



Plants, animals and the logic of development

Elliot M. Meyerowitz

Multicellular plants and animals have evolved independently from a unicellular, last common ancestor. Each lineage started with a common toolkit of functioning genes and evolved to complex, multicellular forms. Comparison of the genes used to serve similar functions shows how organisms can use different genes for similar ends and thereby reveals the principles of development.

The last common ancestor of plants and animals is thought to have been a unicellular eukaryote that had already incorporated its α -proteobacterial (mitochondrial) endosymbiont¹. After the lineages diverged, the one that led to green plants incorporated an additional prokaryotic endosymbiont, the cyanobacterial chloroplast². The divergence of plants and animals has been estimated from sequence data to be around 1.6 billion years ago³, and the earliest known fossils of multicellular plants and animals date to around 570 million years ago⁴. Sometime between 1.6 and 0.6 billion years ago, therefore, multicellular development evolved independently in the lineages that led to present-day plants and animals (Fig. 1).

This presents us with a unique opportunity. While specific developmental mechanisms can be (and are being) revealed in molecular detail by the application of genetics and molecular biology, learning about these mechanisms and comparing them from animal to animal, or from plant to plant, can only tell us about the mechanism of development and its evolution in a single lineage – variations on a theme. Comparing plant development with animal development allows us to learn not only about the mechanisms of development but also about the logic of development – not only how plants or animals develop, but also which aspects of development are lineage specific and which ones are used each time complex multicellular development has evolved. A similar opportunity would accompany the discovery of indigenous life on Mars, allowing an understanding of the logic of life – the two examples, having arisen independently, would permit those features common to all life forms to be separated from those that are present as a result of evolutionary descent in a single lineage. We have not found life on Mars, but plants are abundant and easily studied.

Given that not all of the developmental mechanisms of plants and animals are known yet, it is not yet possible to draw up for comparison complete lists of signal-transduction pathways, transcription-regulating mechanisms and cell–cell communication devices. Genomes and their encoded proteomes can be compared, however, and the available knowledge of gene families and their functions related to plant and animal development can be listed. These comparisons will favour the identification of similarities between plants and animals rather than their uniqueness because less is known about plant development, and many genes that serve plant-specific functions (for example, cell–cell communication via the plasmodesmata that connect plant but not animal cells) have not yet been identified. Even so, comparison of signalling and gene-regulatory mechanisms shows that plants and animals have used their common patrimony of ancestral genes differently in some respects, although similarly in others.

Arabidopsis and *Caenorhabditis*

The plant genome presently available for comparison with those of animals is that of *Arabidopsis thaliana*, the laboratory model organism whose genome project has already yielded two-thirds (more than 85 Mb) of the genomic sequence. At present, the best animal for general comparison is *Caenorhabditis elegans*, the only metazoan whose genome sequence is almost completely known. The initial comparison concerns the gene number. The *C. elegans* genome is estimated to code for ~20 000 different proteins – the present count is 19 141, with the genome sequence nearly complete⁵. A current estimate of the gene-dense (euchromatic) part of the *Arabidopsis* genome is ~120 Mb. Average gene spacing is one gene every 4.3 kb^{6,7}, so the genome codes for ~28 000 proteins – 40% more than the nematode. It seems surprising that a plant, thought to be of simpler organization than animals, would have more genes than an animal – but *C. elegans* is more like a reduced version of an animal, with only 959 cells. However, estimates of gene number in a larger and more complex animal, *Drosophila melanogaster*, are, remarkably, no higher – *Drosophila* is estimated to have 12 000 genes⁸. The *Homo sapiens* genome is estimated to encode as many as 70 000 proteins⁸. Thus, animals have a wide range of gene numbers, and *Arabidopsis* falls within that range.

