The future of evolutionary developmental biology

Peter W. H. Holland

Combining fields as diverse as comparative embryology, palaeontology, molecular phylogenetics and genome analysis, the new discipline of evolutionary developmental biology aims at explaining how developmental processes and mechanisms become modified during evolution, and how these modifications produce changes in animal morphology and body plans. In the next century this should give us far greater mechanistic insight into how evolution has produced the vast diversity of living organisms, past and present.

Evolutionary biology and developmental biology, or embryology, have had a stormy relationship over the past hundred years. At the end of the nineteenth century and the start of the twentieth, they were inseparable; comparisons of the embryonic development of different species were used as evidence for evolution, while evolutionary history was seen as sufficient explanation for almost every structure or process observed in animal development^{1,2}. But as evolutionary theory embraced genetics in the 1920s and 1930s, the study of embryonic development was largely rejected as being insufficiently precise or quantitative to contribute to an increasingly rigorous science.

The past 15 years have seen a timely reconciliation between the two fields, with the vibrant new discipline of evolutionary developmental biology emerging at the interface. This concerns itself with how developmental processes themselves have evolved: how they can be modified by genetic change, and how such modifications produce the past and present diversity of morphologies and body plans. Three main factors have contributed to the emergence and phenomenal growth of evolutionary developmental biology. Ironically, all three depend on genetics — the discipline that split evolution and development apart 60 years earlier.

Man is but a worm?

The first factor, and arguably the most important, was the discovery that animals as different as nematodes, flies and mammals use similar genes for similar developmental purposes, such as controlling the development of spatial organization in the embryo. The ball started rolling with the discovery in 1984 of a shared DNA sequence motif — the homeobox — in a variety of genes that control development in the fruitfly *Drosophila*. This confirmed that genes with distinct

functions in *Drosophila* development were evolutionarily related, that is, they had all derived from the same ancestral gene in some remote and simpler ancestor³⁻⁵.

Even more dramatic was the demonstration that previously unknown genes in other types of animal, including vertebrates, also possessed the homeobox motif⁶. With a few exceptions, developmental biologists were surprisingly cautious in their initial reaction. Most accepted that the homeobox might prove a useful tag for discovering and cloning interesting developmental genes, but largely shied away from the implied suggestion that the homeobox was highlighting developmental mechanisms shared between organisms as distantly related as fruitflies and humans. Clearly, evolutionary and developmental biology were in no rush to form a partnership.

But by 1989 the evidence for common developmental mechanisms was becoming overwhelming. The two principal complexes of homeobox genes in *Drosophila* (the homeotic genes) were proved to be directly related to four clusters of homeobox genes in mammals, and to share with them a similar physical organization and patterns of gene expression. Not only had all these genes derived from the same cluster of genes present in some remote ancestor, but the fruitfly and mammalian gene clusters, the Hox genes, still had a similar and fundamental role in their respective organisms — to specify the identity of different regions along the head-to-tail axis (the anteroposterior axis) of the body⁷.

Evolutionary biologists call this retention of structure and function 'conservation', and other remarkable examples soon followed.

Figure 1 'Man is but a worm' reproduced from Punch's Almanac for 1882. This cartoon originally symbolized the evolution of man, but is pertinent to the striking conservation of developmental mechanisms and genes between vertebrates and bilaterian invertebrates. Reproduced with permission of Punch Ltd.

The *Pax-6* gene turned out to be implicated in eye development virtually throughout the animal kingdom, and a homeobox gene, *tinman*, is involved in heart development in flies and vertebrates. The discovery of conservation permits previously impossible comparisons of the development of organisms with very different body plans, and has stimulated developmental biologists to consider the evolutionary ancestry of developmental mechanisms, often for the first time. So many examples of conservation have now been found that it is no longer considered surprising, and the cartoon in Fig. 1 is even more appropriate now than it was in the 1880s. We can now state with confidence that most animal phyla possess essentially the same genes, and that some (but not all) of these genes change their developmental roles infrequently in evolution.

Constructing the tree

The second crucial factor was the rise of molecular phylogenetics — the comparison of nucleic acid sequences from different organisms and the construction of evolutionary trees from these data. In 1988, Field *et al.* published a pioneering paper that tackled the vast task of constructing the lines of descent — the phylogeny — of the entire animal kingdom using comparisons of ribosomal RNA sequences⁶. At the time, invertebrate zoologists were beginning to appreciate the extent of convergent evolution (that is, when animals with quite different evolutionary histories have similar features), and were thus casting doubt on traditional anatomy-based phylogenetic schemes⁷. Molecular phylogenetics, on the other hand, seemed to provide an objective method of assessing evolutionary relationships. Since 1988, methods of DNA sequence analysis have been improved, the range of species sampled has increased, some potential sources of artefact removed⁸, and complementary molecular data added⁹. The current molecular-based view of invertebrate relationships certainly lacks resolution, but at least it provides a framework within which comparative developmental data can begin to be interpreted (Fig. 2).

