The Origins of Multicellularity and the Early History of the Genetic Toolkit For Animal Development

Antonis Rokas

Vanderbilt University, Department of Biological Sciences, Nashville, Tennessee 37235; email: antonis.rokas@vanderbilt.edu

Annu. Rev. Genet. 2008. 42:235–51

The *Annual Review of Genetics* is online at genet.annualreviews.org

This article's doi: 10.1146/annurev.genet.42.110807.091513

0066-4197/08/1201-0235$20.00

Key Words

cell adhesion, cell-cell signaling, transcriptional regulation, animal phylogeny, choanoflagellate, repeated evolution

Abstract

Multicellularity appeared early and repeatedly in life's history; its instantiations presumably required the confluence of environmental, ecological, and genetic factors. Comparisons of several independently evolved pairs of multicellular and unicellular relatives indicate that transitions to multicellularity are typically associated with increases in the numbers of genes involved in cell differentiation, cell-cell communication, and adhesion. Further examination of the DNA record suggests that these increases in gene complexity are the product of evolutionary innovation, tinkering, and expansion of genetic material. Arguably, the most decisive multicellular transition was the emergence of animals. Decades of developmental work have demarcated the genetic toolkit for animal multicellularity, a select set of a few hundred genes from a few dozen gene families involved in adhesion, communication, and differentiation. Examination of the DNA records of the earliest-branching animal phyla and their closest protist relatives has begun to shed light on the origins and assembly of this toolkit. Emerging data favor a model of gradual assembly, with components originating and diversifying at different time points prior to or shortly after the origin of animals.

INTRODUCTION

Complexity: a problematic term used in a variety of different contexts; here it is used to simply denote increases in numbers of cell types, body size, life-cycle stages, genes, or protein domains

From the simple, undifferentiated bacterial filaments to the macroscopic multicellular forms seen in animals, plants, and fungi, the 25 or so instantiations of multicellularity extant today exhibit a remarkable diversity in genotypic and phenotypic complexity (5, 51) (**Table 1**). For example, the multicellular forms observed in prokaryotes are architecturally and morphologically relatively simple, characterized by the presence, at their most elaborate manifestations, of a few distinct cell types (9). Similar levels of complexity are observed in most cases of eukaryotic multicellularity (7, 9, 103). The independent transitions to multicellularity from unrelated unicellular ancestors offer a unique opportunity for comparative study, especially at the molecular level. We start by identifying the general conditions favoring the emergence of multicellularity, its origins, and its signature in the DNA record.

Most multicellular lineages are characterized by relatively simple architectures and morphologies. However, on a few separate occasions, the transition to multicellularity has burgeoned into macroscopic, architecturally complex body plans (e.g., plants, fungi, and animals) (9). In animals, for example, the evolution of several differentiated cell types generated by the specific expression of a number of cell-type–specific genes (83), and the elaborate coordination of developmental processes, made them stand out as one of the most complex inventions of multicellularity (7, 19, 46, 70). Elucidating the enigma of the origins of multicellularity in animals requires, to a large extent, solving the enigma of the origins of their development.

But what is the genetic basis of animal multicellularity and development? Animal genomes contain thousands of genes involved in carrying out vital routine tasks, such as metabolism and cell division. Many of these genes are shared across eukaryotes and predate the origin of animals per se (23, 60), but some underwent extensive gene duplications and evolved new roles in the construction and patterning of animal bodies. These genes comprise the genetic toolkit for animal development (20, 57), a select set of a few hundred genes from a few dozen gene families involved in three key processes: cell differentiation, cell-cell communication, and cell adhesion. Examples of toolkit components include the Hox transcription factors (35), the cell signaling families of Wnts and receptor tyrosine kinases (53, 62), as well as the gene families of cadherins and integrins, which are involved in cell adhesion (1, 72). Understanding the origins and assembly of the genetic toolkit required for animal multicellularity and development is the second and central focus of this review.

Table 1 The genetic and phenotypic complexity of select, independently evolved, multicellular bacterial and eukaryotic lineages

Lineage[1]	Cell type number	Representative species	Gene number	Genome size (Mb)	References
Actinobacteria	3	*Streptomyces coelicolor*	7825	9	(8)
Cyanobacteria	3	*Nostoc punctiforme*	7432	9	(69)
Myxobacteria	3	*Myxococcus xanthus*	7388	9	(14, 38)
Cellular slime molds	3	*Dictyostelium discoideum*	13,541	34	(7, 32)
Animals	3–122	*Drosophila melanogaster*	13,733	200	(7, 24)
Fungi	3–9	*Coprinus cinereus*[2]	13,544	37.5	(7)
Volvocine green algae	2	*Volvox carteri*[3]	15,544	140	(7)
Plants	5–44	*Arabidopsis thaliana*	25,498	125	(24, 100)

[1] The first three lineages are bacterial; the rest eukaryotic.
[2] Genome unpublished; data retrieved from the Broad Institute (**http://www.broad.mit.edu/**).
[3] Genome unpublished; data retrieved from the Joint Genome Institute (**http://www.jgi.doe.gov/**).

Insights from paleontology, ecology, and phylogenetics provide the temporal, environmental, and historical context within which we can understand the emergence of multicellularity. Similarly, dramatic advances in developmental genetics and comparative genomics are significantly enriching our understanding of the genetic changes associated with multicellular transitions, and of the origins of the animal developmental program in particular. The body of facts now emerging has shed ample light on the tempo and pattern of this pivotal period in life's history and is setting up the framework within which we can understand the origins and assembly of the genetic toolkit for animal multicellularity and development.

THE EVOLUTION OF MULTICELLULARITY: A COMPARATIVE PERSPECTIVE

Why Did Multicellularity Evolve?

It is statistically unlikely that complex phenotypes arise repeatedly by chance (25). Thus, from a comparative perspective, the multiple origins of multicellularity in a wide variety of organisms from distinct evolutionary lineages underscore the notion that key aspects of this phenotype are likely to be, under certain conditions, selectively advantageous. Considerable attention has been devoted to identifying what these aspects or conditions may have been, with a variety of factors implicated as plausible contributors to multicellularity's repeated invention (39, 51). Theoretical work suggests that a multicellular existence could have been advantageous by reducing predation (97), improving the efficiency of food consumption (9), facilitating more effective means of dispersal (9), limiting interactions with noncooperative individuals (71, 77, 78), or dividing labor (71). For example, unicellular lifestyle conflicts, such as the dependence of flagellum-induced motility and mitosis on the same molecular machinery (16, 51), or the requirement for spatial or temporal separation of certain metabolic processes (39, 45), could have been easily resolved in a multicellular setting by functional specialization, at least in principle.

In several instances, theoretical expectations have been put to the test. The results have demonstrated that several reasons typically associated with transitions to multicellularity, such as predation avoidance or higher feeding efficiency, do indeed confer a selective advantage over unicellularity. For example, a number of algal species were able to evolve multicellularity when grown in culture in the presence of predators, thus dramatically reducing their chances of being eaten (11, 47, 66). Similarly, *Volvox* algae (61) and myxobacteria (88) have been shown to be at advantage when multicellular because of their ability to better utilize available nutrients.

