

# *Tempos and modes of collectivity in the history of life*

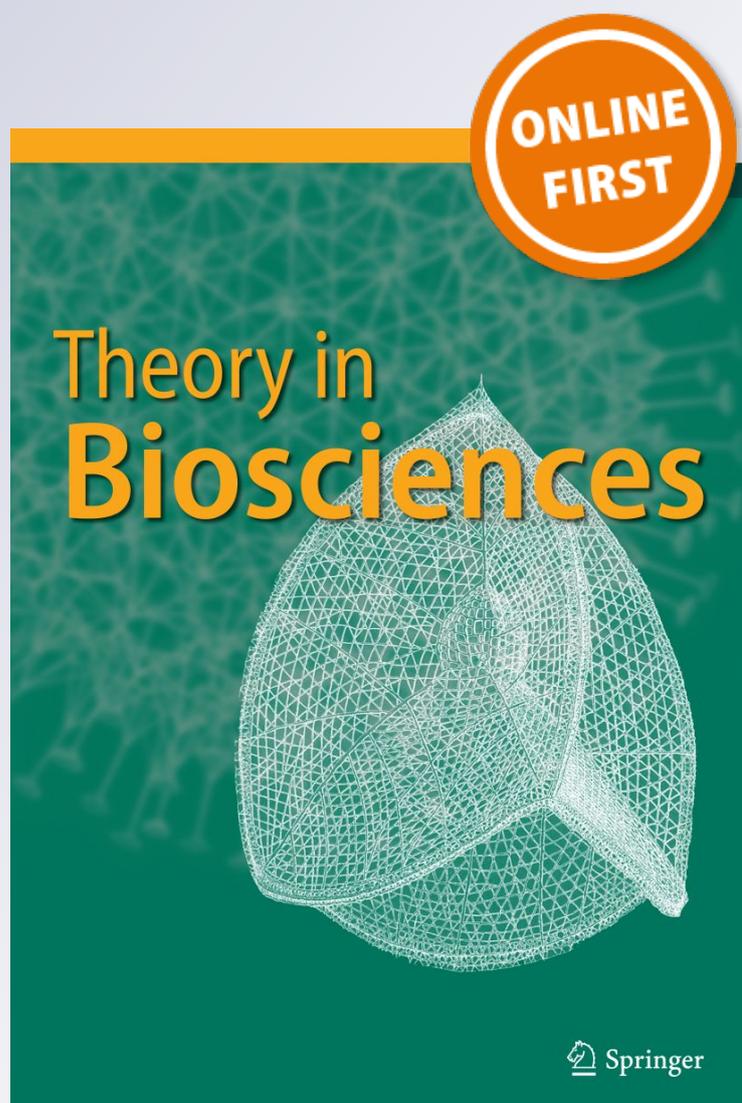
**Douglas H. Erwin**

**Theory in Biosciences**

ISSN 1431-7613

Theory Biosci.

DOI 10.1007/s12064-019-00303-4



 Springer

**Your article is protected by copyright and all rights are held exclusively by This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**



# Tempos and modes of collectivity in the history of life

Douglas H. Erwin<sup>1,2</sup>

Received: 22 August 2018 / Accepted: 4 September 2019

© This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2019

## Abstract

Collective integration and processing of information have increased through the history of life, through both the formation of aggregates in which the entities may have very different properties and which jointly coarse-grained environmental variables (ranging from widely varying metabolism in microbial consortia to the ecological diversity of species on reefs) and through collectives of similar entities (such as cells within an organism or social groups). Such increases have been implicated in significant transitions in the history of life, including aspects of the origin of life, the generation of pangenomes among microbes and microbial communities such as stromatolites, multicellularity and social insects. This contribution provides a preliminary overview of the dominant modes of collective information processing in the history of life, their phylogenetic distribution and extent of convergence, and the effects of new modes for integrating and acting upon information on the tempo of evolutionary change.

