Plants Compared to Animals: The Broadest Comparative Study of Development

Elliot M. Meyerowitz

If the last common ancestor of plants and animals was unicellular, comparison of the developmental mechanisms of plants and animals would show that development was independently invented in each lineage. And if this is the case, comparison of plant and animal developmental processes would give us a truly comparative study of development, which comparisons merely among animals, or merely among plants, do not—because in each of these lineages, the fundamental mechanisms are similar by descent. Evidence from studies of developmental mechanisms in both kingdoms, and data from genome-sequencing projects, indicate that development evolved independently in the lineages leading to plants and to animals.

The Origin of Plants and Animals

We can trace the general course of eukaryotic evolution before and after the existence of the last common ancestor of plants and animals, but only in the crudest terms (Fig. 1). The earliest signs of life on Earth date to 3.8 billion years before the present (Ga) (1). Eukaryotic cells appeared before 2.7 Ga, as indicated by hydrocarbon biomarkers (2); but the earliest fossil evidence for multicellular eukaryotes is 0.6 Ga (3). The last common ancestor of plants and animals appears to have existed around 1.6 Ga, based on rough molecular clock calculations (4)—long after the initial appearance of eukaryotes, and long before a clear fossil record of multicellular eukaryotes. If the fossil record were complete, we could conclude that the last common ancestor of plants and animals was unicellular. But it is far from complete.

Analysis of organelle genomes adds some details, because the endosymbiotic events that established mitochondria and chloroplasts bracket the time of the last common plant/animal ancestor. Sometime after the appearance of the first eukaryotes, but before the last common ancestor of plants and animals, the uptake of the alpha proteobacterium that led to mitochondria occurred; the evidence for this timing is the clear homology of the mitochondrial genomes in plants and animals (4). After the last common ancestor of plants and animals, another endosymbiotic event, the uptake of a cyanobacterium to form the precursor of chloroplasts, occurred only in the plant lineage (6). Therefore, the last common ancestor of plants and animals lived after the alpha-proteobacterial uptake, and before the cyanobacterial—but we do not have dates for these events, nor any knowledge of what sort of eukaryote was involved in either symbiosis. It is simplest to imagine that both uptake events occurred in unicellular eukaryotes, but this is hardly proof that the last common ancestor of the plant and animal lineages was in fact unicellular.

The sequencing of the entire genome of the flowering plant Arabidopsis thaliana (7) and of two animals (Drosophila melanogaster and Caenorhabditis elegans) (8, 9) as well as the impending completion of the sequencing of the human genome (10), along with our knowledge of the function of some of the genes found in these genomes, points strongly to a unicellular (or colonial) last common ancestor. The basic mechanisms of pattern formation and of cell-cell communication in development appear to be independently derived in plants and animals. Nonetheless, there are some surprising similarities in the overall logic of development in the two lineages.

Evidence from Pattern Formation

Pattern formation is one area for comparison between plant and animal development. In Drosophila, as an example of animals, segmental identity is established by the spatially specific transcriptional activation of an overlapping series of master regulatory genes, the HOX homeobox genes. Cognate processes of activation of homologous genes establish the rostral-caudal axis in developing vertebrate brains, and the proximal-distal axis in vertebrate limbs (11). A similar process occurs in the specification of the radial pattern of floral organs in Arabidopsis and other flowering plants. A set of master regulatory genes are transcriptionally activated in a radial pattern of overlapping domains, and the master regulators then specify organ identity in the developing flower (Fig. 2) (12, 13). In both plants and animals, an initial spatially specific pattern of gene transcription is activated and then refined, with the eventual pattern combinatorially specifying segment or organ identity. The plant master regulatory genes identified to date are not, however, members of the homeobox family. Most are members of the MADS box family of transcription factors (14). The homebox and MADS box transcription factor genes are not homologous: There is no detectable similarity in their amino acid sequences, and the protein structures share no resemblance (15–17). Furthermore, both MADS box and homeobox family members are found in plants and in animals, and therefore each family traces its origin to before the last common ancestor (18). Despite the similar use of transcription factors as master regulators of developmental pattern, the plant and animal processes are nonhomologous.

