*-regulatory module was cloned upstream of a sequence encoding the Pitx1 protein and introduced into reduced spine fish, it rescued the spineless phenotype. The *cis*-regulatory module lies in an unstable, repetitive sequence-filled genomic region, possibly accounting for its repeated deletion (Chan et al., 2010). The *pitx1* gene is clearly involved in pattern formation functions upstream of pelvic girdle specification, and in spineless fish there is no *pitx1* expression in the pelvic buds even though the coding region of the gene is intact (Cole et al., 2003; Shapiro et al., 2006). In amniotes *pitx1* operates in the patterning system that organizes the subparts of the appendages developing from the hind limb buds, and forced expression in forelimb buds transforms them into hindlimbs (Logan and Tabin, 1999; Szeto et al., 1999). Thus this gene operates upstream in a portion of the GRN the function of which is to generate the spatial regulatory states that presage the parts of the amniote hindlimb, and also of the pelvis, which is rudimentary in *pitx1*$^{-/-}$ mice (Szeto et al., 1999). Though *pitx1* could execute more downstream roles in pelvic skeletal formation as well, its expression prior to the terminal phases of pelvic skeletogenesis indicates that it also functions in a Box III body part-specific patterning GRN in stickleback fish.

However, rapidly evolving, reduced, or regressive phenotypes can be due to gain of function as well as loss of function mutations. The Mexican cave fish *Astyanax* exists both in riverine surface waters and in various cave populations which were isolated about 10,000 years ago, and the regressively evolved traits of the cave populations have been studied for over a half century. A recurrent change in cave *Astyanax* is degeneration of eyes during larval development. During embryogenesis of cavefish, the eyes initially develop similarly to those of surface conspecifics, including expression of many regulatory genes (Jeffery, 2005, 2009). But then many things go wrong in eye development including apoptotic degeneration of lens and retina. A cause is ectopic spatial expression of *shh* from the normal medial interocular region across the top of the ocular fields in cave fish. As shown experimentally by introduction of *shh* mRNA in surface *Astyanax*, excess Shh causes expression of

transcription repressors (*vax1* and *pax2a*) which interfere with *pax6* expression and thus the downstream *pax6* ocular patterning subcircuit (Jeffery, 2009; Yamamoto et al., 2004) (Baumer et al., 2002). Also, excess Shh indirectly promotes apoptosis in lens and retina. Though yet undefined at the sequence level, in cave *Astyanax* regulatory changes have evidently caused a spatial gain of function in *shh* transcription resulting in regression of the eyes.

The simplest cases of evolutionary trait loss are deleterious mutations in far downstream differentiation genes. Pigmentation is among the regressive traits in cave *Astyanax*. Two pigmentation phenotypes have been shown to be due to mutations in the protein coding sequences of receptors directly involved in pigmentation, *oca2* (Protas et al., 2006) and *mc1r* (Gross et al., 2009). However, in stickleback fishes where there is also loss of pigmentation in lacustrine forms, *cis*-regulatory changes rather than coding region mutations are responsible (Miller et al., 2007). Here the gene responsible encodes Kit Ligand (Steele factor) and this gene has pleiotropic effects, so that total loss of function would be severely deleterious. Loss of function in a single *cis*-regulatory module, on the other hand, has specific effects that under certain conditions are adaptive. Since this is a general feature of *cis*-regulatory vs. coding sequence mutations, it predicts that evolutionary changes in any pleiotropically active gene, as are most regulatory genes, will generally target specific *cis*-regulatory modules (as discussed e.g., by Chan et al., 2010; Miller et al., 2007; Prud'homme et al., 2006). Inverting this argument, we see a powerful evolutionary explanation for the modularity generally typical of the *cis*-regulatory systems controlling expression of regulatory and signaling genes in animal genomes (Davidson, 2001; Davidson, 2006). GRN evolution by regulatory gain and loss of function of expression of these genes would be utterly impossible were these control systems not in general modular, since almost all such genes function in multiple time-space compartments, and in multiple GRNs during development. Physical and functional modularity in the control systems of regulatory genes is thus among the fundamental characteristics of animal genomes that permit, and indeed, that produce, evolution of development by GRN reorganization.

Morphological variation at different taxonomic levels due to single gene regulatory changes—Whereas the foregoing concerns rapidly occurring evolutionary changes in single gene functions which are of adaptive significance, we now face a conundrum. How do we extrapolate from recent evolutionary events to the much more ancient processes by which Order and Class level differences in body plan arose, let alone Phylum level differences?

Recent studies that focused on the adaptive evolution of external traits in and among *Drosophila* species have revealed processes of *cis*-regulatory sequence microevolution. Such processes account for variation in pigmentation patterns due to regulatory changes affecting expression of the *yellow* differentiation gene (Gompel et al., 2005; Rokas and Carroll, 2006); and the *ebony* differentiation gene (Rebeiz et al., 2009). Similarly, *cis*-regulatory evolution in the *shavenbaby* (*ovo*) regulatory gene, which controls trichome differentiation and morphogenesis, determines where this gene is expressed, and thereby the minute pattern differences in trichome distribution distinguishing *Drosophila* species (McGregor et al., 2007). These studies afford multiple real examples of *cis*-regulatory site addition, and quantitative as well as qualitative *cis*-regulatory gain and loss of function due to internal DNA sequence change (cf. Table1). They provide general and specific indication of the flexibility and changeability of *cis*-regulatory modules in local evolution, at the level of function and deployment of differentiation gene batteries, the lowest level in the hierarchy of Fig.3.

Mechanistic studies of intra- and inter-specific evolutionary variation illuminate the next level up as well, that is, evolutionary changes (other than simple loss of function) in the Box III-type pattern formation GRNs that determine the morphological characteristics of given body parts. The results have thus far often resolved into demonstration of alterations in the deployment of signal systems in the development of these parts; i.e., the underlying evolutionary change is in the *cis*-regulatory apparatus controlling time and place of inductive signaling, just as predicted earlier. The causal developmental mechanism underlying the adaptively diverse beak morphologies of Darwin's classic series of Galapagos finch species was solved in these terms by Abzanov et al. (2006, 2004). Species with heavy beaks displayed earlier and higher BMP4 expression in pre-beak neural crest mesenchyme, and species with elongated, pointed beaks were discovered to express Ca Calmodulin at higher levels, indicating that beak length depends on extent of Ca^{++} signaling. Remarkably, experimental over-expression of BMP4 by retroviral gene transfer into developing frontonasal tissues of chicken embryos produced robust beaks, and experimental over-expression of the downstream mediator of Ca^{++} signaling, CaMKII, produced elongated beaks, confirming the causality. To take another example, a recent study shows that short legs in dog breeds such as dachshunds and basset hounds is due to a retrogene encoding FGF4, inserted and evidently controlled by *cis*-regulatory elements carried in non-LTR transposons (Parker et al., 2009).

Changes in upstream patterning apparatus can account for differences in body plan at inter-Ordinal to inter-Class differences, and they are not found in comparing organisms that diverged only a few million or a few thousand years ago or less. For example one of the characters distinguishing bats and rodents, which are of different mammalian Orders and in fact belong to different super-Orders, is the much longer relative length of the forearm skeleton in bats. A candidate regulatory gene known to affect limb skeletal elongation is *prx1* (*mhox*), and in bats this gene is up-regulated after the early limb bud stage compared to mice (Cretekos et al., 2008). The (indirect) causality of this change was then demonstrated by inserting the bat *prx1* limb enhancer into the mouse gene, with the result that the forelimbs of the recipient mouse now develop with relatively longer dimensions. In an essentially similar case, the *tbx5* gene, deeply embedded in the vertebrate heart formation GRN (for review Davidson, 2006), turns out to be regulated differently during heart formation in reptiles than in birds and mammals, a Class-level difference. Expression of this gene is confined to the left ventricle in the developing amniote heart but is expressed across the common ventricle in the three chambered reptile heart (Koshiba-Takeuchi et al., 2009). When uniform *tbx5* expression was forced in the mouse heart, or left ventricle *tbx5* expression was prevented, i.e., if a reptilian *tbx5* spatial regulatory expression was imposed, the mouse developed a three-chambered heart lacking an interventricular septum. Understanding of developmental GRN structure tells us that these examples differ from the foregoing in that they imply the existence of Box III GRN subcircuits in which the targeted genes participate. In contrast, in the peripheral gene examples above, the phenotype is wholly encompassed by changes in a single *cis*-regulatory system.