Technical advances in molecular biology provided the third impetus to evolutionary developmental biology. Low-stringency library screening, the polymerase chain reaction and *in situ* hybridization were each invented or refined in the 1980s, facilitating the cloning and analysis of genes in any species, not just the handful of model species traditionally studied.

In the rest of this article, I outline several directions in which I believe there is a real chance of significant advance in the future. This list is certainly not comprehensive, partly because of my own biases, but also because this area of science is renowned for throwing up surprises.

The limits to conservation

Despite the interest of each new example, simply documenting yet more cases of conservation between *Drosophila*, nematodes and human is likely to add little to our understanding of the actual course of evolution. Attention now needs to be diverted to rigorously determining the limits to conservation in each case.

We need, for example, to establish when and in what type of organism the conserved gene or gene function first appeared; to identify the secondary modifications or losses that have occurred; and, if possible, to detect underlying factors that may have constrained or promoted change. The example of the Hox gene clusters (Box 1) illustrates the general problem.

This example also shows the crucial importance of having a sound phylogeny on which to base comparative genetic and developmental analyses. As is evident from Fig. 2, knowledge of phylogeny is vital if apparent structural or functional similarities are to be interpreted safely as true homologies — similarities due to descent from the same ancestral source — and the limits to conservation of each homology are to be defined precisely.

Genotype into phenotype

The link between the genetic make-up of an organism (its genotype) and its form and function (its phenotype) lies at the heart of

evolutionary developmental biology. A fundamental question to be addressed is how alterations to the genotype as a result of mutation are transformed through the intermediary of development into changes in form. There are several parts to this question. The first is the relative importance of mutations that alter the coding sequence of a gene, and thus the structure and function of the protein it encodes, as against regulatory mutations that affect the site, timing or level of gene expression. We suspect that regulatory mutations will be most important, although both categories of mutation have been documented in developmentally important genes. Dissecting their individual contributions to developmental change is complicated by the fact that both types of mutation can occur in the same gene.

The importance of regulatory mutations can be seen within the vertebrates, where the anterior boundaries of expression of particular Hox genes have altered significantly in different animals. These changes correlate closely with anatomical changes along the anteroposterior axis, such as the location of the junction between neck and thorax^{10,11}. Similarly, the areas of expression of Hox genes responsible for specifying thoracic identity are greatly expanded in python embryos compared with other vertebrates, which mirrors the anatomical extension of thoracic identity along most of the vertebral column in snakes¹².

Box 1 How conserved are Hox genes?

One outstanding question is whether clustered Hox genes specifying regional identity along the anteroposterior axis are present in all multicellular animals and, indeed, whether their origin marks the origins of the animals themselves. Hox gene clusters have been found so far in chordates (the phylum to which vertebrates belong), echinoderms (the starfish, sea urchins and their relatives), arthropods (crustacea and insects), nemertean worms and nematodes. If we map this distribution onto the working framework of animal evolution provided by recent molecular phylogenies^{8,9}, we can infer that a Hox gene cluster existed in the last common ancestor of all extant bilaterians (animals with (usually) bilateral symmetry and embryos with three germ layers) and is designated B in Fig. 2. This organism was thus the ancestor of three major groupings of animals — the deuterostomes (for example, chordates and echinoderms), ecdysozoans (for example, arthropods) and lophotrochozoans (for example, molluscs and annelid worms). We can also safely assume that all phyla descended from this common ancestor possess (or possessed) a Hox gene cluster.

We cannot, however, draw the same conclusion for animal lineages that split off earlier than this common ancestor, notably the cnidarians (corals, sea anemones and jellyfish), ctenophores (comb jellies), sponges and placozoans. Genes described as Hox genes on the basis of their DNA sequence have been reported from each of these phyla, but physical clustering has not been demonstrated. Furthermore, the discovery in bilaterians of ancient clustered genes similar in DNA sequence to Hox genes but with different roles implies that it is inappropriate to designate a gene as Hox on the basis of sequence alone²². Hence, we do not yet know if a Hox gene cluster was present in the common ancestor of all animals (designated A in Fig. 2), or indeed even earlier.