It can be asked whether comparison of animal segment formation and flower development is an appropriate basis on which to consider the evolution of pattern formation in general—that is, it may be that there are pattern-formation processes in plants that use homeodomain proteins just as they are used in animals, and pattern formation processes in animals that rely upon MADS box genes. Studies to date have not implicated the few animal MADS domain proteins as master regulators of pattern formation—there are only two family members each in Drosophila and in Caenorhabditis (19). One of the Drosophila MADS box family members, MEF2, is primarily involved in muscle differentiation (as are its homologs in mammals) (20, 21); the other is involved in wing vein and tracheal development (22). No experiments have shown that mutations in plant homeobox genes have homeotic phenotypes (although many are yet to be studied, and two have been implicated as potential receptors in a signaling process that establishes the upper to lower axis of leaves; see below). Thus, it seems that plants and animals independently evolved the master regulatory processes that serve their logically similar mechanisms of spatial pattern formation, starting with transcription factor families present in the last common ancestor.

Another pair of examples is dorsal-ventral specification in animal embryos, and the specification of the adaxial-abaxial axis in...
leaves. Again, there is surprisingly parallel logic, and in this case, most of the proteins used in animals have no homologs in plants, and vice versa—thereby clearly demonstrating that the processes evolved within each kingdom.

In Drosophila, the dorsal-ventral axis of the embryo is established by a relay of signals between the embryo and its surrounding follicle cells, with ligands traveling through the perivitelline space. Development of the axis appears to begin when the dorsal anterior portion of the embryo produces a transforming growth factor-α (TGF-α)-related protein called Gurken. Gurken acts as a ligand for a receptor of the epidermal growth factor (EGF) type, a receptor tyrosine kinase, in the follicle. Activation of this cascade via Ras activation represses expression of the Pipe protein, a sulfotransferase whose exact function is unknown, on the dorsal side of the embryo, leading to its ventral expression. Ventral Pipe activates production of a protein ligand (Spätzle) for the Toll receptor; activation of Toll leads to proteolysis of Cactus and release of Dorsal from a Cactus-Dorsal complex. This allows Dorsal to enter the nuclei of ventrally located cells in a ventral-dorsal gradient. High nuclear Dorsal activates Twist in the most ventrally located cells; modest levels of nuclear Dorsal activate Short Gastrulation—encoding genes in more dorsally located (lateral) cells, and Twist and Short Gastrulation activity then lead to subsequent differentiative events (23, 24).

The Arabidopsis genome has no relatives of Gurken, no receptor tyrosine kinase homologs or Ras protein, and no genes encoding proteins similar to Pipe. Cactus (homologous with I-kappa B of mammals) has no convincing relatives in the Arabidopsis genome other than the many proteins that share with it ankyrin repeats. Dorsal is a member of the Rel (nuclear factor–kappa B, NF-κB) family of animal transcription factors, with no relatives in Arabidopsis (19); Twist is a basic helix-loop-helix (bHLH) transcription factor family member that shares only weak similarity to the bHLH domain (and no other part of the protein) with Arabidopsis proteins, whereas it shares substantial similarity with proteins from many different animal species. Short Gastrulation is a membrane protein related to Xenopus Chordin, but unrelated to any protein encoded in the Arabidopsis genome.

Toll is a different and interesting story. It acts not only in dorsal-ventral specification in flies, but also in the activation of innate immunity (25). Although there appears to be no plant cognate of Toll acting in pattern formation, there are proteins that act in the plant pathogen response that share with Toll its leucine-rich repeats and its TIR domain (25, 26). These plant proteins apparently do not act via a NF-κB–type of downstream mechanism, at least as far as is known. Thus, it may be that the last common ancestor of plants and animals used some relative of Toll in the pathogen response (a response expected of unicellular as well as multicellular organisms), and that this same system, extensively evolved in the animal lineage for resistance, was also adopted for an animal pattern-formation process.

Adaxial-abaxial axes of leaves in Arabidopsis provide an additional example of pattern formation in plants. The present model for this process begins with the shoot apical meristem at the tip of a growing shoot. Primordia of leaves originate on the flanks of the meristem, and therefore begin with an asymmetric relation to the rest of the plant, with the future adaxial leaf surface adjacent to the meristem, and the abaxial surface distant from it. An unknown signal from the meristem appears to activate members of the REVOLUTA/PHABULOSA/PHAVOLUTA protein family, the RNAs for which are initially uniformly expressed in the leaf primordium. This activation feeds back upon the expression of the genes such that their RNAs are expressed preferentially in the adaxial surface of the primordium; the encoded proteins then specify adaxial fate (27). The REV/PHAB/PHAV family is a plant-specific subfamily of homeobox proteins that also contains a b-zip and a START lipid-binding domain. Because mutations in the START domain create dominantly active versions of the proteins, it is thought that they may be activated by a lipid ligand. The activated form of PHAB or PHAV is thought to act in part by repressing RNA accumulation from the genes KANADI1 and KANADI2 in adaxial regions; their continued abaxial expression represses PHAB and PHAV; in turn (28). The KANADI genes (encoding proteins of the GARP family of transcription factors) then activate members of the YABBY family of transcription factors on the side of the primordia away from the meristem, which leads to abaxial fate (29, 30). Animal genomes as known to date encode no proteins related to members of the GARP family or to those of the YABBY family.