Paradigmatic switches affecting upper level GRN patterning systems: *hox* gene functions—Genes of the trans-bilaterian *hox* complexes have been the subject of a vast amount of phenomenological research, which displays the many and various effects on developmental morphology of *hox* gene knockouts or ectopic *hox* gene expression. The variety of effects precludes any simple interpretation of the functions of these genes in the terms of developmental GRN structure, for the simple reason that they work at diverse levels. Studies of direct *hox* gene targets reveal both other regulatory genes and far downstream genes encoding proteins active, e.g., in apoptosis, cell cycle control, cell adhesion, cell polarity, non-canonical signalling, and cytoskeletal functions (Cobb and Duboule, 2005; Hueber and Lohmann, 2008; Pearson et al., 2005). But *Hox* genes are most

famous for their developmental effects on the placement and the internal organization of body parts. The most important evolutionary and developmental attributes of *hox* gene complex function can be reduced to two statements: first, in organisms in which coherent *hox* complexes exist they are expressed in a vectorial or sequential fashion with respect to the coordinates of the body plan or the body part; and second, they can act as switches which allow (or activate) GRN patterning subcircuits in given locations of the body plan or body part, or alternately they prohibit (or repress) these subcircuits in given locations.

The genomic organization of *hox* gene clusters indicates that distinct mechanisms account for the locations in the body plan where individual *hox* genes are expressed in development. In *Drosophila* a plethora of *cis*-regulatory modules control each aspect of expression of each gene. Particularly well known at the *cis*-regulatory level is the bithorax region (Ho et al., 2009; Maeda and Karch, 2009; Simon et al., 1990). Each specific *hox* gene enhancer responds to local upstream regulatory states which are the product of earlier developmental GRNs, just as in any other developmental process. Similarly, many very well characterized *cis*-regulatory modules controlling very specific spatial and temporal aspects of anterior *hox* gene expression are known in mammals, and often conserved to fish (Tumpel et al., 2009). The prevalence of local *cis*-regulatory hox gene control modules explains how these genes can function in animals which lack large *hox* gene clusters. It is interesting that *hox* genes are not required for embryonic development of organisms that utilize fixed cell lineages for specification (Davidson, 1990), e.g., in *C. elegans* which lacks both a coherent *hox* complex and many *hox* genes (Aboobaker and Blaxter, 2010); in sea urchins (Martinez et al., 1999); or in *Ciona*, which also lacks a coherent *hox* complex (Ikuta et al., 2010). However, in addition to control by local enhancers, another entirely different mechanism that speaks directly to both the evolutionary maintenance of the *hox* gene cluster(s) and the vectorial expression of *hox* genes relative to one another, has come to light in mammals and other tetrapods.

Transcriptional control of the *hoxd* complex has been examined in a unique way over the last decade in an extensive and elegant series of studies carried out on the mouse chromosomal *hoxd* complex by deletions, rearrangements, and insertions of reporter transgenes, including ectopically positioned *hox* genes, at various locations in the complex (Herault et al., 1999; Herault et al., 1998; Kmita et al., 2000; Spitz et al., 2003; Tarchini and Duboule, 2006). To summarize very briefly, early expression in the tetrapod limb bud is controlled not only by local enhancers but also by distant regulatory regions located outside the *hox* gene clusters. One of these operates from the 3' (anterior) end of the cluster and causes the progressive expression of first anterior and then middle *hox* genes in the limb bud region that will give rise to the forearm. Meanwhile the posterior hox genes are repressed by a counteracting locus control region operating from beyond the 5' end of the complex in the anterior cells of the early limb bud, allowing expression of these genes only in the posterior limb bud cells. A second phase of *hoxd* expression is controlled by other complex distant enhancers located 200kb away from the 5' end of the cluster, which are required to pattern the autopod region of the tetrapod limb where the digits form (Tarchini and Duboule, 2006). This “global control region” (GCR) is responsible for a graded expression of the five posterior *hox* genes across the A/P dimension of the autopod. The GCR probably had an ancient role in controlling colinear expression in the CNS, a basal axial organization function that in terms of our Fig.3 would reside somewhere in Box I; part of the active GCR elements are conserved from fish to mammals. However some limb specific elements of the GCR likely evolved in tetrapods, particularly the autopod control device and its patterning GRN, which would make the autopod a novel evolutionary invention with respect to the fish antecedents (Gonzalez et al., 2007; Woltering and Duboule, 2010). More generally, it is an interesting speculation that distant *hox* complex control regions were superimposed during chordate evolution (they are absent from *Drosophila*), and control by local *hox* gene

enhancers was the primal regulatory mode (Spitz et al., 2001). However, since the regulatory landscape to which the local enhancers must respond can be very different in different organisms, they themselves must have evolved in clade specific ways.

Given these systems, deeply conserved and otherwise, by which *hox* gene expression is regionally controlled, we come to their mode of interaction with the GRNs that control development of specific body parts. Sometimes individual *hox* genes act by participating, like any other regulatory genes in patterning GRNs, for example in early hindbrain specification, a Box I function. Together with other important regulatory genes such as *krox* and *Kreisler*, the anterior group *hoxa* and *hoxb* genes establish recursively wired, extremely conserved, rhombomere specific GRNs (Tumpel et al., 2007; 2009). But more often they operate in another, evolutionarily flexible way, such that change in their functions has been directly correlated, in many comparative observations, with evolutionary change in both the positioning and organization of body parts.

Not all body parts require the vectorial patterning function of the hox gene complex, for example they are not expressed in the midbrain or forebrain of vertebrates and they have nothing to do with the specification of the extremely complex regional regulatory states installed during midbrain or forebrain development. Where vectorial inputs are required, *hox* genes intervene in local, mid-development, patterning functions (Box III). Here we can rely on a number of specific examples. These are of immediate evolutionary significance in that the developmental outcomes that they control vary sharply among related clades. For example, the tetrapod limb bud is a “new” evolutionary invention, dating to the emergence of vertebrate forms onto land. Development of the limb depends directly on deployment of *hox* gene expression at several levels of the underlying GRN. The early expression of 5' *hox* genes at the posterior margin of the bud causes expression of the *shh* gene in these cells, ultimately setting up anterior and posterior regulatory states in the limb bud (Zakany et al., 2004). Posterior 5' *hox* gene expression can be thought of as a switch activating the reponsible circuitry. Later, during the autopod expression phase, the GCR responds in turn to graded levels of Shh contributing to the nested pattern of *hox* gene expression in the autopod, and the GCR can be thought of as a node in the patterning network. Another example concerns the axial skeleton in vertebrates, which vary greatly in the distribution of vertebral morphologies, again a developmental function of *hox* gene expression patterns. It is now possible to state just which sets of vertebrae require *hox5*PG, *hox6*PG, *hox9*PG, *hox10*PG, and *hox11*PG (for review, Wellik, 2009). These relationships can all be interpreted in one simple way. This is that for each type of vertebra (cervical, rib-bearing thoracic, lumbar, etc) there is a specific patterning GRN operating at the Box III level, and the products of (often) two adjacent PGs allow it to be activated in the right place along the axis, or may cause it to be activated ectopically when these *hox* genes are activated ectopically. That is to say, these *hox* genes act as regionally active switches that we can imaging sitting on the outside of the boxes containing the morphogenetic patterning GRNs. Switch behavior is particularly easy to perceive when the switch acts negatively: thus the PG10 *hox* genes prevent rib formation, normally used to preclude ribs on the lumbar vertebrae; if expressed ectopically no ribs form and in complete loss of function ribs form almost everywhere (Carapuco et al., 2005; Vinagre et al., 2010; Wellik and Capecchi, 2003). On the other hand *hox6*PG genes promote rib formation. The autonomy of these *hox*-driven switches, as shown by the complete ectopic production of one or another vertebral type in gain of function experiments, implies a useful evolutionary mechanism for variation in axial skeletal proportions. Indeed, comparative observations show that different vertebrate classes have *hox* spatial expression domains that correlate with the axial morphology (examples reviewed in Davidson, 2006). However, the most severe axial changes in tetrapod evolution, those responsible for the body plans of snakes and reptiles, have involved more than merely upstream regulatory changes affecting *hox* gene expression domains (Di-Poi et al., 2010;

Woltering et al., 2009): in addition, the sequences of some of the genes themselves have changed, regulatory linkages between gene expression and effects such as the *hox10* inhibition of rib formation have been broken, and numerous transposon insertions have altered the genomic structure of the posterior *hox* cluster possibly affecting their spatial regulation.

