ELSEVIER

Brain Research Bulletin 66 (2005) 495–509

BRAIN RESEARCH BULLETIN

www.elsevier.com/locate/brainresbull

Recasting developmental evolution in terms of genetic pathway and network evolution . . . and the implications for comparative biology

Adam S. Wilkins *

BioEssays Editorial Office, 10/11 Tredgold Lane, Napier St., Cambridge CB1 1HN, UK

Available online 18 April 2005

Abstract

The morphological features of complex organisms are the outcomes of developmental processes. Developmental processes, in turn, reflect the genetic networks that underlie them. Differences in morphology must ultimately, therefore, reflect differences in the underlying genetic networks. A mutation that affects a developmental process does so by affecting either a gene whose product acts as an upstream controlling element, an intermediary connecting link, or as a downstream output of the network that governs the trait's development. Although the immense diversity of gene networks in the animal and plant kingdoms would seem to preclude any general "rules" of network evolution, the material discussed here suggests that the patterns of genetic pathway and network evolution actually fall into a number of discrete modes. The potential utility of this conceptual framework in reconstructing instances of developmental evolution and for comparative neurobiology will be discussed.

Keywords: Genetic pathways; Genetic networks; Evolution; Development; Homology

1. Introduction

Traditional NeoDarwinian evolutionary biology envisages the genetic basis of the evolution of morphological features as a process of sequential selection events, often involving mutations of minute phenotypic effect, operating via one allele fixation event at a time [24,55,66]. It is essentially a model of evolution through independent and additive genetic effects. While the existence of genetic *interactive* effects has long been recognised in evolutionary genetics, such effects tend to be treated as second-order complications rather than as reflections of a ubiquitous genetic phenomenon that bears fundamental implications for the NeoDarwinian perspective.

The basic genetic model at the heart of NeoDarwinism has deep historical roots. It was formulated over a span of years extending from the late 1920s to the early 1950s [56]. Although it was fully concordant with genetic knowledge of its period, little was then known about genes or how they achieved their effects. In the past 50 years, however, the understanding of genes and gene action has deepened immeasurably. In addition to the revelations of the Watson-Crick model and all that was discovered about gene action in the 1960s, it is now understood that complex and dynamic networks of gene activity underlie developmental processes which, in turn, generate the observable morphological features of organisms [14]. Nevertheless, this insight and the various discoveries relating to network structure have, as yet, hardly touched mainstream evolutionary population genetics [79].

It might be assumed that this conceptual gap between the two fields reflects the sheer novelty of the genetic network concept and the inevitable lag that would precede its absorption into the mainstream evolutionary paradigm. This cannot, however, be the full explanation: the idea that development, hence morphology, is underlain by complex webs of genetic interaction is far from new. It first found expression in a model of cellular differentiation proposed by the originators of the *lac* operon model, employing the then-known principles of bacterial gene regulation [58]. By the late 1960s, gene network concepts were being elaborated either in more general and abstract form [46] or with explicit

* Tel.: +44 1223 355 572; fax: +44 1223 359 761.
E-mail address: awilkins@bioessays.demon.co.uk.

0361-9230/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.brainresbull.2005.04.001

reference to the growing body of knowledge about eukaryotic gene regulation and genome structure [8,9]. Yet, while the general concept of gene networks was widely accepted from the late 1960s onwards, it had little effect on research programmes or on further thinking within developmental biology, let alone evolutionary biology, for more than three decades.

In recent years, however, the importance of networks has received much wider recognition from both the molecular and developmental biology research communities. The principal reason for this change is the growing armament of technical advances that make possible the detailed characterisation of actual networks that underlie a host of cellular and developmental properties. The most thoroughly characterised have been those that underpin basic cellular properties, such as the networks that structure cellular metabolism [43,90]. In addition to these, however, several of the gene networks crucial to specific developmental processes have been described. The first of these was the segmental patterning gene network of the fruit fly *Drosophila melanogaster* (reviewed in [41,69]). Another was the network that gives rise to the development of the vulva in the nematode *Caenorhabditis elegans* [12]. These networks were elucidated primarily through the classical methods of developmental genetics, aided and abetted by molecular methods. More recently, however, other genetic networks for development have been characterized, using a variety of methods but in which molecular strategies have predominated. Some relatively well characterized examples are the networks that govern endomesoderm development in the sea urchin embryo [15], mammalian sex determination [49], and tooth development [84].

The increasing sophistication of network analysis is shown, however, by the elucidation of what might be termed "meta-networks", namely the complete proteome interaction maps of yeast [44], *Drosophila* [26] and of *C. elegans* [52]. These provide the first draft charts of the *total* set of protein interactions, both actual and potential, that take place in these organisms. In a sense, they delineate what might be called the *total network space* of the organisms rather than specific networks that govern particular phenotypic properties. Several of these investigations have, in turn, catalysed some key conceptual advances via graph-theoretic interpretations of network structures (reviewed in [3]). A general conclusion of this work is that many networks show what has been termed a "scale-free" property: the number of connections per intersection or "node" follows a power law distribution for the sum total of connections within the network [2]. When graphed, the striking visual feature of such networks is that they show a few visibly highly connected nodes, so called "hubs", when the value of the negative exponent of the degree of linkage, k, lies between 2 and 3 [3]. Such hubs are not seen either with so-called regular networks (which, by definition, have the same numbers of connections per node) or random networks (whose numerical distribution of degrees of connectivity follows a Poisson distribution) [3,85]. Closer analysis of the structure of several large scale-free networks reveals that they have a modularized, hierarchical structure, which has been superimposed on the generic scale-free property [3].

Beyond the experimental characterisation of specific networks and the theoretical explorations of generic network features, however, a large relatively unexplored area exists: an understanding of the evolutionary changes that have occurred in actual networks. This deficiency is a serious one. If the evolution of morphologies reflects the evolution of developmental processes [11,31] and if each developmental process mirrors the expression of its underlying network, then comprehending the evolution of organisms requires understanding the evolution of their genetic networks [14,92].

Several factors have contributed to this explanatory gap. In the first place, it is intrinsically impossible for the theoretical treatments to fill it. Dealing with generic properties, these approaches are necessarily restricted to providing general pictures rather than specific portraits. Furthermore, the theoretical treatments of network evolution have tended to concentrate on the *internal nodes* of networks and how they become ever more connected over time [2,3,89]. In contrast, in the operation of a biological network, what matters for the organism is the specific set of outputs and how those outputs are triggered in response to a specific set of inputs. It is, after all, the precise spatial and temporal regulation of the particular output activities that determine the biological effect. Ultimately, therefore, to understand the patterns of network evolution in the real world of living things, you have to examine actual organisms and do comparative studies of their genetic networks.

Such comparative analysis, however, presents a formidable challenge. Characterisation of even a single developmental network in even one organism requires a small army of research workers (see, for instance, the list of authors in [15]). Evolutionary insight into the formation of such a network, however, requires comparable analysis of the network of, at least, one related organism and that of an outgroup organism. The amount of work is a direct function of the number of species under comparison and the resulting interpretative constructs, inevitably, will have gaps of unknown extent. Furthermore, the evolutionary interpretations of such comparative work are encumbered by the uncertainties inherent in any phylogenetic reconstruction based solely on comparative studies of living species [34]. In effect, the uncertainties of interpretation will compound as a function, n^x, of the number of species, n, with $x > 1$.

Even if one sets aside such complications, the practical difficulties of fully characterising even one network create pressure for a simplification of such comparative work. The great majority of comparative studies have focussed on changes of employment of single or a few regulatory genes, the phenomenon termed "gene co-option" [64,87] or "gene recruitment" [92]. These studies have proven valuable and

informative but since gene recruitment involves modification of pre-existing networks by addition of a single new functional link, they tend to be the equivalent of a narrow beam of light focussed on only one part of a darkened landscape.

Despite the formidable difficulties in reconstructing the evolution of networks, some useful comparative genetic information is now available for a handful of situations. In addition, the findings of molecular developmental genetics lead to some fairly obvious suggestions of how such networks might evolve step-by-step. By putting these two sources of information together, one can make a start toward assessing the patterns of evolutionary change in networks. And that is the aim of this paper: to provide a provisional systematisation of these patterns. Such categorisation can provide a framework that may be useful in formulating hypotheses of network change in cases where comparative data are sparse.

A starting point for this discussion is that, in principle, genetic networks in development can be visualized as sets of parallel linear pathways connected by links [30]; the connection points are the nodes. Such a conceptual reduction of network structure suggests an approach to analysing network evolution. It involves breaking down the problem of network evolution into three constituent parts: (1) the evolution of genetic pathways; (2) the ways that connecting links can form between pathways to form simple network connections; and (3) the additional events that can generate multi-linked nodes within networks. The main part of this article will be an exploration of network evolution using this approach. The final part will attempt to put this subject within a larger perspective, arguing the importance of understanding network evolution for analysing the issues of comparative neurobiological studies that have been the focus of this meeting.

In the discussion that ensues, the focus will be kept on the patterns of connectivity of gene activities, with relatively little attention paid to the diverse nature of the kinds of molecules and molecular interactions that can occur in networks. For example, the increasing recognition of the importance of non-coding regulatory RNAs (ncRNAs) [54] is certain to have a major impact on thinking about network compositions and operations. Yet, for the purposes of this discussion, and with only one exception (the *Drosophila* sex determination pathway), the realm of molecular detail will be set aside; the general ideas sketched here are applicable irrespective of the molecular details in specific cases. Similarly, the whole quantitative dimension of signalling and interaction within networks will not be explored here though quantitative aspects can determine whether or not a functional link is made or not [76]. Finally, the earlier and influential dichotomous distinction between evolution based on coding sequences versus regulatory changes [48] will also be ignored. It is increasingly apparent, after all, that a large proportion of genes encode "regulatory" functions in some capacity or other, including the enormous number of signalling pathway components. Hence, it follows that many protein-coding sequences are regulatory in nature and that, accordingly, mutations in these sequences have direct regulatory consequences.

