Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Genetics & Development

The molecular origins of multicellular transitions

Antonis Rokas

Multicellularity has evolved multiple times independently from a variety of ancestral unicellular lineages. Past research on multicellularity was focused more on explaining why it was repeatedly invented and less so on the molecular foundations associated with each transition. Several recent comparative functional analyses of microbial unicellular and multicellular genomes have begun to throw considerable light on the molecular commonalities exhibited by independent multicellular transitions. These have enabled the delineation of the likely functional components of the genetic toolkit required for multicellular existence and to surprising discoveries, such as the presence of several toolkit components in unicellular lineages. The study of these toolkit proteins in a unicellular context has begun yielding insights into their ancestral functions and how they were coopted for multicellular development.

Address
Vanderbilt University, Department of Biological Sciences, VU Station B 351634, Nashville, TN 37235, USA

Corresponding author: Rokas, Antonis (antonis.rokas@vanderbilt.edu)

Current Opinion in Genetics & Development 2008, **18**:472–478

This review comes from a themed issue on
Genomes and evolution
Edited by Sarah Teichmann and Nipam Patel

Available online 14th October 2008

0959-437X/$ – see front matter

DOI 10.1016/j.gde.2008.09.004

Introduction

The multiple distinct origins of multicellular lineages have long been viewed as traversing one of the major steps in the evolution of life's complexity [1]. To this date, most of the work and emphasis in the study of multicellularity's origins has been devoted to explaining *why* this phenotype has been repeatedly invented [2]. Although figuring out the precise causes in several cases is challenging because most multicellular origins are very ancient (but see Ref. [3••]), years of theory, modeling, and experiments have identified several potential advantages of multicellular lineages over unicellular ones, which have likely been key catalysts in such transitions [2]. These include selective benefits conferred by larger sizes, such as escape from predation [4] and increase in the efficiency of food consumption [5,6], as well as benefits conferred by allowing for the functional specialization of biological processes that utilize the same molecular machinery [7] or require spatiotemporal separation [8].

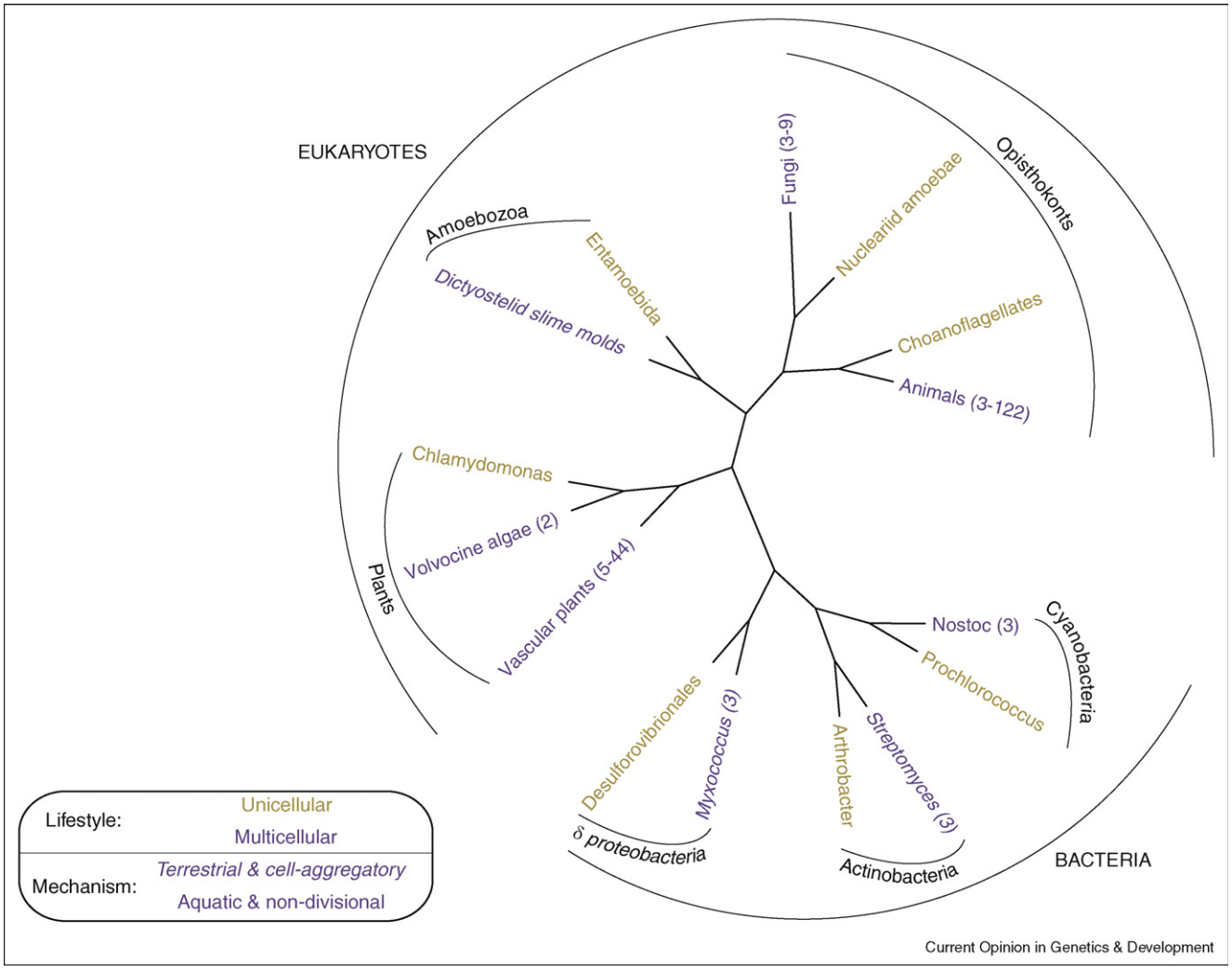
Considerably less attention has been devoted in understanding *how* multicellularity has been invented in each lineage [9], and in identifying the environmental and genetic triggers that have significantly contributed to multicellularity's several manifestations. The changes in the physical and biological environment associated with multicellular lineages must have surely played key roles and abundant geochemical, paleontological, and ecological evidence exists to support such a conclusion [4,10,11]. For example, both an environmental shift imposed by an increase in oxygen levels [12] and a novel ecological pressure placed by the emergence of predation [11] have presumably contributed to the evolution of animal multicellularity [10]. Similarly, the competition for limited supplies of mineral nutrients, such as phosphorus, was likely a driver of multicellularity in volvocine green algae [5,13].