The mechanism by which the *Drosophila ubx* gene represses wing formation in the third thoracic (T3) segment provides the most explicit possible illustration of what it means for a *hox* gene to intervene negatively and switch off a local patterning GRN. In the absence of *Ubx* function in T3, what should be the haltere imaginal disc produces a wing, hence Ed Lewis' famous 4-winged fly (Bender et al., 1983). Thus *Ubx* function is repressive with respect to the wing patterning GRN in the late T3 imaginal disc. The way this works is repression by Ubx and its co-factors of several genes of the wing GRN, as shown by analyses of Ubx^{-} clones in the haltere disc and of Ubx^{+} clones in the wing disc (Galant et al., 2002; Weatherbee et al., 1998). These are direct *cis*-regulatory repressions. There are many arthropod examples not yet examined at the GRN level where the mechanisms of *hox* gene function must be similar. In arthropods the anterior boundaries of expression of the *Ubx/Abd-A* genes vary from Class to Class, and sometimes among orders of the same Class, e.g., among crustaceans, and this boundary is correlated with the type of appendage present on the segment; from these correlations, *Ubx* evidently represses execution of the patterning GRN underlying development of feeding appendages (maxillipeds) and permits development of locomotory thoracic appendages (Averof and Patel, 1997). This inference has been demonstrated, by experimentally decreasing or increasing *Ubx* expression in a shrimp that normally produces one pair of maxillipeds, with the result of producing additional pairs of these appendages or instead only thoracic legs, respectively (Liubicich et al., 2009; Pavlopoulos et al., 2009). *Drosophila* affords many further examples of *hox* gene switches that permit or preclude regional morphogenetic GRN function in body part formation, among the most convincing of which is in heart development (Lo et al., 2002). Further examples of regional *hox* gene control of specific body part identity by *cis*-regulatory intervention are in somatic muscle pair specification. Each muscle develops from founder cells expressing specific transcription factors, i.e., a specific regulatory state (Baylies et al., 1998). There are direct *hox* gene inputs into this process; for example, the alary muscles which connect the aorta of the heart and which require *Ubx* and *AbdA* for their development (Dubois et al., 2007; LaBeau et al., 2009). Throughout the body plan hox genes control clade-specific deployment of organs and structures.

So in summary, the common statement that hox genes "pattern" this or that body part means that they provide negative or positive cis-regulatory inputs into genes which are engaged in the GRN circuits which actually do the work of spatial patterning and body part morphogenesis. Sometimes the *hox* gene inputs form part of the subcircuit itself as when there are feedback linkages between them and other regulatory genes, as in the later limb bud or rhombomere specification circuitry cited above. But in many more cases than those mentioned here the function of these regionally expressed genes is rather to provide a one way switch which provides GO or NO GO instruction to body part-specific GRN patterning circuitry. In evolution the deployment of these switches, and the linkages between them and the body part specific subcircuits, are far more flexible than is the internal structure of these subcircuits. Some of these body part specific GRN structures are in evolutionary terms very ancient indeed.

Conservation and Change in Developmental GRNs

The self described field of "evo-devo" has generated enormous masses of descriptive spatial gene expression data, a frequent object of which is to show evolutionary "conservation" of developmental gene use. Developmental gene use cannot truly be regarded as conserved

unless the regulatory linkages surrounding the genes in the GRN are conserved. Thus gene expression data by themselves are a poor index of evolutionary conservation. Since negative results are uninformative, we learn little of what has changed by looking only at what has not. Unless all forms were “sprung forth fully blown” like Athena from the head of Zeus, the evolution of the diverse body plans of animals requires large scale processes of change in ancestral developmental GRN architecture. Furthermore, what is it that is conserved: is it use of a given gene in a given developmental process? Is it use of a given gene in a given subcircuit in a given process? Here we consider evolutionary conservation and evolutionary change, not of specific individual gene use, but of specific GRN circuitry.

Conservation—The hierarchical Linnean classification system we use, including modern corrections based on molecular phylogenetics, essentially arranges animal body plans on the basis of their evolutionarily shared and derived characters (avoiding convergent associations). Shared body plan characters of given clades ultimately imply conserved developmental regulatory circuitry (Davidson and Erwin, 2009). But other apparently older characters are shared over huge phylogenetic distances across cladistic boundaries, being represented in multiple bilaterian phyla and in diverse body plans. These are particular body parts, such as hearts, and the major domains of brains; and particular cell types, such as muscle and neurons.

Because of their very widespread distribution, some differentiation gene batteries are probably among the oldest features of modern developmental GRNs (Davidson, 2006; Davidson and Erwin, 2009). But just as a cell type is not the same thing as a body part, so a differentiation gene battery is not the same thing as a cell type. During evolution the identity of the effector genes can change radically, whereas the biological function of the cell type remains the same; and in addition, the cell type often has cell biological or morphological characteristics which are not encoded the same way as is activation of sets of effector genes. So we have to consider what GRN structures actually lie at the root of trans-phyletic cell type conservation. A few examples may clarify this issue.

We know many cell types which are present in many types of animals, the specific properties of which depend on conserved differentiation gene batteries including both conserved downstream regulatory states and effector genes. For example, everyone is familiar with pan-eumetazoan (cnidarian plus bilaterian) conservation of striated and smooth muscle. Here the distinctive cellular morphology, the function, and underlying these, the regulatory state consisting of myogenic bHLH factors and MEF2, plus downstream effector genes exemplified by the myosin heavy chain contractile protein are all conserved (Seipel and Schmid, 2005). The same is true of neuronal cell types (e.g., Hayakawa et al., 2004). There are many additional examples where both regulatory state and effector genes are evidently conserved. A comparison between vertebrate and annelid light sensitive non-ocular neurosecretory cell types that produce vasotocin (vasopressin-neurophysin) as well as opsin provides a striking case (Tessmar-Raible et al., 2007). This cell type is located in the forebrains of both a polychaete annelid and zebrafish, as are also very similar chemosensory neurosecretory cell types that produce RF-amide. The vasotocinergic cells of both vertebrate and annelid express similar (Box IV) regulatory states, generated by the *nk2.1*, *rx*, and *otp* genes, as well as a gene producing the MIR-7 microRNA that is also, in both organisms, expressed in the RF-amidonergic cells. Vasotocinergic neurosecretory cells were probably pan-bilaterian cell types, though genes encoding vasotocin have been lost in (sequenced) ecdysozoan lineages. Ocular photoreceptor cells provide another example of a pan-bilaterian cell type in which the Box IV GRN controlling the various subtypes of receptors (rhabdomeric receptors in insects, and rods and cones in vertebrates) operate downstream of regulatory genes of the K50 homeodomain family (Mishra et al., 2010; Ranade et al., 2008). These genes are *otx2* and *crx* in mammals (Corbo et al., 2010; Hennig et al., 2008) and *otd*

in *Drosophila*, (where a paired class regulatory gene, *pph3*, which binds to the same sites as does Pax6 is also utilized in regulation of the same target genes). The transcription factors encoded by *otd*, or *crx* and *otx2* directly activate the *cis*-regulatory control systems of the genes encoding the photoreceptor pigments, in flies and mice. In addition the targets of these regulatory genes, in both flies and mice, include phototransduction genes (rhodopsins, transducins, phosphodiesterase genes, arrestins), and cell morphogenesis genes (Ranade et al., 2008). The mammalian Box IV *crx/otx2* GRN includes a canonical set of six other regulatory genes, interactions among which in mammals determines the photoreceptor subtype (Hennig et al., 2008; Swaroop et al., 2010). That is, in these cell types both downstream effector genes and their immediate regulatory apparatus are deployed in a manner that is widely conserved.

But there is another, profoundly interesting pattern of conservation displayed by pan-bilaterian cell types, in which the downstream effector genes are clade specific, while the definitive upstream regulatory states are conserved across clades. Immune cells provide the most evidence, for as knowledge of the diverse strategies for immune response, both adaptive and non-adaptive, extends beyond mammals, an amazing variety of effector genes is revealed but the same familiar sets of regulatory genes are found to control their expression. Lampreys, for example, have the equivalent of T-cells and B-cells but instead of somatically reassembled T- and B- Ig receptors, they express somatically reassembled variable leucine rich repeat receptors (Guo et al., 2009; Herrin and Cooper, 2010). Yet the T-cell like lamprey cell regulatory state includes factors encoded by familiar T-cell genes, such as *bc11b*, *gata 2/3*, *c-rel*; and like T-cells their development depends on Notch signaling. In *Drosophila* the pathogen activated innate immune response, which deploys a number of antimicrobial effector molecules, depends, as does much of our very different innate immune response, on inducible regulatory factors of the NfKB family (Hoffmann, 2003). And sea urchins, which employ a surprising and unique repertoire of hundreds of receptors of several different classes in their dedicated immune cells (Hibino et al., 2006; Messier-Solek et al., 2010) express in these cells a regulatory state very familiar to students of mammalian hematopoietic systems, such as the factors encoded by the *scl*, *e2a*, *gata1/2/3*, *ikaros*, *runx* genes, and even a *pu.1* like *ets* family gene. Another entirely different system in which a conserved cell type specific regulatory state controls entirely different effector genes, which nonetheless execute the same function, is found in the cells that in development create the outer epidermal barrier against the external world, and which recreate this barrier in wound repair. In vertebrates the barrier is composed of a mixture of cross-linked keratins of diverse kinds, matrix proteins, lipids, special cornified membrane proteins etc; in insects it is composed of cross-linked chitins, plus other proteins and lipids. The structures are entirely non-homologous in molecular identity. In mammals wound repair requires expression (among other proteins) of a cross-linking transglutaminase, while in *Drosophila* it requires expression of dopa decarboxylase and tyrosine hydroxylase, which generates quinones that cross-link chitin and cuticle proteins (Pearson et al., 2009; Ting et al., 2005). But in both flies and mice these functions are directly regulated by genes of the *grainyhead* family which encodes transcription factors that utilizes a unique DNA binding domain (also found in fungi), plus other factors of the *jun/fos* family. The Box IV cell type GRNs are conserved, but the effector genes are entirely diverse.