Most manifestations of multicellularity are relatively simple in architecture, involving only a very small number of cell types (19, 58) (**Table 1**). Cell-type determination typically occurs via the action of a small number of regulatory proteins (49). However, the large number of regulatory proteins present in both prokaryotes and eukaryotes suggests that, from a genomic point of view, these organisms have the potential to generate a much larger number of cell types than those actually observed (19). So why do most multicellular organisms possess so few cell types? Although it is difficult to address this question a posteriori, a plausible explanation may be that there was no selective pressure for early microscopic multicellular organisms to further increase their size, and consequently diversify their pool of cell types beyond a small number (39). Support for this explanation comes from both theory and empirical observations, which indicate that differentiated cell types are generally more likely to evolve in larger multicellular organisms (7, 10, 94, 105).

Any multicellular organism increasing its size is likely to encounter a trade-off between the conflicting selective pressures from escaping predation and avoiding the consumption of the additional energy required to maintain a larger body size. This conflict has been beautifully illustrated by a laboratory experiment where, in the presence of a predator, a culture

Myxobacteria: a group of multicellular δ-proteobacteria, also known as myxomycetes, with a complex life-cycle during which they construct a multicellular fruiting body

of unicellular algae evolved multicellularity in fewer than 100 generations (11). During the course of the experiment, the number of cells per multicellular organism varied between 4 to more than 100, with the population eventually stabilizing to 8-celled bodies despite being much higher in earlier generations. Importantly, an 8-celled body is just big enough to confer escape from predation (11).

When Did Multicellularity Evolve?

Judging from these potential advantages of a multicellular lifestyle over a unicellular one, multicellularity would be expected to appear relatively early in the course of life's evolution. Evidence from the fossil record seems to corroborate this expectation. Simple filamentous manifestations of multicellularity are found in the early fossil records of both bacterial (104) and eukaryotic lineages (58), although the more complex instantiations of multicellularity in both lineages appeared much later.

On the bacterial stem of the tree of life, filamentous cyanobacteria with distinct cell types first appeared approximately 2.5–2.1 billion years ago (bya) (101); their earliest examples were fossilized resting cells that can withstand environmental stress, also known as akinetes, from the genus *Archaeoellipsoides* (4, 101). The fossil record is silent for the other two groups of multicellular bacteria, actinobacteria and myxobacteria, but estimates based on the 16S ribosomal DNA molecular clock offer approximate dates of origin. Actinobacteria appear to be almost as old as differentiated multicellular cyanobacteria, with an estimated date of origin approximately 2.0–1.9 bya (33), whereas multicellular myxobacteria appeared much later in the Proterozoic, close to 1.0–0.9 bya (95).

On the eukaryotic stem, filamentous protists first appear in the fossil record very soon after the appearance of the first unicellular eukaryotes some 1.8 to 1.2 bya, and differentiated multicellular protists appeared no later than 1.2 bya (58). An example of the multicellular complexity exhibited by these early fossils is *Bangiomorpha*, a red algal fossil with at least three distinct cell types (17). Dictyostelid cellular slime molds are thought to have diverged prior to the splitting of fungi and animals (98), but exactly when multicellularity arose in this lineage is unknown. In contrast, the Volvocine green algae, which represent one of the most recent inventions of multicellularity, diverged from their unicellular relatives a mere 0.05 bya (56). Molecular clock estimates place the origin of the complex multicellularity observed in plants, animals, and fungi some time between 1.0–0.4 bya (30), with unambiguous fossils from each of these lineages appearing between 0.6–0.4 bya (50, 102, 109).

Evolution of complex multicellular lineages: too few, too late. Examination of both the bacterial and eukaryotic fossil record strongly indicates that the first experiments in multicellularity were already present much earlier than the emergence of complex multicellularity (19, 58). Examination of Earth's history indicates two major events immediately prior to the origin of complex multicellularity, namely predation (82, 97) and a sharp increase in oxygen levels (42), that may have contributed to its relatively late appearance. For example, the abundance of oxygen in Earth's shallow oceans was an order of magnitude lower than current levels until approximately 0.85 bya (42), and would thus have imposed severe constraints on the evolution of macroscopic bodies with high energy demands. Similarly, multiple lines of evidence argue that it may have been only after the emergence of predators that the selective benefit of a larger size would have been sufficient to drive the evolution of complex multicellular forms (11, 47, 66, 97). Examination of the fossil record suggests that the first predatory eukaryotes appeared approximately 0.75 bya (82, 97), a date strikingly contemporaneous with the emergence of the first ancestors of fungi and animals (30, 82).

How Did Multicellularity Evolve?

Given that multicellularity has evolved repeatedly from independent unicellular lineages,

Cyanobacteria: a group of photosynthetic bacteria that contains unicellular, undifferentiated multicellular (filamentous), and differentiated multicellular species

bya: billion years ago

Actinobacteria: a group of high G+C gram-positive, mostly multicellular bacteria, also known as actinomycetes, that is frequently found in soils

Proterozoic: an era in the geologic time scale that spans from about 2.5 bya to the beginning of the Cambrian period (at 0.54 bya) and during which eukaryotes first appear in the fossil record

Protist: a generic name used to describe any microscopic eukaryotic organism

Green algae: a large and diverse group of unicellular and multicellular of photosynthetic organisms

comparisons of the gene sets of multicellular and unicellular pairs allow us to infer the likely gene set of the unicellular ancestor as well as the changes that have taken place during the evolution of the multicellular species. Although the comparative approach is very powerful, inference of molecular events that have transpired over hundreds of millions of years can be challenging. This is likely to be the case if either of the lineages compared have diverged so long ago that accurate identification of ancestral states or direction of change is difficult (74), or if their genomes have become streamlined as a consequence of adaption to specialized lifestyles (29). Finally, note that not all instantiations of multicellularity are the same, and that they do differ in important details. For example, multicellularity in Volvocine green algae likely evolved as a consequence of incomplete separation after cell division, whereas in Dictyostelid cellular slime molds multicellularity evolved as a consequence of aggregation (104). Thus, any expectation that gene families participating in cell adhesion in the two lineages would show similar trends relative to their unicellular relatives simply because both are multicellular would likely be unfounded.

These caveats notwithstanding, several studies have compared the DNA records of unicellular and multicellular species (8, 38, 45). These first comparisons have investigated a wide variety of characteristics thought to be correlated with transitions to multicellularity, such as the presence of protein domains involved in characteristic multicellular functions (e.g., cell-cell signaling and communication, cell-cell adhesion, and transcriptional regulation) or an increase in their gene family complexity. Data from these comparisons provide three key insights to understanding the origins and assembly of the genetic toolkits associated with transitions to multicellularity. First, many, but not all, of the molecular components of the genetic toolkit are also present in the DNA records of unicellular relatives, which suggests that these components were likely present in their last common (unicellular) ancestor. Second, several of these preexisting components have dramatically diversified in numbers and probably also in function in multicellular lineages. Third, some of the components found in abundance in multicellular lineages are absent from their unicellular relatives and likely represent novel innovations.

Protein domain: polypeptide chains that exhibit structural, functional and evolutionary unity; they are the building block(s) of proteins

A number of studies have examined the independent transitions to multicellularity in the bacterial lineage (8, 38, 69). Comparisons of differentiated multicellular cyanobacteria (e.g., *Nostoc* and *Anabaena*) with their undifferentiated multicellular (e.g., *Trichodesmium*) and unicellular (e.g., *Synechocystis*, *Synechococcus*, and *Prochlorococcus*) relatives revealed large increases in the genes involved in signal transduction and transcriptional regulation (45, 69, 107). For example, whereas the number of transcription factors in differentiated multicellular species ranged between 124 and 172, their number in undifferentiated multicellular or unicellular species ranged between 18 and 64 (107). Evidence for participation of these additional genes in the manifestation of multicellularity comes from analysis of levels of gene expression, which shows that they are up-regulated during the differentiation process (18). A similar trend of an increase in cell-cell signaling and transcriptional regulation genes is seen in comparisons of the multicellular myxobacterium *Myxococcus xanthus* with its unicellular δ-proteobacterial relatives (38). A dramatic increase in regulatory genes is also seen in comparisons of the multicellular actinobacterium *Streptomyces coelicolor* with its unicellular relatives, where the number of σ transcription factors, for example, is approximately fivefold higher (8).