Environmental factors aside, the genetic make-ups of the unicellular ancestors of multicellular lineages must also have been pivotal in facilitating the repeated emergence of the trait. But until recently, the sources and ancestral function of genes fundamental to the multicellular lifestyle were completely unknown, as such genes were thought to be present only within the multicellular lineages to which they were first identified and studied [14,15]. Emerging data from comparative and functional genomics studies in several multicellular lineages and comparisons with their unicellular relatives, as well as experimental studies of 'multicellular' genes in a unicellular context, have begun to provide an increasingly clearer snapshot of the molecular foundations upon which multicellular lineages were founded.

Not all multicellularities are the same

Multicellular lineages have been intermittently sprouting from the tree of life for billions of years. Multicellular forms exist in all three of life's kingdoms, with several representative lineages found in both bacteria and eukaryotes. On the bacterial clade, cyanobacteria, myxobacteria, and actinobacteria are the three main multicellular lineages, whereas on the eukaryotic side these include our familiar plants, animals, and fungi, as well as several lineages of algae (green and brown) and slime molds (dictyostelid and acrasid) (Figure 1). Within several of these lineages, such as the kingdoms of plants and fungi, it is very likely that multicellularity arose more than once [2]. Importantly, multicellularity's repeated inventions have given rise to a remarkable variety of morphologies.

Figure 1



The multiple independent origins of multicellularity. A select set of independently evolved multicellular bacterial and eukaryotic lineages (in blue) and their unicelluar relatives (in tan). Cell-aggregatory multicellular lineages whose origins are terrestrial are shown in italics.

This multiformity of multicellular lifestyles has serious implications for elucidating the molecular underpinnings of several of these multicellular transitions, and consideration of the key differences exhibited between the various multicellular lifestyles is necessary for guiding research on their molecular origins.