**Keywords** Evolutionary novelty · Major evolutionary transitions · Pangenome · Convergence

## Introduction

Life has been characterized by the formation of distributed processes in which interacting entities, often of very dissimilar nature, collectively integrate information and produce a coherent output. Such processes have been described among species within some ecosystems, such as reefs, among individuals within social groups and among differentiated cells within organisms, and even among similar (although not identical) cells within an organ such as the brain (Daniels et al. 2016). In the gene regulatory networks active during development (dGRNs), the developing organism, tissues and cells receive input about the external environment and the developing embryo process those inputs to affect the course of development. The lesser the stereotyped development (in other words, the greater phenotypic plasticity), the greater the influence environmental cues may have on the shape and

development of the organism. Thus, developmental regulation acts as a process of collective information processing for the organism. Information, whether ecological, environmental, developmental or in some other forms, serves as the currency through which biological entities make inferences about the present and future. Paleontologists describe such processes as time-averaging but the concepts discussed here encompass variety of patterns and processes with a variety of terms depending upon the field and the analytical approach.

I will argue here that such distributed or collective information processing has increased in extent and breadth through the history of life with the evolution of new organisms and increases in patterns of integration and expansion in the mechanisms of gene regulation, among other processes. Such biological processes have been broached in a variety of different contexts. Social dynamics have been a primary focus of studies of collective computation (Flack 2012, 2017a; Hein et al. 2015), from quorum sensing in bacteria (Cornforth et al. 2014; Waters and Bassler 2005) to cumulative cultural evolution in humans (Boyd 2018; Smaldino and Richerson 2013). More broadly, however, similar ideas have been developed for studying neural circuits (Hopfield and Tank 1986), physiological processes in plants (Peak et al. 2004) and the migration of neural crest cells during vertebrate development (Szabo and Mayor 2016).

This article is part of the Special Issue on Quantifying Collectivity.

✉ Douglas H. Erwin  
erwind@si.edu

<sup>1</sup> Department of Paleobiology, MRC-121, National Museum of Natural History, Washington, DC 20013-7012, USA

<sup>2</sup> Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Given the range of genetic, developmental, ecological, social and cultural processes that incorporate some degree of distributed or collective information processing, it is hardly surprising that these phenomena have been studied in different ways by different disciplines. For example, niche construction focuses on the fitness effects of changes organisms make to their environment, while those interested in ecosystem engineering have largely focused on the ecological impacts of similar activities. In this contribution, I will survey the range of distributed or collective information processing across the history of life and the context in which they have been studied. One view of evolution is that it is fundamentally a computational process (Bergstrom and Rosvall 2011; Flack 2017a; Valiant 2013) and one dimension of such a computational view of evolution is collective computation.

Collective computation is a mechanistic view of collective behaviors, initially focused on learning but which has been extended to incorporate a wider range of social, ecological and evolutionary processes (Daniels et al. 2016; Flack 2017a, b). Problems of distributed or collective information processing can be investigated as analyzing the solutions nature has found to two problems: how mapping of information is achieved and encoded, and how collective information is integrated to produce coherent output. Finally, a collective computation approach investigates the relationship between this output and the environment.

The growth of collective information processing has been a significant trend in the history of life. Whether collective computation is a useful perspective in understanding this trend depends largely on the questions one is seeking to answer. If the critical question of collectivity is how distributed processes comprising entities with differing and only partial correlated information in aggregate generate novel functions (Daniels et al. 2016), then specific questions for those interested in the history of life are how increases in collective information processing have arisen, whether they have been particularly associated with major episodes of novelty and innovation, whether these processes are unevenly concentrated across the tree of life and whether any have arisen convergently, and what impact the formation of new methods for integration and processing has had upon the tempo(s) of evolutionary change. In this contribution, I begin by addressing the dominant modes, phylogenetic distribution and implication for the tempo of evolution of distributed or collective information processing. In the discussion, I return to collective computation, and how this approach might provide a useful understanding of increased collectivity through the history of life.