We also know relatively little about the extent to which the Hox gene cluster has been secondarily modified in different lineages. There are examples of rapid sequence change in some arthropod Hox genes, cluster breakage in *Drosophila* and nematodes⁹, and loss of some Hox genes in pufferfish²⁸ and possibly barnacles²⁹. Gene loss during evolution is poorly documented but could be common; analysis of developmental genes in parasitic species (which tend to lose functions compared with their non-parasitic relatives) could be useful in this regard. Until the origin of the Hox genes is established and the pattern of their subsequent divergence is clarified, we can no longer be satisfied with statements that simply describe Hox gene clusters as 'conserved'.

The mutations responsible for these alterations have not yet been identified; they could have occurred in the regulatory sequences of the Hox genes themselves, or in other genes that control the activation or stabilization of Hox gene expression. DNA sequence analysis of many more species, chosen with regard to their phylogenetic positions, will be the key to finding candidate mutations, whose effects can then be investigated experimentally. Indeed, such a comparative approach was used recently to identify a mutation in an enhancer of the *Hoxc-8* gene of baleen whales¹³.

How important is gene duplication?

A second question in need of resolution is the importance of gene duplication in the evolution of development. Are there developmental processes that are possible with two copies of a gene but not with one? This is a controversial suggestion, not least because it implies that gene number could impose tight genetic constraints on evolution. But although gene duplications may have been very important in some lineages, as discussed below for the vertebrates, it seems that they are neither necessary nor sufficient for most developmental evolution.

Vertebrates, however, are an interesting case. They have many more Hox genes than invertebrates, as a result of the duplication and reduplication of a single ancestral Hox gene cluster. This was accompanied by an expansion of many other developmentally important gene families¹⁴ and, indeed, the total number of genes in the genome¹⁵ (C in Fig. 2 shows when this occurred). It is tempting to speculate that the origin of the complex vertebrate body plan, with its novel cell types and organ systems, was made possible by the availability of extra genetic raw material, in particular, interacting sets of genes that could be gradually recruited for new developmental roles. This hypothesis predicts a tendency for vertebrate genomes to retain duplicates of developmentally important genes but to lose duplicates of the housekeeping genes that control routine metabolic functions, and predicts that the retained genes would tend to acquire new roles in cell types unique to vertebrates. By testing such predictions, it should be possible to determine whether gene duplication was positively exploited in early vertebrate evolution.

The mechanisms by which duplicate genes acquire new roles also needs resolving. The simplest model suggests that one copy of a duplicate gene retains the ancestral role, while additional copies are free to accumulate mutations and diversify. Current evidence indicates that this is an oversimplification. Members of families of duplicated developmental genes in vertebrates often show overlapping spatial and temporal patterns of expression, and also overlapping functions. This suggests that duplicate genes

Figure 2 Proposed phylogeny of the animal kingdom, based primarily on 18S rDNA sequences^{6,8} and Hox gene cluster composition⁹. All animals can be divided into two basic groups according to the number of layers in the embryonic body wall: the bilaterians (triploblasts) have three, and the diploblasts have two. The names at the tips of the branches indicate the main animal phyla, taxonomic groupings of animals that share the same body plan. The urochordates, cephalochordates and vertebrates comprise the chordate phylum. Several minor phyla are omitted. More than one branch leading from a split (for example at the base of the ecdysozoans or lophotrochozoans) indicates lack of resolution or lack of consensus given the available data. Branches labelled A, B, C refer to the origin of animals, of bilaterians and of vertebrates, respectively; see text for discussion.

may retain the shared ancestral role, but supplement this with new roles.

A very similar picture of duplicate gene expression could, however, be achieved in a totally different way. If each duplicate gene loses some of its suite of original functions, control of a complex developmental process would become divided between the multiple descendants of the original gene¹⁶. The relative importance of these different modes of evolution needs elucidating, and not only with respect to vertebrates.

A new 'New Synthesis'

In all the scenarios outlined above, the creation of genetic variation is seen as fundamentally important to the rate, timing or pattern of evolution. At first sight, this viewpoint does not sit comfortably with the neodarwinian 'New Synthesis', which places more emphasis on the reduction of genetic variation during evolution as a population becomes increasingly adapted to its environment and the desirable genes that bring this about spread throughout the population. If developmental biology is to be fully integrated with evolutionary biology this conceptual gap needs to be bridged.

Some steps have already been taken. First,

it is now clear that we need not think of mutations in developmental control genes (such as Hox genes) as always causing large phenotypic changes. For example, small differences in trichome patterns on the second leg of different *Drosophila* species seem to be the result of subtle differences in the regulation of the Hox gene *Ubx*¹⁷. We need to determine whether numerous mutations of very small phenotypic effect or fewer mutations of larger effect generally contribute to morphological differences between species¹⁸.