Without greatly oversimplifying, extant lineages can be classified by considering three aspects of their multicellular phenotype: its time of origin, its mechanism of origination, and its attained level of complexity (Figure 1). If one considers the undifferentiated filaments of cyanobacteria which are the oldest unambiguous fossils on record as multicellular forms, then the origins of simple multicellular organisms are likely as old as life itself [16]. However, not all multicellular lineages have their origins in deep time; volvocine green algae evolved multicellularity as recently as 0.05 billion years ago (bya) [13]. On the bacterial branch, filamentous cyanobacteria with distinct cell types did not emerge until 2.5–2.1 bya [17] and actinobacteria are thought to have almost as old an origin, whereas myxobacteria appear somewhat later around 1.0–0.9 bya [18]. In the eukaryotic branch, the majority of multicellular lineages likely arose between 1.2 bya, when the earliest red algal differentiated multicellular forms are found, and 0.4 bya, which date the earliest fossil appearance of multicellular plants (reviewed in Ref. [18]).

Multicellularity likely evolved via one of the two distinct mechanisms, either through cell aggregation or through lack of separation following a cell division [9]. Consideration of these mechanisms is crucial for setting up the expectations as to what molecules might have been originally involved in the emergence of multicellular

forms from ancestral unicellular ones in each of these lineages [18]. Examples of 'cell-aggregatory' multicellular lineages are the myxobacteria and the dictyostelid slime molds, whereas 'nondivisional' multicellular lineages are nicely illustrated by fungi and plants. Curiously, all 'nondivisional' multicellular lineages were originally (and some still are) aquatic, whereas lineages which became multicellular through aggregation were (are) terrestrial [9]. Within 'nondivisional', aquatic lineages, there is variation in their cell morphology that is also of considerable significance and relevant to understanding multicellularity's molecular origins. Specifically, organisms such as plants and fungi whose cells are surrounded by cell walls tend to form multicellular linear-like threads called filaments, their unity maintained largely through the rigidness of the cell wall structure, whereas organisms such as animals whose cells lack walls tend to utilize specialized adhesion molecules on the cell surface to achieve coherence [9].

Perhaps the most significant difference between the various multicellular lineages is that of complexity [19,20]. While the complexity of a given multicellular organism may be hard to precisely define and can have a variety of different meanings, a typical measure of complexity that has been employed by many authors is the number of distinct cell types [19–22]. Cell type number is a useful index and a good approximation of both the degree of the division of labor and the complexity of gene regulation for a given multicellular organism. For example, consistent with the notion that animal multicellularity is the most complex in terms of numbers of cell types, metazoans exhibit the greatest diversity of protein domains [22,23] (Figure 2). Measured this way, the complexity of most multicellular organisms, which possess just two or three distinct cell types, may likely be not that different from unicellular ones (several unicellular organisms, such as a variety of spore-forming bacteria, have more than one cell type). This is true for all bacterial multicellular lineages as well as for most eukaryotic ones. In contrast, typical animals and plants are characterized by tens of distinct cell types, whereas fungal species can have up to seven to nine different cell types [19,21].

The genetic toolkits for multicellularity

Genomes typically contain hundreds to thousands of genes that aid organisms to carry out their routine functions, such as metabolism, cellular transport, and protein synthesis; many of the genes participating in these functions are shared across major clades of the tree of life. But multicellular lifestyles also require the existence of functions that are distinct from or nonexistent in unicellular ones [24,25••]. For example, multicellular organisms frequently require biological molecules with adhesive properties so that they maintain their multicellular cohesion. Similarly, movement or any complex action necessitates that cells within multicellular organisms possess cell–cell signaling pathways that enable communication and coordination. Finally, the differentiation of genetically identical germ cells into a number of functionally diverse cell types critically depends on a complex network of transcriptional regulators.