## Dominant modes of collectivity in the history of life

Here, I can only sketch the nature of integrative information gathering and processing in the history of life and provide limited examples. I find it useful to consider the enormous variety of such processes in two classes: aggregates of dissimilar entities, such as microbial consortia, pangenomes and ecological communities, and collectives of similar entities, particularly in multicellular development. Many other aspects of the history of life also illustrate the extension of collective efforts, including the development of social and cultural strategies, but these will not be addressed here.

## Aggregational modes

Throughout the history of life, aggregates of dissimilar entities have generated the ability to integrate disparate and partial information sources to produce a coherent output, with some of the most potent examples coming from the behavior of microbes. Of the many examples that could be provided, I will highlight just three here: microbial consortia forming distinctive sedimentary structures known as stromatolites, the origin of the eukaryotic cell and sharing gene sequences. The coherent output of the collective differs in each example, in illustrative ways.

Stromatolites are layered sedimentary structures formed by microbial assemblages and are the oldest visible evidence of life on Earth at about 3.5 billion years ago (Gya) (Allwood et al. 2006). Perhaps surprisingly, as they are formed by diverse communities of microbes, stromatolites form a variety of distinctive macrostructures: domes, columns and branching columns. At a finer level, the laminae within a stromatolite reflect microbial interactions with sediment. Living stromatolites show a layered structure with oxygen-producing phototrophs in the surficial layers. A sharp attenuation of light and oxygen with depth allows formation of layers of progressively more redox-restricted heterotrophic bacteria including sulfate reducers and other metabolisms (Grotzinger and Knoll 1999; Reid et al. 2003). Although stromatolites are not as widespread as they were a billion or more years ago, they occur today in restricted marine settings such as Shark Bay in Western Australia. The aggregate collective interaction of microbial communities with the environment generates the distinctive stromatolite morphologies.

If stromatolites represent the sedimentary products of microbial consortia jointly confronting the environment, symbiosis between cells can be seen as an extension of feedbacks between organisms, where one microbial lineage takes up residence and contributes to a host cell. Symbioses form at many levels within cells and have played an important role

in facilitating metabolic diversity (with the chloroplast only the most prominent example) (Douglas 2014). The origin of eukaryotic cells was the most significant such event, combining an archeal host cell, probably related to the Asgard clade of archaea, and an  $\alpha$ -proteobacteria which formed the mitochondria (Eme et al. 2017; Lopez-Garcia and Moreira 2015; McInerney et al. 2014). The origin of the eukaryotic cell appears to have been a singular event in evolution, albeit one with incredible consequences for the history of life. Eukaryotic cells have greater internal complexity than bacterial or archaeal cells (the other two major clades of life), can reach larger sizes as a single cell and can generate a wide range of multicellular forms, including plants, animals and fungi.

Many microbes are capable of adapting rapidly to new conditions, and one of the techniques enabling this rapid evolution is the presence of a large suite of genes that are not present in every genome. The classic example is *Escherichia coli* which has a core of 2085 genes found in upward of 95% of the 2085 genomes sampled. Individual genomes of *E. coli* have 4500–5000 genes, and the additional 2500 or so genes beyond the conserved, core genome come from representatives of another 90,000 gene families composing the accessory genome (McInerney et al. 2017). Genes from the accessory genome are found in individuals depending on recent or existing environmental conditions. Pangenomes are not restricted to bacteria, but are found in eukaryotes as well. Through the accessory genome within a species and horizontal gene transfer (HGT), some estimates suggest that there may be an almost limitless supply of genes from which microbes can sample (Lapierrel and Gogarten 2009). Both the pangenome and HGT challenge the traditional ‘tree-thinking’ of many biologists (McInerney et al. 2011), but more importantly they establish the capacity of many species to adapt to changing environmental conditions by integrating imperfect information about shifts in their environment and adjusting their genomic composition, and thus their phenotypic response, accordingly. (The term ‘pan-genome’ has also been applied to the combined genomes of host and microbial elements in a holobiont, but this is a very different meaning of the term.)