Thus, again we find similar processes controlled by nonhomologous genes, indeed by animal genes that for the most part are not found in plants, and plant genes without detectable relatives in animals.

Chromatin

In both plants and animals the establishment and maintenance of the spatial patterns of expression of master regulatory transcription factors involves chromatin. Chromatin processes seem to be conserved between plants and animals. A striking example of conservation of developmental processes at the level of chromatin is maintenance of the pattern of expression of homeotic genes. In Drosophila, after establishment of the pattern of activation of the Hox genes that establish segmental pattern by transient activators and repressors, their state of expression is maintained by cell-heritable chromatin states of activation and repression (31, 32). For example, the protein coded by Enhancer of zeste [E(z)], a Polycomb group protein, maintains repression of Hox genes such as Ultrabithorax (33). Mutations in E(z) cause ectopic expression of the Hox genes, and consequent homeotic phenotypes.

A homologous pathway appears to regulate floral homeotic genes in plants. Mutants of the Arabidopsis gene CURLY LEAF (CLF) show ectopic expression of the floral-pattern gene AGAMOUS, which acts in the specification of stamen and carpel identity, and also misregulation of APETALA3, a master regulator of petal and stamen identity. The CLF gene codes for an E(z) homolog (34). Thus, flies and plants both appear to use Polycomb class gene products to maintain decisions established by homeotic genes, except that the homeotic genes acted upon by E(z) in flies are homeobox genes, and those acted upon by CLF in Arabidopsis are MADS box genes.

Other chromatin-mediated processes also appear to be conserved between plants and animals. A survey of the Arabidopsis genome shows genes coding for homologs of the typical animal chromatin-regulatory proteins, such as histones, histone-modifying proteins, SWI/SNF2-like chromatin remodeling aden-
osine triphosphatases, Trithorax group proteins, and additional Polycomb group proteins (7, 35, 36). Indeed, discovery of the striking similarity between pea and cow histones was the first result of plant and animal protein sequence comparisons (37).

**Cell-Cell Signaling**

Cell-cell signaling represents a different but equally fundamental developmental process. The general picture is that most aspects of cellular signaling have evolved independently in plants and animals, although there are some common features. An example is cell-cell signaling in the shoot apical meristem (SAM), a collection of undifferentiated cells that act as a reservoir of stem cells for the growth of each shoot. The SAM is divided into a set of zones, each with cells characterized by different division behaviors, and expression of different genes. Communication between two of the zones (the apical central zone and the underlying rib meristem) is mediated by production of a small protein ligand (CLAVATA3) in the central zone that activates a transmembrane receptor kinase (CLAVATA1) in the rib meristem cells, thereby causing a reduced rate of cell division relative to differentiation of cells at the bottom face of the rib meristem (38, 39). Described at this level, the process seems little different from animal signaling with protein ligands [for example, Boss in R8 cells of Drosophila eyes (40), or Gurken in oocytes] to transmembrane receptor kinases [such as Sevenless in R7 cells (16, 41), or the EGF receptor Gurken receptor]. However, Sevenless and the Gurken receptor are receptor tyrosine kinases, and there are no members of this receptor family in Arabidopsis. Both work via activation of Ras, for which there are no Arabidopsis homologs (7). CLAVATA1 is a transmembrane receptor serine/threonine kinase (42). Animals, too, have receptor serine/threonine kinases, the TGF-β receptors, but these do not resemble CLAVATA1, and in addition the TGF-β receptors signal through cellular SMADS proteins, which have not been found in the Arabidopsis genome. The CLAVATA1 protein has protein domains familiar in animals—an extracellular domain of leucine-rich repeats, and a cytoplasmic serine/threonine kinase domain. But these domains are not found attached to each other (so far, at least) in animals. Thus, the protein domains found in this plant system are presumably from the plant/animal ancestor, but they have been assembled in a novel fashion to provide a communication process in plants.