2. A first step: charting the evolution of linear segments of genetic architecture, namely genetic pathways

2.1. General considerations

Genetic pathways are the simplest form of "genetic architecture", namely the diverse functional patterns of connectivity between different genes and gene products. A genetic pathway can be defined specifically as a linear sequence of gene activities, each one affecting or making possible the sequence of the next. In all pathways, it is conventional to refer to the first or early steps as "upstream" and the final or later steps as "downstream" ones.

The first genetic pathways to be characterised were those underlying biochemical and metabolic sequences (see [92], pp. 99–108). The delineation of pathways of biochemical change began in the early 1900s, and constituted the major programme of activity in biochemistry from the 1930s through the 1960s. Today, it remains an important, though perhaps less central, component of research in biochemistry. The cardinal characteristics of biochemical pathways are that they involve sequences of conversions of successive substrates – the product of one reaction becomes the substrate of the next. In these conversions, the gene products themselves (the enzymes) frequently do not interact (though, in some pathways, they form multi-enzyme complexes). In these pathways, each step is essential. If one blocks an early, or "upstream", step by any means (biochemical or genetic), the pathway soon ceases to produce new product.

The genetic pathways that underlie segments of developmental processes differ from such metabolic pathways in two fundamental respects. Often, they involve sequences of direct molecular interactions between gene products or between segments of genes and the immediate upstream gene products. One consequence is that they can be represented differently from biochemical pathways. For metabolic pathways, depiction should include both the sequence of substrates and the names of the enzymes responsible for each step (usually written over the arrows indicating the conversion step). In genetic pathways for development, however, it is often sufficient to simply denote the genes (or their gene products) connected by arrows since it is the sequence of gene product interactions (or interactions of gene products with gene regulatory sequences) that constitute the pathway of events. This schematisation will be the convention adopted in this article. A second consequence of the directness of gene interactions that characterises genetic pathways for development is that it is easier to uncouple upstream from downstream events. Mutations that activate downstream events, independently of the occurrence of upstream events, can occur, with the result that upstream events are often intrinsically less crucial than downstream ones, relative to biochemical pathways. In effect, downstream events can be uncoupled by mutational events from upstream events, which is generally not possible in biochemical pathways.

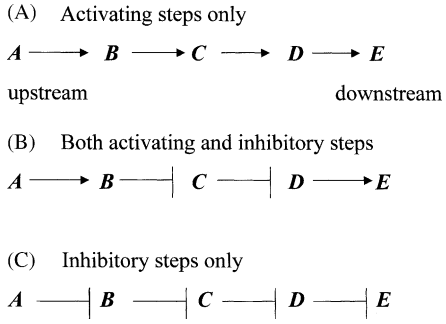


Fig. 1. Three kinds of genetic pathway. (A) A pathway consisting solely of positive-control (activating) steps. (B) A pathway consisting of a mix of positive-control and negative-control (inhibitory) steps. (C) A pathway composed solely of negative control steps.

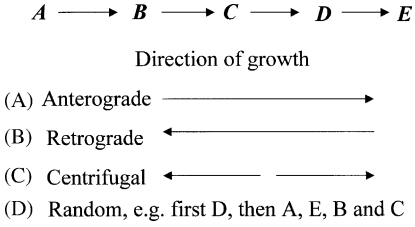


Fig. 2. Four possible patterns of step-by-step pathway evolution, relative to the direction of pathway operation. (A) Anterograde growth: the pathway evolved by addition of genes in the same temporal order as shown in its mode of operation. (B) Retrograde evolution: the pathway evolved in the reverse order to its mode of operation today. (C) Centrifugal: the pathway evolved by addition of elements in both directions, proceeding from an element that is internal in the contemporary pathway. (D) Random: there was no temporal correspondence between the addition of elements and the structure of the present-day pathway.

A diagram of three generic sorts of genetic pathway for developmental processes is given in Fig. 1; they differ in terms of their mix of positive or activating steps (shown with arrows) and inhibitory steps (shown with bars). The first and simplest kind is a sequence of activations (Fig. 1A). A second, slightly more complicated kind involves a mix of positive and negative signals (Fig. 1B). The third and final kind, which is undoubtedly the least common, is a sequence consisting solely of inhibitory steps (Fig. 1C).

All pathways, regardless of their structure, must be themselves products of evolution; hence, every pathway has an evolutionary history. Furthermore, each must have arisen in step-wise fashion since any pathway consisting of three or more elements is too complicated to have arisen in a single mutational event. Given this consideration, one can conceive of four general patterns of origination of a linear, causal sequence of gene activities (Fig. 2). The first pattern is that of evolution by step-wise growth from upstream to downstream. Such a pattern would mirror the present-day sequence of activities of the pathway and could be termed forward or "anterograde evolution". The second possible pattern is the reverse: growth from the downstream-most upwards, or "retrograde evolution". The third kind of pattern that can be imagined would be growth from somewhere in the interior of the pathway outwards toward both upstream and downstream; such a pattern might be termed "centrifugal evolution". The fourth conceivable pattern is one in which the sequence of evolutionary additions to the pathways bears no systematic relationship to the structure of the pathway as it exists today; such a pattern of non-regular growth of the pathway can be termed "random".

To determine which evolutionary mode seems most probable, one must start with a known genetic pathway, examine its structure and then either make deductions from its structure as to which explanation seems most probable or, preferably, use comparative data from different organisms to decide the issue. An unexpected difficulty is the relative scarcity of true (linear) pathways. Many initial descriptions of the genetic basis of developmental processes are framed as pathway interpretations but then morph into networks as further details of the actual genetic architecture come to light [53,92].

Nevertheless, a handful of biological systems appear to follow a true linear, that is pathway, organisation. Of these, the best characterised appear to be the sex determination pathways of the nematode *C. elegans* and of the fruit fly *D. melanogaster*, whose initial descriptions of these pathways were both produced in 1980. Given their apparent unrelatedness at the organisational and compositional levels [39,40], they should furnish independent test cases for genetic pathway evolution.

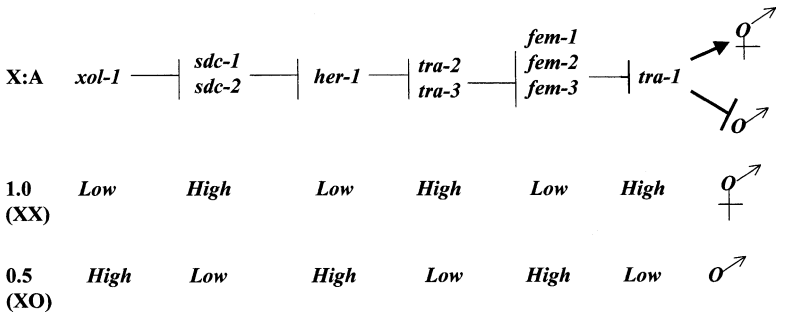


Fig. 3. A schematic of the sex determination pathway of the nematode *Caenorhabditis elegans*. The sequence of gene steps is shown at the top; the two sequences of activity in the hermaphrodite- (female) and male-determining pathways are shown at the bottom. Relative activities are shown as "High" and "Low" though it is possible that the latter corresponds essentially to zero activity.

2.2. The case for retrograde addition as a major mode of pathway evolution

From its structure, the *C. elegans* pathway looks the more unusual of the two sex determination pathways: it consists of a negative series of steps [36], as schematized in Fig. 1C. It is depicted in Fig. 3 along with the pattern of gene activities seen in female and male development (see [50] for a review of the molecular biology). The striking feature of a pathway that consists of a sequence of inhibitory steps is that the pattern of gene expression events is a series of alternating high and low activities. If an activity is high, the immediately downstream gene activity that it regulates will necessarily be low. Males and females have reciprocal patterns of high and low activity because the first gene activity in the pathway, *xol-1* (XO-lethal) is differentially regulated by the difference between the two sexes in the ratio of X chromosomes to autosome sets (the X:A ratio). Females, possessing two X chromosomes and therefore an X:A ratio of 1, repress *xol-1* activity, while males with only a single X, and an X:A ratio of 0.5 do not repress this gene's expression and therefore have high *xol-1* activity. That initial difference is relayed through an alternating series of high and low activities to the final difference between the two sexes, that of *tra-1*, with females having high *tra-1* activity and males low *tra-1*.

That specific difference in gene activity is the crucial one: it is the final downstream gene activity difference, for *tra-1*, that triggers the onset of one pathway of sexual differentiation or the other (reviewed in [38]). This was shown in an elegant experiment by Jonathan Hodgkin, in which the sex determination switch was made to "run off" an allelic difference in *tra-1*, using females heterozygous for a hypermorphic *tra-1* allele and a null *tra-1* allele and XX phenotypic males homozygous for the null [37]. This experiment strongly suggests that the entire functional *raison d'etre* of the pathway is to create a single gene activity difference, namely the two states of *tra-1* activity. In the light of that conclusion, however, the length and structure of the pathway presents a paradox: it is far more complex than it needs to be to achieve its simple function of controlling *tra-1* activity [91]. It should, therefore, not be seen as some sort of economical "design" engineered by evolution for economy and efficiency but rather as the product of a much messier process of evolutionary tinkering or "bricolage" [16,42]. In this process, an entity considerably more elaborate than one dictated by considerations of economy and efficiency is generated by evolutionary processes.