So we see that ancient cell type specific functions, that were utilized in the lineage ancestral to all bilaterians, are essentially defined by specific regulatory states, that is to say by genomically encoded GRN cassettes that produce cell type specific regulatory states. Sometimes the effector gene sets which these regulatory states animate are at least partially conserved, sometimes not. In evolutionary terms, the genomic repository of basic bilaterian (or eumetazoan) functions such as immunity, wound repair, contraction, photoreception, was

built into these cell type specific regulatory cassettes, and they have ever since retained their identity.

Some body parts are also conserved across the cladistic boundaries. This implies that there is something in the genetic programs for development of these body parts that is also conserved. However, in cases where the final structures are diverse, and develop via very diverse pattern formation and morphogenetic mechanisms, it may be that only the Box II GRN circuits encoding the initial establishment of the progenitor field from which the body part will be built are conserved, plus the final deployment of conserved cell types. Comparative GRN analysis is beginning to reveal “kernels” (Davidson and Erwin, 2006), in which regulatory genes wired together in certain conserved linkages execute upstream regulatory functions in development of given body parts. These circuits are characterized by extensive feedback wiring, and where tested, interference with expression of any of their genes results in developmental catastrophe. These features, and developmental canalization due to the upstream position of such kernels in the body part GRN, explain their exceptional evolutionary conservation. Examples include what may be a pan-bilaterian (i.e., from flies to mice) kernel for heart specification (Davidson, 2006); and an (at least) pan-echinoderm kernel underlying mesoderm specification in both sea urchin and sea star development ((McCauley et al., 2010); these lineages have not shared a common ancestor since the end of the Cambrian). Similarly, a fundamental Box II subcircuit may underlie mesoderm specification in vertebrate embryogenesis (Swiers et al., 2010). A recursively wired triple feedback circuit has been proposed as a kernel underlying the pluripotential state of endothelial/hematopoietic precursors that arise in vertebrate development (Pimanda et al., 2007). There are also many less coherent observations, not yet at the level of an explicit GRN, in which detailed patterning similarities plus some gene interaction data strongly suggest the existence of GRN kernels that yet await elucidation. One convincing example is the brain, where a large amount of work has illuminated striking similarities in both A/P and medio-lateral patterns of regulatory gene expression as well as homologous gene interactions between *Drosophila* and mouse (Davidson, 2006; Denes et al., 2007; Lowe et al., 2003; Seibert and Urbach, 2010; Tessmar-Raible et al., 2007).

Evolutionary change in GRN architecture—Evolutionary rewiring of GRN architecture by means of *cis*-regulatory co-optive change of given linkages among regulatory genes is the most common upper level evolutionary mechanism by which developmental process is altered. That is, the GRN of a common ancestor is the source structure for diverse alterations in that structure in the derived descendants. But of course, not all parts of the structure are equally accessible to change, for the reasons we have tried here to point out. Sometimes the contrast between the conserved and non-conserved parts of a given GRN is quite dramatic, as in a comparison between sea urchin and sea star endomesodermal GRNs that revealed an extremely conserved, 5-gene subcircuit, surrounded by linkages not one of which had survived the half billion years since divergence without change (Hinman and Davidson, 2007). This was an inter-Class comparison; for visualization of the process of evolutionary GRN rewiring the inter-Ordinal comparison between developmental GRNs in *Drosophila* and *Tribolium* is illuminating.

Remarkable examples of architectural GRN rewiring have come to light in comparisons of the segmentation GRNs of various insects. Since the short germ band mode of development appears to be pleisiomorphic for insects and their sister group the crustaceans, the linkages seen in the early A/P patterning GRN of *Tribolium,* for example, may be closer to the ancestral linkages than the derived linkages of *Drosophila* GRNs. This is supported by a vast literature on many other insects and crustaceans as well (see following citations for references to work on other species). Every major aspect of A/P patterning analyzed at the gene interaction level appears to include some different linkages in *Drosophila* compared to

Tribolium. For example, it had been thought that the absence of the *bicoid* gene outside of higher Diptera was compensated in other groups by a similar function of the anteriorly expressed *otd* gene, which encodes a regulator with a Bcd-like homeodomain and target specificity. But recent work shows that in *Tribolium otd* functions very differently from *bcd* in *Drosophila*, in that it operates through different downstream linkages (Kotkamp et al., 2010). Unlike *bcd* in *Drosophila*, it controls D/V patterning, by repressing *sog* expression; it affects *zen* expression; and it contributes no spatial A/P input to the patterning process. Similarly it is clear that some of the GRN wiring downstream of the *hunchback* gene differs, for in *Tribolium hb* apparently does not directly control primary pair rule genes, and does not repress, but rather activates *giant* (Choe et al., 2006; Marques-Souza et al., 2008). On the other hand, in both species *hb* sets the anterior boundary of *Ubx* expression and provides an activating input into the *kruppel* regulatory system (Marques-Souza et al., 2008). The architecture of pair rule GRNs in *Drosophila* and *Tribolium*, which are composed of largely the same genes is very different (Choe and Brown, 2009; Jaynes and Fujioka, 2004), but, amazingly, they generate the same downstream outcomes, the expression of *wg* and *en* across each parasegment border. Some linkages in the pair rule GRNs are the same, but many are entirely rewired: e.g., *eve* directly represses *wg* in *Tribolium* while *eve* indirectly represses *en* in *Drosophila* (via *slp*; (Choe and Brown, 2009), and upstream of this, in *Tribolium* is a possibly pleisiomorphic kernel-like segmentation subcircuit, consisting of mutual interconnections among *eve*, *runt*, and *odd*, which runs sequentially to pattern the forming segments (Choe et al., 2006). Another example of extensive rewiring among the same genes engaged in the same developmental process since divergence between Coleoptera and Diptera is in eye specification. A comparison between the relatively well known eye specification GRN of *Drosophila* (for reviews, Friedrich, 2006; Kumar, 2009)) with that governing larval and adult eye specification in *Tribolium* (Yang et al., 2009) displays remarkable differences. The genes at the top of the *Drosophila* hierarchy, e.g., *toy* and *ey* (*pax6* orthologues), are not even needed for adult eye development in *Tribolium*, where another gene in the same network, *dachshund*, operates redundantly with *pax6*, rather than being located downstream of the *pax6* genes in the network.

Co-option of subcircuits—Considering what we know of how developmental GRNs are constructed, it is not surprising that successful developmental programs are used repeatedly, plugged into various positions in the GRN hierarchy. One example of a subcircuit-level co-option event has been discovered in sea urchins, in which a Class-specific evolutionary modification has caused the acquisition of an embryonic skeleton not present in other echinoderms. The shared feature of echinoderms is an endoskeleton in the adult organism. A comparison of regulatory gene expression in embryonic and adult skeletogenic precursor cells revealed a large overlap at least at the nodes of these GRNs which very likely extends also to the linkages between them (Gao and Davidson, 2008; Oliveri et al., 2008). Regulatory genes exclusive to the embryonic GRN are those determining the embryonic location in which the skeletogenic GRN is activated. Thus, by modifying probably only a small number of *cis*-regulatory sequences, the skeletogenic GRN subcircuit was re-deployed such that it is activated both in the embryo and in the adult, most likely by use of the exact same genomic sequences.