Cell–cell communication

What about specific gene families and gene functions? One example is provided by the cellular receptors for external information, which play a crucial role in cell–cell communication during development as well as in environmental responses. The last common ancestor of *Arabidopsis* and *Caenorhabditis* was unicellular; therefore, any sensor found in both would descend from a sensor of the external environment, whereas others would be specific to one lineage. The *C. elegans* genome has 270 nuclear hormone receptors (which are related to mammalian steroid receptors⁹); the *Arabidopsis* genome, so far, has none. Thus, such receptors themselves are not a necessity for complex multicellular development. The *Arabidopsis* genome is already known to contain >100 receptor serine/threonine kinases. These fall into approximately ten different families (based on the nature of their extracellular domain^{10–14}), some acting as sensors of the developmental state of nearby cells¹⁵, others probably as sensors of external pathogens¹⁶. No homologues of these kinases are found in *C. elegans*. Thus, in each organism, there are hundreds of novel hormones or ligands to discover, but, in the plant, many will bind to extracellular domains of serine/threonine kinases, and, in the worm, to nuclear hormone receptors.

Arabidopsis has (at least) eight genes for histidine kinase two-component receptors – five act as receptors for the plant hormone

Elliot M.
Meyerowitz
meyerow@cco.
caltech.edu

Division of Biology
156-29, California
Institute of
Technology,
1200 East California
Bldv, Pasadena,
CA 91125, USA.

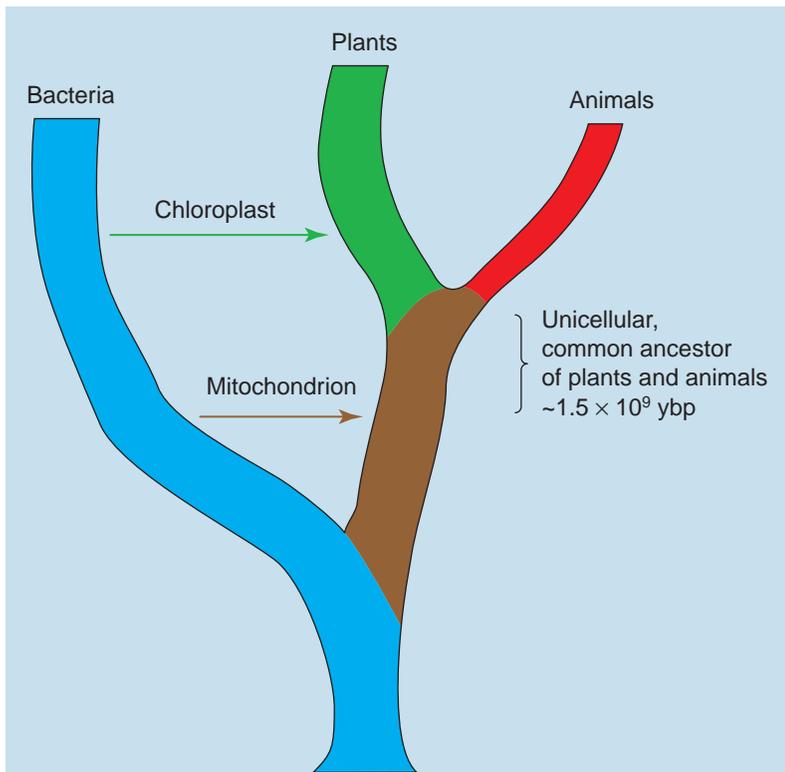


FIGURE 1. Simplified diagram showing events in the evolution of multicellular plants and animals from a unicellular, eukaryotic common ancestor. Blue shows prokaryotes; brown, the eukaryotes ancestral to plants and animals; red, animals; and green, plants. Abbreviation: ybp, years before present.