From that perspective, a simple, if incomplete, explanation of the evolution of the pathway suggests itself. The idea begins with the proposition that the earliest ancestral form of the pathway was much simpler, conceivably just *tra-1* regulated by a simple switch or an allelic difference, as in the Hodgkin (1983) [37] experiment. If that were the case, then the evolution of the pathway would have consisted of the sequential selection for and addition of inhibitory steps, moving upstream at each step [91]. Such a step-wise construction, proceeding in the reverse direction to that in which the present day pathway operates, is that of retrograde addition (Fig. 2B). Although such an explanation is incomplete because it neglects the nature of the selective forces that might have driven the process, it accounts for the structure and apparently unnecessary complexity of the pathway. It also comports with the common sense idea – though such ideas, admittedly, are not always the most reliable base for inference – that since the "business end" of a pathway is its downstream output, evolution would surely have selected this most essential part of a pathway before elaborating its regulatory superstructure.

A hypothesis only has worth, of course, if it can be tested. This particular hypothesis, in fact, makes a testable prediction. If the *C. elegans* pathway grew by successive recruitment events of new upstream gene activities, starting from *tra-1*, and if this involved selection for inhibition at each step, then, in principle, different patterns of inhibitory gene recruitment could have taken place in different lineages. The prediction follows from the fact that for any gene activity, there are numerous ways, and numerous other gene products, that can inhibit its gene activity. A gene activity, after all, can be inhibited at any multiple levels: transcriptional, RNA splicing, export of the message from the nucleus, translational steps, mRNA degradation controls or post-translational modification of products. At each level, many gene products should be able to carry out such an inhibition. Thus, if the selectional pressures are simply for recruitment of inhibitory activity at each step, it is probable that, in different organismal lineages, different gene activities will have been brought into play. The prediction, therefore, for the model of retrograde addition is that of preferential functional conservation of downstream elements relative to upstream ones. In the particular case of the *C. elegans* sex determination pathway, the prediction is that if one surveys the nematodes, one should find widespread usage (functional conservation) of *tra-1* but differences in composition of upstream regulators amongst the different nematode lineages. The extent of divergence of composition of the upstream regulators should be roughly proportional to evolutionary time and phylogenetic divergence.

Unfortunately, phylogenetically wide-ranging comparative studies of nematode sex determination have not yet been carried out, hence the specific case for which the hypothesis of retrograde addition was proposed [91] remains untested. Yet, the prediction that retrograde addition should be reflected in preferential functional conservation of downstream elements in any set of related pathways should apply generally, including those pathways involving solely positive elements (Fig. 2A) or a mix of positive and negative elements (Fig. 2B). The reasoning is similar to that employed above: just as there are numerous ways to inhibit any particular gene activity, there are usually multiple ways to boost a gene activity. Hence, pathways growing upwards by means of addition of activation steps should also have the potential to grow in various ways, with addition of different gene activities. Downstream regulatory gene activities would be expected to be those showing greatest functional conservation because those downstream activities are the closest to the cell dif-

ferentiation functions that are the "output" of the pathway, hence its ultimate biological function.

Fortunately, comparative studies of sex determination in both vertebrates and insects have been carried out and permit a test of the general prediction of the hypothesis of retrograde addition. Take the vertebrate pathway first. In eutherian mammals, the reference pathway deduced from genetic and molecular analysis of mice and humans is (reviewed in [92], pp. 186–187):

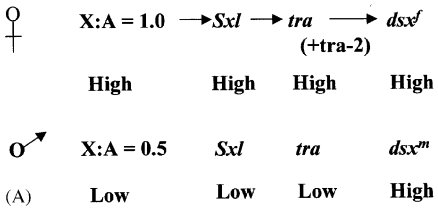
$$Sry \dashv Wnt4 \rightarrow Dax1 \dashv Sf1 \rightarrow (Amh, Sox9, DMRT\ genes)$$

In this pathway, the upstream gene, *Sry*, is the key Y chromosome gene that initiates the cascade and produces maleness in XY offspring [28,81]. If an egg is fertilized not by a Y-bearing sperm but by an X-bearing one, the resulting zygote will lack an *Sry* gene and the resulting zygote will be set on the pathway of female development. The critical downstream genes in the pathway (*Amh*, *Sox9*, and one-to-three genes of the *Dmrt* gene family) set male development in train; the pathway diagrammed above ensures that they are "on" if *Sry* is present and "off" if *Sry* activity is absent.

How much, and which parts, of this pathway are shared by other vertebrates? The answer is that the downstream male-determining genes are widely shared amongst the tetrapod vertebrates – and perhaps fish – and *Dax1* is perhaps similarly utilized [80]. In contrast, *Sry* is utilized only in the eutherian and metatherian mammals but not the monotremes [27] and not even in all the eutherian mammals [45]. This general phylogenetic pattern of difference in upstream elements in combination with shared downstream elements is as predicted by the retrograde addition model.

The second set of sex determination pathways for which there are comparative data are those of insects. Here, the reference pathway is the major sex determination pathway of the fruit fly, *D. melanogaster*, which governs the development of the secondary sexual traits of this animal and which has been characterised in exquisite detail (see review by Schutt and Nöthiger (2000) [70]). The pathway sequence is schematised in Fig. 4A while its actual molecular details are summarised in Fig. 4B. In contrast to the *C. elegans* pathway, this pathway consists of a sequence of activation steps, which take place in the female embryos, and a corresponding set of default steps in the males (which occurs in the absence of that sequence of activation steps).

The series of female-specific events that constitutes the pathway begins with a highly specific transcriptional activation step, which takes place at a particular promoter of the first gene in the pathway, *Sex lethal* (*Sxl*). The defining feature of this pathway, however, is that its main sequence consists of a regulated sequence of sex-specific differential alternative RNA splicing steps. In contrast to the molecularly heterogeneous nature of the events in the *C. elegans* pathway [50], the *Drosophila* sex determination pathway is essentially an RNA splicing cascade. In the female, the key events are the splicing out of sections of coding sequence containing stop codons from the transcripts of both the *Sxl* and *tra* genes. The result is that a functional TRA protein, in combination with the constitutively expressed gene product of the gene *tra-2*, carries out the splicing of the downstream-most gene, *doublesex* (*dsx*) to produce a female-specific form of the DSX transcription factor, DSX-F. The latter activates female-specific genes and represses male-specific genes. In males, in contrast, the upstream splicing events fail to occur, with the result that the stop codons of *Sxl* and *tra* are retained in the transcripts and the further consequence that only highly truncated, inactive fragments of SXL and TRA are produced. The consequence is that splicing of the *dsx* transcript takes place by the "default" mode, to give the male-specific transcription factor, DSX-M.



Events in female embryos

1. High X:A ratio (sis loci activity) activates early promoter, Pe, of *Sxl*
2. Early burst of *Sxl* expression from Pe allows splicing, small amount of active SXL
3. Spontaneous activation of maintenance promoter, Pm, of *Sxl*; SXL allows splicing out of exon 3, with its stop codon. Result: More SXL and positive feed-back loop
4. SXL blocks 5' splice site of exon 2 of *tra*, forces alternative splicing of exon 2, elimination of portion with stop codon. Result: Active TRA protein
5. Active TRA, with constitutive TRA2 protein, splices *dsx* to give DSX-F protein. Result: activation of female-specific genes, repression of male genes

Events in male embryos

1. Low X:A ratio means insufficient SIS products to activate Pe of *Sxl* gene
2. Spontaneous activation of Pm of *Sxl*, with exon 3, containing stop codon. Result: production of truncated, inactive SXL
3. In absence of active SXL, *tra* transcribed with complete exon 2, containing stop codon. Result: inactive, truncated TRA protein produced
4. In absence of active TRA, default splicing of *dsx* takes place, to produce DSX-M protein. Result: activation of male-specific genes, repression of male genes

(B)

Fig. 4. (A) A schematic of the sex determination pathway of the fruit fly, *Drosophila melanogaster*, showing the differences in gene activity for the three key genes (Sex lethal, transformer and double sex), for female- and male-determination. See text for description. (B) The divergent molecular events in the alternate pathway sequences of events for female and male embryos of *Drosophila*. A full description of these details can be found in Schutt and Nöthiger (2000) [70].

What do comparative studies reveal about the evolution of this pathway? There is no equivalent mutational analysis of sex determination systems in other insects but there is an alternative, molecular method for making comparisons: one looks for sex-specific alternative splicing of the upstream and downstream genes, *Sxl* and *dsx*, respectively. A variety of insects have been examined in this way in recent years and the answer seems clear (Table 1). Only in the drosophilids is *Sxl* employed as a sex determining gene while in all the species examined, *dsx* is used as the downstream control gene,

Table 1
Sex specific splicing of *Sxl* and *dsx* in insects

Genus or species/order	*Sxl*	*dsx*	Reference
Drosophila (Diptera)	Yes	Yes	[7]
Ceratitis capitata (Diptera)	No	Yes	[68]
Musca domestica (Diptera)	No	Yes	[35,57]
Megaselia scalaris (Diptera)	No	Yes	[77]
Chrysomya rufifacies (Diptera)	No	n.d.	[59]
Batrocera tryoni (Diptera)	n.d.	Yes	[75]
Bombyx mori (Lepidoptera)	n.d.	Yes	[86]
Apis mellifera (Hymenoptera)	n.d.	Yes	Cited in [4]

n.d.: not determined.

with recognisable DSX-F and DSX-M forms, similar to their *Drosophila* counterparts.

The unavoidable implication is that the *Drosophila* pathway evolved by some form of retrograde addition, in which *Sxl* (and presumably its specific X chromosome activators) was recruited to an ancestral pathway that already contained *dsx* and which utilized sex-specific alternative splicing of *dsx*. Yet, this conclusion itself raises a formidable difficulty: the structure of the pathway precludes the kind of relatively simple successive recruitment of upstream genes envisaged for the *C. elegans* pathway, discussed above. The alternative splicing of the *dsx* primary transcript to yield the dsx^f mRNA product takes place *only* if a highly precise pattern of molecular interactions, starting with the regulated transcriptional start of *Sxl*, has preceded that final splicing event (Fig. 4B).