A similar interpretation has been applied to the apparent conservation of proximodistal patterning mechanisms in entirely non-homologous bilaterian appendages (Lemons et al., 2010). Thus, *Drosophila* and vertebrate leg progenitor fields express the same set of regulatory genes in the same sequential order along the proximodistal axis, although as a result of different regulatory interactions. Interestingly, this very same sequence of regulatory gene expressions is observed also along the anteroposterior axis in the head neuroectoderm of *Drosophila* and *Saccoglossus*, a hemichordate lacking appendages. McGinnis et al. therefore propose that the similarity of patterning observed in these non-

homologous bodyparts might be the result of independent co-options of a subcircuit with conserved function. A relatively recent co-option of an entire body part has occurred in teleost fish, resulting in the formation of a secondary jaw in the same location where ancient pharyngeal teeth developed (Fraser et al., 2009). Malawi Cichlids, which possess both oral and pharyngeal jaws and teeth, show very similar expression of signaling molecules and transcription factors in tooth forming cells in both locations, supporting the hypothesis that tooth development in oral and pharyngeal jaws is driven by the same tooth GRN. However, one substantial difference exists, which is the expression of a set of *hox* genes in the pharyngeal but not the oral jaw. In mouse pharyngeal arches, *hox* genes are expressed in all but the first pharyngeal arch, which gives rise to the oral jaw. Mutation of genes in the *hoxa* cluster resulted in formation of ectopic jaw-like skeletal structures and *hox* genes are therefore thought to prevent the development of jaws in caudal pharyngeal arches (Minoux et al., 2009). In other words, the jaw-forming patterning GRN would be expressed in more posterior pharyngeal arches were it not for the repressive *hox* gene switch. Therefore, if the same was true in the teleost ancestor, a prerequisite for the evolutionary co-option of the jaw GRN to a causal pharyngeal arch in Cichlids would have been the uncoupling of this developmental program from the the repressive *hox* input. These examples may display the co-optive redeployment of developmental GRNs, and the switch-like function of *hox* genes in controlling spatial utilization of GRNs.

CONCLUSIONS

To make sense of the physical mechanisms that underlie the origin of animal body plans (Davidson and Erwin, 2009), and during and since the Cambrian their innovative diversification, we must consider how change in DNA sequence can affect development of the body plan at the system level. For development of the body plan is a heritable regulatory system process, which we can represent and manipulate and comprehend only in terms of genomically encoded GRN architecture. The evidence that comes to us from evo-devo comparisons of gene expression patterns, from detailed studies of regulatory changes in single genes, from direct comparative GRN analysis, from evolutionary conservation and from evolutionary innovation, and from the fossil record, can only be integrated in a mechanistic way by resolving the meaning of this evidence in terms of its import for developmental GRN architecture. This is the path to demystification of body plan evolution. This project cannot be approached, except in an indirect exemplary sense, by looking at change in single *cis*-regulatory modules or single proteins, nor in ignorance of the regulatory gene interactions that constitute the architecture of developmental GRNs. The theory of evolution by change in GRN architecture also generates the path to experimental validation of evolutionary process by synthetic changes in developmental GRNs. This approach is already beginning to be applied, as we review above. The genomic control of the developmental process itself can only be understood in terms of the genomic regulatory system, and so must time based change in that regulatory system, the basis of body plan evolution.

Acknowledgments

We gratefully acknowledge support for this work from NIH Grant HD-037105, and from NSF Grant IOS-0641398.

References

Aboobaker A, Blaxter M. The nematode story: Hox gene loss and rapid evolution. Adv Exp Med Biol. 2010; 689:101–110. [PubMed: 20795325]

Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature. 2006; 442:563–567. [PubMed: 16885984]

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. Bmp4 and morphological variation of beaks in Darwin's finches. Science. 2004; 305:1462–1465. [PubMed: 15353802]

Averof M, Patel NH. Crustacean appendage evolution associated with changes in Hox gene expression. Nature. 1997; 388:682–686. [PubMed: 9262403]

Balhoff JP, Wray GA. Evolutionary analysis of the well characterized endo16 promoter reveals substantial variation within functional sites. Proc Natl Acad Sci USA. 2005; 102:8591–8596. [PubMed: 15937122]

Baumer N, Marquardt T, Stoykova A, Ashery-Padan R, Chowdhury K, Gruss P. Pax6 is required for establishing naso-temporal and dorsal characteristics of the optic vesicle. Development. 2002; 129:4535–4545. [PubMed: 12223410]

Baylies MK, Bate M, Ruiz Gomez M. Myogenesis: a view from Drosophila. Cell. 1998; 93:921–927. [PubMed: 9635422]

Bender W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P, Lewis EB, Hogness DS. Molecular Genetics of the Bithorax Complex in Drosophila melanogaster. Science. 1983; 221:23–29. [PubMed: 17737996]

Bolouri H, Davidson EH. The gene regulatory network basis of the “community effect,” and analysis of a sea urchin embryo example. Dev Biol. 2010; 340:170–178. [PubMed: 19523466]

Britten RJ. Mobile elements inserted in the distant past have taken on important functions. Gene. 1997; 205:177–182. [PubMed: 9461392]

Britten RJ, Davidson EH. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q Rev Biol. 1971; 46:111–138. [PubMed: 5160087]

Cameron RA, Davidson EH. Flexibility of transcription factor target site position in conserved cis-regulatory modules. Dev Biol. 2009; 336:122–135. [PubMed: 19766623]

Carapuco M, Novoa A, Bobola N, Mallo M. Hox genes specify vertebral types in the presomitic mesoderm. Genes Dev. 2005; 19:2116–2121. [PubMed: 16166377]

Chan YF, Marks ME, Jones FC, Villarreal G Jr, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, et al. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science. 2010; 327:302–305. [PubMed: 20007865]

Choe CP, Brown SJ. Genetic regulation of engrailed and wingless in Tribolium segmentation and the evolution of pair-rule segmentation. Dev Biol. 2009; 325:482–491. [PubMed: 19028487]

Choe CP, Miller SC, Brown SJ. A pair-rule gene circuit defines segments sequentially in the short-germ insect Tribolium castaneum. Proc Natl Acad Sci USA. 2006; 103:6560–6564. [PubMed: 16611732]

Cobb J, Duboule D. Comparative analysis of genes downstream of the Hoxd cluster in developing digits and external genitalia. Development. 2005; 132:3055–3067. [PubMed: 15944189]

Cole NJ, Tanaka M, Prescott A, Tickle C. Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. Curr Biol. 2003; 13:R951–952. [PubMed: 14680650]

Corbo JC, Lawrence KA, Karlstetter M, Myers CA, Abdelaziz M, Dirkes W, Weigelt K, Seifert M, Benes V, Fritsche LG, et al. CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. Genome Res. 2010; 20:1512–1525. [PubMed: 20693478]

Cretekos CJ, Wang Y, Green ED, Martin JF, Rasweiler JJt, Behringer RR. Regulatory divergence modifies limb length between mammals. Genes Dev. 2008; 22:141–151. [PubMed: 18198333]

Davidson EH. How embryos work: a comparative view of diverse modes of cell fate specification. Development. 1990; 108:365–389. [PubMed: 2187672]

Davidson, EH. Genomic Regulatory Systems: Development and Evolution. San Diego: Acdemic Press; 2001.

Davidson, EH. Gene Regulatory Networks in Development and Evolution. San Diego, CA: Academic Press/Elsevier; 2006. The Regulatory Genome.

Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans. Science. 2006; 311:796–800. [PubMed: 16469913]

Davidson EH, Erwin DH. An integrated view of precambrian eumetazoan evolution. Cold Spring Harb Symp Quant Biol. 2009; 74:65–80. [PubMed: 20375317]

Davidson EH, Erwin DH. Evolutionary innovation and stability in animal gene networks. J Exp Zool B Mol Dev Evol. 2010; 314:182–186. [PubMed: 19937660]

Denes AS, Jekely G, Steinmetz PR, Raible F, Snyman H, Prud'homme B, Ferrier DE, Balavoine G, Arendt D. Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. Cell. 2007; 129:277–288. [PubMed: 17448990]

Dermitzakis ET, Bergman CM, Clark AG. Tracing the evolutionary history of Drosophila regulatory regions with models that identify transcription factor binding sites. Mol Biol Evol. 2003; 20:703–714. [PubMed: 12679540]

Di-Poi N, Montoya-Burgos JI, Miller H, Pourquie O, Milinkovitch MC, Duboule D. Changes in Hox genes' structure and function during the evolution of the squamate body plan. Nature. 2010; 464:99–103. [PubMed: 20203609]

Dubois L, Enriquez J, Daburon V, Crozet F, Lebreton G, Crozatier M, Vincent A. Collier transcription in a single Drosophila muscle lineage: the combinatorial control of muscle identity. Development. 2007; 134:4347–4355. [PubMed: 18003742]

Elgar G, Vavouri T. Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. Trends Genet. 2008; 24:344–352. [PubMed: 18514361]

Erwin DH, Davidson EH. The evolution of hierarchical gene regulatory networks. Nat Rev Genet. 2009; 10:141–148. [PubMed: 19139764]

