THE EVOLUTION OF SIGNALLING PATHWAYS IN ANIMAL DEVELOPMENT

André Pires-daSilva and Ralf J. Sommer

Despite the bewildering number of cell types and patterns found in the animal kingdom, only a few signalling pathways are required to generate them. Most cell–cell interactions during embryonic development involve the Hedgehog, Wnt, transforming growth factor-β, receptor tyrosine kinase, Notch, JAK/STAT and nuclear hormone pathways. Looking at how these pathways evolved might provide insights into how a few signalling pathways can generate so much cellular and morphological diversity during the development of individual organisms and the evolution of animal body plans.

> Embryologists now recognize receptors and signal transducing molecules as components of the competence apparatus that enables certain cells to respond to specific inducers. If macroevolution involves changing morphological features, then the altering of signal transduction pathways becomes critical for any discussion of large scale evolution.[1]

Embryonic induction was one of the first general principles of development to be determined by early developmental biologists. Spemann and Mangold, for example, observed that two-headed salamanders could be generated by transplanting a specific piece of embryonic tissue into another embryo[2]. The transplanted tissue was able to induce the fate of neighbouring cells in the host embryo, indicating that cells might communicate with each other through secreted signals. Research in the past two decades has yielded important advances towards the identification of the molecules that are involved in signalling processes[3–9].

After millions of years of evolution, signalling pathways have evolved into complex networks of interactions. Surprisingly, genetic and biochemical studies revealed that only a few classes of signalling pathways are sufficient to pattern a wide variety of cells, tissues and morphologies. The specificity of these pathways is based on the history of the cell (referred to as the 'cell's competence'), the intensity of the signal and the cross-regulatory interactions with other signalling cascades. Many studies have been dedicated to analysing the specific details of signalling pathways in a few model organisms. By contrast, only a few studies have been initiated to understand how these pathways can be modified to generate new morphologies, how new components are integrated into existing signalling pathways and how the pathways themselves evolve.

Here, we review the evolution of signalling pathways from three perspectives. First, we review the genetic repertoires of signalling systems seen in present-day organisms. Second, we analyse how the evolution of signalling processes might be involved in creating morphological novelties. Finally, we address the problem of how signalling pathways evolve, using nematode sex determination as a case study. Although this review concentrates on the evolution of signalling pathways in animal development, it is important to note that signal transduction is also important in other biological processes, such as physiological control, adaptive immunity and neurobiology.

Properties of signalling pathways

Cell–cell interactions through signal-transduction pathways are crucial in the coordination of embryonic development. Typically, signalling pathways are activated by

Department of Evolutionary Biology, Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 37–39, D-72076 Tübingen, Germany.
Correspondence to R.J.S.
e-mail: ralf.sommer@tuebingen.mpg.de
doi:10.1038/nrg977

© 2002 Nature Publishing Group

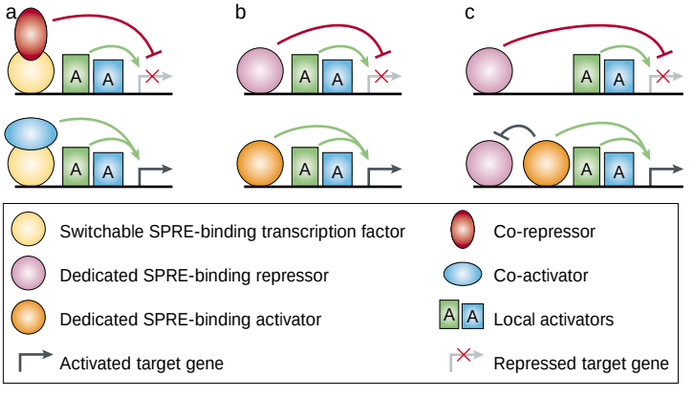
the binding of a ligand to a transmembrane receptor, which in turn leads to the modification of cytoplasmic transducers. Subsequently, these transducers activate transcription factors that ultimately alter gene expression. One of the most surprising findings about signalling processes is that only a few pathways are involved in and are responsible for most of animal development (BOX 1): Hedgehog (Hh), wingless related (Wnt), transforming growth factor-β (TGF-β), receptor tyrosine kinase (RTK), Notch, Janus kinase (JAK)/signal transducer and activator of transcription (STAT) and nuclear hormone pathways are used repeatedly throughout the development of individuals and throughout the evolution of metazoans (for reviews, see REFS 3–9). These seven pathways are a subset of a larger group of conserved signalling pathways summarized by Gerhart[10]. However, most of the other pathways, such as signalling through G-protein-coupled receptors, are not repeatedly used in development. The observation that the same set of seven pathways is used many times in development indicates that signalling systems are highly flexible in generating distinct responses in different tissues and species. In fact, it is becoming apparent that signal-transduction pathways are not based on the linear sequential activation of signalling components, but have the potential to branch at many steps of a cascade. For example, RTK signalling has been shown in several systems to branch at the level of the RAF kinase[11].

Given the flexibility of signalling pathways, research in the past decade has concentrated on the question of how specificity is achieved in any signalling response. There is now clear evidence that the specificity of cellular responses can be achieved by at least five mechanisms, which in some cases act in combination, highlighting the network properties of signalling pathways in living cells[11,12] (FIG. 1). First, the same receptor can activate different intracellular transducers in different tissues. One example for this mechanism is lethal-23 (LET-23) RTK signalling in *Caenorhabditis elegans*, which is transduced by inositol-1,4,5-trisphosphate ($Ins(1,4,5)P_3$) in the gonad[13] and RAS/mitogen-activated protein kinase (MAPK) in the vulva and other tissues[14]. Second, differences in the kinetics of the ligand or receptor might generate distinct cellular outcomes. The strength of receptor affinity for the Wnt ligand, for example, results in the activation of alternative pathways, leading either to cytoskeleton reorganization through Rho and Jun amino-terminal kinase (JNK) or to the regulation of gene expression through β-catenin[15]. Third, combinatorial activation by signalling pathways might result in the regulation of specific genes. Several signalling pathways can be integrated either at signalling proteins or at enhancers of target genes. One example is the muscle and heart enhancer of the *Drosophila even-skipped* (*eve*) gene, which contains functional binding sites for response to Wnt, TGF-β and RTK[16]. Mutation at any of these sites abolishes *eve* expression. Fourth, cells that express distinct transcription factors might respond differently when exposed to the same signals. In *C. elegans*, RTK stimulation results in vulva induction only in the hypodermis, owing to the tissue-specific expression of lineage-31 (LIN-31) (the downstream transcription factor)[17]. In some cases, tissue-specific factors might antagonize signalling pathways or modulate the target gene specificities. One example is the secreted molecule Cerberus that inhibits the trunk-inducing Wnt and TGF-β pathways and that allows the formation of head structures in the vertebrate embryo[9,18]. Fifth,

Box 1 | Types of transcriptional control

Seven major cell–cell signalling pathways can be distinguished in animal systems: the wingless related (Wnt), transforming growth factor-β (TGF-β), Hedgehog (Hh), receptor tyrosine kinase (RTK), Janus kinase (JAK)/signal transducer and activator of transcription (STAT), Notch and nuclear receptor pathways. These pathways have been shown to act repeatedly during animal development and they are diverse with regard to their biochemical mechanisms and the complexity of the individual components that are involved. What these pathways have in common is the activation of specific target genes by the regulation of signal-dependent transcription factors. In the case of the Wnt, Notch, Hh and nuclear receptor signalling pathways, the signal-dependent transcription factors function as repressors in the absence of signalling, but turn into activators on ligand signalling. Barolo and Posakony[91] have recently called this type of regulation 'default repression' and this type of transcriptional control 'type I'. In type I transcriptional switching, the signalling pathway response elements, as well as the transcription factors, are identical for default repression and transcriptional activation. The transcription factors are Suppressor of Hairless (Su(H)) in the Notch pathway, Tcf/Lef (Pangolin) in the Wnt pathway, Gli/Ci (Cubitus interruptus) in the Hh pathway, and the nuclear receptors themselves. The fact that the nuclear hormone receptors function directly as transcriptional regulators indicates the strongest reduction in signalling pathway complexity. In contrast to these four cases, the TGF-β and RTK signalling pathways achieve transcriptional switching by separate repressors and activators, which can bind to similar (type II) or distinct (type III) DNA binding sites. In addition, the TGF-β and RTK pathways might also work by type I regulation, turning default repressors into activators, as seen for the Wnt, Hh, Notch and nuclear receptor pathways. The fact that the TGF-β and RTK pathways often use separate repressor and activator proteins provides several sites of regulation.