How can one reconcile the comparative evidence, which indicates that *Sxl* was recruited specifically in the drosophilids with the requisite orchestrated sequence of molecular interactions leading to correct sex-specific splicing of *dsx*, which only occurs in these species if *Sxl* is present? Pomiankowski et al. (2004) [62] have proposed a scheme of pathway evolution that can, in principle, account for these seemingly contradictory facts. It depends on the fixation within the lineage of a sequence of mutations, most of which have the property of favouring fidelity of signal for one sex while reducing fidelity of signal for sexual development for the other. In this hypothetical scheme, the putative ancestral state involved a segregating allelic difference at *dsx* (much as in the *C. elegans* experiment of Hodgkin (1983) [37]) but with control passed successively upwards along the pathway in a retrograde direction. Thus, control, in this hypothesis, passes first from *dsx* to *tra*, then to *Sxl* and, finally, to the X chromosome regulators of *Sxl* [62]. This constitutes another form of retrograde pathway evolution but one that is more complicated than a sequence of gene additions. It involves one such addition (*Sxl*) but the over-all character is that of a sequential fixation of mutations in the reverse direction to that of the molecular interactions in the pathway.

Finally, there is the possibility that *all* metazoan sex determination pathways evolved in retrograde fashion from a gene related to *dsx*, which is a member of the so called DMRT gene family (mentioned above in connection with mammalian sex determination) [65,94]. In *C. elegans*, the *dmrt* gene is called *mab-3*; it is switched on in male embryos, which do not have *tra-1* activity, and is required for male development [73,65]. The acronym DMRT stands for *dsx mab-3* related transcription factors. In light of the discovery of shared DMRT genes in sex determination pathways, the initial conclusion that different animals have completely different genetic pathways of sex determination [39,40] has to be reclassified as a premature generalization. It was based on insufficient knowledge at the time of the molecular structure of *mab-3*.

So far, we have concentrated on the pattern of evolution of sex determination pathways. It can be argued, however, that sex determination pathways are special in their evolutionary history and that their properties may have little relevance to the evolution of other kinds of pathways. Does retrograde pathway (or network) evolution, in fact, take place outside of the arena of sex determination? The evidence is fragmentary but suggestive and is summarized in Table 2. The findings (see listed references for details) suggest that in processes as diverse as segmental patterning, left–right patterning in chordates, formation of dopaminergic neurons and metazoan germ line development, as well as sex determination, there is preferential functional conservation of downstream elements. This, in turn, suggests that those elements provided the foundations of their respective pathways. A similar conclusion, reached by a different route, is the idea that pathways for certain sensory capabilities may have originated with the primordial basic cell differentiative capacities, with complex regulatory superstructures subsequently added during evolution ([14], Chapter 9; [21]). The general conclusion is that, both from evidence and argument, retrograde pathway evolution appears to be a common and even fairly general mode of pathway evolution.

Table 2
Preferential conservation of downstream functions

Pathway/network	Phylogenetic group	Conserved function(s)	Reference
Sex determination	Insects	*Dsx*	Table 1, this paper
Sex determination	Vertebrates	*DMRT/Sox9/SF1/Amh*	[80]
Segmental patterning	Arthropods	*engrailed* (*en*)	([92], Chapter 7)
Left–right patterning	Vertebrates	*nodal*, *PitX2*	[22]
Dopaminergic CNS neurons	Vertebrates	Dopaminergic biochemistry	P. Vernier, this issue
Germ line development	Metazoa	*vasa*, germ cell cytology	[19]

2.3. But not all pathway evolution involves retrograde addition

Yet, while the evidence in favor of retrograde addition as a pattern of genetic pathway evolution is accumulating, there are pieces of circumstantial evidence and some persuasive general reasons for thinking that it is not the sole mode of pathway evolution.

One piece of evidence comes from the *Drosophila* sex determination system itself. The pathway described above, shown in Fig. 4A, which determines the externally visible secondary sexual characteristics, is often referred to as "the" sex determination pathway of the fruit fly. Analysis, however, suggests that the fruit fly also possesses three variant pathways, which govern more specialized sexually dimorphic features (see [70] for review). These involve: the appearance of a particular, male-specific, abdominal muscle, induced by a neural signal; the system of dosage compensation, which activates the single-X of the male to produce the same over-all gene activity as the two X's of the female, and; the development of the female germ line. They are outlined schematically in Fig. 5. What all these pathways have in common is the involvement of *Sxl*. Two of them, however, have different downstream targets while the last (the germline pathway) has both novel regulators upstream of *Sxl* and new downstream target genes.

Interpreting these variant pathways in evolutionary terms requires a hypothesis as to which form of the pathway came first. If the evolutionary hypothesis proposed by Pomiankowski et al. [62] is substantially correct, then the main sex determination pathway, which governs secondary sexual characteristics, was the initially evolved form. Its creation involved recruitment of *Sxl* to bind the *tra* transcript and redirect its splicing, and the subsequent evolution of modifiers on the *Sxl*-bearing chromosome, as that chromosome evolved into the X. From this perspective, the other *Sxl*-employing pathways shown in Fig. 5B–D evolved subsequently. If so, then the variant CNS pathway (Fig. 5B) evolved through substitution and replacement of the downstream target gene, *dsx*, by other gene products (those of *fru* and *dsf*). In contrast, the dosage compensation pathway (Fig. 5C) could have evolved, with even greater simplicity, simply by directly utilizing the ancestral biochemical function of *Sxl*, its RNA-binding capacity [47] to repress expression of one of the key (male-specific) dosage compensation activation genes (*msl-2*). (When *msl-2* is not active, the whole dosage compensation pathway in males, which serves to activate the single X to activity levels achieved by two X chromosomes in the females, shuts down.) In principle, all that would have been required for the evolution of this variant pathway was the acquisition and fixation of a mutation, in either *Sxl* or *msl-2* to promote the binding of SXL protein to *msl-2* transcript. Finally, the evolution of the oogenesis pathway would have had to involve a minimum of two substitution-replacement events, one upstream and one downstream. The occurrence of an upstream substitution-replacement event is also indicated in the case of the two vole species that lack *Sry* [45]. There is, however, a second possibility for the evolution of the *Drosophila* oogenesis pathway: the independent recruitment of *Sxl* to a pre-existing germ line pathway. The latter may seem unlikely but cannot be dismissed. A relatively newly recruited gene, which is now expressed in a tissue or cell type that it had not been expressed in before, might have an enhanced probability of experiencing further gene recruitments, as will be discussed later.

Although the idea of substitution-addition events in pathway evolution seems probable in the case of the *Drosophila* sex determination system, there is no reason to think that this system would be unique in experiencing such changes. Similar variations-upon-a-theme, suggestive of replacement-substitution events are as apparent in the germ-line sex determination pathways in *C. elegans* [51]. Indeed, just as ret-

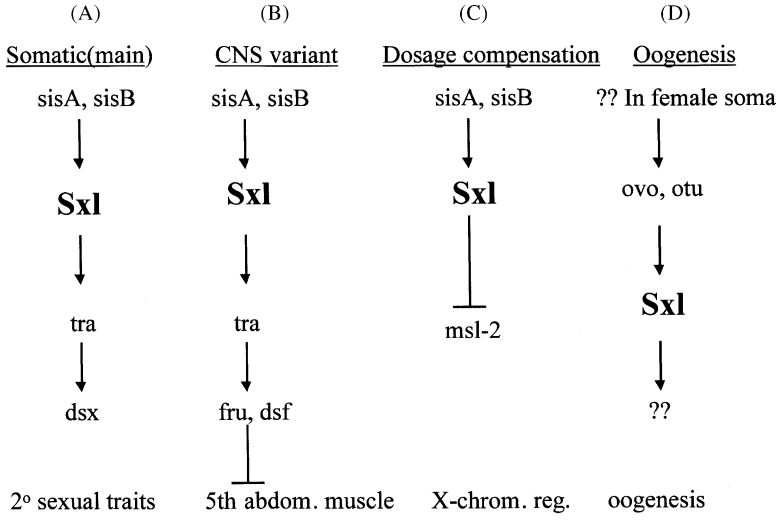


Fig. 5. Diagrammatic representation of the four sex determination pathways that govern all sexually dimorphic features in *Drosophila*. The main pathway (left) probably evolved first, with the three variant pathways arising as evolutionary derivatives; see text for discussion.

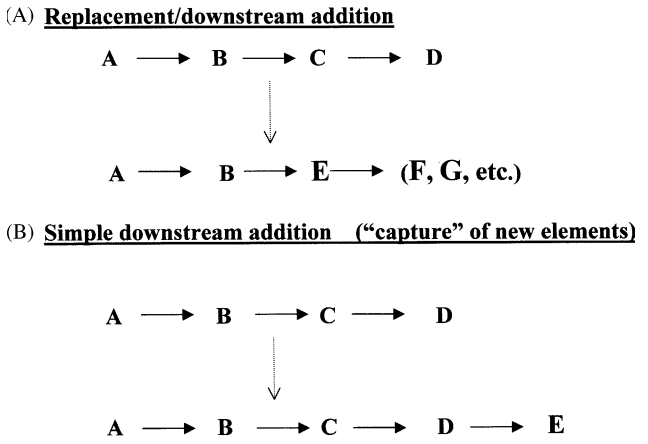


Fig. 6. Two modes of addition of downstream elements (anterograde evolution). (A) A substitution event that simultaneously truncates downstream events of the initial pathway and adds new downstream elements. (B) Simple addition of a new downstream element (gene *E*). In principle, a single mutational event in gene *B* could create *B*'s new activity on element *E* while eliminating its interaction with *C*.

rograde pathway evolution does not appear to be restricted to sex determination systems, there is no a priori reason why these other patterns should not be fairly common events in pathway evolution in general. In addition to substitution-replacement events occurring in pathway evolution (Fig. 6A), either downstream or possibly upstream, there should be the possibility of simple downstream addition (Fig. 6B). The latter process would occur as new target genes come under the control of major individual transcription factors in pathways ([14], Chapter 9; [21,92]). Indeed, there seems no a priori reason why there should not be slow evolutionary turnover amongst the downstream target genes, involving both addition and subtraction events, that are governed by key coordinating transcription factors, such as that encoded by *Pax-6* for visual capacities, in diverging lineages ([92], pp. 155–169).