Fraser GJ, Hulsey CD, Bloomquist RF, Uyesugi K, Manley NR, Streelman JT. An ancient gene network is co-opted for teeth on old and new jaws. PLoS Biol. 2009; 7:e31. [PubMed: 19215146]

Friedrich M. Ancient mechanisms of visual sense organ development based on comparison of the gene networks controlling larval eye, ocellus, and compound eye specification in Drosophila. Arthropod Struct Dev. 2006; 35:357–378. [PubMed: 18089081]

Galant R, Walsh CM, Carroll SB. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development. 2002; 129:3115–3126. [PubMed: 12070087]

Gao F, Davidson EH. Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution. Proc Natl Acad Sci USA. 2008; 105:6091–6096. [PubMed: 18413604]

Garza D, Medhora M, Koga A, Hartl DL. Introduction of the transposable element mariner into the germline of Drosophila melanogaster. Genetics. 1991; 128:303–310. [PubMed: 1649067]

Gogvadze E, Buzdin A. Retroelements and their impact on genome evolution and functioning. Cell Mol Life Sci. 2009; 66:3727–3742. [PubMed: 19649766]

Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. Nature. 2005; 433:481–487. [PubMed: 15690032]

Gonzalez F, Duboule D, Spitz F. Transgenic analysis of Hoxd gene regulation during digit development. Dev Biol. 2007; 306:847–859. [PubMed: 17448461]

Gross JB, Borowsky R, Tabin CJ. A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish Astyanax mexicanus. PLoS Genet. 2009; 5:e1000326. [PubMed: 19119422]

Guo P, Hirano M, Herrin BR, Li J, Yu C, Sadlonova A, Cooper MD. Dual nature of the adaptive immune system in lampreys. Nature. 2009; 459:796–801. [PubMed: 19474790]

Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB. Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. PLoS Genet. 2008; 4:e1000106. [PubMed: 18584029]

Hayakawa E, Fujisawa C, Fujisawa T. Involvement of Hydra achaete-scute gene CnASH in the differentiation pathway of sensory neurons in the tentacles. Dev Genes Evol. 2004; 214:486–492. [PubMed: 15378362]

Hennig AK, Peng GH, Chen S. Regulation of photoreceptor gene expression by Crx-associated transcription factor network. Brain Res. 2008; 1192:114–133. [PubMed: 17662965]

Herault Y, Beckers J, Gerard M, Duboule D. Hox gene expression in limbs: colinearity by opposite regulatory controls. Dev Biol. 1999; 208:157–165. [PubMed: 10075849]

Herault Y, Beckers J, Kondo T, Fraudeau N, Duboule D. Genetic analysis of a Hoxd-12 regulatory element reveals global versus local modes of controls in the HoxD complex. Development. 1998; 125:1669–1677. [PubMed: 9521905]

Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, Tanaka M, Schartl M. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet. 2010; 6:e1000844. [PubMed: 20169179]

Herrin BR, Cooper MD. Alternative adaptive immunity in jawless vertebrates. J Immunol. 2010; 185:1367–1374. [PubMed: 20660361]

Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, Buckley KM, Brockton V, Nair SV, Berney K, et al. The immune gene repertoire encoded in the purple sea urchin genome. Dev Biol. 2006; 300:349–365. [PubMed: 17027739]

Hinman VF, Davidson EH. Evolutionary plasticity of developmental gene regulatory network architecture. Proc Natl Acad Sci USA. 2007; 104:19404–19409. [PubMed: 18042699]

Ho MC, Johnsen H, Goetz SE, Schiller BJ, Bae E, Tran DA, Shur AS, Allen JM, Rau C, Bender W, et al. Functional evolution of cis-regulatory modules at a homeotic gene in Drosophila. PLoS Genet. 2009; 5:e1000709. [PubMed: 19893611]

Hoffmann JA. The immune response of Drosophila. Nature. 2003; 426:33–38. [PubMed: 14603309]

Hong JW, Hendrix DA, Papatsenko D, Levine MS. How the Dorsal gradient works: insights from postgenome technologies. Proc Natl Acad Sci USA. 2008; 105:20072–20076. [PubMed: 19104040]

Hueber SD, Lohmann I. Shaping segments: Hox gene function in the genomic age. Bioessays. 2008; 30:965–979. [PubMed: 18798525]

Ikuta T, Satoh N, Saiga H. Limited functions of Hox genes in the larval development of the ascidian Ciona intestinalis. Development. 2010; 137:1505–1513. [PubMed: 20335361]

Jaynes JB, Fujioka M. Drawing lines in the sand: even skipped et al. and parasegment boundaries. Dev Biol. 2004; 269:609–622. [PubMed: 15110723]

Jeffery WR. Adaptive evolution of eye degeneration in the Mexican blind cavefish. J Hered. 2005; 96:185–196. [PubMed: 15653557]

Jeffery WR. Regressive evolution in Astyanax cavefish. Annu Rev Genet. 2009; 43:25–47. [PubMed: 19640230]

Jimenez-Delgado S, Pascual-Anaya J, Garcia-Fernandez J. Implications of duplicated cis-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty. Brief Funct Genomic Proteomic. 2009; 8:266–275. [PubMed: 19651705]

Johnson R, Samuel J, Ng CK, Jauch R, Stanton LW, Wood IC. Evolution of the vertebrate gene regulatory network controlled by the transcriptional repressor REST. Mol Biol Evol. 2009; 26:1491–1507. [PubMed: 19318521]

Kazazian HH Jr. Mobile elements: drivers of genome evolution. Science. 2004; 303:1626–1632. [PubMed: 15016989]

Kmita M, van Der Hoeven F, Zakany J, Krumlauf R, Duboule D. Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. Genes Dev. 2000; 14:198–211. [PubMed: 10652274]

Koshiba-Takeuchi K, Mori AD, Kaynak BL, Cebra-Thomas J, Sukonnik T, Georges RO, Latham S, Beck L, Henkelman RM, Black BL, et al. Reptilian heart development and the molecular basis of cardiac chamber evolution. Nature. 2009; 461:95–98. [PubMed: 19727199]

Kotkamp K, Klingler M, Schoppmeier M. Apparent role of Tribolium orthodenticle in anteroposterior blastoderm patterning largely reflects novel functions in dorsoventral axis formation and cell survival. Development. 2010; 137:1853–1862. [PubMed: 20431120]

Kumar JP. The molecular circuitry governing retinal determination. Biochim Biophys Acta. 2009; 1789:306–314. [PubMed: 19013263]

LaBeau EM, Trujillo DL, Cripps RM. Bithorax complex genes control alary muscle patterning along the cardiac tube of Drosophila. Mech Dev. 2009; 126:478–486. [PubMed: 19272319]

Lemons D, Fritzenwanker JH, Gerhart J, Lowe CJ, McGinnis W. Co-option of an anteroposterior head axis patterning system for proximodistal patterning of appendages in early bilaterian evolution. Dev Biol. 2010; 344:358–362. [PubMed: 20435033]

Liberman LM, Stathopoulos A. Design flexibility in cis-regulatory control of gene expression: synthetic and comparative evidence. Dev Biol. 2009; 327:578–589. [PubMed: 19135437]

Liubicich DM, Serano JM, Pavlopoulos A, Kontarakis Z, Protas ME, Kwan E, Chatterjee S, Tran KD, Averof M, Patel NH. Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. Proc Natl Acad Sci USA. 2009; 106:13892–13896. [PubMed: 19666517]

Lo PC, Skeath JB, Gajewski K, Schulz RA, Frasch M. Homeotic genes autonomously specify the anteroposterior subdivision of the Drosophila dorsal vessel into aorta and heart. Dev Biol. 2002; 251:307–319. [PubMed: 12435360]

Logan M, Tabin CJ. Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. Science. 1999; 283:1736–1739. [PubMed: 10073939]

Lowe CJ, Wu M, Salic A, Evans L, Lander E, Stange-Thomann N, Gruber CE, Gerhart J, Kirschner M. Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. Cell. 2003; 113:853–865. [PubMed: 12837244]

Ludwig MZ, Bergman C, Patel NH, Kreitman M. Evidence for stabilizing selection in a eukaryotic enhancer element. Nature. 2000; 403:564–567. [PubMed: 10676967]

Maeda RK, Karch F. The bithorax complex of Drosophila an exceptional Hox cluster. Curr Top Dev Biol. 2009; 88:1–33. [PubMed: 19651300]

Marques-Souza H, Aranda M, Tautz D. Delimiting the conserved features of hunchback function for the trunk organization of insects. Development. 2008; 135:881–888. [PubMed: 18216167]

Martinez P, Rast JP, Arenas-Mena C, Davidson EH. Organization of an echinoderm Hox gene cluster. Proc Natl Acad Sci USA. 1999; 96:1469–1474. [PubMed: 9990047]