The most obvious biological aggregates are ecological assemblages, from microbial communities to tropical forests and coral reefs to the soil in your garden. The next section will discuss the ways in which individual species may be involved in collective ecological processes through niche construction and social interactions, but whether ecological assemblages and communities exhibit collective computational properties depends upon the nature of these assemblages. If they are largely random assortments of species responding to similar environmental conditions, then ecological assemblages are less likely to exhibit the properties of collective computation. Some assemblages, however, may exhibit the collective integration of information and

generation of coherent output that characterizes collective computation. This dichotomy roughly corresponds to the distinction between the individualistic, Gleasonian view of ecological communities versus the more tightly integrated Clementsian view familiar from any ecology textbook. Recent work in ecology has shown that ecological assemblages are more complex than the simple Gleasonian versus Clementsian dichotomy, with some more closely matching one and others more closely matching another. Theoretical studies of modularity and interconnectedness of ecosystems suggest that both attributes are likely to increase through time (Nordbotten et al. 2018) (at least between major biotic disturbances, which may reset ecosystem dynamics: Wagner et al. 2006). Here, I simply want to signpost the structure of ecological communities as a promising avenue for future explorations of collective computation.

### Collective modes

Collective entities have probably figured throughout the history of life, but their extent and impact have expanded greatly over the past billion years with the origin and expansion of clades with complex, differentiated multicellularity: animals, fungi and the green algae and land plants. Social interactions are found in many microbial clades and likely date back billions of years, but have also increased in scope with the diversification of a range of animal groups with social behavior. Social insects and various mammalian groups display a greater range of social behaviors than the animal clades that dominated the Paleozoic, for instance. Here, I briefly discuss four examples of collective activities: multicellular development, particularly in animals; the role of nervous systems in internal coordination versus external sensing; social cooperation; and niche construction. Many other collective modes could be studied and which likely have been important, including cellular collectives (Nandell et al. 2014), increased modularity (Melo et al. 2016; Wagner et al. 2007), canalization and evolvability (Hallgrímsson et al. 2019; Pigliucci 2008; Wagner and Altenberg 1996) and evolution and learning (Watson et al. 2016; Watson and Szathmáry 2016).

Multicellularity has evolved in dozens of different eukaryotic lineages, although only plants, fungi and animals have acquired highly differentiated cell types and complex developmental controls (Knoll 2011). The multicellularity guided by dGRNs is the type of multicellularity most commonly encountered, but life has effectively achieved specialization and division of labor via two other routes as well: effective multicellularity of microbial consortia (Shapiro 1988) and the various levels of endosymbiosis within eukaryotic cells (Delwiche 1999). The eukaryotic cell is most prominent example of the importance of symbiosis, but each of these three strategies has required the evolution of means of

integrating information, enabling cooperation and managing conflict. These three strategies have also enabled greater complexity and varying degrees of independence from the environment.

Building from multi-functional single-celled ancestors (Arendt 2008), multicellular clades acquired the tools for complex spatial and temporal patterning of gene activity. No overall design or control of animal development exists, even in the genome. Rather, metazoan development involves interaction between the logical structure of regulatory interactions encoded in the genome, the developing embryo and the environment. Multicellular patterning has been best studied in animals, where there are a number of examples of intensively studied gene regulatory networks (Peter and Davidson 2015). Davidson used a Boolean network approach to dGRNs, illustrating how the interactions between different components of the regulatory genome (transcription factors, enhancers, signaling pathways) generate regional patterning and cell specification. The Boolean approach has been validated experimentally (Peter et al. 2012) demonstrating the logical structure of regulatory patterning. The evolution of this control system involves elaboration and co-option of regulatory subcircuits as well as the expansion of the developmental networks (Erkenbrack and Davidson 2015; Erwin and Davidson 2009; Peter and Davidson 2017; Rebeiz et al. 2015). Although we know that the complexity of developmental processes expanded greatly with the origin and early evolution of Metazoa (Sebe-Pedros et al. 2016; Tweedt and Erwin 2015), with the growth of transcription factor families (Schmitz et al. 2016) and regulatory RNAs (Morris and Mattick 2014), in most cases, there is too little comparative GRN data to provide a more detailed view of their evolution. One system in which there is considerable detail on the origin and elaboration of GRNs and the regulatory genome are animal nervous systems.