3. From pathways to networks

In proceeding from patterns of pathway evolution to thinking about the modalities of network evolution, a brief working definition of the term genetic "network", as applied to development, should be given. It will be defined here as "the particular set of genes and the pattern of their interactions over time required for development of a specific phenotypic property, such as a cell or tissue type or a surface pattern of elements (e.g. bristles, colours)". Networks are reticulated structures (by definition) and highly dynamic ones (by observation). They usually involve a capability to respond to multiple input signals and display multiple potential outputs. The dynamism guarantees that the network operating early in a developmental process is not the same, in composition or structure, to the one that governs the final events of that process. Indeed, "the" network for a developmental process is, in reality, a continuum of changing networks.

In molecular terms, networks typically involve interactions, via signal transduction networks, between different cell types [30] as well as transcriptional cascades that operate both within and between cells. Some networks, however, such as that which governs the first stages of segmental patterning in holometabolous insect embryos consists primarily of interacting transcription factors wholly within one cytoplasm [41].

A critical feature of networks is their complex relationships between inputs and outputs. While the defining *structural* characteristic of a typical pathway is its linearity, its essential *functional* property is the fixed relationship between input and output. A pathway may have only two states, either "on" or "off" or alternative end-products, as in the main *Drosophila* sex determination pathway but for any strictly linear (that is non-branched) pathway, there will be a fixed output product/activity for a particular input signal.

Networks, in contrast, are intrinsically cross-connected structures and, as such, the relationships that they exhibit between input signals and outputs are more complex. In effect, the cross-connections can serve to channel the results of input signals into novel outcomes. At the simplest formal level, networks can be visualized as composed of distinct linear segments (pathways) in which certain elements are cross-connected by certain functional links [30]. This simplicity is not observed in the "meta-networks" of metabolism and protein interactions but these, especially the latter, involve abstracted sets of total potential interactions, many of which are not seen in specific cell types.

Yet, even visualizing networks as pathways cross-connected by functional links should not obscure some of the actual complexities seen in developmental genetic networks. These include multiple cross connections involving specific components and positive feedback loops. At the molecular level, a further set of complexities becomes apparent. Functional interactions can either be direct and involve physical interactions between components of the two pathways or can be mediated by a sequence of steps between them. For many of the deduced functional interactions, the number of intervening steps is unknown. A recently discovered example of a direct interaction, however, is that between components of the TGF-β and Notch signalling pathways, which serves to inhibit myogenic differentiation [13].

Furthermore, interactions between pathways can be either positive or negative, as shown in highly schematic fashion in Fig. 7. If positive, the interaction can couple the effect of an input signal for one pathway to the production of the output signal of a second pathway; if negative, the activation of one pathway can inhibit the production of the output of another activated pathway. These functional links serve to coordinate and integrate developmental responses in response to complex sets of incoming signals, such integration being essential to ensure proper development of the organism. Although both prokaryotes [74] and simple eukaryotes [29,20] utilize networks, a great elaboration of networks was almost certainly an accompaniment of, and prerequisite for, the evolution of

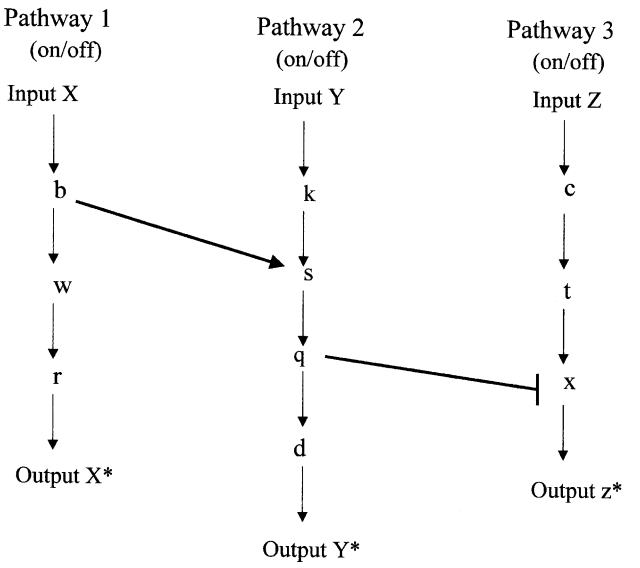


Fig. 7. Linking pathways to form networks. A positive (activating) interaction links pathways 1 and 2 while a negative (inhibitory) interaction between pathways 2 and 3 is also diagrammed. Both forms of interaction alter input–output relationships between the pathways.

multicellular organisms, with their far greater developmental repertoires than prokaryotes or unicellular eukaryotes.

An intrinsic feature of networks is the *node*, a point at which two or more signals connect. In genetic networks, nodes are molecules that interact with two or more other molecules. A schematized, and simplified, version of an actual network illustrating this feature is shown in Fig. 8. This network controls flowering time and serves to integrate responses to a variety of different signals (temperature, day-length, autonomous developmental programs) to ensure the onset of flowering under appropriate environmental conditions. These pathways converge on the downstream targets, *LFY* and *AP1*, which act as the immediate control point for turning on flowering (see review in [63]). In contrast to this example, most of the signalling in developmental networks in animals involves internally generated signals/inputs. These can be regarded either as inputs from other networks or as upstream elements within the networks themselves. The clean demarcation of elements as "upstream" or "downstream", however, is often more difficult for networks than for pathways and is impossible when the same gene performs multiple roles in the same network (see below).

An important question about the genetic networks that underlie development is whether they are scale free. At present, there is too little information to judge. To do so, one would probably need a network with >1000 links, yet no developmental genetic network has been characterized to that level of detail. The ubiquity of scale-free networks in biological systems, however, makes this seem a likely possibility. Furthermore, the known modularity of developmental systems [6,64] would be consistent with the possibility that their underlying genetic networks have the structure of modularized scale-free systems, as do metabolic networks [3]. If the scale-free form of organization is found to apply to genetic networks underlying development, then hubs – nodes with exceptionally large numbers of connections – should also be present. Numerous multiply connected molecular nodes are now known in various developmental systems while some of the downstream elements of the major signalling pathways are undoubtedly hubs.

4. Evolution of networks

It is futile to ask how the first genetic networks originated. As noted above, networks are found in prokaryotic cells and

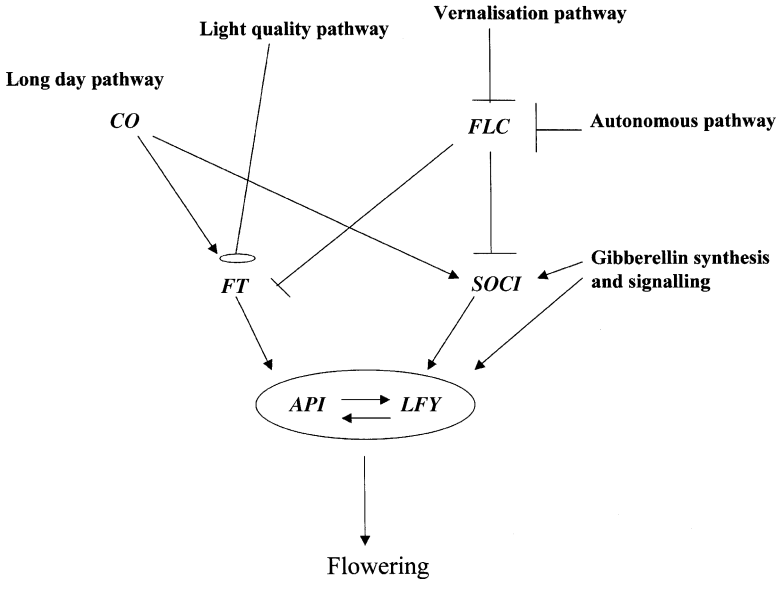


Fig. 8. The network that governs initiation of flowering in *Arabidopsis*. The immediate signal that triggers flowering is provided by the AP1/LFY node, near the bottom, while other nodes higher (earlier) in the network integrate signals from four distinguishable pathways (after Putterill et al., [63]).

have almost certainly existed from the time that the first cells evolved. Hence, the picture, drawn above, of the evolution of networks from multiple independent pathways through the addition of functional cross-linkages of some kind is not intended to represent an actual historical progression. It is simply a heuristic device, an abstract depiction of the process by which formerly distinct regulatory sequences can become functionally linked. Instead of asking how networks originate, the sensible evolutionary questions concern the ways in which networks become increasingly complex.

As pointed out by Gerhart and Kirschner [25], many networks show a high proportion of inhibitory steps. Such steps should be readily selectable, whenever there is a biological gain to be made by damping down a particular set of outputs at a particular time or place in a developing organism. Activator steps, in contrast, would be selected when there was selective pressure for producing a particular output at a particular time or place. In either case, all that is needed is a mutation that either activates expression of a gene activity that already possesses the activity or a mutation in a component that is present (at the right concentration) that creates the new inhibitory/activating property. Even if the mutation produces a weak, but positively, selected effect, it might be retained in the population, becoming amplified in frequency, providing an opportunity for further mutations that enhance its effect to be selected. Such subsequent, optimizing mutations might well be difficult to separate from the original one, barring the kind of detailed analysis that may be prohibitively expensive in time or materials or both.