A detailed overview of these signalling pathways and their mode of action has recently been provided by Barolo and Posakony[91]. In type I signalling (shown in panel **a**), the transcription factor that regulates the signalling pathway response element (SPRE) is converted from a repressor to an activator of transcription. In type II signalling (shown in panel **b**), an activator replaces the repressor, both of which bind to the same enhancer element. And, in type III signalling (shown in panel **c**), the repressor and the activator recognize different binding sites. In the presence of signalling, the activator releases the repressor activity (see figure). These authors also propose that, despite the different primary mechanisms that are involved in achieving transcriptional control, three common functional properties can be deduced from all seven pathways. Activator insufficiency, cooperative activation and default repression emerge as the common principles behind signalling pathways in metazoans[91]. Figure reproduced with permission from REF. 91 © (2002) Cold Spring Harbor Press.



© 2002 Nature Publishing Group

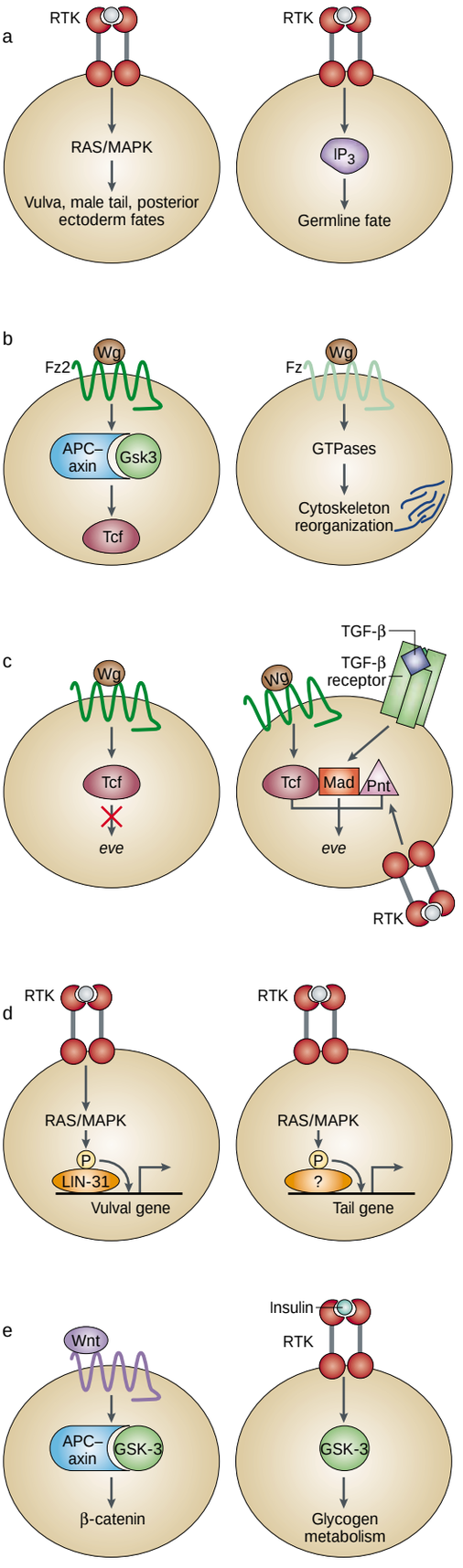


Figure 1 | **Mechanisms of signalling specificity. a** | The differential expression of distinct transducers leads to cell-specific responses using the same receptor tyrosine kinase (RTK). In the vulva, male tail and posterior ectoderm of *Caenorhabditis elegans*, the RTK lethal-23 (LET-23) activates the RAS/mitogen-activated protein kinase (MAPK) signalling cascade, whereas in the germline, the inositol-1,4,5-trisphosphate (Ins(1,4,5)P_3; IP_3 in the figure) signalling pathway is activated in response to the same signal. **b** | The *Drosophila* Frizzled 2 (Fz2) and Frizzled (Fz) receptors, although structurally similar, have different kinetic properties. The ligand-binding domain of Fz2 has a tenfold higher affinity for Wingless (Wg) than does Fz, inducing the activation of either the transcription factor (Tcf) or reorganization of the cytoskeleton, respectively. **c** | Expression of the transcription factor *even-skipped* (*eve*) in *Drosophila* muscle and in heart precursor cells depends on signalling through three pathways: Wingless (Wg), transforming growth factor-β (TGF-β) and RTK. Transcription factors that are activated by these cascades — Tcf, Mothers against dpp (Mad) and Pointed (Pnt) — and cell-specific transcription factors act synergistically at the *eve* enhancer to activate its transcription. A failure to activate any of these pathways leads to the loss of *eve* enhancer activity. **d** | The RTK–RAS–MAPK signalling pathway specifies distinct cell fates in *C. elegans*, depending on the presence of different transcription factors in a given tissue. Activation of these DNA-binding proteins (lineage-defective-31 (LIN-31) in the vulva and an unknown factor (?) in tail precursor cells) results in the transcriptional regulation of tissue-specific genes. **e** | Glycogen synthase kinase-3 (GSK-3), when sequestered by the cytoplasmic protein complex adenomatous polyposis coli (APC)–axin, mediates Wnt signalling by activation of β-catenin. In the absence of these specific protein–protein interactions (for example, in the insulin–RTK signalling pathway in vertebrates), GSK-3 phosphorylates glycogen synthase to regulate glycogen metabolism. Modified with permission from REF. 11.

compartmentalization of the signal in the cell can contribute to specificity. The recruitment of components into protein complexes prevents cross signalling between unrelated signalling molecules or targets multifunctional molecules to specific functions. Glycogen synthase kinase-3 (GSK-3), for example, is involved in Wnt signalling and glycogen metabolism[19]. Part of its functional specificity relies on compartmentalization in the cell by cytosolic protein complexes, for example during Wnt signalling[20,21].

To understand the full complexity of signalling pathways and networks, it is important to complement genetics and biochemistry with theoretical studies. Modelling of the system properties (such as ligand–receptor interactions and diffusion rate) of signalling pathways can help to predict which parameters are important for the flexibility and for the potential to change during evolution. Gene network models that are based on gene interactions for segmentation and neurogenetic networks in *Drosophila* embryos, for example, were found to give the correct spatial expression of Hh/Wnt and Notch pathways, respectively[22,23]. These modelling experiments led to the interesting observation that changes in many variables (~50), such as ligand concentration, the K_M of ligand-receptor interaction and the half-life of transducers, can result in the same pattern of gene expression[22,23]. So, although signalling systems are flexible to generate diverse outcomes, they are also robust.

K_M
The substrate concentration at which the reaction rate of an enzyme is half maximal, also known as the Michaelis–Menten constant.

© 2002 Nature Publishing Group

Box 2 | Gene duplications

Gene duplications seem to be relatively common — on average 0.01 duplications occur per gene per million years[92]. Duplicated genes, however, are expected to be lost or to be mutated into pseudogenes in a relatively short time period of a few million years[93]. It has been estimated that ~60% of the duplications are either lost or mutate into pseudogenes[94]. There are two mechanisms that might maintain gene duplicates: one of the copies acquires a new function or both copies accumulate partial loss-of-function mutations that complement each other, but jointly retain the full set of subfunctions that are present in the original ancestral gene[95] (subfunctionalization). It has been proposed that subfunctionalization is a common mechanism for functionally related proteins[96] (such as components of a signalling pathway). The co-evolution between functionally related proteins could facilitate the maintenance of gene duplications. Duplications of genes that encode ligands often correlate with duplications of receptor-encoding genes, as is seen in the case of the insulin–NGF gene family and its receptors. As a result, the divergence of the two interacting proteins would correlate with each other, eventually creating new functional interactions.

BILATERIANS
Multicellular animals that have a real body cavity (coelom) and a primary bilateral symmetry. They include all multicellular organisms except for the sponges, cnidarians and ctenophorans.

PROTOSTOMES
Animals whose development is characterized by the formation of a single opening. The protostomial phyla are subdivided into the ecdysozoans and the lophotrochozoans.

ECDYSOZOA
A bilaterian clade that is characterized by external cuticles that are shed during stages of development. It includes the insects and nematodes.

LOPHOTROCHOZOA
A bilaterian clade that is characterized by a lophophore (a specific morphological structure) or a trochophore larval stage. Well-known members include the molluscs and the annelids.

DEUTEROSTOMES
A bilaterian clade that is characterized by the formation of distinct mouth and anal openings.

ASCERTAINMENT BIAS
An error that is introduced with a biased sampling scheme.