The role of animal nervous systems in sensing the external environment, transmitting information and controlling responses is well known, and there is a voluminous literature on computation as a model for neural activity. Here, I am less concerned with neural activities in humans, or even mammals, than with the origin and early evolution of the nervous system (Arendt et al. 2016; Erwin 2015). Nervous systems play equally important roles in internal coordination within an organism. As Jékeley and colleagues have noted, coordinating muscle activity is not a simple task, and whether the external input–output function or internal control was the initial function of early nervous systems is unresolved (Jekely et al. 2015). They argue that the two models must (and can) be reconciled to understand how early animals controlled behavior, physiology and development. The origin, initial establishment and subsequent evolution of animal nervous systems has become an area of intensive interest over the past decade as comparative developmental studies

with a broader range of organisms provide new insights. Because of this work, studies of neural circuit evolution provide considerable insight into the evolution of dGRNs and of collective modes of collective computation more generally. For example, Tosches (2017) identified six modes of neural circuit evolution, discussed the types of genes involved, the postulated developmental and regulatory mechanisms of activity and provided some specific examples. Because network evolvability is a critical issue for collective computation, one conclusion from Tosches is of considerable relevance. She suggests that the topology of different neural circuits (modularity, feedback, connectivity and other properties) generates different functional properties, with only a portion of the possible changes to the circuit yielding viable outcomes. She concludes that there may be an inherent bias imposed on circuit topology and its expression in the genome imposed by these limitations.

Much previous work on collective activities has focused on social cooperation, particularly in vertebrates (e.g., Flack 2012), and social insects (Feinerman and Korman 2017), as well as flocking behavior and other forms of coordination (Couzin 2009; Hein et al. 2015). Social collective computation has expanded with the diversification of placental mammals through the Cenozoic and with the expansion of various clades of eusocial insects since the Cretaceous. But collective activities are found across Eumetazoa. Coordinated broadcast spawning occurs in many marine groups, from cnidarians to molluscs. The coordination of behaviors through stigmergy is often explained with ants, but applies equally to a variety of burrowing organisms which can be recorded as trace fossils (Plotnick 2003; Seilacher 1986). The burst of new burrowing behaviors which marks the appearance of animals in the fossil record (Jensen et al. 2005) expresses increased behavioral complexity, often with overtones of interaction, across many marine clades. The evolution of neural circuits discussed in the previous paragraph will be relevant for the evolving possibility space for social cooperation, but this is an issue that awaits further exploration.

Although many biologists once viewed organisms as responding passively to environmental challenges, Darwin knew that this was not true: His final book was on earthworms (Darwin 1881). Many organisms actively construct their own niche, influencing their fitness and often that of other species. The construction of beaver dams is a classic example of niche construction (Odling-Smee et al. 2003) or ecosystem engineering (Jones et al. 1997). But effect of a single organism on its environment is generally limited. Rather, niche construction is the consequence of the collective impact of organisms, often over many generations, on their environment. But collectively these activities can significantly impact physical, chemical and social or cultural resources. Moreover, although the evidence remains

rudimentary, niche constructing activities may have been important in the Ediacaran–Cambrian diversification of animals (Erwin and Tweedt 2011) and seem to have increased over time (Erwin 2008). These largely impressionistic surveys leave considerable room for more rigorous and quantitative analysis of the history of niche construction through the Phanerozoic and its evolutionary impact.