The elaboration of network complexity in evolution may well resemble the results of experimental approaches that developmental geneticists carry out in the laboratory. A common strategy in developmental genetics for elucidating the workings of the genetic architecture underlying a developmental phenomenon, e.g. eye development, is to take a mutant that is affected in that process and then select for mutations that alter the magnitude of the effect. These either make the mutant effect stronger ("enhancer" mutations) or diminish or eliminate it ("suppressor" mutations). If the original mutation is a leaky (hypomorphic) allele, the modifying mutations will often be within the normal pathway that the original (mutant) gene is part of. This strategy has helped elucidate pathways of eye development in *Drosophila* [78] and vulval development in *C. elegans* [33]. In some cases, however, the modifying mutation is a component that is not part of the standard pathway. Such alteration of pathway activity can be said to be due to "lateral" modifiers and these can, in principle, be either inhibitory or activator activities. Furthermore, they can either act directly on a component of the original pathway or indirectly, though a sequence of molecular activities. Such lateral modifiers, when themselves part of pre-existing regulatory structures (as virtually all will be), are precisely the kind of mutations that would provide functional cross-links in networks (Fig. 7).

In principle, any gene product in an existing pathway or network can act as a site for new linkages. Yet, not all should be equipotential in this respect. A gene product that has, in its evolutionary history, already experienced incorporation in many networks has that history of potential interactions built into its sequence. Hence, a new functional linkage in that protein, created by a mutational event, brings a whole suite of potential new functional linkages to the pathway/network in which the new interacting gene product is embedded. Which ones are actually used at any one time will, of course, depend on many factors (e.g. cell type, the signals being received by that cell, etc.). A gene product with many sites that are already part of the total gene product "interactome" of that organism, however, is a priori more likely to be recruited to a new use than a protein with few such interactions. In effect, it has an abundance of potentially-activatable sites, relative to a gene product that is relatively depauperate in such sites. Those sites might involve not only different molecular partners but different forms of biochemical function, affecting regulation in different manners. An example would be the HnRNA-K protein, which appears to have acquired multiple, new functions in the evolution of complex eukaryotes from unicellular ancestors [5].

In turn, the probability of a gene product acquiring multiple sites of interaction, to become a linkage-rich node, should be, at least in part, a function of its evolutionary age. In effect, the "work load" of a gene, that is its number of biological roles, as measured by its degree of connectedness in total network space, should reflect its evolutionary age [16]. An example would be the Hox genes, which are at least as old as the Bilateria. Many of these display an increasing number of roles as one proceeds from simpler to more complex metazoans, with their more elaborate developmental processes and anatomies. Another predisposing factor to acquisition of multiple molecular linkages would be large gene product size, for the simple reason that larger molecules will have more mutational targets than smaller ones. In light of such considerations, one can begin to glimpse plausible and concrete molecular and evolutionary grounds for the existence of highly connected nodes, that is hubs, in genetic networks for development.

A particular feature of genetic networks in development is that many genetic elements have multiple roles within the network. This was first shown most strikingly for the segmental patterning gene network in *Drosophila*, whose principal early components are all transcription factors. Not only do most of the gene products interact with numerous other members of the network but the specific interaction, whether activating or inhibiting, is a function of concentration for many [41]. Multiple roles for several regulatory factors are also seen in the mammalian sex determination network [49] and mammalian (vertebrate) tooth development [84]. Such multiple usage of specific gene products in specific networks is far more frequent than one would expect by chance employment of those factors and, ultimately, there are only two probable explanations. The first is that the multiple usages reflect multiple evolutionary modifications of a single initial pathway, as is probably the case in the divergent *Sxl*-utilizing pathways for

sex determination in *Drosophila* (Fig. 5). The other possibility is that once a gene has been recruited for a particular usage, and is being expressed in a particular cell or tissue type, its probability of further recruitments is substantially enhanced. If one accepts the idea that each inter-product interaction is a selection for either inhibition or enhancement of one partner or the other, there are probably many gene products that can be effective in one capacity or the other for any given partner. This is certainly not implausible for inhibitory interactions; if two macromolecules taken at random have any significant binding capacity, it is not improbable that the binding will inhibit of one or the other or both, to some degree. In effect, inhibition does not require a high degree of specificity, if binding between two molecules is non-negligible. The high frequency of inhibitory interactions in networks [25] in itself argues that mutational events creating such are not rare. More surprisingly, new activation steps, at least in transcription, may also not require exceptional specificity. Much evidence along these lines can be found in the observations that form the basis of Mark Ptashne's "acid blob" model of gene activation [23].

Regardless of how such multiple usage arises in evolution, it has an interesting genetic consequence. A null mutation in such a gene is more likely to abolish development of the trait determined by the network than it would if the gene were restricted to a single role, at least one positioned relatively downstream. In effect, the phenotypes of null mutants in such multiple-usage genes seem to suggest that the respective wild-type gene exerts a single, over-all form of control ("master gene" qualities) when they are, in reality, participating in several ways at several points, often with other genes.

5. How does thinking about networks connect to issues of comparative biology?

At first glance, it might seem that the patterns of genetic pathway and network evolution, as outlined above, are too general and too abstract to have much relevance for the problems of morphological relatedness and difference that are the province of comparative neurobiology. The question of relevance, however, needs to be divided in two. The first issue concerns the *theoretical* applicability of network concepts to matters of comparative biology. The second question concerns the matter of *practical utility*: does a network perspective promise to illuminate and help resolve specific questions in comparative biology?

Take conceptual relevance first. If one agrees that morphology is the outcome of developmental processes and that the latter are underlain by the genetic networks that drive them, the phenomenon of genetic networks is unarguably relevant to comparative work. In particular, it has a direct bearing on the core concept of all comparative biological studies, that of homology. In comparative neurobiology, for instance, some of the longest-standing controversies concern questions of homology of different brain regions between mammals and birds or reptiles. Here, the term "homology" is being used in the classic Darwinian sense, namely "sameness" of the structure in question, in different organisms, by reason of evolutionary descent from the same structure in the common ancestor of those organisms. The evaluation of whether a given structure is homologous or not between the subject organisms is primarily based on visible similarity of morphological features and secondarily on spatial relationships during development. For closely related animals, e.g. members of the same order or class, the morphological similarity is usually sufficiently great such that there is little or no dispute. For instance, the forelimb homologies of bat wings, whale flippers and human hands are well accepted. For more distantly related animals, however, such as members of different vertebrate classes, the evolved divergences can obscure or seemingly out-weigh the similarities. For example, the long-standing controversies over putative homologous relationships between brain regions between Aves and Mammalia (see Reiner, this issue) are in this category. As long as the grounds of argument remain rooted in visible morphology and developmental process, controversies about homology tend to involve circular arguments and, hence, are incapable of definitive resolution.

When viewed in terms of network structure, however, there is, at least, the potential to untangle them – though doing so does require a new perspective. The basic consideration is that differences in structure between putative homologs reflect differences in the structure of their networks. Every time that a new network connectivity arises or is lost in evolution, a mutationally-based functional discontinuity occurs. Such discontinuities have no representation in standard population genetics models of evolutionary change. Furthermore, network evolution also involves a fairly sharp departure from traditional notions of homology. At the level of the genetic network, homology is partial [1,10,93]. From the basic considerations outlined in this paper, however, one can go even further. Visible similarity of putative homologs must involve a high degree of identity of outputs (downstream elements) of those networks. Conversely, the more ambiguous the morphological resemblance, the greater the degree of divergence in outputs there must be. In contrast, variations in upstream controls (inputs) can alter timing and/or placement of the structures without altering visible morphological similarity of the putative homologous structures. From this perspective, the problem of serial homology, which has always been contentious [32] in terms of the classical definition of homology as similarity-by-virtue-of-descent, disappears: serially homologous structures, such as insect legs, must utilize conserved downstream modules whose expression is activated by (somewhat) different upstream controls.

Thus, recasting the problem of homology in terms of network similarity/difference can, in principle, break the circularity of arguments that rely wholly on morphology. Yet, to be genuinely useful to comparative biologists, one has to go further into specifics. For every dispute about homology, one would ideally like to know the basic network structure for

the morphological features under comparison and how they differ. Yet, as discussed in the Introduction, empirical characterization of genetic networks for developmental processes is difficult and expensive, in terms of both resources and time. The increasing sophistication of modern molecular and bioinformatic methods, however, are making the task a more and more tractable one. If one can experimentally analyze the development of the structure in one or both organisms under study, one can use microarray techniques or proteomics to identify many of the molecular players involved. One can then use a combination of experimental and bioinformatic methods to help define both the immediate and longer-range functional linkages of these genes [26,44,52]. Alternatively, sometimes a key gene has been implicated by genetic means [18]. This is a particularly useful entry point when the developmental biology cannot be directly investigated (as in questions of primate brain structure). One can then use various techniques to ascertain that gene's network relationships. These methods involve both experimental methods, if at least fresh post-mortem samples are available [17], and bioinformatics techniques, even if experimental analysis is difficult or impossible. The protein interactome maps [26,44,52] provide an important example of the kinds of approach that can be used. A detailed discussion of these strategies would be out of place here; the essential point is that the detailed elucidation of networks is steadily increasing in feasibility.