A genomewide survey indicates that this scenario is general. The Arabidopsis genome contains over 150 genes that code for leucine rich–repeat transmembrane receptor kinases, and over 400 genes—nearly 2% of all Arabidopsis genes—that code for members of the transmembrane receptor serine/threonine kinase family with various putative extracellular domains. The closest genes to these in

---

**Fig. 2.** Homeotic mutants of Arabidopsis. (A) Wild-type flower, with four whorls of organs: (from periphery to center) sepals, petals, stamens, and carpels. (B) apetala2 mutant, with carpels, stamens, stamens, and carpels. The mutated gene is a member of the AP2-EREBP plant-specific transcription factor family (19). (C) pistillata mutant, lacking activity of a MADS box transcription factor and thus having sepals, petals, carpels, and carpels. (D) agamous mutant, lacking activity of a MADS box family member, with a repeating sepal, petal, petal structure. Photos courtesy of J. L. Riechmann [panels (A) and (D) from (60), with permission].

---

There are five genes that code for five related ethylene receptors in the Arabidopsis genome. The receptors act in a curious, negative way, in that loss-of-function mutations lead to a constitutive ethylene response (51). The receptors act in air lacking ethylene to repress the normal responses to the hormone (which include effects on seedling growth and on senescence, as well as induction of genes thought to be involved in the pathogen response). When ethylene-bound, the receptors are inactivated, and the ethylene response ensues. All five gene products resemble bacterial two-compartment receptors, and at least one of the encoded proteins, ETR1, has been shown in vitro to be a histidine kinase, as have the bacterial proteins (52). The only domains in the sequence databases related to the NH₂-terminal ethylene-binding domain of the ethylene receptors are other putative plant ethylene receptors and proteins (of unknown organinal function) encoded in cyanobacterial genomes (53). Thus, it appears that plants obtained the ethylene receptor genes at the time of the uptake of the photosynthetic plastid and, like most genes from the endosymbiont, they have been transferred to the nucleus. The ethylene receptors function, at least in part, by interaction with a plant Raf (mitogen-activated protein kinase kinase kinase) homolog, a typically eukaryotic signal transduction protein (54). Thus, these receptors would appear to have entered the eukaryotic genome from the
cyanobacterium, but became the upstream portion of a typically eukaryotic signal transduction cascade.

Sequence analysis of the phytochromes tells a similar story. Phytochromes are red and far-red light receptors with a tetrapyrrole chromophore, related to two-component receptors, that constitute one of the several classes of plant photoreceptors. There are five phytochrome genes in Arabidopsis (7). These genes, like the ethylene receptors, have homologs in cyanobacteria as well as in plants; the cyanobacterial genes encode light receptors (55). Once again it appears that the bacterial receptors became adapted to a eukaryotic signal transduction pathway, because the downstream components of the signal transduction pathway, as far as they are known, are not found encoded in bacterial genomes (56). Indeed, it has been shown that the plant phytochromes, despite their homology to cyanobacterial histidine kinases, act as serine/threonine kinases—and therefore, that part of the process of adaptation to a eukaryotic signaling process involved evolution to a new enzymatic specificity (57).

A more general look at the Arabidopsis genome implies that these examples are not unique. Several estimates of the number of Arabidopsis nuclear genes that derive from the cyanobacterial symbiont have been made, based on a variety of different assumptions, and the numbers range up to 2200 (7, 58, 59). Around 650 to 900 code for proteins found in chloroplasts (58), indicating that as many as 1550 nuclear genes of Arabidopsis, or 6%, may have been derived from the protochloroplast and have been, like ethylene receptors and phytochromes, put to some use outside the chloroplast.

In summary, whole-genome sequencing studies and the evidence of experiments in developmental genetics indicate that plants and animals have evolved development independently. Although the logic underlying many developmental processes is similar, the molecules that carry out the logical plan are unrelated, or represent novel arrangements of ancient protein domains. Plants therefore represent the proper comparison to animals in any truly comparative study of development.

References and Notes
15. W. J. Gehring et al., Cell 78, 211 (1994).
59. I thank J. Bowman, E. Haswell, and T. Ito for comments on the manuscript; J. L. Riechmann for the photographs in Fig. 2; and the National Institutes of Health, National Science Foundation, and Department of Energy for support of my laboratory's work.