### Phylogenetic distribution of collective computation

Evolutionary biologists inevitably view almost any problem from a ‘tree-thinking’ perspective, evaluating the phylogenetic distribution of some feature and considering whether it could have arisen multiple times independently (convergence). Convergence is a ubiquitous feature of evolution, but robust analysis of it requires considerable specificity. Insects, birds, bats all fly as did pterosaurs during the Mesozoic, but the specific features of their wings and powered flight are unique to each clade, and thus, they cannot be considered to be convergent except in the grossest sense. Similarly, niche construction occurs in so many different ways that it is too broad a concept with which to evaluate convergence. More useful, for example, was the discovery of a Miocene whale in Peru, *Odobenocetops*, where the skull is highly convergent upon walruses (de Muizon 1993). Walruses are restricted to the Northern Hemisphere, but evidently grubbing through sediment with your tusks for bivalves was a sufficiently attractive way of making a living that this whale evolved the same adaptations. Specific types of niche construction (disturbance of the sediment as a feeding strategy for molluscs) can be associated with the evolution of specific clades. More broadly, nervous systems and the role of embodied cognition are specific to animals (even if there is some controversy over the number of times that nervous systems evolved). With some knowledge of the phylogenetic distribution of different types of collective activities, two further questions can be addressed: Was a particular mode evolutionarily unique, or sufficiently common that it was almost certain to evolve? And what impact did the appearance of collective activities have on the clade which possessed it: Did it lead to diversification of species, greater ecological impact or improved performance?

These two questions are important tests of the importance of collectivity. If it provides significant advantages to clades, then we should expect those types of collectivity which are relatively difficult to acquire to be found in clades which, all other things being equal, are more successful than sister clades lacking them. Such tests are now a standard part of the repertoire of comparative phylogenetics. In contrast, aspects of collectivity which are

more evolutionarily accessible, including many types of niche construction, colony formation and developmental modularity, among others, should be expected to have arisen independently in many different clades (although the specific implementations may be very different) and may often exhibit convergence.

In general, because aggregational forms of collectivity involve the assembly of different entities, they are more likely to arise independently in various circumstances through time and space. Thus, reefs, as means of localizing productivity in a high energy, often tropic environment, have arisen repeatedly over at least the past 2 billion years, but with great variety in the primary organisms responsible for forming the reefs, from microbes to sponges, bryozoans, bivalves to the corals which dominate reefs today (Kiessling et al. 2010; Wood 1999). Many other ecological associations have similar properties, including microbial consortia. As discussed previously, the phylogenetic extent of pangenomes is currently unclear, but they represent a viable strategy to balance the drive for small genomes in individual microbes (to accelerate growth) with the need to maximize genetic variation for changing environmental conditions. The considerations outlined above would suggest that pangenomes are likely to have evolved many times among various microbial groups. There is an obvious exception to the generalization that aggregational collectivity is likely to be relatively easy to acquire: the origin of the eukaryotic cell. There is no evidence that eukaryotes evolved more than once, although secondary and tertiary symbioses are widespread (Delwiche 1999).

The evolution of development is one area of collective modes where there is clear evidence of a phylogenetic signal both across eukaryotes and within Metazoa. As noted above, although multicellularity is found in many different eukaryotic clades, complex differentiation has appeared in only three clades. Taking either the number of differentiated cell types or McShea’s more generalized metric of parts within parts (McShea 1991), these lineages have experienced increases in collectivity and in complexity both relative to other eukaryotic lineages and within each clade. Within Metazoa, the complexity of developmental processes has expanded considerably since the clades diverge from the closest living outgroups (choanoflagellates and filastrians). Processes of regulatory control have expanded within Metazoa, including the development of distal enhancers (Sebe-Pedros et al. 2011), expansion of transcription factor families and signaling pathways (Degan et al. 2009; Sebe-Pedros and de Mendoza 2016) which increased the combinatorics possibilities of *cis*-regulatory gene regulation, and in the regulatory processes, they control (Davidson and Erwin 2010; Erwin and Davidson 2009; Rebeiz et al. 2015). Regulatory control via various small RNAs such as microRNA has also increased (Berezikov