PARALOGUES
Homologous genes that have originated by gene duplication.

DIPLOBLASTS
A group of ancestral animals, such as the cnidarians and the porifera, that do not develop mesoderm.

Apart from flexibility, robustness is a second general property of signalling pathways that is not readily apparent. Both theoretical and experimental data postulate that two types of general network architecture might account for robustness: positive- and negative-feedback loops[24]. In positive-feedback loops, the formation of the ligand is often enhanced, thereby amplifying, stabilizing or prolonging signalling. The enhancement of signalling pathways by positive feedback can guarantee the commitment of a cell to a particular developmental decision[25]. Positive-feedback regulation is used in developmental contexts, such as the maintenance of *Ultrabithorax* (*Ubx*) expression in the *Drosophila* gut by Tgf-β and Wnt signalling, but is also seen in aberrant circumstances, such as the autocrine stimulation in tumour cells[26,27]. Negative-feedback loops are used to inhibit and/or limit signalling. One example is the limitation of Hh signalling to the posterior compartment of the *Drosophila* wing disc; anterior cells that receive Hh ligand activate transcription of Patched, a membrane protein that sequesters Hh, thereby preventing its diffusion[28]. By using artificial gene networks, it has been shown that negative-feedback loops also have the property to stabilize fluctuations of biochemical parameters[29].

In summary, signalling pathways are nonlinear, highly integrative biological modules with robust properties that ensure reproducible outcomes of developmental processes. At the same time, however, they are flexible enough to allow changes in the signalling response during development and evolution.

An *in silico* analysis

Once the basic components of signalling systems had been identified in model organisms such as *C. elegans* and *Drosophila*, a first overview of the evolutionary alterations of signalling pathways was revealed by looking at the signalling gene repertoire in other species. However, a comprehensive analysis of the evolution of signalling pathways is hampered by the disparity of the sequence availability for different organisms. Whole-genome sequences are available for some medically important organisms (such as *Plasmodium* and *Anopheles*)[30,31], but not for those organisms that represent the key taxa of metazoan evolution (see below). Nonetheless, several important conclusions emerge from both gene-by-gene analysis in phylogenetically informative taxa and genome-wide comparisons in model organisms.

The pre-genome era. Until recently, studies of the evolution of signalling pathways were carried out by searching for the individual components of these pathways in those organisms considered to represent taxa that were phylogenetically informative for metazoan evolution. Key transitions in animal evolution are the invention of multicellularity, the occurrence of the BILATERIANS and the, still debatable, relationship between the principal metazoan groups, such as the PROTOSTOMES (ECDYSOZOA and LOPHOTROCHOZOA) and the DEUTEROSTOMES. Gene-by-gene searches in different organisms tend to show a strong ASCERTAINMENT BIAS because only those components that are already known in model organisms, such as yeast, flies and worms, can be searched for. However, more unbiased approaches, such as forward genetic approaches in the zebrafish, have often re-identified signalling pathways that are familiar in other systems (for example, see REFS 32,33). One of the most important findings in the evolution of signalling pathways is that they evolved before the occurrence of the bilaterians. Although no comprehensive searches for signalling molecules have been carried out in any metazoan phyla, it is fair to say that the basic components of most pathways have been found in all studied bilaterian taxa. However, different numbers of PARALOGOUS molecules were identified in organisms of divergent animal phyla. So, a second major conclusion in the evolution of signalling systems is the importance of gene duplication and subsequent protein sequence divergence (BOX 2; for a review, see REF. 34).

If bilaterians share most signalling pathways, what about animals with a simple body plan, such as the DIPLOBLASTS and the PROTISTS? Recent work in the CNIDARIANS and the PORIFERA points to the fact that even these groups contain several signalling systems. By contrast, the limited searches that have been carried out in different unicellular organisms do not provide evidence for advanced signalling pathways. For example, genes that encode members of Wnt, TGF-β, Hh, Notch and nuclear receptor signalling were found in diploblastic animals[35], but were not found in unicellular eukaryotes or plants[36]. The recently published genomes of the protists *Plasmodium falciparum* and *P. yoelii* show that these protists have only a few signalling molecules, such as a Ras family GTPase, a cAMP-dependent protein kinase and a 14-3-3-like molecule, the latter being a kinase binding protein that functions as an adaptor protein for signalling networks[30,37]. However, this negative evidence has to be considered with care: as *Plasmodium* is a highly specialized parasite with an evolutionary lineage of ~150 million years (Myr)[37], many important genes might have been lost.

Was the evolution of signalling pathways a prerequisite for the occurrence of animal multicellularity? The

© 2002 Nature Publishing Group

PROTISTS
Unicellular heterotrophic eukaryotes.

CNIDARIANS
A simple and ancient phylum of multicellular animals, such as jellyfish or corals, found mainly in marine environments.

PORIFERA
A phylum of multicellular animals with only two cell layers, the ectoderm and the endoderm, that are separated by an acellular mesogloea.

CHOANOFLAGELLATES
A group of protists that contain one flagellum at some stage of their life history.

PARAZOANS
An animal subkingdom that includes the porifera and the placozoa, the latter of which contains only one species (*Trichoplax*).

EUMETAZOANS
An animal subkingdom that includes the cnidarians, the ctenophorans and the bilaterians.

CYCLOSTOMES
A group of ancestral jawless fishes, including the lampreys.

GNATHOSTOMES
The group of higher fishes, all of which are characterized by the presence of jaws.

PSEUDOCOELOMATES
Animals, such as the nematode, that do not have a body cavity that is fully lined with mesodermal cells.

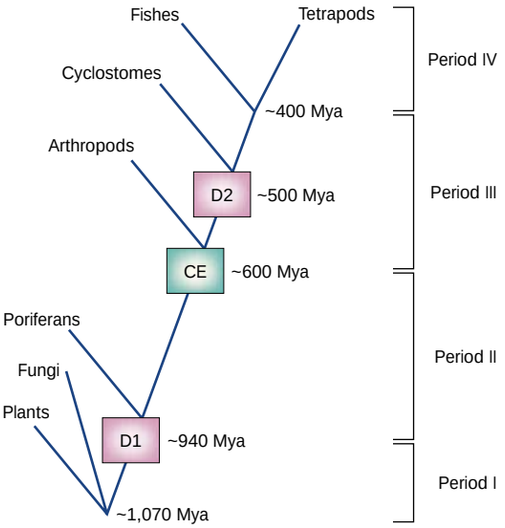


Figure 2 | **Gene duplication events do not correlate with the origin of the principal animal groups.** In the period of separation of plants, fungi and animals, extensive gene duplications have occurred (D1). No major duplication event was found in period II, which is the time between the parazoan–eumetazoan and the protostome–deuterostome splits. Only after the Cambrian explosion (CE), at the split of the cyclostome–gnathostome lineage, has another burst of gene duplications occurred (D2). Mya, million years ago. Modified with permission from REF. 41.

analysis of what is thought to be the last common protist ancestor of the metazoans could provide evidence for this hypothesis. On the basis of morphological and molecular studies, it has been indicated that the CHOANOFLAGELLATES might form a sister group of the metazoans[38,39]. Recent studies in the choanoflagellate *Monosiga brevicollis* identified the first receptor tyrosine kinase found outside the metazoans, indicating that the evolution of some signalling molecules might have preceded multicellularity[40]. Future research in this species might therefore help to provide a better picture of signalling systems in unicellular eukaryotes and might provide a link to the origin of the metazoans.

As far as the evolution of the principal animal groups is concerned, is there any correlation between the burst of gene duplications of signalling components and the phylogenetic diversification of these taxa? Some recent results indicate that extensive duplications of genes that encode signalling molecules might have occurred at least twice, once before the divergence of the PARAZOANS and the EUMETAZOANS (~1,070 Myr ago) and again around the divergence of the CYCLOSTOMES and the GNASTHOSTOMES (~500 Myr ago)[41] (FIG. 2). According to this analysis, the Cambrian explosion, the period ~520 Myr ago during which many phyla are thought to have originated was, however, not accompanied by gene duplication.

Finally, it should be noted that the interpretation of the pattern of signalling pathway evolution strongly depends on the phylogenetic placement of the animals under consideration. For example, some recent molecular studies indicate that nematodes and other PSEUDOCOELOMATES probably did not branch from the base of the bilaterian tree as in the traditional view, but form one CLADE (the Ecdysozoa) with moulting protostomes, such as *Drosophila*[42]. Given these new phylogenetic relationships, the absence of Hh and JAK/STAT signalling components in the *C. elegans* genome represents a gene loss, as these two pathways are present in *Drosophila* and in deuterostomes. So, the phylogenetic placement of organisms has important consequences on how the evolutionary pattern of signalling pathways is interpreted.