McCauley BS, Weideman EP, Hinman VF. A conserved gene regulatory network subcircuit drives different developmental fates in the vegetal pole of highly divergent echinoderm embryos. Dev Biol. 2010; 340:200–208. [PubMed: 19941847]

McGregor AP, Orgogozo V, Delon I, Zanet J, Srinivasan DG, Payre F, Stern DL. Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature. 2007; 448:587–590. [PubMed: 17632547]

Messier-Solek C, Buckley KM, Rast JP. Highly diversified innate receptor systems and new forms of animal immunity. Semin Immunol. 2010; 22:39–47. [PubMed: 20022762]

Miller CT, Beleza S, Pollen AA, Schluter D, Kittles RA, Shriver MD, Kingsley DM. cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. Cell. 2007; 131:1179–1189. [PubMed: 18083106]

Minoux M, Antonarakis GS, Kmita M, Duboule D, Rijli FM. Rostral and caudal pharyngeal arches share a common neural crest ground pattern. Development. 2009; 136:637–645. [PubMed: 19168678]

Mishra M, Oke A, Lebel C, McDonald EC, Plummer Z, Cook TA, Zelhof AC. Pph13 and orthodenticle define a dual regulatory pathway for photoreceptor cell morphogenesis and function. Development. 2010; 137:2895–2904. [PubMed: 20667913]

Oda-Ishii I, Bertrand V, Matsuo I, Lemaire P, Saiga H. Making very similar embryos with divergent genomes: conservation of regulatory mechanisms of Otx between the ascidians Halocynthia roretzi and Ciona intestinalis. Development. 2005; 132:1663–1674. [PubMed: 15743880]

Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, MacIsaac KD, Rolfe PA, Conboy CM, Gifford DK, Fraenkel E. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. Nat Genet. 2007; 39:730–732. [PubMed: 17529977]

Ohno, S., editor. Evolution by Gene Duplication. New York: Springer-Verlag; 1970.

Ohshima K, Hattori M, Yada T, Gojobori T, Sakaki Y, Okada N. Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates. Genome Biol. 2003; 4:R74. [PubMed: 14611660]

Oliveri P, Tu Q, Davidson EH. Global regulatory logic for specification of an embryonic cell lineage. Proc Natl Acad Sci USA. 2008; 105:5955–5962. [PubMed: 18413610]

Ostertag EM, Kazazian HH Jr. Biology of mammalian L1 retrotransposons. Annu Rev Genet. 2001; 35:501–538. [PubMed: 11700292]

Parker HG, VonHoldt BM, Quignon P, Margulies EH, Shao S, Mosher DS, Spady TC, Elkahloun A, Cargill M, Jones PG, et al. An expressed fgf4 retrogene is associated with breed-defining chondrodysplasia in domestic dogs. Science. 2009; 325:995–998. [PubMed: 19608863]

Pavlopoulos A, Kontarakis Z, Liubicich DM, Serano JM, Akam M, Patel NH, Averof M. Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. Proc Natl Acad Sci USA. 2009; 106:13897–13902. [PubMed: 19666530]

Pearson JC, Juarez MT, Kim M, Drivenes O, McGinnis W. Multiple transcription factor codes activate epidermal wound-response genes in Drosophila. Proc Natl Acad Sci USA. 2009; 106:2224–2229. [PubMed: 19168633]

Pearson JC, Lemons D, McGinnis W. Modulating Hox gene functions during animal body patterning. Nat Rev Genet. 2005; 6:893–904. [PubMed: 16341070]

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, et al. In vivo enhancer analysis of human conserved non-coding sequences. Nature. 2006; 444:499–502. [PubMed: 17086198]

Peter IS, Davidson EH. Modularity and design principles in the sea urchin embryo gene regulatory network. FEBS Lett. 2009; 583:3948–3958. [PubMed: 19932099]

Pimanda JE, Ottersbach K, Knezevic K, Kinston S, Chan WY, Wilson NK, Landry JR, Wood AD, Kolb-Kokocinski A, Green AR, et al. Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. Proc Natl Acad Sci USA. 2007; 104:17692–17697. [PubMed: 17962413]

Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffery WR, Zon LI, Borowsky R, Tabin CJ. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat Genet. 2006; 38:107–111. [PubMed: 16341223]

Prud'homme B, Gompel N, Rokas A, Kassner VA, Williams TM, Yeh SD, True JR, Carroll SB. Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. Nature. 2006; 440:1050–1053. [PubMed: 16625197]

Ranade SS, Yang-Zhou D, Kong SW, McDonald EC, Cook TA, Pignoni F. Analysis of the Otd-dependent transcriptome supports the evolutionary conservation of CRX/OTX/OTD functions in flies and vertebrates. Dev Biol. 2008; 315:521–534. [PubMed: 18241855]

Rastegar S, Hess I, Dickmeis T, Nicod JC, Ertzer R, Hadzhiev Y, Thies WG, Scherer G, Strahle U. The words of the regulatory code are arranged in a variable manner in highly conserved enhancers. Dev Biol. 2008; 318:366–377. [PubMed: 18455719]

Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB. Stepwise modification of a modular enhancer underlies adaptation in a Drosophila population. Science. 2009; 326:1663–1667. [PubMed: 20019281]

Rokas A, Carroll SB. Bushes in the tree of life. PLoS Biol. 2006; 4:e352. [PubMed: 17105342]

Ruvkun G, Wightman B, Burglin T, Arasu P. Dominant gain-of-function mutations that lead to misregulation of the C. elegans heterochronic gene lin-14, and the evolutionary implications of dominant mutations in pattern-formation genes. Dev Suppl. 1991; 1:47–54. [PubMed: 1742500]

Seibert J, Urbach R. Role of en and novel interactions between msh, ind, and vnd in dorsoventral patterning of the Drosophila brain and ventral nerve cord. Dev Biol. 2010; 346:332–345. [PubMed: 20673828]

Seipel K, Schmid V. Evolution of striated muscle: jellyfish and the origin of triploblasty. Dev Biol. 2005; 282:14–26. [PubMed: 15936326]

Shapiro MD, Bell MA, Kingsley DM. Parallel genetic origins of pelvic reduction in vertebrates. Proc Natl Acad Sci USA. 2006; 103:13753–13758. [PubMed: 16945911]

Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005; 15:1034–1050. [PubMed: 16024819]

Simon J, Peifer M, Bender W, O'Connor M. Regulatory elements of the bithorax complex that control expression along the anterior-posterior axis. EMBO J. 1990; 9:3945–3956. [PubMed: 1979031]

Spitz F, Gonzalez F, Duboule D. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell. 2003; 113:405–417. [PubMed: 12732147]

Spitz F, Gonzalez F, Peichel C, Vogt TF, Duboule D, Zakany J. Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. Genes Dev. 2001; 15:2209–2214. [PubMed: 11544178]

Swaroop A, Kim D, Forrest D. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. Nat Rev Neurosci. 2010; 11:563–576. [PubMed: 20648062]

Swiers G, Chen YH, Johnson AD, Loose M. A conserved mechanism for vertebrate mesoderm specification in urodele amphibians and mammals. Dev Biol. 2010; 343:138–152. [PubMed: 20394741]

Szeto DP, Rodriguez-Esteban C, Ryan AK, O'Connell SM, Liu F, Kioussi C, Gleiberman AS, Izpisua-Belmonte JC, Rosenfeld MG. Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. Genes Dev. 1999; 13:484–494. [PubMed: 10049363]

Tarchini B, Duboule D. Control of Hoxd genes' collinearity during early limb development. Dev Cell. 2006; 10:93–103. [PubMed: 16399081]

Tessmar-Raible K, Raible F, Christodoulou F, Guy K, Rembold M, Hausen H, Arendt D. Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. Cell. 2007; 129:1389–1400. [PubMed: 17604726]

Ting SB, Caddy J, Hislop N, Wilanowski T, Auden A, Zhao LL, Ellis S, Kaur P, Uchida Y, Holleran WM, et al. A homolog of Drosophila grainy head is essential for epidermal integrity in mice. Science. 2005; 308:411–413. [PubMed: 15831758]

Tumpel S, Cambronero F, Ferretti E, Blasi F, Wiedemann LM, Krumlauf R. Expression of Hoxa2 in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon Hoxb1. Dev Biol. 2007; 302:646–660. [PubMed: 17113575]

Tumpel S, Wiedemann LM, Krumlauf R. Hox genes and segmentation of the vertebrate hindbrain. Curr Top Dev Biol. 2009; 88:103–137. [PubMed: 19651303]