2011). Expansions of the combinatorics complexity have also occurred within lineages, with perhaps the best-studied example being butterfly wing patterns, where enhancer shuffling and a modular design have enabled the generation of wide variation even within a single genus (Gompel et al. 2005; Van Belleghem et al. 2017).

### Effects of collectivity on the tempo of evolutionary change

The extent of collectivity through the history of life raises the question of its impact on both the tempo and the mode of evolutionary dynamics. The second of these questions is the easier to answer, as collectivity is associated with several of the major evolutionary transitions, each of which established new entities upon which selection can act.

The origin of the eukaryotic cell was one of eight events in the history of life which have been described as major evolutionary transitions (METs; Maynard Smith and Szathmáry 1995) and evolutionary transitions in individuality (Calcott and Sterelny 2011; Michod 1999). These events involve the formation of new evolutionary individuals via cooperation and integration of entities which were previously independent. Of the eight METs, the first three were concerned with the origin of life, while the others were the origin of the eukaryotic cell, origin of sex, sexual differentiation, coloniality and the origin of human societies and language. In his review of Maynard Smith and Szathmáry's book, Queller suggested that some transitions were egalitarian, such as the origin of the eukaryotic cell where both entities continue to reproduce but divide tasks in the new entity, while others were fraternal, including cells in a multicellular organism or social insects (Queller 1997). This egalitarian vs fraternal dichotomy roughly parallels the aggregational vs collective distinction used here. From the standpoint of collectivity, METs are noteworthy because the formation of new evolutionary individuals requires means to ensure cooperation and manage the conflict between previously distinct entities. But METs also generate opportunities. A promising avenue of future work would be exploring the role collectivity may have played in both facilitating the METs themselves, and in exploiting the new opportunities created. But to the extent that evolutionary dynamics such as the METs influence modes of evolution, they will also impact the extent of collectivity. In cases such the origins of eukaryotes, multicellularity and language, three of the canonical METs, expansion of collectivity is inherently part of these transitions.

Canalization of development to buffer variability and ensure faithful growth is a clear example of collectivity across the developing embryo (Takahashi 2018). Evaluating the importance of this across evolutionary timescales

is difficult, but as Webster has recently described, robustness against developmental noise can be assessed in the fossil record by measuring the fluctuating asymmetry of bilaterally symmetrical organisms such as trilobites (Webster 2018).

Rigorously evaluating the tempo of evolutionary change, particularly in deep time, is generally challenging and requires either a robust fossil record with excellent age control or an equally robust phylogenetic framework with molecular clock estimates of divergence times (Erwin 2006). Of course, all of these would be even more useful. In a general sense, the influence of collectivity on tempo of evolutionary change is currently difficult to evaluate. The one area in which this may not be true is for aspects of social cooperation, particularly where cultural evolution has evolved. By introducing an additional dynamic, the rate of some aspects of evolution may have increased.

Thus, collectivity is *associated* with changes in tempo and mode and there are circumstantial reasons to believe that it could increase rates of evolutionary change where there is a direct fitness effect, for example, in niche construction, the pangenome and in regulatory networks (whether developmental, ecological or social). This correlation leaves unresolved whether increased collectivity itself drove the increased rates of change, or whether it was a consequence of evolutionary success due to other factors will require more fine-grained and specific studies.