A large fraction of the additional genes associated with cell-cell signaling and transcriptional regulation observed in these multicellular-unicellular comparisons can be accounted for by gene duplication (8, 38). For example, genomic analysis of *M. xanthus* identified more than 1500 duplications that occurred during the transition to multicellularity, and determined that cell-cell signaling and regulatory genes underwent 3 to 4 times as many duplications as would be expected by chance (38). Although the origins of

many of these genes predate multicellularity, their function in the unicellular relatives is not always obvious. Take, for example, the gene cluster identified in the differentiated multicellular cyanobacterium *Anabaena* to regulate differentiation and pattern formation of heterocysts (110). The cluster is conserved across both differentiated and undifferentiated multicellular cyanobacteria, but absent from unicellular ones, suggesting that its ancestral role (likely still present in undifferentiated filamentous species) was a more general one in filamentation (110). The study of gene families with key roles in multicellularity in unicellular relatives will be critical for understanding the genes' ancestral functions and their cooption to the multicellular developmental program.

ORIGINS AND EVOLUTION OF THE GENETIC TOOLKIT FOR ANIMAL MULTICELLULARITY AND DEVELOPMENT

Important clues to the origins and assembly of the genetic toolkit may be gleaned through careful comparisons of the DNA records of extant animal phyla and their closest, mostly unicellular, protist relatives. Notwithstanding a major expansion of the genetic toolkit during early chordate evolution (43), examination of the DNA record of protostomes (such as nematodes, fruitflies, and mollusks) and deuterostomes (such as echinoderms, tunicates, and vertebrates) shows that the genomes of bilaterally symmetrical animals are characterized by fairly similar toolkit gene sets (57, 73). Thus, the toolkit's essential components were probably already in place by the origin of bilaterian animals near the end of the Proterozoic (59).

The presence of the genetic toolkit in the bilaterian ancestor has two serious implications. The first is that the bewildering diversity of bilaterian body plans was generated by further tinkering of the basic genetic toolkit, especially via the modification of patterns of gene expression through the evolution of *cis*-regulatory elements, as well as via the acquisition and subsequent functional diversification of new genes through sequence duplications. This topic has been examined in great depth (20, 27) and is not discussed further here. The second implication, pertinent to the scope of this review, is that if we wish to retrace the early evolution of the genetic toolkit, examination of the DNA records of bilaterians alone is not likely to suffice. We have to look deeper into life's evolutionary genealogy to seek its origins in the DNA records of the morphologically simplest animal phyla (such as poriferans, ctenophores, placozoans, and cnidarians), or even further back in time, in the closest protist relatives of animals (such as choanoflagellates and ichthyosporeans). To do so requires that we first reconstruct the origin and evolutionary diversification of major animal groups, with special emphasis on resolving relationships among early-branching lineages and on identifying the protist relatives of animals.

Examination of the fossil record reveals a Precambrian origin of sponge, cnidarian, and bilaterian body fossils, whereas the first fossil occurrences of the uniquely distinct bilaterian body plans of phyla such as arthropods, chordates, mollusks, echinoderms, and annelids are found in Cambrian-age rock strata (102). While fossils are our only direct window to the past, their utility in reconstructing the evolutionary diversification of animals may be limited. For example, the fossil record is silent regarding the earliest appearances of unicellular and colonial relatives of animals (58). Perhaps more importantly, fossils can only impose lower bounds on divergence estimates because recognizable body fossils always appear after the cladogenetic events that give rise to distinct lineages (57), whereas the time interval between these two events is unknown (15). Thus, reconstructing the evolutionary diversification of animals and their relatives requires that we turn our attention to the DNA record of extant representatives of these lineages.

The DNA record has proven exceptionally useful for charting the tempo and pattern of life's evolutionary history, and has helped to clarify the tempo and mode of an enormous number of key evolutionary events (26). Contrary to the progress observed in the resolution

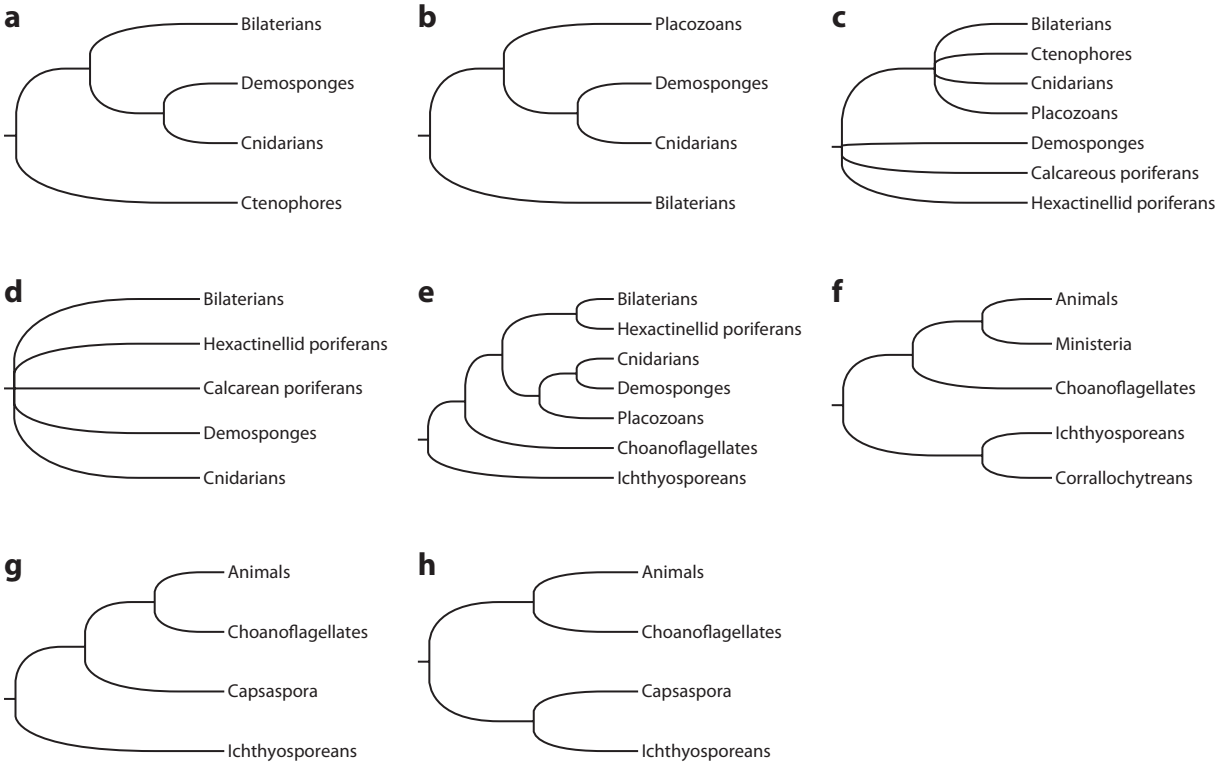

Figure 1

A representative sample of alternative phylogenetic scenarios among early-branching animals and the closest protist relatives of animals. Phylogenies from (*a*) (31), (*b*) (28), (*c*) (79), (*d*) (86), (*e*) (40), (*f*) (98), (*h*) mitochondrial genome phylogeny from (90), and (*g*) nuclear gene phylogeny from (90).

of innumerable other branches of the tree of life, the early history of animal diversification has proven recalcitrant to resolution (**Figure 1**). Below, we review the state of knowledge in two parts of the animal tree that are critical to this review, namely the early branching animal lineages, and the closest protist relatives of animals.