The post-genome era. The availability of whole-genome sequences of *Saccharomyces cerevisiae*, *C. elegans*, *Drosophila melanogaster* and the human marked a new era in the study of the evolution of signalling pathways. Complete genome sequences reveal a nearly full picture of the total number of signalling components (or their potential absence) in a given species. How many components of individual signalling pathways are there in worms, flies and mammals, and are there any general trends between the complexity of signalling pathways and body plans? Do all pathways evolve in similar ways, such as by an increase of paralogous proteins through gene duplications? It was observed early on that the protein families such as Wnt, TGF-β and RTK were greatly expanded in vertebrates relative to invertebrates[43–46]. For example, 29 ligands with TGF-β domains were found in humans, but only 6 in *Drosophila* and 4 in *C. elegans*[43–46] (TABLE 1). However, marked gene expansions are not unique to vertebrates; the *C. elegans* genome encodes at least 270 nuclear hormone receptors, whereas there are only 25 in *Drosophila* and 59 in humans[43–46]. In general, disparities in the abundance of certain protein families seem to be the rule rather than the exception. Therefore, the complexity of an organism cannot simply be deduced from gene numbers and the abundance of signalling pathways. Gene numbers can be affected markedly in one other way — as a result of whole-genome duplication. For example, there is now clear evidence that, during the evolution of vertebrates, there was a doubling or even quadrupling of the complete genome[34]. Several current studies are trying to evaluate whether genome duplication alone can account for the greater gene numbers seen in vertebrate species.

With every new eukaryotic genome being published, the understanding of the evolution of signalling pathways is enhanced substantially. All of the primary literature provides new insight into the evolution of gene families and functional units such as signalling pathways[43–46]. Although these genome sequences provide a platform for evolutionary analyses, two types of limitation restrict the findings solely on the basis of *in silico* analysis. The most important limitation in comparing genomes and their proteins on a one-to-one level stems from the modular nature of proteins, most of which contain two or more domains. These domains can occur in different combinations in different proteins and organisms. So, it is often impossible

© 2002 Nature Publishing Group

Table 1 | Numbers of signalling molecules in selected pathways

Signalling molecules	Species			
	Human	Fly	Worm	Yeast
Ligand				
RTK	48	3	4	0
TGF-β	29	6	4	0
Wnt	18	7	5	0
Notch	3	2	2	0
STAT	7	1	1	0
Receptor				
RTK	25	6	1	0
Wnt	12	6	5	0
NHR	59	25	270	1

NHR, nuclear hormone receptor; RTK, receptor tyrosine kinase; STAT, signal transducer and activator of transcription; TGF-β, transforming growth factor-β; Wnt, wingless related. The table contains selected entries from REF. 44.

to decide what the function of individual proteins might be solely based on *in silico* analysis. Many components of signalling pathways are multi-domain proteins, such as receptors that cross the cell membrane and that combine extracellular domains involved in ligand binding and also specific inhibitors, as well as intracellular domains involved in signal transduction in the cell. It is interesting to note that, when compared with *C. elegans*, the *Drosophila* genome, and to an even greater extent the human genome, contain higher numbers of multi-domain proteins[43,44,46]. However, a proper assignment of most of these proteins awaits functional analysis. In some cases, studies of genes, which at the sequence level resemble signalling molecules, did not provide any evidence for their involvement in signalling processes. In *C. elegans*, for example, 28 RTK receptors were found in the genome, but only 6 of them can be shown genetically to be involved in RTK signalling[47].

A second limitation stems from the ability of gene-prediction programs to detect divergent gene sequences. An important recent example is the *C. elegans pry-1* (*polyray-1*) gene. PRY-1 is a component of the WNT pathway with weak sequence similarity to vertebrate axin. Although the sequence similarity is limited, *pry-1* is functionally equivalent to vertebrate axins and can substitute for the zebrafish gene *masterblind*[48].

Overall, the *in silico* comparison of the evolution of signalling pathways allows several important conclusions. First, the seven signalling systems that are present in higher animals are older than the bilaterians and might, to a large extent, also be older than the metazoans. Therefore, the evolution of these signalling systems might have been a prerequisite for the evolution of animal multicellularity. At the same time, the exact role of the evolution of signalling pathways for the emergence of multicellularity and for the signalling repertoire of the phylogenetically important protists remains unknown. Second, metazoan phyla differ in their number of signalling genes. Nonetheless, the importance of the burst of gene duplications in the evolution of new body plans is not known, as morphological complexity is often unrelated to the number of signalling components found in a given organism. Third, the availability of whole-genome sequences provides a framework for the analysis of pathway evolution. The combination of experimental and theoretical studies will allow important insights into the network properties of signalling systems and into these properties for the evolution of complexity.

CLADE
A taxon or other group of organisms that share a closer common ancestor with one another than with members of any other clade.

Co-option of signalling pathways

From what we have described so far, it is evident that, despite the small number of signalling pathways, these pathways are flexible enough to be used in generating morphologies as diverse as worms and sea urchins. Even during the formation of completely new body plans, such as the secondary radial symmetry seen in the echinoderm lineage, the same developmental pathways have been recruited to specify new cell types, tissues and morphologies. This process of re-using the existing genetic units has been termed 'co-option'[49]. Co-option was originally recognized as a general principle for the evolution of transcription factors, such as *hairy*, *engrailed* and others that are involved in *Drosophila* segmentation[49]; however, it also holds true for signalling pathways.

The use of signalling cascades to generate new structures can be exemplified by the formation of wing eyespots in butterflies (for a review, see REF. 50). Hh signalling is used in the *Drosophila* wing disc to organize anterior–posterior patterning[3], a function that is apparently conserved in the butterfly *Precis coenia*[51] (FIG. 3). In addition to this ancestral function, Hh signalling has been recruited in the butterfly wing to cells that form the eyespot foci, an evolutionary novelty among insects[51]. Specifically, the activation of Hh signalling components, such as the Hh receptor encoded by *patched* and the signal transducer and transcription factor encoded by *cubitus interruptus* occurs in the posterior compartment of the butterfly wing disc in cells that are adjacent to those that express Hh (FIG. 3). These observations support the view that cells receiving Hh signalling will differentiate into eyespot foci. But, how are entire pathways co-opted into new developmental processes? In the case of the butterfly eyespots, it has been suggested that, once repression (in this case by the Engrailed protein) is relieved, the components of the signalling cascade are transcribed, which results in the activation of the Hh pathway[51]. Increasing transcription levels of one of the signalling components might be enough to regulate transcription of other genes from the same pathway[52]. However, the simple activation of a cascade does not necessarily result in the activation of specific target genes. How Hh signalling forces the formation of an eyespot instead of something else remains to be understood.

The turtle shell is a second example of a morphological novelty that is caused by the HETEROTOPIC expression of a signalling pathway. The turtle shell is not homologous to any other structure in other reptiles and is induced by the outgrowth of the dorsal body wall. The induction of