ethylene¹⁷, and one seems to act in the signalling pathway of another plant hormone class, cytokinins¹⁸. Although bacteria have abundant members of this receptor family, and yeast has members as well, none has yet been found in animals, including *C. elegans*. Animals, including *C. elegans*, have large numbers of receptor tyrosine kinases¹⁹, whereas plants have none, so far. The *C. elegans* genome is characterized by its enormous numbers (>1000, around 5% of the genome) of G-protein-coupled receptors, many thought to be chemoreceptors²⁰. So far, only one member of the G-protein-coupled receptor family has been identified in the *Arabidopsis* genome, matching the single heterotrimeric G-protein α subunit found so far^{21,22}. Thus, when the molecules that cells use to sense their environments are compared, the situation at the level of specific protein families is very different between the plant and animal kingdoms. Nonetheless, the developmental logic appears to be the same in both kingdoms – both plant and animal cells have multiple and surprisingly numerous receptors for external information, both developmental and environmental. It would appear that cell–environment signalling and numerous signals are a requirement for complex multicellular life – but that the exact mechanism of signalling is not crucial and is determined by history, not necessity.

In the nucleus

Signalling pathways generally extend from the external environment to the nucleus. In the nucleus, plants and animals seem, so far, much more similar. Both use multiple families of DNA-binding proteins as transcription factors. Some of these transcription factor families are unique to each lineage. For example, *Arabidopsis* has a large family of transcription factors used in environmental response and in developmental decisions called the EREBP/AP2 family (unrelated to the animal protein AP2), with a conserved

DNA-binding domain not found in the *C. elegans* (or any other animal) genome²³. However, both lineages also contain many of the same DNA-binding protein families – both plants and animals have zinc-finger transcription factors, homeobox proteins, b-zip factors and numerous Myc homologues. Even more similar, so far, are chromatin proteins – the *Arabidopsis* genome contains genes homologous to members of the SWI–SNF complex, the Trithorax complex and other protein complexes known in fungi and animals to be involved in chromatin-level gene regulation and mediator complexes, and, as far as it is known, the plant proteins serve functions similar to the animal ones^{24–27}. In fact, recognition of the great similarity between plant (pea) and animal (calf) histone H4 entailed, historically, the first comparison of plant and animal proteomes²⁸. In the nucleus, there are differences, but the effect of common ancestry is evident – plants and animals are not evolutionarily distant enough to have evolved their gene-regulation mechanisms separately.

Cytoplasmic pathways

What about processes that connect the divergent cell-surface receptors to the similar proteins involved in transcriptional control? Not nearly so much is known about the cytoplasmic pathways from cell surface to nucleus in plants in comparison with those in animals, and still much remains unknown in animals. Proteomic comparison shows similarity between the available proteins in plants and animals as well as unique features in each. Plants have not yet been shown to have receptor tyrosine kinases or Ras orthologues, but there is a well-studied Raf homologue in *Arabidopsis* that acts downstream of a histidine kinase two-component ethylene receptor^{29,30}. The *Arabidopsis* genome has many mitogen-activated protein (MAP) kinases, and MAP kinase kinases³¹, some of which might act downstream of the two-component receptors. If so, a type of receptor not found in animals will connect to a cytoplasmic pathway common to plants and animals (and fungi) – a situation already demonstrated in yeast³². Some well-studied animal signal-transduction pathways have not yet been found in plants; for example, the heteromeric receptor for TGF- β and similar growth factors (such as BMP2/4). These act as serine/threonine kinases³³, but, despite hundreds of such kinases in the *Arabidopsis* genome, there are no proteins in the database that are similar to the TGF- β receptor subunits. The downstream components of TGF- β signalling, the Smad proteins, while represented in many related sequences in *C. elegans* and other animals, have not yet been found in the *Arabidopsis* genome.

Outside the plasma membrane

If there is a gradient of differences from the nucleus to the cytoplasm, with plant and animal cells being most similar in their nuclei and mechanisms of transcriptional control, and most different in the molecular nature (although not the developmental function) of their environmental receptors, then one might expect maximal novelty in the extracellular matrix and in cell–cell connections. This might be true – land plants, at least, have a cell wall largely consisting of the carbohydrate cellulose; there are many proteins associated with it, such as arabinogalactans³⁴, that have no known animal counterparts. The extracellular matrix for animal cells is largely protein, with *C. elegans* showing particular artistry in the use of collagens (with ~170 such genes^{5,35,36}), which are animal specific. The cell-surface features, like the environmental sensor proteins, appear to have diverged as multicellular development evolved and, similarly, represent alternative ways of achieving the same ends.