The Diversification of Early-Branching Animals

Most attempts to reconstruct early animal history raise intriguing questions about the evolution of animal development (**Figure 1*a*–*e***). For example, several molecular (6, 12) and morphological (13) studies have identified poriferans as the earliest-diverging branch of the animal tree, a placement in agreement with observations that poriferans are the first animals to appear in the fossil record (13, 37, 75). Further support for this placement comes from the remarkable cytological similarities shared between choanocytes, the feeding cells of sponges, and a phylum of unicellular and colonial protists known as the choanoflagellates (48, 64) (see below). In contrast, other molecular studies point to a clade of early-branching animals that group with bilaterians. In these studies, *Trichoplax adherens*, the single representative of the enigmatic phylum of placozoans, features as the earliest branching phylum on the sister clade to bilaterians (28). Placozoans exhibit a very simple body plan characterized by just four cell types, an absence of organs, and axis of symmetry (7, 28). Other more complex scenarios

have also been proposed that include, for example, poriferan paraphyly (12, 75), cnidarians as the sister group to bilaterians (75), or ctenophores as the bilaterian sister group (76). More radical placements have also been put forward. For example, a recent analysis of extensive molecular data identified ctenophores as the earliest branching phylum of the animal tree (31). Given that ctenophores are morphologically and developmentally much more complex than either poriferans or placozoans (67), their placement would require either loss of this complexity in the placozoan and poriferan lineages or its independent gain in ctenophores (31).

How we reconcile these sharply contrasting views of early animal history remains an open question. The lack of inclusion of representative taxa from key lineages frequently makes comparisons between studies problematic. For example, neither *Trichoplax* nor representatives of two of the three poriferan classes (31) were included in the study that identified ctenophores as the earliest branching lineage (**Figure 1**). Another puzzling feature of several of these studies is that their (contradictory) conclusions are strongly supported. Unfortunately, concatenations of large gene numbers will almost always yield high clade support values, even if the underlying support for one topology over another is marginally better (84). Thus, high clade support values do not always guarantee that the topology obtained is correct (80, 84, 87, 99). The list of studies reporting absolute support for alternative conflicting animal phylogenies has grown in recent years, a result most likely attributable to the increased data.

On the basis of experimental and simulation analyses, we have proposed that early animal evolution was likely an evolutionary radiation (86). This result is compatible with the fossil record (102), and can explain the conflicting conclusions reached by other studies as short-stemmed, long-branched phylogenies are notoriously difficult to resolve (34). The implications of a radiation during early animal evolution for understanding the origins and assembly of the toolkit of animal development and multicellularity are profound (84). If the origin of animals were compressed in time (73, 86), more than 600 million years later it might matter little to know the exact relationships between most phyla to understand the evolution of the molecular tool kit that enabled the evolution of the body plans of the 35 or so animal phyla.

The Search for the Protist Relatives of Animals

Which are the closest extant relatives of animals? Several studies have pointed to five eukaryotic lineages: the *Ministeria* clade, the *Capsaspora* clade, the corallochytreans, the choanoflagellates, and the ichthyosporeans (also known as mesomycetozoans) (89, 90, 98). Although a consensus view of their evolutionary affiliations and placement with respect to animals has yet to emerge, these studies have evinced that all these lineages have deep origins (89, 90, 98). These five protist lineages exhibit a wide variety of lifestyles: *Capsaspora* and ichthyosporeans are parasitic, whereas choanoflagellates, corallochytreans, and *Ministeria* are all free-living (68, 98). Differences are also observed in their morphological characteristics: Corallochytreans and *Ministeria* lack flagellae, but choanoflagellates are flagellated (68, 98).

In the absence of precise phylogenetic knowledge, identifying which of these protist lineages may offer the best comparison with animals requires further examination of their biology and lifestyle. The study of *Ministeria* and corallochytreans presents practical challenges because both groups are difficult to culture, especially in bacteria-free environments (89). In contrast, comparisons of the multicellular and unicellular lifestyle based on the genetic makeup of ichthyosporeans and *Capsaspora* present analytical challenges, as their DNA records are likely to have been influenced by the parasitic lifestyles of these organisms (68). Nevertheless, under certain conditions, ichthyosporeans form multicellular structures (89), suggesting that their genomes may indeed offer vital clues to the molecular origins of multicellularity.

A number of attributes indicate that the most valuable lineage for comparative purposes may be the phylum of choanoflagellates. Cell structure in choanoflagellates, a bulbous cell body surrounded by a protoplasmic, apical collar that encircles their single flagellum, is thought to be ultrastructurally remarkably similar to that of choanocytes, the feeding cells of sponges (48, 64). This similarity has given rise to multiple suggestions that their cellular morphology may be reminiscent of that of the unicellular ancestor of animals (65, 93). Importantly, several recent phylogenetic studies have elucidated the relationships between poriferans and choanoflagellates. Several lines of evidence indicate that choanoflagellates are very close relatives of animals, counter to the hypothesis that they may be a lineage secondarily derived and simplified from poriferan ancestors (51, 55, 85, 86).

All 125 choanoflagellate species known to date have retained a free-living lifestyle, and representatives of each of the three families in the phylum exhibit considerable phenotypic diversity, mostly associated with external cell ornamentations and covers (52). Importantly, a number of choanoflagellate species form multicellular (colonial) structures. An interesting example is offered by *Proterospongia*, a choanoflagellate with a two-phase life cycle, of which one is multicellular, and with a total of four distinct cell morphologies (65). The multicellular stage has the shape of a gelatinous mass, with collared cells on the surface and collarless ones at its interior (44).

The Origins and Assembly of the Genetic Toolkit

An early genomic comparison of the unicellular yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* with humans, flies, and worms found only three highly-conserved genes present in animals that did not have homologs in unicellular yeasts (106). However, when protein domains rather than genes are used as the units of comparison, large-scale differences in content become apparent. For example, a comparison of *Caenorhabditis elegans* with *S. cerevisiae* revealed the presence of several novel domains involved in transcriptional regulation and extracellular adhesion in the worm proteome, as well as an enrichment in domains shared by both organisms (22). In agreement with inferences from studies on bacterial transitions to multicellularity, the transition to multicellularity in animals may not have required the evolution of new genes but rather an increase of complexity of certain gene families, either through the evolution of novel domains or the further shuffling of the domain set already available.

We propose three models to explain the origin and assembly of the animal genetic toolkit, preanimal, pan-animal, and within-animal (**Figure 2**). According to the preanimal model, the origins of the toolkit predate the origin of animals with some, if not all, components of the toolkit present in protist relatives of animals. The pan-animal model argues for an explosive origin of the toolkit; the toolkit is absent in the close relatives of animals but all components are present in even the earliest-branching animals. Finally, the within-animal model suggests that the genetic toolkit was incrementally assembled during early animal evolution, with some, but not all, components of the toolkit present in early-branching animals.