© 2002 Nature Publishing Group

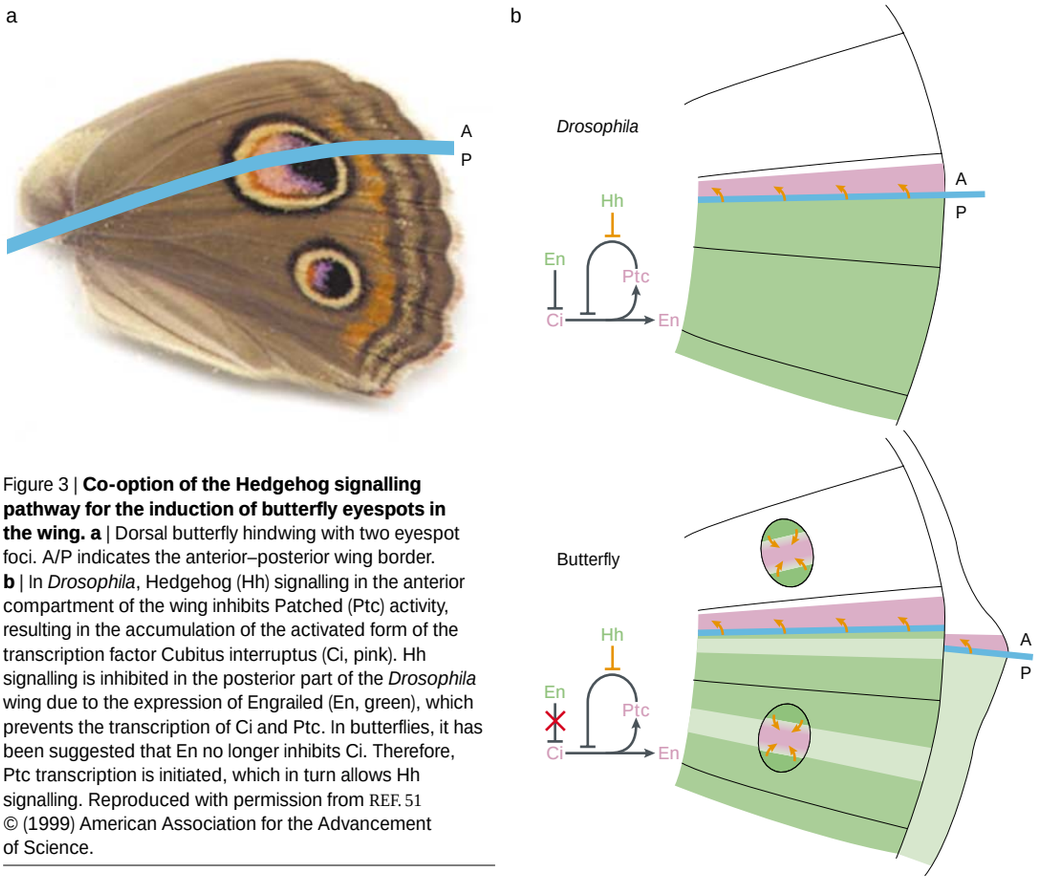


Figure 3 | **Co-option of the Hedgehog signalling pathway for the induction of butterfly eyespots in the wing. a** | Dorsal butterfly hindwing with two eyespot foci. A/P indicates the anterior–posterior wing border. **b** | In *Drosophila*, Hedgehog (Hh) signalling in the anterior compartment of the wing inhibits Patched (Ptc) activity, resulting in the accumulation of the activated form of the transcription factor Cubitus interruptus (Ci, pink). Hh signalling is inhibited in the posterior part of the *Drosophila* wing due to the expression of Engrailed (En, green), which prevents the transcription of Ci and Ptc. In butterflies, it has been suggested that En no longer inhibits Ci. Therefore, Ptc transcription is initiated, which in turn allows Hh signalling. Reproduced with permission from REF. 51 © (1999) American Association for the Advancement of Science.

the shell outgrowth correlates with the expression of fibroblast growth factor 10 (FGF-10, a component of the RTK pathway)[53], a molecule that is important in the induction of another outgrowth, the limb[54].

The examples of the butterfly eyespot and the turtle shell indicate that the co-option of signalling pathways is involved in the generation of morphological novelties. More generally, it has to be emphasized that all developmental processes that are involved in the generation of new structures require co-option events. Whenever signalling pathways were involved in such developmental processes, these pathways also had to be co-opted. So, the co-option of signalling pathways is a common principle in animal evolution. However, it is important to note that the finding of co-option itself does not "provide a clear picture of the precise [molecular] mechanism involved"[55]. The formation of new regulatory linkages can occur in many ways by different types of mutation. But, what is the nature of those mutations that change the expression of signalling components, and how many are necessary to recruit a signalling pathway to new tissues? Although it might still be too early to draw any firm conclusions, a larger body of work has begun to indicate that many 'evolutionary mutations' are caused by alterations in the promoter and enhancer regions of regulatory genes or their downstream targets[56]. It has to be stressed, however, that most of these studies concentrate on the evolution of transcription factors rather than signalling components[57]. It will be interesting to see if future work indicates similar patterns for the components of signalling pathways.

In contrast to transcription factors, many components of signalling pathways act together and have to be co-expressed to regulate the development of cells and tissues. One interesting observation in this context is the existence of synexpression groups[58]. Synexpression groups refer to sets of genes that are co-activated in similar expression patterns and that might function in the same process. One example is TGF-β signalling in *Xenopus*, in which seven genes (*Bone morphogenic protein receptor* II (*Bmp-2*), *Smad-6*, *Smad-7*, *XVent 2*, *Bmp-4*, *Bambi* and *Bmp-7*) are co-transcribed in the dorsal part of the eyes, the ventral branchial arches and the posterior dorsal region of the fin[59–62]. A simple mechanism for the activation of a set of genes in a complex spatial pattern is to regulate their transcription by a transcription factor that binds to common promoter elements. *Cis*-regulatory elements have a modular structure; that is, independent elements can act and evolve independently of each other[63]. Therefore, duplications of enhancer elements can result in the novel tissue-specific transcription of an entire signalling pathway. Alternatively, the regulation of a transcription factor can be altered by duplication and divergence of its enhancer.

HETEROTOPY
The displacement of the development of an organ in space.

© 2002 Nature Publishing Group

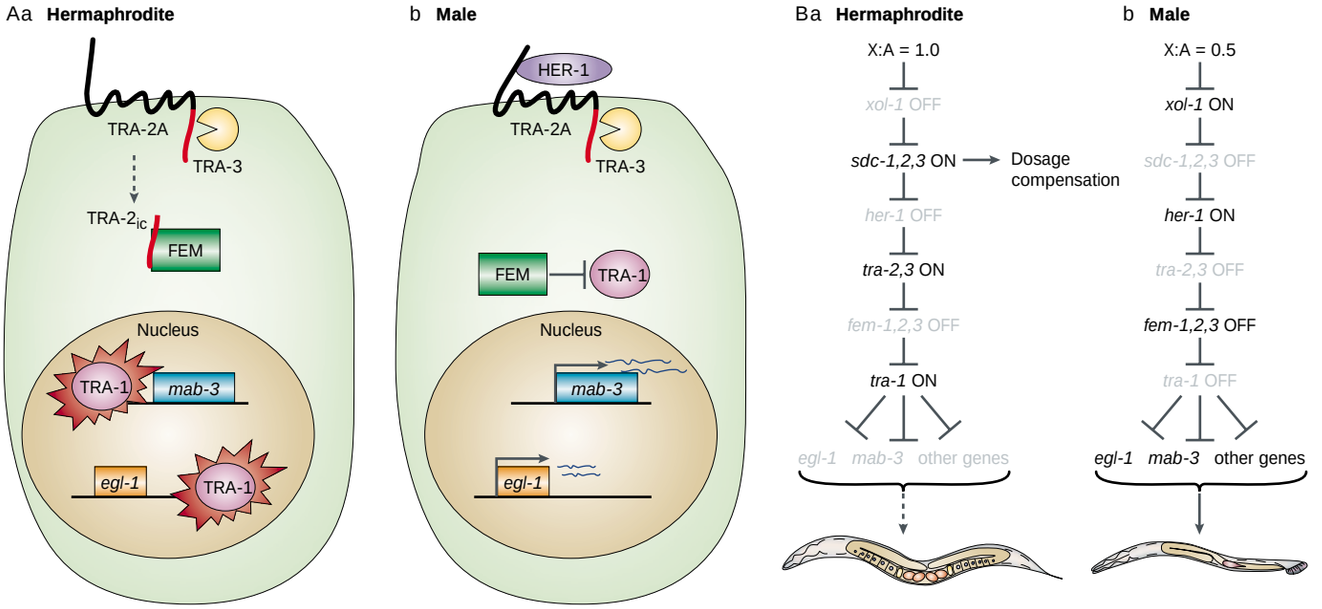


Figure 4 | **Somatic sex determination in *Caenorhabditis elegans*. A** | Molecular model. **Aa** | In hermaphrodites, the activity of the zinc-finger-containing transcription factor TRA-1 is released by the inhibition of feminization (FEM) proteins and is able to activate the transcription of its targets. FEM proteins are inhibited by interacting with the intracellular domain of the membrane protein TRA-2 ($TRA\text{-}2_{IC}$), which is cleaved by the calpain protease TRA-3. **Ab** | In males, a novel protein hermaphrodite-1 (HER-1) binds to TRA-2, thereby preventing the cleavage of $TRA\text{-}2_{IC}$ by TRA-3. FEM proteins are therefore able to inhibit TRA-1 activity. **B** | Genetic model. A series of negative regulatory interactions triggered by the ratio of the set of sex chromosomes to autosomes (X:A) results in high TRA-1 activity in hermaphrodites and low TRA-1 activity in males. TRA-1 regulates the transcription of various sex-specific genes, such as *egg-laying defective-1* (*egl-1*) and *male abnormal-3* (*mab-3*). The genes *sdc-1*, *sdc-2* and *sdc-3* control DOSAGE COMPENSATION as well as sex determination. Part **A** reproduced with permission from REF. 97 © (2002) Elsevier Science, part **B** reproduced with permission from REF. 98 © (2001) Macmillan Magazines Ltd.