6. Conclusions

Traditional NeoDarwinian conceptions of morphological evolution have been based on the premise that such evolution is based on the sequential fixation of mutations of individually small phenotypic effect, whose effect is cumulative [24,55,66]. In the past decade or so, however, this view has come in for new scrutiny and revision. In particular, the postulate that only mutations of small phenotypic effect are involved in morphological evolution has been re-evaluated. The potential importance of mutations of individually larger effect has been advocated both on observational grounds [61,67,72,88] and theoretical ones [60]. Recent examples would be attempts to explain complex traits such as language ability [18], skull shape [82] and pelvic development [72] in terms of particular genes. Yet, even this revision, with its emphasis on single gene-based phenotypic effects, is still largely rooted in the past because it tends to ignore the network context-dependence of such effects. In particular, when the single gene to whom crucial transformative effects are being attributed encodes a transcriptional regulator gene [18,72], it must be the case that a complete explanation requires elucidation of the network in which that gene acts.

Ultimately, the properties of networks, and the phenotypic consequences of mutational alterations of connectivity patterns in networks will have to be incorporated into evolutionary genetic models. Biological development, after all, is underlain by genetic networks and that fact is crucial to understanding the long-known and ubiquitous phenomenon of "genetic background", in which the expressivity and or penetrance of a mutant gene can be greatly influenced by other factors in the genotype. This phenomenon, in turn, is crucial to thinking about how a particular new genetic difference in a population may initially be affected by selection [71,83]. Although the idea of genetic networks underlying development, hence morphology, is beginning to enter evolutionary biology texts [66], it has not yet significantly influenced the standard evolutionary genetic models.

Similarly, the concept of partial homology is at odds with traditional notions of homology, which underlie a good deal of comparative biology. Yet this idea is now inescapable [1,10,93].

Thus, while the concepts of pathway and network evolution outlined in this paper are neither particularly abstract nor difficult, they constitute a challenge to traditional thinking and experimental analyses in both evolutionary and comparative biology. Accordingly, their incorporation into the standard thinking of these fields might well proceed slowly.

Nevertheless, the perspective offered in this paper may provide the outlines of a framework for interpreting the evolutionary origins of those differences in structures that are the focus of comparative morphological studies. Perhaps most importantly, this framework may even be useful for making predictions. The reason is that the kinds of pathway and network change that have been inferred to take place, and which have been the central subject of this article, actually form a reasonably small discrete set of patterns. These may be summed up quickly. Thus, linear segments of networks or whole integrated linear sequences (genetic pathways) appear to grow from downstream to upstream in many cases, while downstream additions can provide fine-tuning. In addition, substitutions at any point in the chain may take place, sometimes truncating a pathway and producing new outputs. On the other hand, interconnecting linkages, either activating or inhibiting, can form between linear segments and can alter relationships between input signals and downstream events. Further mutational events can strengthen those linkages while new selective events can, in principle, act to amplify genetic changes that sever pre-existing connections. Gene recruitment events can commandeer segments of networks or whole networks while, once a gene has been recruited, it may be subject to further recruitment events within the evolving network.

These kinds of change do not exhaust the possibilities but they probably account for the majority of events. As new networks and their evolutionary changes are explored, this provisional categorization of the kinds of events that modify connectivity patterns in networks can serve as a rough guide to the various genetic events that have shaped the evolutionary events of particular networks. High degrees of morphological similarity between structures point to fairly similar network outputs (downstream events) while variations in timing or positioning or tissue provenance would be indicative of comparable shifts of activity in relatively more upstream elements. As more and more networks are provisionally characterized,

these sorts of consideration may have utility in permitting rough predictions, from the observed phenotypic difference, about the kinds and approximate placement of the alterations in the network(s) that may have occurred. The ever-expanding armoury of experimental and bioinformatic methods for exploring network connections can then be exploited to test such predictions.

References

[1] E. Abouheif, Developmental genetics and homology: a hierarchical approach, Trends Ecol. Evol. 12 (1997) 405–408.
[2] A.-L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[3] A.-L. Barabasi, Z.N. Oltvai, Network biology: understanding the cell's functional organization, Nat. Rev. Genet. 5 (2004) 101–113.
[4] M. Beye, M. Hasselmann, M.K. Fondrik, R.E. Page Jr., S.W. Omholt, The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein, Cell 114 (2004) 419–429.
[5] K. Bomstyk, O. Denisenko, J. Ostrowski, hnRNP K: one protein, multiple functions, BioEssays 26 (2004) 629–638.
[6] J.T. Bonner, The Evolution of Complexity, Princeton University Press, Princeton, 1988.
[7] D. Bopp, G. Calhoun, J.I. Horabin, M. Samuels, P. Schedl, Sex-specific control of sex-lethal is a conserved mechanism for sex determination in the genus *Drosophila*, Development 22 (1996) 971–982.
[8] R.J. Britten, E.H. Davidson, Gene regulation for higher cells: a theory, Science 165 (1969) 349–357.
[9] R.J. Britten, E.H. Davidson, Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty, Q. Rev. Biol. 46 (1971) 111–133.
[10] A.B. Butler, W.M. Seidel, Defining sameness: historical, biological and generative homology, BioEssays 22 (2000) 111–125.
[11] S.B. Carroll, J. Grenier, S.D. Weatherbee, From DNA to Diversity, Blackwell Science, Malden, 2001.
[12] T.R. Clandinin, W.R. Katz, P.W. Sternberg, *Caenorhabditis elegans* HOM-C genes regulate the response of vulval precursor cells to inductive signals, Dev. Biol. 182 (1997) 150–161.
[13] C. Dahlqvist, A. Blokzijl, G. Chapman, A. Falk, K. Dannaeus, C.F. Ibanez, U. Lendahl, Functional Notch signalling is required for BMP4-induced inhibition of myogenic differentiation, Development 130 (2003) 689–699.
[14] E.H. Davidson, Genomic Regulatory Systems, Academic Press, San Diego, 2001.
[15] E.H. Davidson, J.P. Rast, P. Oliveri, A. Ransick, C. Calestani, C.-H. Yuh, T. Minokawa, G. Amore, V. Hinman, C. Arenas-Mena, et al., A genomic regulatory network for development, Science 295 (2002) 1669–1678.
[16] D. Duboule, A.S. Wilkins, The evolution of bricolage, Trends Genet. 14 (1998) 54–59.
[17] W. Enard, P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giavalisco, K. Nieswelt-Struwe, E. Muchmore, A. Varki, R. Ravid, et al., Intra- and interspecific variation in primate gene expression patterns, Science 296 (2002) 340–343.
[18] W. Enard, M. Przeworski, S.E. Fisher, C.S.L. Lai, V. Wiebe, T. Kitano, A.P. Monaco, S. Paabo, Molecular evolution of FOXP2, a gene involved in speech and language, Nature 418 (2002) 869–872.
[19] C. Extavour, M. Akam, Mechanisms of germ cell specification across the metazoans: epigenesis and preformation, Development 130 (2003) 5869–5884.
[20] D.E. Featherstone, K. Broadie, Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network, BioEssays 24 (2002) 267–274.
[21] D.E.K. Ferrier, Homeobox, in: M. Pagel (Ed.), Encyclopaedia of Evolution, Oxford University Press, Oxford, 2002, pp. 473–475.
[22] A. Fischer, C. Viebahn, M. Blum, FGF8 acts as a right determinant during establishment of the left-right axis in the rabbit, Curr. Biol. 12 (2002) 1807–1816.
[23] J.A. Fischer, E. Giniger, T. Maniatis, M. Ptashne, GAL4 activates transcription in *Drosophila*, Nature 332 (1988) 853–856.
[24] R.A. Fisher, The Genetical Theory of Natural Selection, Dover Publications, New York, 1958.
[25] J. Gerhart, M. Kirschner, Cells, Embryos and Evolution, Blackwell Science, Malden, MA, 1997.
[26] L. Giot, et al., A protein interaction map of *Drosophila melanogaster*, Science 302 (2003) 1727–1736.
[27] J.A.M. Graves, Evolution of the testis-determining gene – the rise and fall of SRY, in: D. Chadwick, J. Goode (Eds.), The Genetics and Biology of Vertebrate Sex Determination, Novartis Symposium 244, John Wiley & Sons, Chichester, 2002, pp. 86–101.
[28] J.J. Gubbay, J. Colignon, N. Vivian, P. Goodfellow, R. Lovell-Badge, A gene mapping to the sex-determination region of the mouse Y chromosome is a member of a novel family of embryologically expressed genes, Nature 346 (1991) 245–249.
[29] N. Guelzim, S. Bottani, F. Kepes, Topological and causal structure of the yeast transcriptional regulatory network, Nat. Genet. 31 (2002) 60–63.
[30] M.S. Halfon, A.M. Michaelson, Exploring genetic regulatory networks in metazoan development: methods and models, Physiol. Genomics 10 (2002) 131–143.
[31] B.J. Hall, Evolutionary Developmental Biology, second ed., Kluwer Academic Publishers, Dordrecht, 2000.
[32] B.J. Hall, Descent with modification: the unity underlying homology and homoplasy as seen through the analysis of development and evolution, Biol. Rev. 78 (2003) 409–433.
[33] M. Han, P.W. Sternberg, let-60, a gene that specifies vulval cell fates during *C. elegans* vulval induction, encodes a ras protein, Cell 63 (1990) 921–931.
[34] P.H. Harvey, M.D. Pagel, The Comparative Method in Evolutionary Biology, Oxford University Press, Oxford, 1991.
[35] M. Hediger, G. Burghardt, C. Siegenthaler, N. Buser, D. Hilfiker-Kleiner, A. Dubendorfer, D. Bopp, Sex determination in *Drosophila melanogaster* and *Musca domestica* converges at the level of the terminal regulator doublesex, Dev. Genes Evol. 214 (2004) 29–42.
[36] J.A. Hodgkin, More sex determination mutants of *Caenorhabditis elegans*, Genetics 96 (1980) 649–664.
[37] J. Hodgkin, Two types of sex determination in a nematode, Nature 304 (1983) 267–268.
[38] J. Hodgkin, Sexual dimorphism and sex determination, in: W.B. Wood (Ed.), The Nematode *Caenorhabditis elegans*, Cold Spring Harbor Press, Cold Spring Harbor, 1988.
[39] J. Hodgkin, Sex determination compared in *Drosophila* and *Caenorhabditis*, Nature 344 (1990) 721–728.
[40] J. Hodgkin, Genetic sex determination mechanisms and evolution, BioEssays 14 (1992) 253–261.
[41] M. Hulskamp, D. Tautz, Gap genes and gradients – the logic behind the gaps, BioEssays 13 (1991) 261–268.
[42] F. Jacob, Evolution and tinkering, Science 196 (1977) 1161–1166.
[43] H.P. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabasi, The large-scale organization of metabolic networks, Nature 407 (2000) 651–654.
[44] H.P. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Lethality and centrality in protein networks, Nature 411 (2001) 41–42.
[45] W. Just, et al., Absence of Sry in species of the vole *Ellobius*, Nature Genetics 11 (1995) 117–118.
[46] S.A. Kauffman, Metabolic stability and epigenesis in randomly connected nets, J. Theor. Biol. 22 (1969) 437–467.
[47] R.L. Kelley, J. Wang, L. Bell, M.I. Kuroda, Sex lethal controls dosage compensation in *Drosophila* by a non-splicing mechanism, Nature 387 (1997) 195–199.