Data emerging from several studies strongly indicate that different components of the genetic toolkit originated and diversified at different time points during the transition to animal multicellularity (1, 51, 53, 54, 63, 72), suggesting that more than one of the proposed models may be required to explain its origins and assembly (**Figure 3**). This inference was recently validated by the genome sequencing and analysis of the unicellular choanoflagellate *Monosiga brevicollis* (55). Comparisons of the choanoflagellate genome with animal and fungal genomes suggest that most cell-adhesion gene families clearly predate animal origins, thus conforming to a preanimal model, whereas most cell-cell signaling and differentiation gene families postdate animal origins, which supports either a within-animal or a pan-animal model (62, 72).

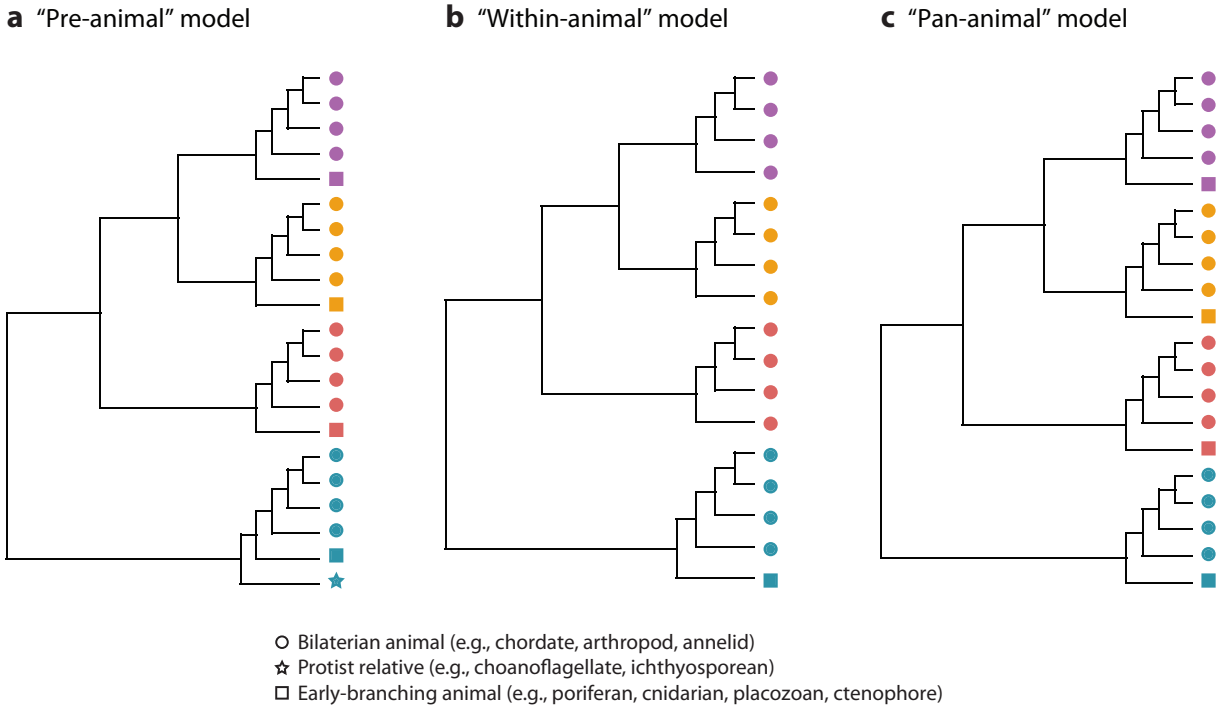

Figure 2

Three alternative models for the evolution of gene family complexity of the genetic toolkit for animal multicellularity and development. (*a*) The pre-animal model, (*b*) the within-animal model, and (*c*) the pan-animal model. The different colors represent different members of the same gene family, whereas the different shapes correspond to the different clades in which protein members are found (e.g., bilaterians, early-branching animals, protists). For example, in the pre-animal model four proteins of the same protein family are present in both bilaterian (*circles*) and early-branching animals (*squares*), but only one member of the protein family—the most basal—is present in eukaryote relatives (*star*).

For example, whereas the cell adhesion family of cadherins is very diverse in choanoflagellates (1), and thus likely to have been similarly so in the unicellular common ancestor shared by choanoflagellates and animals, beta integrins or Wnts are entirely absent from choanoflagellates (55).

The indelible stamp of lowly origin of the cell adhesion machinery. The adhesion of animal cells to their neighbors and the extracellular matrix is a fundamental aspect of animal multicellularity. A few major classes of genes such as the cadherins, the integrins, the selectins (e.g., C-type lectins), and the immunoglobulin superfamily (e.g., fibronectin type III domains) play a key role in mediating adhesion in animal cells. Examination of the choanoflagellate proteome suggests that the gene machinery participating in adhesion in animals was likely well developed in the unicellular ancestor of animals and choanoflagellates. Most of the domains typically found in animals are present in choanoflagellates, including those of cadherins, C-type lectins, immunoglobulins, and α integrins (1, 54, 55). However, what is the function of such a diverse set of adhesion molecules in a unicellular organism that is not known to form cell-cell connections? Examination of the extracellular localization of two choanoflagellate cadherins reveals their presence, and colocalization with actin, at the organism's apical collar (1). The choanoflagellate collar serves as a food-catching device onto which bacteria are latched and transferred toward the cell (44), raising the

Cadherins	127	32	17		46			23
b integrins	11	8	2	2	5	3		0
Wnts	19	13	8	5	12			0
Homeobox	290	87	177	118	154	36	31	2
T-box	17	9	8	21	13			0

Figure 3

Different components of the genetic toolkit originated and diversified at different time points during the transition to animal multicellularity. For example, whereas cadherins are as diverse in choanoflagellates as they are in flies, several other gene families are either absent (e.g., T-box) or less diverse (e.g., Homeobox) in choanoflagellates relative to animals. It is not known whether all phyla within the other major clades exhibit similar levels of gene family complexity. Data for cadherins from (1), for Wnt from (62), for poriferan homeobox genes from (63), and for T-box genes from (55, 108). All other numbers were calculated by searching the proteomes of representative species with the corresponding domains as constructed by the PFAM database (36), using an E-value cut-off of 10^{-5}.

possibility that the origins of this major animal cell adhesion gene family may lie in molecules originally invented for prey capture (1).

Several genes participating in the formation of the extracellular matrix are also well conserved and predate animal origins, including collagen, laminins, and fibronectins (55). Perhaps the most spectacular example of the deep, preanimal origins of some of these gene families is offered by collagens, the most abundant protein family in the mammalian body (41), homologs of which are found not only in choanoflagellates, but also in the animal sister kingdom, the fungi (21). However, integrins, one of the major receptors of collagen, are not found in fungi (41). Furthermore, whereas in animals integrins are functional as heterodimers constructed out of α and β subunits (41), the choanoflagellate genome contains only α integrins (55). This finding suggests that the interaction between integrins and collagen in choanoflagellates may differ from their interaction in animals, and that its study may yield important insights about the evolution of animal cell adhesion to the extracellular matrix.