In summary, the co-option of signalling pathways and transcription factors represents a principle of animal evolution. Future selected case studies will hopefully provide a comprehensive picture of the mechanisms that surround co-option events.

Searching for case studies

So far, we have examined the evolution of signalling pathways from two points of view. From molecular and morphological perspectives, we have looked at the signalling gene repertoire in present-day organisms and we have analysed how the evolution of signalling might be involved in creating morphological novelty. In between, however, are the pathways themselves. One limitation of our current understanding of signalling pathways in any given species is that, whereas the core pathways might be similar, it remains unclear how many more components (often species and/or cell specific) exist. The question of how important such factors are for the evolution of signalling pathways and for the generation of morphological novelty is an interesting one, and brings us to the question: how do the pathways themselves evolve? Understanding the logic of the genetic, molecular and biochemical activities in greater detail, this problem can start to be addressed. Although there are no good case studies for any of the 'canonical' signalling pathways, sex determination is an informative substitute as it shares many of the characteristics of the seven pathways that have been discussed so far. Below, we consider nematode sex determination as a case study and discuss the challenges in understanding how signalling pathways evolve.

Sex determination has long intrigued evolutionists, geneticists and developmental biologist. A large body of evidence indicates that the process evolves rapidly (for reviews, see REFS 64,65). In *Drosophila* and *C. elegans*, sex determination is triggered by differences in the ratio of the set of X chromosomes to autosomes. Animals with two X chromosomes develop as females or, in the case of *C. elegans*, as self-fertilizing hermaphrodites, but develop as males if only one X chromosome is present. Genetic and molecular studies revealed that *C. elegans* sex determination relies on a signalling pathway. Genetically, it involves a cascade of negatively acting factors that, at a molecular level, consists of a signalling pathway that triggers the zinc-finger transcription factor transformer-1 (TRA-1)[66] (FIG. 4).

To understand the logic of *C. elegans* somatic sex determination, we have to take a closer look at the molecular mechanism (FIG. 4). TRA-1 has been shown to act as a repressor of male characteristics by binding to the promoter of two male-specific genes in the hermaphrodite[67,68]. So, the male fate can be regarded as the ground state that is executed if TRA-1 is absent. But why is TRA-1 only inhibiting male-specific genes

DOSAGE COMPENSATION
A mechanism that regulates the expression of sex-linked genes that differ in dose between females and males.

© 2002 Nature Publishing Group

in hermaphrodites and not in males? TRA-1 activity is regulated by a signalling cascade: in males, the hermaphrodite-1 (HER-1) protein binds the transmembrane protein TRA-2 extracellularly and thereby inhibits the binding and processing of the intracellular domain of TRA-2 by TRA-3 (REFS 69–71). The *tra-3* gene encodes a protease that processes the intracellular domain of TRA-2 only in the hermaphrodite[71]. As the intracellular domain of TRA-2 is not processed in the male, the *fem* gene products are active and inhibit the function of the transcription factor TRA-1 (REF. 72). In the hermaphrodite, *her-1* is not expressed. Therefore, the intracellular domain of TRA-2 is processed by TRA-3 and inhibits the function of the FEM proteins. As a result, the TRA-1 protein is active and inhibits the male fate in the hermaphrodite. This sophisticated biochemical pathway immediately raised the question of its evolutionary origin. In addition, comparative studies in closely related nematodes indicate that sex determination genes evolve rapidly, showing low sequence similarity in orthologues of *C. elegans*, *C. briggsae* and *C. remanei*[73–77]. However, despite significant sequence divergence, the *C. briggsae* and *C. remanei* orthologues have crucial roles in somatic sex determination in these species[73–78].

Recent findings provided a new twist to the evolution of sex determination. Despite the fast evolution of most molecules in the sex determination machinery, other sex determination genes are relatively conserved. *Drosophila* and *C. elegans* share genes involved in sex determination that are both structurally and functionally similar[79]. *C. elegans male abnormal-3* (*mab-3*) and *Drosophila doublesex* (*dsx*) mutants fail to develop certain sex-specific structures. Like *dsx* in *Drosophila*, epistatic analysis of *mab-3* has shown that it is the most downstream gene in the cascade[80]. Both genes encode proteins that contain a non-classical zinc-finger DNA-binding motif named the DM (Doublesex and Mab-3) domain[79,81]. Both the male-specific isoform of Dsx (Dsx^M) in *Drosophila* and that of MAB-3 in *C. elegans* repress transcription of genes that encode proteins of similar function (but with different molecular structure), namely the yolk proteins[68,82,83]. They are also involved in the development of male sensory structures (sex combs in *Drosophila* and rays in *C. elegans*) and in mating behaviour. More recently, on the basis of expression-pattern and mutational analyses, vertebrate DM-containing genes have been implicated in sexual dimorphism as well[84–88].

Given the framework of fast-evolving sex determination genes and of the sophisticated signalling mechanism, Wilkins[89] proposed a scenario that explains how this diverse mechanism in *C. elegans* might have evolved. The central argument in his hypothesis is that the pathway evolved in reverse order from the final step in the cascade, the transcription factor, up to the first step, the receptor and the ligand. With regard to the biochemical mechanism, a bottom-up evolution can, for example, proceed by the sequential addition of upstream components that regulate the activity of the transcription factor, resulting in the mechanism that is seen in *C. elegans*. However, from an evolutionary perspective, it is much more complicated to imagine how other components were added upstream. In particular, how selection favours the addition of upstream molecules remains unknown. Wilkins[89] proposed that, under sub-optimal sex ratios, frequency-dependent selection for the minority sex drove the acquisition of new genetic switches. These switches would than act as NEOMORPHIC and dominant-negative units that reverse the function of the previous step.

The evolution of nematode sex determination is an excellent case study for understanding how signalling pathways evolve *per se*. Wilkins' hypothesis, the bottom-up evolution of *C. elegans* sex determination, is testable by studying other nematodes. According to this hypothesis, the downstream transcription factors TRA-1 and MAB-3 should be the most functionally conserved components of the signalling system. Therefore, the transcription factors, but not the upstream components, might be involved in sex determination of nematodes that are distantly related to *C. elegans*. The fact that sex determination genes, such as *tra-1*, *tra-2*, *tra-3* and *feminization-2* (*fem-2*), evolved fast at the sequence level does not directly imply that the functional relationship of the components changed with a similar speed. Recent studies indicate that the co-evolution of individual pathway components might, in part, be responsible for the fast evolution of sex determination genes in the genus *Caenorhabditis*[78,90]. Therefore, only a study of more distantly related species, for which genetic tools are available that are comparable with those of *C. elegans*, can answer the question of how complex signalling pathways might have evolved.

Conclusions

We have addressed three aspects of the evolution of signalling pathways; namely, the genetic repertoire seen in present-day species, the importance of the co-option of signalling pathways for the generation of morphological novelty and the evolution of signalling pathways *per se*. It is evident that, in all of these aspects, our understanding is far from complete. Nonetheless, substantial progress has been made in the past ten years, providing fascinating insights into the evolution of animal development. Evolutionary developmental biology, which is not restricted to signalling pathways, is starting to offer an enhanced picture of the molecular repertoire of the last common ancestor of multicellular and bilaterian animals.

Many interesting new questions are associated with these observations. Is there any correlation between the number of signalling genes and the diversification of body plans? Does the number of types of signalling pathway hamper morphological evolution? Is the flexibility of signalling systems sufficient to explain the novelties of body plans? How are new components integrated into existing networks, and how does this change the behaviour of a signalling network? How do signalling systems really evolve at the microevolutionary level, that is, what type of mutations occur and how are such changes

NEOMORPHIC
A qualitatively new feature of a phenotype that is produced by a mutant allele.

© 2002 Nature Publishing Group

fixed in natural populations? And finally, why do some pathways evolve faster than others? We have reviewed studies that are beginning to provide answers to some of these questions. Although we are far from answering them completely, there is light at the end of the tunnel: whole-genome analysis, in combination with the generation of functional toolkits for some non-model organisms in informative phylogenetic positions (such as non-dipteran insects and nematodes other than *C. elegans*), offers the opportunity to address several of these questions in the years to come.