Second, it is now clear that populations can harbour extensive genetic variation with the potential to cause morphological change but which is only revealed under particular conditions. For example, perturbation of the *Drosophila* heat-shock protein Hsp90 by mutation or changes in environmental conditions uncovers phenotypic variation generated by otherwise hidden genetic variation¹⁹.

These findings may point the way towards a logical framework for the 'microevolution' of development — the generation of small genetically determined differences in development that lead to the relatively minor variations in morphology. We can assume that mutations of small phenotypic consequence

arise frequently, and are either exposed immediately to selection pressures or sheltered temporarily in cryptic form by interactions with other genes. Mutations with larger phenotypic effects can also arise, but it is unclear how frequently they contribute to morphological evolution. An extreme view is that the largest-effect mutations serve only to stabilize morphological change already produced by the gradual accumulation of small changes²⁰.

The huge challenge for the future is to convert this conceptual framework into a quantitative model, with parameters such as magnitude of phenotypic effect (and its heterogeneity), number of genes involved, mutation rates, effects of genetic recombination, gene additivity and effects of the environment. It will then be important to examine the model's behaviour in relation to variables such as population size or degree of fragmentation, selection pressure and genetic drift.

It remains to be seen whether the lessons learnt from the study of microevolution will be sufficient to explain the much greater morphological and physiological differences between higher taxa such as phyla. I suspect that radical alterations to genetic systems (for example, duplication of the whole genome or major alterations in the mechanism of gene regulation) will need to be included if we are to explain some truly major transitions, such as the origin of multiple germ layers in the ancestor of the bilaterians (B in Fig. 2) or the emergence of the vertebrates.

Vermes of the Vendian

Fifteen years ago, few developmental biologists would have heard of the Ediacaran fauna, and few palaeontologists would have confessed to an interest in *Drosophila* genetics. Much has changed. Palaeontologists and developmental biologists are now regularly combining data to tackle key questions in evolutionary developmental biology. Consider timescales. Palaeontology is vital to estimating when a particular evolutionary change occurred; it can also reveal whether such a change was correlated with other evolutionary events or with environmental change. The Cambrian explosion of animal phyla is the classic example, but remains controversial. Palaeontology clearly records a rapid increase in the abundance and diversity of animal fossils at the base of the Cambrian, but disagreement abounds as to whether this reflects increases in body size, the origin of shells and skeletons, or a true rapid diversification of body plans. Even if the last explanation is accepted, did this burst of evolution occur in the ancestors of all multicellular animals, or just among the bilaterian lineage (A versus B in Fig. 2)?

Resolving these questions is important, not least because it would indicate where (and whether) to search for possible internal

genetic or external environmental triggers for animal diversification. Several lines of evidence need to be combined to move the debate forward. The continued study of Precambrian (Vendian) fossils is clearly important. If any of these can convincingly be shown to be allied to extant bilaterian phyla, this would cast doubt on the idea of the rapid diversification of bilaterian lineages in the Cambrian. Further analysis of developmental control genes in cnidarians and ctenophores will also be useful, as this might reveal whether bilaterians indeed have unique developmental characteristics^{21,22}.

Another major contribution of palaeontology is in the recognition of ancestral character states and extinct character combinations. These can help us deduce the actual path of developmental evolution, given the genetic and developmental data at our disposal from extant organisms. Our understanding of the early evolution of vertebrates²³, the radiation of the lophotrochozoans²⁴ and the evolution of the tetrapod limb²⁵ have all benefited from such data. Evolutionary developmental biology will continue to benefit from fossil evidence, provided that the science of palaeontology continues to be widely appreciated.

Embracing genomics

The sequencing of complete genomes from multicellular organisms promises to revolutionize the biological sciences. What are the implications for developmental biology? Animal developmental biologists will probably have to be content in the foreseeable future with a nematode or two, a couple of insects, human, mouse and two rather advanced fish. The opportunities will still be huge. The complete genome sequence of the nematode *Caenorhabditis elegans* has already yielded surprises, including some previously undiscovered Hox genes, secondary loss of the Hedgehog signalling molecule and one of its receptor components, and an exceptionally large number of genes for steroid-hormone receptors^{9,26}. As each genome is sequenced, it will yield its own set of lineage-specific expansions and reductions within families of homologous genes. Eventually, a crude but complex picture should emerge of the general pathways of genome diversification during evolution. It will then be a major task to determine which genomic changes relate to modifications in developmental control or morphology. The pattern of differences should, however, give immediate clues as to which sorts of gene families, or genetic pathways, are prone to change, and which may be evolutionarily constrained.