Vavouri T, Walter K, Gilks WR, Lehner B, Elgar G. Parallel evolution of conserved non-coding elements that target a common set of developmental regulatory genes from worms to humans. Genome Biol. 2007; 8:R15. [PubMed: 17274809]

Vinagre T, Moncaut N, Carapuco M, Novoa A, Bom J, Mallo M. Evidence for a myotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. Dev Cell. 2010; 18:655–661. [PubMed: 20412779]

Walters J, Binkley E, Haygood R, Romano LA. Evolutionary analysis of the cis-regulatory region of the spicule matrix gene SM50 in strongylocentrotid sea urchins. Dev Biol. 2008; 315:567–578. [PubMed: 18262514]

Wang J, Lee AP, Kodzius R, Brenner S, Venkatesh B. Large number of ultraconserved elements were already present in the jawed vertebrate ancestor. Mol Biol Evol. 2009; 26:487–490. [PubMed: 19052148]

Weatherbee SD, Halder G, Kim J, Hudson A, Carroll S. Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the Drosophila haltere. Genes Dev. 1998; 12:1474–1482. [PubMed: 9585507]

Wellik DM. Hox genes and vertebrate axial pattern. Curr Top Dev Biol. 2009; 88:257–278. [PubMed: 19651308]

Wellik DM, Capecchi MR. Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. Science. 2003; 301:363–367. [PubMed: 12869760]

Woltering JM, Duboule D. The origin of digits: expression patterns versus regulatory mechanisms. Dev Cell. 2010; 18:526–532. [PubMed: 20412768]

Woltering JM, Vonk FJ, Muller H, Bardine N, Tuduce IL, de Bakker MA, Knochel W, Sirbu IO, Durston AJ, Richardson MK. Axial patterning in snakes and caecilians: evidence for an alternative interpretation of the Hox code. Dev Biol. 2009; 332:82–89. [PubMed: 19409887]

Yamamoto Y, Stock DW, Jeffery WR. Hedgehog signalling controls eye degeneration in blind cavefish. Nature. 2004; 431:844–847. [PubMed: 15483612]

Yang X, Zarinkamar N, Bao R, Friedrich M. Probing the Drosophila retinal determination gene network in Tribolium (I): The early retinal genes dachshund, eyes absent and sine oculis. Dev Biol. 2009; 333:202–214. [PubMed: 19324029]

Zakany J, Kmita M, Duboule D. A dual role for Hox genes in limb anterior-posterior asymmetry. Science. 2004; 304:1669–1672. [PubMed: 15192229]

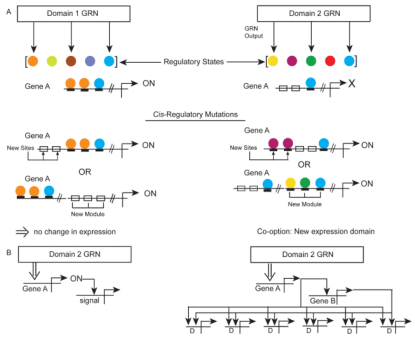

Fig. 1. *Cis*-regulatory mutations resulting in co-optive change in domain of expression of a regulatory gene, and examples of possible consequences at the GRN level

(A) Co-option event: The gene regulatory networks operating in spatial Domains 1 and 2 produce different regulatory states (colored balls, representing diverse transcription factors). A *cis*-regulatory module of Gene A, a regulatory gene, has target sites for factors present in the Domain 1 regulatory state and so Gene A and its downstream targets are expressed in Domain 1, but not in Domain 2 where only one of the three sites can be occupied. Two alternative types of *cis*-regulatory mutations are portrayed: appearance of new sites within the module by internal nucleotide sequence change; and transposition into the DNA near the gene of a module from elsewhere in the genome bearing new sites. While these gain of function changes do not affect the occupancy of the *cis*-regulatory sites of Gene A in Domain 1, the new sites allow Gene A to respond to the regulatory state of Domain 2, resulting in a co-optive change in expression so that Gene A is now active in Domain 2 (modified from Davidson and Erwin, 2010). **(B) Gain of function changes in Domain 2 GRN architecture caused by co-option of Gene A:** Gene A might control expression of an inductive signaling ligand, which could alter the fate/function of adjacent cells now receiving the signal from Domain 2 (left); Gene A might control expression of Gene B, another regulatory gene, and together with it cause expression of a differentiation (D) gene battery, which in consequence of the cooption is now expressed in Domain 2 (right).

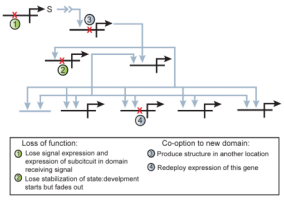

Fig. 2. Functional evolutionary consequences of *cis*-regulatory mutations depend on their location in GRN architecture

A GRN circuit encoding the control system of a differentiation gene battery (bottom tiers) activated in response to a signal (S) from adjacent cells (top tier); linkages shown in gray and genes in black. The double arrow indicates signal reception and transduction causing gene expression in the recipient cells. Note that the middle tier of circuitry consists of a dynamic feedback stabilization subcircuit. The numbered red “x” symbols denote mutational changes in the *cis*-regulatory modules controlling expression of these genes, keyed by number to the functional consequences listed in the box below. Loss of function mutations (1 and 2) are indicated in green, and co-optive gain of function mutations (3 and 4) resulting in expression of the affected gene in a new domain, as in Fig.1A, are indicated in gray (modified from Erwin and Davidson, 2009).

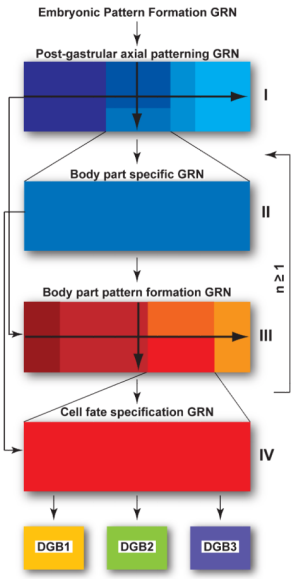

Fig. 3. Symbolic representation of hierarchy in developmental GRNs

The developmental process begins with the onset of embryogenesis at top. The outputs of the initial (i.e., pregastrular) embryonic GRNs ("Embryonic pattern formation GRN") are used after gastrulation to set up the GRNs which establish regulatory states throughout the embryo, organized spatially with respect to the embryonic axes (axial organization and spatial subdivision are symbolized by orthogonal arrows and colored patterns). These spatial domains divide the embryonic space into broad domains occupied by pluripotential cell populations already specified as mesoderm, endoderm, future brain, future axial neuroectoderm, non-neural ectoderm, etc. The GRNs establishing this initial mosaic of postgastrular regulatory states, including the signaling interactions that help to establish domain boundaries, are symbolized as **Box I**. Within Box I domains the progenitor fields for the future adult body parts are later demarcated by signals plus local regulatory spatial information formulated in Box I, and given regulatory states are established in each such field by the earliest body part specific GRNs. Many such progenitor fields are thus set up during postgastrular embryogenesis, and a GRN defining one of these is here symbolized as **Box II**. Each progenitor field is then divided up into the subparts that will together constitute the body part, where the subdivisions are initially defined by installation of unique GRNs producing unique regulatory states. These "sub-body part" GRNs are symbolized by the oriented patterns of **Box III.** Since some body parts are ultimately of great complexity, the process of patterned subdivision and installation of successively more confined GRNs may be iterated, like a "do-loop", symbolized here by the upwards arrow from Box III to Box II, labeled n≥1. Towards the termination of the developmental process in each region of the late embryo, the GRNs specifying the several individual cell types and deployed in each subpart of each body part, are symbolized here as **Box IV.** Post-embryonic generation of specific cell types (from stem cells) is a Box IV process as well. At the bottom of the diagram are indicated several differentiation gene batteries ("DGB1,2,3"), the final outputs of each cell type. Morphogenetic functions are also programmed in each cell type (not shown). For discussion and background see text and Davidson, 2001; 2006.

Table 1

Evolutionary alterations in cis-regulatory modules and their possible functional consequences (if any)

Effect of change at sequence level	LOF	Quantitative output change	Input gain/loss within GRN	GOF; Cooptive redeployment to new GRN
(Int.): Appearance of new target site(s)	X	X	X	X
(Int.): Loss of old target site(s)	X	X	X	X
(Int.): Change in site no.		X		
(Int.): Change in site spacing	X	X	X	
(Int.): Change in site arrangement	X	X	X	
(Cxt.): translocation of module to new gene	X		X	X
(Cxt.): Module deletion	X			
(Cxt.): New tethering function	X		X	X
(Cxt.): Duplication, subfunctionalization				X

Int., internal change in *cis*-regulatory module sequence; Cxt., change affecting genomic context of *cis*-regulatory