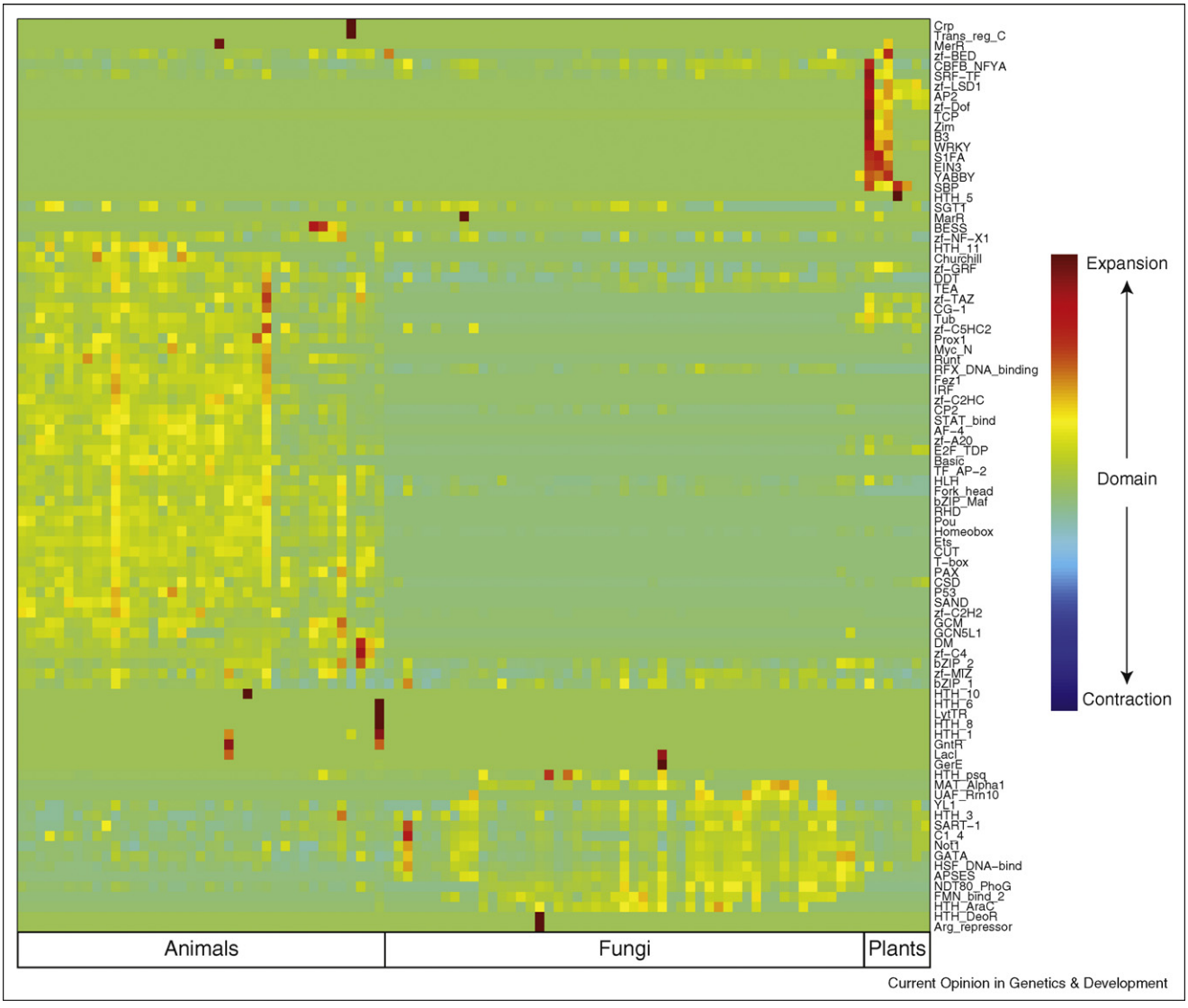
Naturally, the extent to which such a genetic toolkit for multicellular development will be enriched in proteins involved in cell adhesion, cell–cell signaling, and cell differentiation will be strongly dependent on the origin, style, and complexity of the multicellular lineage in question. Given multicellularity's repeated invention and the wide diversity of forms that have evolved as a consequence, whether one could identify such trends which are shared by multicellular transitions at the genomic level remained, until recently, an open question. However, analyses on a number of independently evolved pairs of unicellular and multicellular relatives suggest that such enrichments in cell adhesion, cell–cell signaling, and cell differentiation molecular machinery are not only very commonly observed in multicellular transitions [22,23,25••,26–28], but also further serve as telltale signs of the influence of the multicellular lifestyle on genome content.

The origins and evolution of genetic toolkit components

Until recently, protein families thought to be involved in multicellularity were frequently thought to be present only within the multicellular lineages where they were first identified. For example, before their discoveries in the unicellular choanoflagellate *Monosiga brevicollis*, molecules such as cadherins, integrins, receptor tyrosine kinases, laminins, and immunoglobulins, as well as Myc and Sox/TCF transcription factors, were all thought to be animal-specific [14,15,24,25••]. In rare cases, the diversity of particular toolkit components in unicellular relatives has been found to be as high as or higher than in their multicellular counterparts. For example, a total of 23 cadherins were identified in *M. brevicollis*, 6 more than the number identified in *Drosophila melanogaster* [29•], whereas this choanoflagellate's tyrosine kinase signaling machinery was found to apparently be more diverse not only than the *Drosophila* proteome but also than the human one [30••,31••].