The early animal origins of the cell-cell signaling machinery. Cell communication is critical for the generation and maintenance of multicellularity in animals, and a handful of core signaling pathways, such as nuclear hormone receptors, Hedgehog, Wnt, TGFβ, Notch, and receptor tyrosine kinases, are involved in its materialization (81). In contrast to the preanimal origin of most of the gene machinery associated with cell adhesion, the origins of signaling pathways were an animal innovation (55). Several of the pathways (e.g., Wnt and TGFβ) are absent from choanoflagellates, although they appear to be present in early-branching animals (2, 55). Perhaps surprisingly, Wnts exhibit remarkable gene family complexity in early-branching animals; the cnidarian *Nematostella vectensis* contains gene representatives for at least 11 of the 12 recognized Wnt subfamilies (62) (**Figure 3**). This complexity of Wnts in early-branching animals argues for an episodic, pan-animal origin of this gene family, although the sudden increase in complexity may be an artifact of the lack of thorough sampling for these genes in placozoans, poriferans, or ctenophores.

Nonetheless, distinct domains of certain pathways are discernible in the choanoflagellate genome (e.g., Notch, Hedgehog, and MAPK), suggesting that animal signaling molecules may have evolved, at least partially, through the shuffling and co-option of pre-existing

TF: transcription factor

domains. The evolutionary origin of the Hedgehog protein offers a telling example of the likely importance of this process and its potential role in the genesis of the genetic toolkit. Bilaterian Hedgehog proteins are composed of two domains, aptly known as the hedge and the hog (3). Choanoflagellates have only the hog domain, whereas poriferans and cnidarian proteomes contain both domains but as parts of distinct proteins, suggesting that the Hedgehog protein likely first evolved through domain shuffling in an early animal ancestor (3, 55, 96).

The emergenece of novel transcriptional regulation machinery in the animal lineage. Transcriptional regulation is of crucial importance in the manifestation of animal multicellularity and development (20, 27). Here is where the protist heritage of the choanoflagellate proteome is most fully exposed, as its proteome contains the standard set of transcriptions factors (TFs) observed across eukaryotes, with most of the well-known animal TFs absent (1, 55). In contrast, examination of the proteomes of early-branching animals shows an appreciable increase in TF family complexity, with both poriferans and cnidarians containing several representatives of the Fox, T-box, Paired, and POU families (63, 108). However, transcription factor family complexity among early-branching animals is not equal; cnidarians are qualitatively (e.g., Hox class homeobox genes are present only in cnidarians) and quantitatively more complex relative to poriferans and placozoans (55, 91) (**Figure 3**). Further examination of the proteomes of early-branching animal phyla is likely to be crucial in understanding the origins of animal transcription factors.

CONCLUSION

In summary, examination of the DNA record of several multicellular lineages has already identified several important molecular trends associated with transitions to multicellularity. On the animal front, the comparison of choanoflagellates with early-branching and bilaterian animals has already yielded important insights into the tempo and mode of the genesis of the genetic toolkit and the likely functions of the gene machinery that predated but was co-opted for multicellularity in the time antecedent to the transition.

Questions about deep origins and major evolutionary transitions were once thought to be, for all practical purposes, imponderable. Important advances in our understanding of how to read and make sense of Earth's early life and environmental history, the theory and experiments associated with transitions in individuality, the genetics of animal development, and finally the DNA record of a multitude of creatures have changed all this. Our understanding of the life and weather in Proterozoic oceans is continuously improving, the theoretical and practical conditions for unicellular to multicellular transitions are being worked out, at the same time as comparisons of several independent evolutions of multicellularity are revealing telltale molecular changes in key parts of the molecular machinery.

Much, however, remains to be understood. If the origins of some of the gene machinery that makes us multicellular can be found in our unicellular relatives, how did it get there in the first place and what was its original function? How are we to reconcile the conflicting evolutionary scenarios of relationships among early-branching animals with the genesis and early evolution of the genetic toolkit? Was the genetic toolkit causal in the evolution of animal multicellularity or simply its product? What was the relative contribution of extrinsic (ecological and environmental) and intrinsic (genetic) factors in the origins of animal multicellularity? If what has been achieved so far is any guide for how future work will progress, the prospects could not be more promising. To quote the great embryologist Hans Spemann (92): "What has been achieved is but the first step; we still stand in the presence of riddles, but not without hope of solving them. And riddles with the hope of solution—what more can a scientist desire?"

SUMMARY POINTS

1. Multicellularity has repeatedly evolved, at different times, in several prokaryotic and eukaryotic lineages.
2. Several different ecological, environmental, and genetic factors have likely contributed to the emergence of most multicellular lineages.
3. Examination of the DNA record of several lineages suggests that multicellular transitions are frequently characterized by increases in gene family complexity of molecules involved in one of three key processes for multicellular growth and differentiation: cell adhesion, cell-cell signaling, and transcriptional regulation.
4. Bilaterally symmetrical animals, which represent the majority of animal lineages, possess a genetic toolkit for animal development, a select set of gene families involved in adhesion, cell communication, and differentiation.
5. Increasing evidence indicates that early animal history was an evolutionary radiation, suggesting that the exact relationships between early-branching phyla may be less important in understanding the origin and assembly of the genetic toolkit.
6. Five protist lineages are the closest relatives to animals, with the choanoflagellates, a clade of unicellular and colonial organisms, the most suitable for comparative purposes.
7. Examination of the DNA record of choanoflagellates and its comparison with that of early-branching, and bilaterian animals supports a model of gradual origins and assembly of the genetic toolkit, with different components originating and expanding at different time points prior to or soon after the origin of animals.

DISCLOSURE STATEMENT

The author is not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I am grateful to Nicole King and Sean B. Carroll for introducing me to choanoflagellates and the origins of multicellularity. Research in the Rokas lab is supported by the Searle Scholars Program and Vanderbilt University.

LITERATURE CITED

1. **Abedin M, King N. 2008. The premetazoan ancestry of cadherins. *Science* 319:946–48**
2. Adamska M, Degnan SM, Green KM, Adamski M, Craigie A, et al. 2007. Wnt and TGF-beta expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS ONE* 2:e1031
3. **Adamska M, Matus DQ, Adamski M, Green K, Rokhsar DS, et al. 2007. The evolutionary origin of Hedgehog proteins. *Curr. Biol.* 17:R836–37**
4. Amard B, BertrandSarfati J. 1997. Microfossils in 2000 Ma old cherty stromatolites of the Franceville Group, Gabon. *Precambrian Res.* 81:197–221
5. Baldauf SL. 2003. The deep roots of eukaryotes. *Science* 300:1703–6
6. Baurain D, Brinkmann H, Philippe H. 2007. Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? *Mol. Biol. Evol.* 24:6–9

Presents tantalizing evidence that cadherins may mediate bacterial prey capture in choanoflagellates.

A concise and clear explanation of the origin and early evolution of bilaterian Hedgehog proteins.