Caveat emptor

An example of a set of cytoplasmic proteins involved in signal transduction that is better-known in plants but has animal

homologues is a machine for transportation of transcriptional regulators in and out of the nucleus in response to environmental cues. In *Arabidopsis*, this protein complex regulates responses to red and far-red light. A homologous complex, with homologous protein components, exists in mammals, although its function in mammals is unknown³⁷. This complex would seem to be a standard example of similarity in plant and animal function – except for one point: the genes coding for the proteins in this complex are not recognizable in the *C. elegans* proteome. This implies that the complex existed in the common ancestor of plants and animals but has been lost in the worm lineage after the divergence of the mammalian and nematode lineages. Thus, it is inadequate to compare a single animal and a single plant – either might have suffered deletion of genes for proteins that in fact are found in other plants or in other animals – and thus, until many plant and animal genomes have been sequenced, the conclusion that one protein is found in one lineage, and not in the other, is preliminary. This same point has been proved frequently in studies of bacterial proteins, because many proteins are found in some bacteria but have no homologues evident in others³⁸. The aggregate proteomes of plants, animals, fungi and bacteria are the proper material for comparison, and they will only become available for sequence comparisons gradually, as more and more organismal genomes are sequenced and annotated. Finding homologous proteins is a firm result, but failing to find them is always preliminary.

A future beyond compare

The examples given here are only random selections from a broad array of possibilities – whole-genome comparisons of *C. elegans* and *A. thaliana* are yet to be carried out and analysed, and comparisons of the aggregate plant and animal genomes are objectives for the even more distant future. Even so, some preliminary conclusions can be ventured: plants and animals resemble each other, at the level of molecular mechanism, in basic cell-biological processes, particularly intranuclear processes – although there are many unique features. When one examines cell–cell communication, however, different protein parts are found to serve logically identical functions. At this level, it appears that plants and animals have evolved novel and independent mechanisms – as might be expected, given their separate evolution of multicellularity. The fact that such different machines serve such similar roles is an indication that complex



TCB•TIBS•TIG

FIGURE 2. *Arabidopsis thaliana* and *Caenorhabditis elegans*. This double portrait features an *unc46* mutant nematode and a wild-type flower.

*As is the bud bit with an envious worm
Ere he can spread his sweet leaves to the air,
Or dedicate his beauty to the sun.*

William Shakespeare, *Romeo and Juliet* Act I, Sc. 1.

multicellular development requires remarkably numerous cell–cell signals. This is a real principle of development – one of many that will be revealed by continued comparison of animals and plants (Fig. 2).