From the perspective of understanding how animal body plans evolve, it is unfortunate that the genome sequences most likely to be completed first do not encompass a particularly wide range of the body plans present within the animal kingdom. Obvi-

ous examples to add include a cnidarian, a lophotrochozoan (for example a mollusc or annelid worm), an echinoderm, and a urochordate (a tunicate) or cephalochordate (amphioxus).

Complete genomes provide more than simply a catalogue of genes. For example, because of the roles of chromatin structure and nuclear architecture in gene regulation²⁷, neighbouring genes could be subject to coordinated regulation. We might then expect the position of a gene on a chromosome to be functionally relevant and conserved in some cases. Once several complete genomes are available the hypothesis of conserved gene position can be tested. It seems quite possible that the correspondence between a Hox gene's position in a gene cluster and its expression along the anteroposterior axis, so fundamental to patterning the bilaterian body plan, may be just the tip of an iceberg.

Peter W. H. Holland is at the School of Animal and Microbial Sciences, University of Reading, Whiteknights, PO Box 228, Reading RG6 6AJ, UK. e-mail: p.w.h.holland@reading.ac.uk

1. Darwin, C. *The Origin of Species by Means of Natural Selection* (John Murray, London, 1859).
2. Haeckel, E. *The Evolution of Man: A Popular Exposition of the Principal Points of Human Ontogeny and Phylogeny* (Appleton, New York, 1896).
3. McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308**, 428–433 (1984).
4. Scott, M. P. & Weiner, A. J. *Proc. Natl Acad. Sci. USA* **81**, 4115–4119 (1984).
5. McGinnis, W. *Genetics* **137**, 607–611 (1994).
6. Field, K. G. et al. *Science* **239**, 748–753 (1988).
7. Willmer, P. G. *Invertebrate Relationships: Patterns in Animal Evolution* (Cambridge Univ. Press, 1990).
8. Aguinaldo, M. A. et al. *Nature* **387**, 489–493 (1997).
9. de Rosa, R. et al. *Nature* **399**, 772–776 (1999).
10. Burke, A. C., Nelson, C. E., Morgan, B. A. & Tabin, C. *Development* **121**, 333–346 (1995).
11. Gaunt, S. J., Dean, W., Sang, H. & Burton, R. D. *Mech. Dev.* **82**, 109–118 (1999).
12. Cohn, M. J. & Tickle, C. *Nature* **399**, 474–479 (1999).
13. Shashikant, C. S. et al. *Proc. Natl Acad. Sci. USA* **95**, 15446–15451 (1998).
14. Holland, P. W. H. & García-Fernández, J. *J. Dev. Biol.* **173**, 382–395 (1996).
15. Simmen, M. W., Leitgeb, S., Clark, V. H., Jones, S. J. M. & Bird, A. *Proc. Natl Acad. Sci. USA* **95**, 4437–4440 (1997).
16. Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. & Postlethwait, J. *Genetics* **151**, 1531–1545 (1999).
17. Stern, D. L. *Nature* **396**, 463–466 (1998).
18. Mackay, T. F. C. *BioEssays* **18**, 113–121 (1996).
19. Rutherford, S. L. & Lindquist, S. *Nature* **396**, 336–342 (1998).
20. Budd, G. E. *BioEssays* **21**, 326–332 (1999).
21. De Robertis, E. M. & Sasai, Y. *Nature* **380**, 37–40 (1996).
22. Brooke, N. M., García-Fernández, J. & Holland, P. W. H. *Nature* **392**, 920–922 (1998).
23. Aldridge, R. J. & Purnell, M. A. *Trends Ecol. Evol.* **11**, 463–468 (1996).
24. Conway Morris, S. & Peel, J. S. *Phil. Trans. R. Soc. Lond. B* **347**, 305–358 (1995).
25. Coates, M. I. *Development* (Suppl.) 169–180 (1994).
26. Ruvkun, G. & Hobert, O. *Science* **282**, 2033–2041 (1998).
27. Cockell, M. & Gasser, S. M. *Curr. Opin. Gen. Dev.* **9**, 199–205 (1999).
28. Aparicio, S. et al. *Nature Genet.* **16**, 79–83 (1997).
29. Mouchel-Vielh, E., Rigolot, C., Gilbert, J.-M. & Deutsch, J. S. *Mol. Phylogenet. Evol.* **9**, 382–389 (1998).

Acknowledgements. I thank M. Cohn, S. Shimeld and A. Holland for comments on the manuscript, B. Okamura for the bryozoan photograph in Fig. 2, and B. Cohen, M. Telford and other colleagues for helpful discussions. I hope that non-animal biologists will excuse my zoocentric selection of examples.