More commonly, multicellularity genetic toolkit components are found in unicellular relatives but their diversity is not as high as in multicellular organisms. Several transcription factor families, including the MADS-box and homeodomain-containing ones, have undergone significant expansions in green plants relative to the unicellular green alga *Chlamydomonas* [28]. Similarly, immunoglobulins, helix–loop–helix and C_2H_2 zinc finger transcription factors, as well as a variety of other protein families are numerically much more abundant in metazoan proteomes relative to the choanoflagellate

Figure 2



Transcription factors of independently evolved multicellular lineages exhibit distinct patterns of protein domain expansions and contractions. Heatmap rows correspond to the PFAM transcription factors from animals, fungi, and plants from the DNA-binding domain (DBD) database, whereas columns correspond to animal, fungal, and plant species from the DBD [23]. Expansions of domains are indicated by the spectrum of colors toward red, whereas contractions of domains by the spectrum of colors toward blue. Note that the number of animal-specific transcription factors is much higher than that of fungal-specific or plant-specific ones. The transcription factor occurrence values were kindly provided by the DBD database [23] and were normalized for both organism proteome size and protein family.

one [25••]. These trends also hold in independent comparisons between multicellular and unicellular lineages among the prokaryotes [8,26,27,32,33]. Specifically, proteins involved in transcriptional control and regulation [26,27,32], as well as those associated with cell–cell signaling and the signal transduction machinery [8,26,27] are greatly enriched in the multicellular bacterial lineages when compared with their closest unicellular relatives.

Occasionally, some but not all of the domains that define protein families present in a given multicellular lineage are found in the proteomes of their unicellular relatives. For example, only the hog domain of animal hedgehog proteins has been found in choanoflagellates [25••,34]. Intriguingly, early-branching animals such as poriferans and cnidarians contain both the hedge and hog domains, but as parts of distinct proteins [35], suggesting that the hedgehog protein found in bilaterian animals likely arose via domain shuffling during early animal evolution [34,35]. Alternatively, as in the case of the tyrosine kinase signaling machinery in choanoflagellates and animals, several proteins in the unicellular lineage contain domains in combinations never before observed in the proteins from the multicellular lineage [25••,30••,36••].

This suggests that these shared protein domains were likely present in the last common ancestor of the two lineages but that following their divergence, they were independently utilized as building blocks to generate a diverse set of signaling proteins presumably tailored to each organism's lifestyle [30••,36••].

The protein families that are part of genetic toolkits for multicellularity have been most often recruited into multicellularity-relevant processes independently from distinct unicellular ancestors. This recruitment appears to have been partially built upon a set of protein domains that appear to be present across a wide variety of unicellular and multicellular lineages, such as the basic helix–loop–helix, homeodomain, and C_2H_2 zinc finger domains [28], but which were subsequently employed and expanded for multicellularity and development-related purposes in certain multicellular lineages. For example, homeodomain proteins are found in plants, fungi, and animals as well as in other eukaryotes [28], but it is in animals where their contribution to developmental body-building is most significant [37]. Plant homeodomain proteins also participate in development and have expanded relative to their unicellular relatives too [28], but here the body-building role is mostly undertaken by the MADS-box family of proteins [38]. Interestingly, MADS-box proteins are also found in fungi, animals, and other eukaryotes but their number and developmental role in these lineages is much more limited [28,39]. The patterns exhibited by homeodomain and MADS-box proteins are consistent with data from several other transcription factor families, with individual lineages showing unique and characteristic patterns of protein domain expansions and contractions [22,23] (Figure 2). But lineage-specific transcription factors with key roles in multicellularity are also found. One such example is the APSES transcription factors which are involved in the regulation of cellular differentiation in fungi [40].

The function of genetic toolkit components in a unicellular context

One of the most surprising results generated by the comparative analysis of several unicellular eukaryotic genomes has been the extent to which their genomes contain proteins which were previously only known from multicellular organisms [25••,28]. Their presence in the unicellular relatives of multicellular lineages suggests that these proteins were likely present in their last common unicellular ancestor, immediately before the emergence of multicellularity. Not surprisingly, however, our knowledge of the function of such proteins mainly stems from studies in multicellular organisms. Thus, the examination of the function of proteins with key developmental roles in multicellular lineages in a unicellular context has the potential to shed light on what was their original function before their recruitment for multicellular purposes and what was their contribution, if any, to the emergence of multicellularity.