References

- Muller, M. and Martin, W. (1999) The genome of *Rickettsia prowazekii* and some thoughts on the origin of mitochondria and hydrogenosomes. *BioEssays* 21, 377–381
- Delwiche, C.F. and Palmer, J.D. (1997) The origin of plastids and their spread via secondary symbiosis. *Plant Syst. Evol. (Suppl.)* 11, 53–86
- Wang, D.Y.C. *et al.* (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. London B Biol. Sci.* 266, 163–171
- Xiao, S.H. *et al.* (1998) Three-dimensional preservation of algae and animal embryos in a Neoproterozoic phosphorite. *Nature* 391, 553–558
- C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 282, 2012–2018
- Asamizu, E. *et al.* (1998) Structural analysis of *Arabidopsis thaliana* chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and TAC clones. *DNA Res.* 5, 379–391
- Bevan, M. *et al.* (1998) Analysis of 1.9 mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 391, 485–488
- Miklos, G.L.G. and Rubin, G.M. (1996) The role of the genome project in determining gene function: Insights from model organisms. *Cell* 86, 521–529
- Chervitz, S.A. *et al.* (1998) Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 282, 2022–2027
- Walker, J.C. (1994) Structure and function of the receptor-like protein kinases of higher plants. *Plant Mol. Biol.* 26, 1599–1609
- Thummler, F. *et al.* (1995) Differential accumulation of the transcripts of 22 novel protein kinase genes in *Arabidopsis thaliana*. *Plant Mol. Biol.* 29, 551–565
- Deeken, R. and Kaldenhoff, R. (1997) Light-repressible receptor protein kinase: a novel photo-regulated gene from *Arabidopsis thaliana*. *Planta* 202, 479–486
- Hervé, C. *et al.* (1999) Characterization of the *Arabidopsis* lec-RK-a genes: Members of a superfamily encoding putative receptors with an extracellular domain homologous to legume lectins. *Plant Mol. Biol.* 39, 671–682
- He, Z.H. *et al.* (1999) A cluster of five cell wall-associated receptor kinase genes, Wak1–5, are expressed in specific organs of *Arabidopsis*. *Plant Mol. Biol.* 39, 1189–1196
- Fletcher, J.C. *et al.* (1999) Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. *Science* 283, 1911–1914
- Song, W. *et al.* (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* 270, 1804–1806
- Hua, J. *et al.* (1998) EIN4 and ERS2 are members of the putative ethylene receptor gene family in *Arabidopsis*. *Plant Cell* 10, 1321–1332
- Kakimoto, T. (1996) CK11, a histidine kinase homolog implicated in cytokinin signal transduction. *Science* 274, 982–985

Acknowledgements
My laboratory's work on plant development is supported by the US National Science Foundation, the US NIH, the US Dept of Energy and the Human Frontier Science Program.

19 Hunter, T. (1987) A thousand and one protein-kinases. *Cell* 50, 823–829

20 Bargmann, C.I. (1998) Neurobiology of the *Caenorhabditis elegans* genome. *Science* 282, 2028–2033

21 Ma, H. (1994) GTP-binding proteins in plants – New members of an old family. *Plant Mol. Biol.* 26, 1611–1636

22 Plakidou-Dymock, S. et al. (1998) A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. *Curr. Biol.* 8, 315–324

23 Riechmann, J.-L. and Meyerowitz, E.M. (1998) The AP2/EREBP family of plant transcription factors. *J. Biol. Chem.* 379, 633–646

24 Goodrich, J. et al. (1997) CURLY LEAF: A Polycomb group gene that regulates homeotic gene expression in *Arabidopsis*. *Nature* 386, 44–51

25 Grossniklaus, U. et al. (1998) Maternal control of embryogenesis by MEDEA, a Polycomb-group gene in *Arabidopsis*. *Science* 280, 446–450

26 Luo, M. et al. (1999) Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 296–301

27 Ohad, N. et al. (1999) Mutations in FIE, a WD Polycomb group gene, allow endosperm development without fertilization. *Plant Cell* 11, 407–415

28 DeLange, R.J. et al. (1969) Calf and pea histone IV. III. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. *J. Biol. Chem.* 244, 5669–5679

29 Kieber, J.J. et al. (1993) CTR1, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the raf family of protein-kinases. *Cell* 72, 427–441

30 Clark, K.L. et al. (1998) Association of the *Arabidopsis* CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5401–5406

31 Ichimura, K. et al. (1998) Isolation of ATMEKK1 (a MAP kinase kinase kinase) – Interacting proteins and analysis of a MAP kinase cascade in *Arabidopsis*. *Biochem. Biophys. Res. Commun.* 253, 532–543

32 Posas, F., Witten, E.A. and Saito, H. (1998) Requirement of STE50 for osmotic stress-induced activation of the STE11 mitogen-activated protein kinase kinase kinase in the high-osmolarity glycerol response pathway. *Mol. Cell. Biol.* 18, 5788–5796