[48] M.-C. King, A.C. Wilson, Evolution at two levels in humans and chimpanzees, Science 188 (1975) 107–116.
[49] P. Koopman, Sry, Sox9 and mammalian sex determination, in: G. Scherer, M. Schmid (Eds.), Genes and Mechanisms in Vertebrate Sex Determination, Birkhauser-Verlag, Basel, 2001.
[50] P.E. Kuwabara, J. Kimble, Molecular genetics of sex determination in *C. elegans*, Trends Genet. 8 (1992) 164–168.
[51] P.E. Kuwabara, M.D. Perry, It ain't over till it's ova: germline sex determination in *C. elegans*, BioEssays 23 (2001) 596–604.
[52] S. Li, et al., A map of the interactome network of the metazoan, *C. elegans*, Science 303 (2003) 540–543.
[53] A. Martinez-Arias, A.M.C. Brown, K. Brennan, Wnt signalling: pathway or network? Curr. Opin. Genet. Dev. 9 (1999) 447–454.
[54] J. Mattick, Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms, BioEssays 25 (2003) 930–939.
[55] J. Maynard Smith, Evolutionary Genetics, Oxford University Press, Oxford, 1989.
[56] E. Mayr, W.B. Provine, The Evolutionary Synthesis, Harvard University Press, Cambridge, 1980.
[57] M. Meise, et al., Sex-lethal, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*, Development 125 (1998) 1487–1494.
[58] J. Monod, F. Jacob, General conclusions: teleonomic mechanisms in cellular metabolism, growth and differentiation, Cold Spring Harb. Quant. Biol. 26 (1962) 389–401.
[59] F. Muller-Holtkamp, The sex-lethal gene homologue in *Chrysomya rufifacia* is highly conserved in sequence and exon-intron organisation, J. Mol. Evol. 41 (1995) 467–477.
[60] H.A. Orr, The population genetics of adaptation: the distribution of factors fixed during adaptive evolution, Evolution 52 (1998) 935–949.
[61] M.F. Palopoli, N.H. Patel, Neo-Darwinian developmental evolution: can we bridge the gap between pattern and process? Curr. Opin. Genet. Dev. 6 (1996) 502–508.
[62] A. Pomiankowski, R. Nothiger, A.S. Wilkins, The evolution of the *Drosophila melanogaster* sex determination pathway, Genetics 166 (2004) 1761–1773.
[63] J. Putterill, R. Laurie, R. Macknight, It's time to flower: the genetic control of flowering time, BioEssays 26 (2004) 363–373.
[64] R. Raff, The Shape of Life, University of Chicago Press, Chicago, 1996.
[65] C.S. Raymond, et al., Evidence for evolutionary conservation of sex-determining genes, Nature 391 (1998) 691–695.
[66] M. Ridley, Evolution, Blackwell Scientific, Malden, MA, 2004.
[67] M. Ronshaugen, N. McGinnis, W. McGinnis, Hox protein mutation and macroevolution of the insect body plan, Nature 415 (2002) 914–917.
[68] G. Saccone, I. Peluva, D. Artiaco, E. Giordano, D. Bopp, L. Polito, The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene sex-lethal is structurally conserved but not sex-specifically regulated, Development 125 (1998) 1495–1500.
[69] D. St. Johnstone, C. Nusslein-Volhard, The origin of pattern and polarity in the *Drosophila* embryo, Cell 68 (1992) 201–209.
[70] C. Schutt, R. Nöthiger, Structure, function and evolution of sex-determining systems in dipteran insects, Development 127 (2000) 667–677.
[71] I. Schmalhausen, Factors of Evolution, Blakiston Company, Philadelphia, 1949.
[72] M.D. Shapiro, M.E. Marks, C.L. Peichel, B.K. Blackman, K.S. Nereng, B. Jonsson, D. Schluter, D.M. Kingsley, Genetic and developmental basis of evolutionary pelvic reduction in three-spine sticklebacks, Nature 428 (2004) 717–723.
[73] M.M. Shen, J.A. Hodgkin, mab-3, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*, Cell 54 (1988) 1019–1031.
[74] S.S. Shen-Orr, R. Milo, S. Mangan, U. Alon, Network motifs in the transcriptional regulation network of *Escherichia coli*, Nat. Genet. 31 (2002) 64–68.
[75] D.A.C. Shearman, M. Frommer, The *Batrocera tryoni* homologue of the *Drosophila melanogaster* sex determination gene doublesex, Insect Mol. Biol. 7 (1998) 1–12.
[76] R.M. Shymko, P. de Meyts, R. Thomas, Logical analysis of timing-dependent receptor signalling specificity: application to the insulin receptor metabolic and mitogenic signalling pathways, Biochem. J. 326 (1997) 463–469.
[77] V. Sievert, S. Kubu, W. Trout, Expression of the sex-determining cascade genes sex-lethal and doublesex in the phorid fly *Megaselia scalaris*, Genome 40 (1997) 211–214.
[78] M.A. Simon, D.A. Bowtell, G.S. Dodson, T.R. Lavesty, G.M. Rubin, Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signalling by the sevenless protein tyrosine kinase, Cell 67 (1991) 701–716.
[79] R.S. Singh, Darwin to DNA, molecules to morphology: the end of classical population genetics and the road ahead, Genome 46 (2003) 938–942.
[80] A. Sinclair, C. Smith, P. Western, P. McClive, A comparative analysis of vertebrate sex determination, in: D. Chadwick, J. Goode (Eds.), The Genetics and Biology of Sex Determination, Novartis Foundation Symposium 244, John Wiley & Sons, Chichester, 2002, pp. 102–114.
[81] A.H. Sinclair, et al., A gene from the human sex-determining region of the mouse Y chromosome encodes a protein with homology to a conserved DNA-binding motif, Nature 346 (1990) 240–244.
[82] H.H. Stedman, B.W. Kozyak, A. Nelson, D.M. Thesier, L.T. Su, D.W. Low, C.R. Bridges, J.B. Schrager, N. Minugh-Purvis, M.A. Mitchell, Myosin gene mutation correlates with anatomical changes in the human lineage, Nature 428 (2004) 415–418.
[83] C. Stern, Gene and character, in: G.L. Jepsen, E. Mayr, G.G. Simpson (Eds.), Genetics, Paleontology and Evolution, Princeton University Press, Princeton, 1949, pp. 13–23.
[84] D. Stock, The genetic basis of modularity in the development and evolution of the vertebrate dentition, Philos. Trans. R. Soc. Lond. B. 355 (2001) 1633–1653.
[85] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[86] M.G. Suzuki, F. Ohbayashi, K. Mita, T. Shimada, The mechanism of sex-specific splicing at the doublesex gene is different between *Drosophila melanogaster* and *Bombyx mori*, Insect Biochem. Mol. Biol. (2001).
[87] J. True, S.B. Carroll, Gene co-option in physiological and morphological evolution, Annu. Rev. Cell Dev. Biol. 18 (2002) 53–80.
[88] S.R. Voss, H.B. Shafer, Evolutionary genetics of metamorphic failure using wild-caught vs. laboratory axolotls (*Ambystoma mexicanum*), Mol. Ecol. 9 (2000) 1401–1407.
[89] A. Wagner, How the global structure of protein interaction networks evolves, Philos. R. Soc. Lond. B 270 (2003) 457–466.
[90] A. Wagner, D.A. Fell, The small world inside large metabolic networks, Proc. R. Soc. Lond. B 268 (2001) 1803–1810.
[91] A.S. Wilkins, Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway, BioEssays 17 (1995) 71–77.
[92] A.S. Wilkins, The Evolution of Developmental Pathways, Sinauer Associates, Sunderland, MA, 2002.
[93] G.A. Wray, Evolutionary dissociations between homologous genes and homologous structures, in: G.R. Bock, G. Cardew (Eds.), Homology, Novartis Symposium no. 222, John Wiley & Sons, Chichester, 1999.
[94] D. Zarkower, Invertebrates may not be so different after all, in: D. Chadwick, J. Goode (Eds.), The Genetics and Biology of Sex Determination, Novartis Symposium 244, John Wiley & Sons, Chichester, 2002, pp. 115–135.