The examination of the function of the newly discovered homologs of key multicellular proteins in unicellular lineages has begun yielding exciting results [29•,41,42••,43]. For example, plant homeodomain transcription factors belonging to the KNOX and BELL families regulate shoot apical meristem formation [44]. However, a recent study of two transcription factors from these two protein families in the unicellular alga *Chlamydomonas* found that their function in a unicellular context is to regulate the haploid–diploid transition, a process that involves the regulation of gamete differentiation, zygote development, and meiosis [42••]. Intriguingly, homeodomain proteins in several fungi participate in the sexual cycle in much the same way (e.g. [45]), suggesting that original function of homeodomain proteins may have been in the initiation of the sexual cycle [42••], and that their involvement in plant development may have emerged within this context. Similarly, recent work in choanoflagellates suggests that the original function of some of the machinery involved in cell–cell adhesion in multicellular animals may have been in bacterial prey capture [29•].

Conclusions

Several novel insights have emerged from the first handful of genome-scale comparisons and interesting hypotheses about the likely ancestral functions of genes involved in multicellularity and development have been generated and functionally tested. The increasing elucidation of eukaryotic history [46], coupled with the deciphering of genomes from key branches of the tree of life [47•] and the creation of novel unicellular microbe model systems for the study of multicellular gene function [48,49], suggests that many more startling and fundamental discoveries lie ahead. However, a deeper understanding of the origins of multicellular transitions will rest not only on additional genome comparisons and functional tests, but also on the potential integration of this body of work with the theoretical framework developed for the study of cooperation [50] and with the further exploration of the lifestyle, phenotypes, and social behavior of microbes in their natural environments [51–53].

Acknowledgements

I thank Nipam Patel and Sarah Teichmann for the invitation to contribute to this volume, and Derek Wilson for providing the raw data for Figure 2. Research in AR's lab is supported by the Searle Scholars Program and Vanderbilt University.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Smith JM, Szathmáry E: *The Major Transitions in Evolution*. Oxford/New York: W.H. Freeman Spektrum; 1995.

2. Grosberg RK, Strathmann RR: **The evolution of multicellularity: a minor major transition?** *Annu Rev Ecol Evol System* 2007, **38**:621-654.

3. •• Herron MD, Michod RE: **Evolution of complexity in the volvocine algae: transitions in individuality through Darwin's eye**. *Evolution* 2008, **62**:436-451.
The volvocine algae are a great model system for the study of the origin of multicellularity because they are closely related to unicellular algae and species within the group and exhibit a range of multicellular phenotypes. This paper describes the evolutionary reconstruction of a series of characters contributing to multicellularity across the lineage and shows with characteristic clarity how a series of small character changes, including reversals, can lead to a wide variety of complex multicellular phenotypes.

4. Boraas ME, Seale DB, Boxhorn JE: **Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity**. *Evol Ecol* 1998, **12**:153-164.

5. Koufopanou V, Bell G: **Soma and germ — an experimental approach using *Volvox***. *Proc R Soc Lond Ser B-Biol Sci* 1993, **254**:107-113.

6. Rosenberg E, Keller KH, Dworkin M: **Cell density-dependent growth of *Myxococcus xanthus* on casein**. *J Bacteriol* 1977, **129**:770-777.

7. Buss LW: *The Evolution of Individuality*. Princeton, N.J.: Princeton University Press; 1987.

8. Kaiser D: **Building a multicellular organism**. *Annu Rev Genet* 2001, **35**:103-123.

9. Bonner JT: **The origins of multicellularity**. *Integr Biol* 1999, **1**:27-36.

10. Valentine JW: **Prelude to the Cambrian explosion**. *Annu Rev Earth Planet Sci* 2002, **30**:285-306.

11. Stanley SM: **An ecological theory for the sudden origin of multicellular life in the late Precambrian**. *Proc Natl Acad Sci U S A* 1973, **70**:1486-1489.

12. Holland HD: **The oxygenation of the atmosphere and oceans**. *Philos Trans R Soc B-Biol Sci* 2006, **361**:903-915.

13. Kirk DL: **A twelve-step program for evolving multicellularity and a division of labor**. *Bioessays* 2005, **27**:299-310.

14. King N, Carroll SB: **A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution**. *Proc Natl Acad Sci U S A* 2001, **98**:15032-15037.

15. King N, Hittinger CT, Carroll SB: **Evolution of key cell signaling and adhesion protein families predates animal origins**. *Science* 2003, **301**:361-363.