33 Padgett, R.W. et al. (1998) TGF-beta signaling, Smads, and tumor suppressors. *BioEssays* 20, 382–390

34 Pennell, R. (1998) Cell walls: structures and signals. *Curr. Opin. Plant Biol.* 1, 504–510

35 Cox, G.N., Kramer, J.M. and Hirsh, D. (1984) Number and organization of collagen genes in *Caenorhabditis elegans*. *Mol. Cell. Biol.* 4, 2389–2395

36 Kramer, J.M. (1994) Genetic analysis of extracellular matrix in *C. elegans*. *Annu. Rev. Genet.* 28, 95–116

37 Wei, N. and Deng, X.W. (1999) Making sense of the COP9 signalosome – a regulatory protein complex conserved from *Arabidopsis* to human. *Trends Genet.* 15, 98–103

38 Koonin, E.V. et al. (1997) Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. *Mol. Microbiol.* 25, 619–637



The origin and evolution of segmentation

Gregory K. Davis and Nipam H. Patel

Arthropods, annelids and chordates all possess segments. It remains unclear, however, whether the segments of these animals evolved independently or instead were derived from a common ancestor. Considering this question involves examining not only the similarities and differences in the process of segmentation between these phyla, but also how this process varies within phyla, where the homology of segments is generally accepted. This article reviews what is known about the segmentation process and considers various proposals to explain its evolution.

To most of us who are segmented, the advantages of a segmental body plan are perhaps not immediately obvious. However, segments are of fundamental evolutionary and developmental importance to several metazoan phyla, including our own (see Glossary). The history of metazoan life – should we ever be up to the task of writing it – would surely be deficient without at least a chapter on the evolution of segments and the ways in which they are made.

Homology of segmentation and phylogeny

Before considering the evolution of segmentation, we first need to specify what we mean by a 'segment'. True segmentation, or metamerism, is usually considered to be the repetition along the anterior–posterior axis of a structural unit that comprises a suite of characters involving the entire body¹. Animals and plants, which evolved multicellularity independently, possess simple serial repetition of structure, and serial repetition is also a feature of some animal appendages, such as insect antennae, suggesting that it evolved multiple times in many contexts. Traditionally, however, it is the body segments of arthropods, annelids and chordates that have been accorded special significance as examples of true metamerism.

Historically, intuitive ideas concerning the evolutionary origins of segmentation in arthropods, annelids and chordates have strongly influenced our picture of the evolutionary relationships among the

bilaterally symmetrical metazoans (the bilaterians). At the close of the 19th century, the segmentation observed in these three phyla was commonly held to be homologous, that is, derived from a segmented common ancestor. Thus, segmentation was often used to unite these groups within a single clade. However, the protostome–deuterostome distinction², made at the beginning of the 20th century, asserted that most bilaterian phyla are more closely related to either chordates or annelids plus arthropods than these two groups are to each other. Since then, the deep phylogenetic separation of chordates and annelids plus arthropods has been retained and confirmed, leading many to regard the segmentation in these two groups as having evolved independently³. Similarly, true segmentation traditionally has been regarded as a shared, derived character of annelids and arthropods, uniting these phyla in a clade to the exclusion of unsegmented phyla, such as the molluscs⁴. However, recent analyses of morphological⁴ and molecular^{5,6} data have suggested that these two segmented phyla are actually more closely related to several unsegmented phyla than they are to each other (Fig. 1).

If this most recent version of metazoan phylogeny⁷ is correct, then it represents a direct challenge to the supposed shared, derived characters that previously united annelids and arthropods, segmentation being primary among them. Essentially, we are now faced with three different hypotheses for the evolution of segmentation (Fig. 1). While it is true that mere phylogenetic separation does

Gregory K. Davis
gdavis@midway.
uchicago.edu

Committee on
Developmental
Biology, University
of Chicago, Chicago,
IL 60637, USA.

Nipam H. Patel
npatel@midway.
uchicago.edu

Dept of Organismal
Biology and
Anatomy, and the
Howard Hughes
Medical Institute,
University of
Chicago,
MC 1028, N-101,
5841 S. Maryland
Ave, Chicago,
IL 60637, USA.