The evolutionary analysis of animal development is still at its beginning and, therefore, there are more open questions than solid answers. With several case studies on their way, the emergence of general patterns is, however, only a question of time. The conservation of signalling pathways was only established as a general concept after developmental processes had been studied in flies, worms and vertebrates. In a similar way, general principles in evolutionary developmental biology will emerge after a sophisticated analysis of several case studies has been carried out in selected animals.

1. Gilbert, S. F. & Bolker, J. A. in *Homologies of Process and Modular Elements of Embryonic Construction* (ed. Wagner, G. P.) 435–454 (Academic, San Diego, California, 2001).
2. Spemann, H. & Mangold, H. Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *W. Roux. Arch. Entw. Organ.* **100**, 599–638 (1924) (in German).
3. Ingham, P. W. & McMahon, A. P. Hedgehog signalling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059–3087 (2001).
4. Cadigan, K. M. & Nusse, R. Wnt signalling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
5. Moon, R. T., Bowerman, B., Boutros, M. & Perrimon, N. The promise and perils of Wnt signalling through β-catenin. *Science* **296**, 1644–1646 (2002).
6. Massague, J. & Chen, Y. G. Controlling TGF-β signalling. *Genes Dev.* **14**, 627–644 (2000).
7. Mumm, J. S. & Kopan, R. Notch signalling: from the outside in. *Dev. Biol.* **228**, 151–165 (2000).
8. Castelli-Gair Hombria, J. & Brown, S. The fertile field of *Drosophila* JAK/STAT signalling. *Curr. Biol.* **12**, R569 (2002).
9. McKenna, N. J. & O'Malley, B. W. Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–474 (2002).
10. Gerhart, J. 1998 Warkany lecture: signaling pathways in development. *Teratology* **60**, 226–239 (1999).
11. Tan, P. B. & Kim, S. K. Signalling specificity: the RTK/RAS/MAP kinase pathway in metazoans. *Trends Genet.* **15**, 145–149 (1999).
12. Dumont, J. E., Pecasse, F. & Maenhaut, C. Crosstalk and specificity in signalling. Are we crosstalking ourselves into general confusion? *Cell Signal.* **13**, 457–463 (2001).
13. Clandinin, T. R., DeModena, J. A. & Sternberg, P. W. Inositol trisphosphate mediates a RAS-independent response to LET-23 receptor tyrosine kinase activation in *C. elegans*. *Cell* **92**, 523–533 (1998).
14. Aroian, R. V. & Sternberg, P. W. Multiple functions of *let-23*, a *Caenorhabditis elegans* receptor tyrosine kinase gene required for vulval induction. *Genetics* **128**, 251–267 (1991).
15. Rulifson, E. J., Wu, C. H. & Nusse, R. Pathway specificity by the bifunctional receptor Frizzled is determined by affinity for Wingless. *Mol. Cell* **6**, 117–126 (2000).
16. Halfon, M. S. *et al.* Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74 (2000).
17. Tan, P. B., Lackner, M. R. & Kim, S. K. MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31 WH transcription factor complex during *Caenorhabditis elegans* vulval induction. *Cell* **93**, 569–580 (1998).
18. De Robertis, E. M., Larrain, J., Oelgeschlager, M. & Wessely, O. The establishment of Spemann's organizer and patterning of the vertebrate embryo. *Nature Rev. Genet.* **1**, 171–181 (2000).
19. Cohen, P. & Frame, S. The renaissance of GSK3. *Nature Rev. Mol. Cell Biol.* **2**, 769–776 (2001).
20. Dajani, R. *et al.* Crystal structure of glycogen synthase kinase 3β structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* **105**, 721–732 (2001).
21. Frame, S., Cohen, P. & Biondi, R. M. A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation. *Mol. Cell* **7**, 1321–1327 (2001).
22. von Dassow, G., Meir, E., Munro, E. M. & Odell, G. M. The segment polarity network is a robust developmental module. *Nature* **406**, 188–192 (2000).
23. Meir, E., von Dassow, G., Munro, E. & Odell, G. M. Robustness, flexibility, and the role of lateral inhibition in the neurogenic network. *Curr. Biol.* **12**, 778–786 (2002).
24. Freeman, M. Feedback control of intercellular signalling in development. *Nature* **408**, 313–319 (2000).
25. Meinhardt, H. & Gierer, A. Pattern formation by local self-activation and lateral inhibition. *Bioessays* **22**, 753–760 (2000).
26. Bienz, M. & Tremml, G. Domain of *Ultrabithorax* expression in *Drosophila* visceral mesoderm from autoregulation and exclusion. *Nature* **333**, 576–578 (1988).
27. Thuringer, F. & Bienz, M. Indirect autoregulation of a homeotic *Drosophila* gene mediated by extracellular signaling. *Proc. Natl Acad. Sci. USA* **90**, 3899–3903 (1993).
28. Chen, Y. & Struhl, G. Dual roles for Patched in sequestering and transducing Hedgehog. *Cell* **87**, 553–563 (1996).
29. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
30. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
31. Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002).
32. Schulte-Merker, S., Lee, K. J., McMahon, A. P. & Hammerschmidt, M. The zebrafish organizer requires *chordino*. *Nature* **387**, 862–863 (1997).
33. Fisher, S. & Halpern, M. E. Patterning the zebrafish axial skeleton requires early *chordin* function. *Nature Genet.* **23**, 442–446 (1999).
34. Prince, V. E. & Pickett, F. B. Splitting pairs: the diverging fates of duplicated genes. *Nature Rev. Genet.* **3**, 827–837 (2002).
35. Steele, R. E. Developmental signaling in *Hydra*: what does it take to build a "simple" animal? *Dev. Biol.* **248**, 199–219 (2002).
36. McCarty, D. R. & Chory, J. Conservation and innovation in plant signaling pathways. *Cell* **3**, 201–209 (2000).
37. Carlton, J. M. *et al.* Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii*. *Nature* **419**, 512–519 (2002).
38. Lang, B. F., O'Kelly, C., Nerad, T., Gray, M. W. & Burger, G. The closest unicellular relatives of animals. *Curr. Biol.* **12**, 1773–1778 (2002).
A molecular phylogenetic study that confirms the earlier morphological proposal (see reference 39) that metazoans evolved from a choanoflagellate-like ancestor.
39. Clark, J. On the spongiae ciliatae as infusoria flagellata; or observations on the structure, animality and relationship of *Leucosolenia botryoides*, Bowerbank. *Annu. Mag. Nat. Hist.* **1**, 133–142 (1868).
40. King, N. & Carroll, S. B. A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. *Proc. Natl Acad. Sci. USA* **98**, 15032–15037 (2001).
41. Miyata, T. & Suga, H. Divergence pattern of animal gene families and relationship with the Cambrian explosion. *Bioessays* **23**, 1018–1027 (2001).
42. Aguinaldo, A. M. *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489–493 (1997).
43. The International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
44. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
45. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
46. Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).
47. Ruvkun, G. & Hobert, O. The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* **282**, 2033–2041 (1998).
48. Korswagen, H. C. *et al.* The axin-like protein PRY-1 is a negative regulator of a canonical Wnt pathway in *C. elegans*. *Genes Dev.* **16**, 1291–1302 (2002).
The authors show that, despite low sequence similarity, *Caenorhabditis elegans* PRY-1 can rescue a zebrafish *axin* mutant. This has implications for genomic studies, in which sequencing annotation must be complemented with functional studies.
49. Raff, R. A. *The Shape of Life: Genes, Development and the Evolution of Animal Form* (Chicago Univ. Press, Illinois, 1996).
50. Beldade, P. & Brakefield, P. M. The genetics and evo–devo of butterfly wing patterns. *Nature Rev. Genet.* **3**, 442–452 (2002).
51. Keys, D. N. *et al.* Recruitment of a Hedgehog regulatory circuit in butterfly eyespot evolution. *Science* **283**, 532–534 (1999).
52. Hepker, J., Wang, Q. T., Motzny, C. K., Holmgren, R. & Orenic, T. V. *Drosophila cubitus interruptus* forms a negative feedback loop with *patched* and regulates expression of Hedgehog target genes. *Development* **124**, 549–558 (1997).
53. Loredo, G. A. *et al.* Development of an evolutionarily novel structure: fibroblast growth factor expression in the carapacial ridge of turtle embryos. *J. Exp. Zool.* **291**, 274–281 (2001).
Together with reference 51, this study indicates that the recruitment of a signalling pathway into an ectopic position is involved in the origin of a morphological novelty.
54. Sekine, K. *et al.* Fgf10 is essential for limb and lung formation. *Nature Genet.* **21**, 138–141 (1999).
55. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer Associates, Sunderland, Massachusetts, 2002).
56. Davidson, E. H. *Genomic Regulatory Systems* (Academic, San Diego, California, 2001).
57. Tautz, D. Evolution of transcriptional regulation. *Curr. Opin. Genet. Dev.* **10**, 575–579 (2000).
58. Niehrs, C. & Pollet, N. Synexpression groups in eukaryotes. *Nature* **402**, 483–487 (1999).
An insightful review that proposes a mechanism for the rapid recruitment of entire signalling pathways to new tissues.
59. Gawantka, V. *et al.* Gene expression screening in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech. Dev.* **77**, 95–141 (1998).
60. Bhushan, A., Chen, Y. & Vale, W. Smad7 inhibits mesoderm formation and promotes neural cell fate in *Xenopus* embryos. *Dev. Biol.* **200**, 260–268 (1998).
61. Frisch, A. & Wright, C. V. XBMPRII, a novel *Xenopus* type II receptor mediating BMP signaling in embryonic tissues. *Development* **125**, 431–442 (1998).
62. Hata, A., Lagna, G., Massague, J. & Hemmati-Brivanlou, A. Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev.* **12**, 186–197 (1998).
63. Arnone, M. I. & Davidson, E. H. The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**, 1851–1864 (1997).