16. Waggoner BM: **Eukaryotes and multicells: origin**. *Encyclopedia of Life Sciences*. Chichester: John Wiley & Sons, Ltd.; 2001http://www.els.net/.

17. Tomitani A, Knoll AH, Cavanaugh CM, Ohno T: **The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives**. *Proc Natl Acad Sci U S Am* 2006, **103**:5442-5447.

18. Rokas A: **The origins of multicellularity and the early history of the genetic toolkit for animal development**. *Annu Rev Genet* 2008, 42: in press.

19. Bonner JT: *The Evolution of Complexity by Means of Natural Selection*. Princeton, N.J.: Princeton University Press; 1988.

20. Carroll SB: **Chance and necessity: the evolution of morphological complexity and diversity**. *Nature* 2001, **409**:1102-1109.

21. Bell G, Mooers AO: **Size and complexity among multicellular organisms**. *Biol J Linn Soc* 1997, **60**:345-363.

22. Vogel C, Chothia C: **Protein family expansions and biological complexity**. *PLoS Comput Biol* 2006, **2**:e48.

23. Wilson D, Charoensawan V, Kummerfeld SK, Teichmann SA: **DBD — taxonomically broad transcription factor predictions: new content and functionality**. *Nucleic Acids Res* 2008, **36**:D88-92.

24. King N: **The unicellular ancestry of animal development**. *Dev Cell* 2004, **7**:313-325.

25. •• King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I *et al.*: **The genome of the choanoflagellate *Monosiga brevicollis* and the origins of metazoan multicellularity**. *Nature* 2008, **451**:783-788.
The definitive analysis of the content of the genome of the choanoflagellate *Monosiga brevicollis* and how it differs from animals with respect to molecular machinery implicated in multicellularity and development. Choanoflagellates are one of the closest unicellular relatives of animals. Their morphological similarity to the feeding cells of poriferans, one of the earliest branching shoots of the animal tree, has been interpreted by many as suggestive of what the last unicellular ancestor of animals looked like.

26. Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen JA, Ronning CM, Barbazuk WB, Blanchard M, Field C *et al.*: **Evolution of sensory complexity recorded in a myxobacterial genome**. *Proc Natl Acad Sci U S A* 2006, **103**:15200-15205.

27. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D *et al.*: **Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)**. *Nature* 2002, **417**:141-147.

28. Riano-Pachon DM, Correa LG, Trejos-Espinosa R, Mueller-Roeber B: **Green transcription factors: a *Chlamydomonas* overview**. *Genetics* 2008, **179**:31-39.

29. • Abedin M, King N: **The premetazoan ancestry of cadherins**. *Science* 2008, **319**:946-948.
Cadherins play key roles in animal cell adhesion and signaling. This study examines the diversity of cadherins in the choanoflagellate *Monosiga brevicollis*, and finds that this unicellular protist contains a set of cadherins as diverse as that of several animals. Experiments on two choanoflagellate cadherins show that these molecules appear to associate with the actin cytoskeleton and localize at the cell's apical collar, indicating that the ancestral, premulticellular, function of cadherins may have been in bacterial prey capture.

30. •• Manning G, Young SL, Miller WT, Zhai Y: **The protist, *Monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan**. *Proc Natl Acad Sci U S A* 2008, **105**:9674-9679.
Along with Refs. [31••,36••] this article provides a superb comparative functional analysis of the tyrosine kinase signaling machinery in the choanoflagellate *Monosiga brevicollis*. Not only is the diversity of certain components higher in the choanoflagellate genome than in any animal lineage, but also the data suggest that several components independently coexpanded in both choanoflagellates and animals.

31. •• Pincus D, Letunic I, Bork P, Lim WA: **Evolution of the phosphotyrosine signaling machinery in premetazoan lineages**. *Proc Natl Acad Sci U S A* 2008, **105**:9680-9684.
See annotation to Ref. [30••].

32. Wu JY, Zhao FQ, Wang SQ, Deng G, Wang JR, Bai J, Lu JX, Qu J, Bao QY: **cTFbase: a database for comparative genomics of transcription factors in cyanobacteria**. *Bmc Genomics* 2007, **8**:104.

33. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R: **An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium**. *Photosynth Res* 2001, **70**:85-106.

34. Snell EA, Brooke NM, Taylor WR, Casane D, Philippe H, Holland PW: **An unusual choanoflagellate protein released by Hedgehog autocatalytic processing**. *Proc R Soc Lond Ser B-Biol Sci* 2006, **273**:401-407.