7. Bell G, Mooers AO. 1997. Size and complexity among multicellular organisms. *Biol. J. Linn. Soc.* 60:345–63
8. **Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, et al. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417:141–47**
9. Bonner JT. 1988. *The Evolution of Complexity by Means of Natural Selection*. Princeton, NJ: Princeton Univ. Press. 260 pp.
10. Bonner JT. 2003. On the origin of differentiation. *J. Biosci.* 28:523–28
11. **Boraas ME, Seale DB, Boxhorn JE. 1998. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evol. Ecol.* 12:153–64**
12. Borchiellini C, Manuel M, Alivon E, Boury-Esnault N, Vacelet J, Le Parco Y. 2001. Sponge paraphyly and the origin of Metazoa. *J. Evol. Biol.* 14:171–79
13. Botting JP, Butterfield NJ. 2005. Reconstructing early sponge relationships by using the Burgess Shale fossil *Eiffelia globosa*, Walcott. *Proc. Natl. Acad. Sci. USA* 102:1554–59
14. Brenner DJ, Krieg NR, Garrity GM, Staley JT. 2005. *Bergey's Manual of Systematic Bacteriology*. Vol. II, *The Proteobacteria*. New York: Springer 1388 pp.
15. Budd GE, Jensen S. 2000. A critical reappraisal of the fossil record of the bilaterian phyla. *Biol. Rev.* 75:253–95
16. Buss LW. 1987. *The Evolution of Individuality*. Princeton, NJ: Princeton Univ. Press. 201 pp.
17. Butterfield NJ. 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26:386–404
18. Campbell EL, Summers ML, Christman H, Martin ME, Meeks JC. 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. *J. Bacteriol.* 189:5247–56
19. Carroll SB. 2001. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* 409:1102–9
20. Carroll SB, Grenier JK, Weatherbee SD. 2001. *From DNA to Diversity.* New York: Blackwell Sci. 214 pp.
21. Celerin M, Ray JM, Schisler NJ, Day AW, StetlerStevenson WG, Laudenbach DE. 1996. Fungal fimbriae are composed of collagen. *EMBO J.* 15:4445–53
22. Chervitz SA, Aravind L, Sherlock G, Ball CA, Koonin EV, et al. 1998. Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 282:2022–28
23. Chothia C, Gough J, Vogel C, Teichmann SA. 2003. Evolution of the protein repertoire. *Science* 300: 1701–3
24. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203–18
25. Conway Morris S. 2003. *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge, UK/New York: Cambridge Univ. Press. 464 pp.
26. Cracraft J, Donoghue MJ, eds. 2004. *Assembling the Tree of Life*. Oxford: Oxford Univ. Press. 576 pp.
27. Davidson EH. 2001. *Genomic Regulatory Systems; Development and Evolution.* New York: Academic. 261 pp.
28. Dellaporta SL, Xu A, Sagasser S, Jakob W, Moreno MA, et al. 2006. Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum. *Proc. Natl. Acad. Sci. USA* 103:8751–56
29. Doolittle RF. 2002. Biodiversity: Microbial genomes multiply. *Nature* 416:697–700
30. Douzery EJ, Snell EA, Bapteste E, Delsuc F, Philippe H. 2004. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. USA* 101:15386–91
31. Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, et al. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452:745–49
32. Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sucgang R, et al. 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435:43–57
33. Embley TM, Stackebrandt E. 1994. The molecular phylogeny and systematics of the actinomycetes. *Annu. Rev. Microbiol.* 48:257–89
34. Felsenstein J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.* 27:401–10

The genome analysis of a model actinobacterium, one of the three bacterial multicellular lineages.

A remarkable experiment where multicellularity was evolved in a laboratory culture following the introduction of a predator.

35. Ferrier DE, Holland PW. 2001. Ancient origin of the Hox gene cluster. *Nat. Rev. Genet.* 2:33–38
36. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, et al. 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* 34:D247–51
37. Gehling JG, Rigby JK. 1996. Long expected sponges from the Neoproterozoic Ediacara fauna of South Australia. *J. Paleontol.* 70:185–95
38. Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, et al. 2006. Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci. USA* 103:15200–5
39. Grosberg RK, Strathmann RR. 2007. The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.* 38:621–54
40. Haen KM, Lang BF, Pomponi SA, Lavrov DV. 2007. Glass sponges and bilaterian animals share derived mitochondrial genomic features: a common ancestry or parallel evolution? *Mol. Biol. Evol.* 24:1518–27
41. Heino J. 2007. The collagen family members as cell adhesion proteins. *BioEssays* 29:1001–10
42. Holland HD. 2006. The oxygenation of the atmosphere and oceans. *Philos. Trans. R. Soc. London Ser. B* 361:903–15
43. Holland PWH. 1999. Gene duplication: past, present and future. *Semin. Cell Dev. Biol.* 10:541–47
44. Hyman LH. 1940. *The Invertebrates: Protozoa through Ctenophora*. New York: McGraw-Hill. 726 pp.
45. Kaiser D. 2001. Building a multicellular organism. *Annu. Rev. Genet.* 35:103–23
46. Kamada T. 2002. Molecular genetics of sexual development in the mushroom *Coprinus cinereus*. *BioEssays* 24:449–59
47. Kampe H, Konig-Rinke M, Petzoldt T, Benndorf J. 2007. Direct effects of *Daphnia*-grazing, not infochemicals, mediate a shift towards large inedible colonies of the gelatinous green alga *Sphaerocystis schroeteri*. *Limnologica* 37:137–45
48. Karpov SA, Leadbeater BSC. 1998. Cytoskeleton structure and composition in choanoflagellates. *J. Eukaryot. Microbiol.* 45:361–67
49. Kauffman SA. 1987. Developmental logic and its evolution. *BioEssays* 6:82–87
50. Kenrick P, Crane PR. 1997. The origin and early evolution of plants on land. *Nature* 389:33–39
51. King N. 2004. The unicellular ancestry of animal development. *Dev. Cell* 7:313–25
52. King N. 2005. Choanoflagellates. *Curr. Biol.* 15:R113–14
53. King N, Carroll SB. 2001. A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. *Proc. Natl. Acad. Sci. USA* 98:15032–37
54. King N, Hittinger CT, Carroll SB. 2003. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science* 301:361–63
55. King N, Westbrook MJ, Young SL, Kuo A, Abedin M, et al. 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origins of metazoan multicellularity. *Nature* 451:783–88
56. Kirk DL. 2005. A twelve-step program for evolving multicellularity and a division of labor. *BioEssays* 27:299–310
57. Knoll AH, Carroll SB. 1999. Early animal evolution: emerging views from comparative biology and geology. *Science* 284:2129–37
58. Knoll AH, Javaux EJ, Hewitt D, Cohen P. 2006. Eukaryotic organisms in Proterozoic oceans. *Philos. Trans. R. Soc. London Ser. B* 361:1023–38
59. Knoll AH, Walter MR, Narbonne GM, Christie-Blick N. 2004. Geology. A new period for the geologic time scale. *Science* 305:621–22
60. Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, et al. 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5:R7
61. Koufopanou V, Bell G. 1993. Soma and germ—an experimental approach using *Volvox*. *Proc. R. Soc. London Ser. B* 254:107–13
62. Kusserow A, Pang K, Sturm C, Hrouda M, Lentfer J, et al. 2005. Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* 433:156–60
63. Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM. 2008. Genesis and expansion of metazoan transcription factor gene classes. *Mol. Biol. Evol.* 25:980–96
64. Leadbeater B, Kelly M. 2001. Evolution of animals, choanoflagellates and sponges. *Water Atmos.* 9:9–11
65. Leadbeater BSC. 1983. Life-history and ultrastructure of a new marine species of *Proterospongia* (Choanoflagellida). *J. Mar. Biol. Assoc. U.K.* 63:135–60

An insightful examination of a myxobacterial genome providing evidence for the importance of gene duplication of select gene families in multicellularity transitions.

A landmark investigation of the evolution of major components of the genetic toolkit for animal development and multicellularity.

A masterful synthesis of the fossil evidence for eukaryotic life in the Proterozoic.