© 2002 Nature Publishing Group

64. Bull, J. J. *Evolution of Sex Determining Mechanisms* (Benjamin/Cummings, Menlo Park, California, 1983).
65. Marin, I. & Baker, B. S. The evolutionary dynamics of sex determination. *Science* **281**, 1990–1994 (1998).
66. Zarkower, D. & Hodgkin, J. Molecular analysis of the *C. elegans* sex-determining gene *tra-1*: a gene encoding two zinc finger proteins. *Cell* **70**, 237–249 (1992).
67. Conradt, B. & Horvitz, H. R. The TRA-1A sex determination protein of *C. elegans* regulates sexually dimorphic cell deaths by repressing the *egl-1* cell death activator gene. *Cell* **98**, 317–327 (1999).
68. Yi, W., Ross, J. M. & Zarkower, D. *mab-3* is a direct *tra-1* target gene regulating diverse aspects of *C. elegans* male sexual development and behavior. *Development* **127**, 4469–4480 (2000).
69. Perry, M. D. *et al.* Molecular characterization of the *her-1* gene suggests a direct role in cell signaling during *Caenorhabditis elegans* sex determination. *Genes Dev.* **7**, 216–228 (1993).
70. Kuwabara, P. E., Okkema, P. G. & Kimble, J. *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway. *Mol. Biol. Cell* **3**, 461–473 (1992).
71. Sokol, S. B. & Kuwabara, P. E. Proteolysis in *Caenorhabditis elegans* sex determination: cleavage of TRA-2A by TRA-3. *Genes Dev*. **14**, 901–906 (2000).
72. de Bono, M., Zarkower, D. & Hodgkin, J. Dominant feminizing mutations implicate protein–protein interactions as the main mode of regulation of the nematode sex-determining gene *tra-1*. *Genes Dev.* **9**, 155–167 (1995).
73. de Bono, M. & Hodgkin, J. Evolution of sex determination in *Caenorhabditis*: unusually high divergence of *tra-1* and its functional consequences. *Genetics* **144**, 587–595 (1996).
74. Haag, E. S. & Kimble, J. Regulatory elements required for development of *Caenorhabditis elegans* hermaphrodites are conserved in the tra-2 homologue of *C. Remanei*, a male/female sister species. *Genetics* **155**, 105–116 (2000).
75. Hansen, D. & Pilgrim, D. Molecular evolution of a sex determination protein. FEM-2 (pp2c) in *Caenorhabditis*. *Genetics* **149**, 1353–1362 (1998).
76. Chen, P. J., Cho, S., Jin, S. W. & Ellis, R. E. Specification of germ cell fates by *fog-3* has been conserved during nematode evolution. *Genetics* **158**, 1513–1525 (2001).
77. Stothard, P., Hansen, D. & Pilgrim, D. Evolution of the PP2C family in *Caenorhabditis*: rapid divergence of the sex-determining protein FEM-2. *J. Mol. Evol.* **54**, 267–282 (2002).
78. Haag, E. S., Wang, S. & Kimble, J. Rapid coevolution of the nematode sex-determining genes *fem-3* and *tra-2*. *Curr. Biol.* **12**, 2035–2041 (2002).
Despite the functional conservation of FEM-3 and TRA-2 in *Caenorhabditis* sex determination, the domain of interaction between these two proteins is evolving rapidly at the sequence level. This study shows that TRA-2 and FEM-3 can interact in three species of nematodes, but that this interaction is species specific.
79. Raymond, C. S. *et al.* Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691–695 (1998).
80. Shen, M. M. & Hodgkin, J. *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* **54**, 1019–1031 (1988).
81. Erdman, S. E. & Burtis, K. C. The *Drosophila* Doublesex proteins share a novel zinc finger related DNA binding domain. *EMBO J.* **12**, 527–535 (1993).
82. Coschigano, K. T. & Wensink, P. C. Sex-specific transcriptional regulation by the male and female Doublesex proteins of *Drosophila*. *Genes Dev.* **7**, 42–54 (1993).
83. An, W. & Wensink, P. C. Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer. *Genes Dev.* **9**, 256–266 (1995).
84. Raymond, C. S., Kettlewell, J. R., Hirsch, B., Bardwell, V. J. & Zarkower, D. Expression of *Dmrt1* in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev. Biol.* **215**, 208–220 (1999).
A landmark paper describing a gene that is in common in the sex determination pathway of distantly related phyla.
85. Matsuda, M. *et al. DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* **417**, 559–563 (2002).
86. Guan, G., Kobayashi, T. & Nagahama, Y. Sexually dimorphic expression of two types of DM (*Doublesex*/*Mab-3*)-domain genes in a teleost fish, the tilapia (*Oreochromis niloticus*). *Biochem. Biophys. Res. Commun.* **272**, 662–666 (2000).
87. Kettlewell, J. R., Raymond, C. S. & Zarkower, D. Temperature-dependent expression of turtle *Dmrt1* prior to sexual differentiation. *Genesis* **26**, 174–178 (2000).
88. De Grandi, A. *et al.* The expression pattern of a mouse *doublesex*-related gene is consistent with a role in gonadal differentiation. *Mech. Dev.* **90**, 323–326 (2000).
89. Wilkins, A. S. Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *Bioessays* **17**, 71–77 (1995).
The author proposes a mechanism for the evolution of a regulatory cascade of inhibitory interactions from bottom to top.
90. Wang, S. & Kimble, J. The TRA-1 transcription factor binds TRA-2 to regulate sexual fates in *Caenorhabditis elegans*. *EMBO J.* **20**, 1363–1372 (2001).
91. Barolo, S. & Posakony, J. W. Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev.* **16**, 1167–1181 (2002).
92. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
93. Lynch, M. & Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473 (2000).
94. Nadeau, J. H. & Sankoff, D. Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* **147**, 1259–1266 (1997).
95. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
In this paper, the authors present an interesting mechanism for the retention of duplicated genes.
96. Fryxell, K. J. The coevolution of gene family trees. *Trends Genet.* **12**, 364–369 (1996).
97. Goodwin, E. B. & Ellis, R. E. Turning clustering loops: sex determination in *Caenorhabditis elegans*. *Curr. Biol.* **12**, 111–120 (2002).
98. Zarkower, D. Establishing sexual dimorphism: conservation amidst diversity? *Nature Rev. Genet.* **2**, 175–185 (2001).

Acknowledgements

We apologize to those scientists whose work has not been cited due to space restrictions. We thank D. Rudel and J. Srinivasan for critically reading the manuscript.

Online links

DATABASES

The following terms in this article are linked online to:

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink

axin | *dsx* | *engrailed* | *eve* | Fz | Fz2 | Gli/Ci | *hairy* | Hh | JAK | Mad | Patched | Pnt | STAT | Su(H) | Tcf/Lef | TGF-β | *Ubx* | Wg | Wnt

WormBase: http://www.wormbase.org

β-catenin | *egl-1* | *fem-2* | GSK-3 | HER-1 | JNK | LET-23 | LIN-31 | *mab-3* | *pry-1* | *sdc-1* | *sdc-2* | *sdc-3* | TRA-1 | TRA-2 | TRA-3

ZFIN: http://zfin.org

masterblind

FURTHER INFORMATION

Ralf J. Sommer's lab: http://www.eb.tuebingen.mpg.de

Access to this interactive links box is free online.

© 2002 Nature Publishing Group