35. Adamska M, Matus DQ, Adamski M, Green K, Rokhsar DS, Martindale MQ, Degnan BM: **The evolutionary origin of hedgehog proteins**. *Curr Biol* 2007, **17**:R836-R837.

36. •• Mayer BJ: **Clues to the evolution of complex signaling machinery**. *Proc Natl Acad Sci U S A* 2008, **105**:9453-9454.
See annotation to Ref. [30••].

37. Carroll SB, Grenier JK, Weatherbee SD: *From DNA to Diversity*. Blackwell Science; 2001.

38. Meyerowitz EM: **Plants compared to animals: the broadest comparative study of development**. *Science* 2002, **295**:1482-1485.

39. Becker A, Theissen G: **The major clades of MADS-box genes and their role in the development and evolution of flowering plants**. *Mol Phylogenet Evol* 2003, **29**:464-489.

40. Shelest E: **Transcription factors in fungi**. *FEMS Microbiol Lett* 2008, **286**:145-151.

41. Segawa Y, Suga H, Iwabe N, Oneyama C, Akagi T, Miyata T, Okada M: **Functional development of Src tyrosine kinases during evolution from a unicellular ancestor to multicellular animals**. *Proc Natl Acad Sci U S Am* 2006, **103**:12021-12026.

42. •• Lee JH, Lin H, Joo S, Goodenough U: **Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family**. *Cell* 2008, **133**:829-840.

This paper is one of the first studies to infer the likely ancestral function of proteins that later evolved to assume key roles in plant multicellular development. The authors examine the role of two homeodomain-containing proteins, Gsp1 and Gsm1, in the unicellular green alga *Chlamydomonas* and find that their physical interaction is required to combinatorially control meiosis and zygote development. Thus, the function of homeodomain proteins in the plant unicellular ancestor may have been regulating the meiotic sexual cycle.

43. Li W, Young SL, King N, Miller WT: **Signaling properties of a non-metazoan Src kinase and the evolutionary history of Src negative regulation**. *J Biol Chem* 2008, **283**:15491-15501.

44. Hake S, Smith HM, Holtan H, Magnani E, Mele G, Ramirez J: **The role of knox genes in plant development**. *Annu Rev Cell Dev Biol* 2004, **20**:125-151.

45. Hull CM, Boily MJ, Heitman J: **Sex-specific homeodomain proteins Sxi1alpha and Sxi2a coordinately regulate sexual development in *Cryptococcus neoformans***. *Eukaryot Cell* 2005, **4**:526-535.

46. Yoon HS, Grant J, Tekle YI, Wu M, Chaon BC, Cole JC, Logsdon JM, Patterson DJ, Bhattacharya D, Katz LA: **Broadly sampled multigene trees of eukaryotes**. *Bmc Evol Biol* 2008, **8**: 14.

47. • Ruiz-Trillo I, Burger G, Holland PW, King N, Lang BF, Roger AJ, Gray MW: **The origins of multicellularity: a multi-taxon genome initiative**. *Trends Genet* 2007, **23**:113-118.

The availability of unicellular protist genomes has substantially contributed to understanding the molecular origins of multicellular lineages. This paper describes perhaps the most exciting genome sequencing initiative for the study of multicellularity to date. The initiative aims to provide the research community with the genomes of several relatively unknown, but key from a phylogenetic perspective, unicellular and colonial lineages at fungal–animal boundary.

48. King N: **Choanoflagellates**. *Curr Biol* 2005, **15**:R113-R114.

49. Vallon O, Dutcher S: **Treasure hunting in the *Chlamydomonas* genome**. *Genetics* 2008, **179**:3-6.

50. West SA, Griffin AS, Gardner A: **Evolutionary explanations for cooperation**. *Curr Biol* 2007, **17**:R661-R672.

51. Aguilar C, Vlamakis H, Losick R, Kolter R: **Thinking about *Bacillus subtilis* as a multicellular organism**. *Curr Opin Microbiol* 2007, **10**:638-643.

52. Webb JS, Givskov M, Kjelleberg S: **Bacterial biofilms: prokaryotic adventures in multicellularity**. *Curr Opin Microbiol* 2003, **6**:578-585.

53. Palkova Z: **Multicellular microorganisms: laboratory versus nature**. *EMBO Rep* 2004, **5**:470-476.