The deciphering of the poriferan DNA record is the next major advance, and this is the first systematic analysis of their transcription factors in the context of early animal evolution.

66. Lurling M. 2003. Phenotypic plasticity in the green algae *Desmodesmus* and *Scenedesmus* with special reference to the induction of defensive morphology. *Ann. Limnol.-Int. J. Lim.* 39:85–101
67. Martindale MQ. 2005. The evolution of metazoan axial properties. *Nat. Rev. Genet.* 6:917–27
68. Medina M, Collins AG, Taylor JW, Valentine JW, Lipps JH, et al. 2003. Phylogeny of Opisthokonta and the evolution of multicellularity and complexity in Fungi and Metazoa. *Int. J. Astrobiol.* 2:203–11
69. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, et al. 2001. An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynth. Res.* 70:85–106
70. Meyerowitz EM. 2002. Plants compared to animals: the broadest comparative study of development. *Science* 295:1482–85
71. Michod RE. 2007. Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl. Acad. Sci. USA* 104:8613–18
72. Nichols SA, Dirks W, Pearse JS, King N. 2006. Early evolution of animal cell signaling and adhesion genes. *Proc. Natl. Acad. Sci. USA* 103:12451–56
73. Ohno S. 1996. The notion of the Cambrian pananimalia genome. *Proc. Natl. Acad. Sci. USA* 93:8475–78
74. Penny D, Hendy MD, Poole AM. 2003. Testing fundamental evolutionary hypotheses. *J. Theor. Biol.* 223:377–85
75. Peterson KJ, Butterfield NJ. 2005. Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record. *Proc. Natl. Acad. Sci. USA* 102:9547–52
76. Peterson KJ, Eernisse DJ. 2001. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol. Dev.* 3:170–205
77. Pfeiffer T, Bonhoeffer S. 2003. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proc. Natl. Acad. Sci. USA* 100:1095–98
78. Pfeiffer T, Schuster S, Bonhoeffer S. 2001. Cooperation and competition in the evolution of ATP-producing pathways. *Science* 292:504–7
79. Philippe H, Telford MJ. 2006. Large-scale sequencing and the new animal phylogeny. *Trends Ecol. Evol.* 21:614–20
80. Phillips MJ, Delsuc FD, Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. *Mol. Biol. Evol.* 21:1455–58
81. Pires-daSilva A, Sommer RJ. 2003. The evolution of signalling pathways in animal development. *Nat. Rev. Genet.* 4:39–49
82. Porter SM, Meisterfeld R, Knoll AH. 2003. Vase-shaped microfossils from the Neoproterozoic Chuar Group, Grand Canyon: a classification guided by modern testate amoebae. *J. Paleontol.* 77:409–29
83. Raff RA. 1996. *The Shape of Life: Genes, Development and the Evolution of Animal Form*. Chicago, IL: Univ. Chicago Press. 520 pp.
84. Rokas A, Carroll SB. 2006. Bushes in the tree of life. *PLoS Biol.* 4:e352
85. Rokas A, King N, Finnerty J, Carroll SB. 2003. Conflicting phylogenetic signals at the base of the metazoan tree. *Evol. Dev.* 5:346–59
86. Rokas A, Kruger D, Carroll SB. 2005. Animal evolution and the molecular signature of radiations compressed in time. *Science* 310:1933–38
87. Rokas A, Williams BL, King N, Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425:798–804
88. Rosenberg E, Keller KH, Dworkin M. 1977. Cell density-dependent growth of *Myxococcus xanthus* on casein. *J. Bacteriol.* 129:770–77
89. Ruiz-Trillo I, Burger G, Holland PW, King N, Lang BF, et al. 2007. The origins of multicellularity: a multi-taxon genome initiative. *Trends Genet.* 23:113–18
90. Ruiz-Trillo I, Roger AJ, Burger G, Gray MW, Lang BF. 2008. A phylogenomic investigation into the origin of metazoa. *Mol. Biol. Evol.* 25:664–72
91. Ryan JF, Mazza ME, Pang K, Matus DQ, Baxevanis AD, et al. 2007. Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS ONE* 2:e153
92. Sander K, Faessler PE. 2001. Introducing the Spemann-Mangold organizer: experiments and insights that generated a key concept in developmental biology. *Int. J. Dev. Biol.* 45:1–11
93. Saville-Kent W. 1880–1882. *A Manual of the Infusoria*. London: David Bogue

The definitive genomic comparison between differentiated multicellular cyanobacteria with their undifferentiated multicellular and unicellular relatives.

A lucid description of unicellular relatives of animals and fungi and an announcement of plans to sequence the genomes of several of them.

94. Schaap P, Winckler T, Nelson M, Alvarez-Curto E, Elgie B, et al. 2006. Molecular phylogeny and evolution of morphology in the social amoebas. *Science* 314:661–63
95. Shimkets LJ. 1990. Social and developmental biology of the myxobacteria. *Microbiol. Rev.* 54:473–501
96. Snell EA, Brooke NM, Taylor WR, Casane D, Philippe H, Holland PW. 2006. An unusual choanoflagellate protein released by Hedgehog autocatalytic processing. *Proc. R. Soc. London Ser. B* 273:401–7
97. Stanley SM. 1973. An ecological theory for the sudden origin of multicellular life in the late Precambrian. *Proc. Natl. Acad. Sci. USA* 70:1486–89
98. Steenkamp ET, Wright J, Baldauf SL. 2006. The protistan origins of animals and fungi. *Mol. Biol. Evol.* 23:93–106
99. Takezaki N, Figueroa F, Zaleska-Rutczynska Z, Takahata N, Klein J. 2004. The phylogenetic relationship of tetrapod, coelacanth, and lungfish revealed by the sequences of 44 nuclear genes. *Mol. Biol. Evol.* 21:1512–24
100. The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
101. Tomitani A, Knoll AH, Cavanaugh CM, Ohno T. 2006. The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc. Natl. Acad. Sci. USA* 103:5442–47
102. Valentine JW. 2004. *On the Origin of Phyla*. Chicago: Univ. Chicago Press. 614 pp.
103. Valentine JW, Collins AG, Meyer CP. 1994. Morphological complexity increase in metazoans. *Paleobiology* 20:131–42
104. Waggoner BM. 2001. Eukaryotes and multicells: origin. In *Encyclopedia of Life Sciences*: Chichester: Wiley. **http://www.els.net/**
105. Willensdorfer M. 2008. Organism size promotes the evolution of specialized cells in multicellular digital organisms. *J. Evol. Biol.* 21:104–10
106. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, et al. 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415:871–80
107. Wu JY, Zhao FQ, Wang SQ, Deng G, Wang JR, et al. 2007. cTFbase: a database for comparative genomics of transcription factors in cyanobacteria. *BMC Genomics* 8:104
108. Yamada A, Pang K, Martindale MQ, Tochinai S. 2007. Surprisingly complex T-box gene complement in diploblastic metazoans. *Evol. Dev.* 9:220–30
109. Yuan X, Xiao S, Taylor TN. 2005. Lichen-like symbiosis 600 mya. *Science* 308:1017–20
110. Zhang W, Du Y, Khudyakov I, Fan Q, Gao H, et al. 2007. A gene cluster that regulates both heterocyst differentiation and pattern formation in *Anabaena* sp strain PCC 7120. *Mol. Microbiol.* 66:1429–43

Annual Review of Genetics

Volume 42, 2008

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at http://genet.annualreviews.org/errata.shtml