### Discussion and conclusions

The brief survey here suggests that collectivity, both in aggregate and in collectives, has increased in frequency and diversity through the history of life, even leaving aside contributions from a highly encephalized ape over the past million years. Certain aspects of evolution such as the new evolutionary individuals constructed during METs appear to have had significant impacts on the extent of collective activity. More intriguingly, there is also suggestive evidence from developmental GRNs, and particularly from work on neural circuits, that the topology of such networks may influence the possibilities for subsequent evolutionary change.

In this contribution, I have used the terms 'collective information processing' and 'collectivity' to encompass a wide range of organismal activities, but the focus of this special issue (and the meeting from which these papers originated) was more narrowly focused on collective computation. Inferring that the processes described here reflect collective computation raises a number of issues. Most importantly perhaps is whether we can attribute such activities in the deep past to computation on the basis of comparison to experimental studies of living taxa. Paleobiologists have generally steered clear of invoking fitness

differences, for example, because it is rarely possible to rigorously assess such differences in deep time. Hence, in examining the macroevolutionary dynamics of niche construction and ecosystem engineering, I have largely used the latter term (Erwin 2008). Beyond this, the term ‘collective computation’ naturally raises the question of just what the various processes discussed here actually ‘compute.’ One could consider developmental GRNs to be computing organismal phenotype and social circuits to be computing behavior, for example. I prefer to view computation as a metaphor, however, as it is unclear that any of the cases described here actually involve algorithmic computation. There are many biological processes, including some described here, which can be *modeled* as computation: GRNs have been effectively modeled as both Boolean networks (Bolouri and Davidson 2002; Peter and Davidson 2011) and by differential equations. Each approach captures something of the behavior of gene regulatory networks and is useful in addressing different questions. That computational algorithms provide informative models does not mean, however, that these systems here are computing algorithmically. This may not matter, however, if the metaphor of collective computation helps articulate a research program into important aspects of the history of life.

Flack (2017a) has argued that increased understanding of the role of collective computation in biology will help in identifying laws or generalities. Over the past decade paleobiologists, other evolutionary biologists and philosophers of biology have been engaged in a spirited debate over the relative significance of deterministic versus contingent aspects of the history of life, with the former arguing that evolutionary trajectories and outcomes have been more limited than might otherwise seem apparent (Conway Morris 2010; Erwin 2016; McGhee 2011, 2016). Investigations of collective computation provide an interesting window into this debate. Recall Tosches’ conclusion that the topology of neural networks influenced permitted pathways of subsequent change. This is likely to be true of GRNs more generally and possibly other aspects of the regulatory genome, and this could be one avenue constraining phenotypic evolution. Whether this could also be true of the many other networks involved in collective computation is unknown at this point, but certainly greater appreciation of the tempos and modes of collective computation may also help in adjudicating the relative contributions of deterministic and contingent processes.

**Acknowledgements** I appreciate the opportunity extended by the organizers to participate in the Santa Fe Institute working group meeting on collective computation (B. Daniels, J. Flack and M. Laubichler) and to contribute to this special issue. The discussions among the participants in the meeting also greatly influenced the resulting paper. I was supported by NASA through National Astrobiology Institute (Grant No. NNA13AA90A) to the MIT node.

## References

- Allwood AC, Walter MR, Kamber BS, Marshall CP, Burch IW (2006) Stromatolite reef from the Early Archaean era of Australia. *Nature* 441:714–718
- Arendt D (2008) The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet* 8:868–882
- Arendt D, Tosches MA, Marlow H (2016) From nerve net to nerve ring, nerve cord and brain: evolution of the nervous system. *Nat Rev Neurosci* 17:61–72. <https://doi.org/10.1038/nrn.2015.15>
- Berezikov E (2011) Evolution of microRNA diversity and regulation in animals. *Nat Rev Genet* 12:846–860
- Bergstrom CT, Rosvall M (2011) The transmission sense of information. *Biol Philos* 26:159–176. <https://doi.org/10.1007/s10539-009-9180-z>
- Bolouri H, Davidson EH (2002) Modeling transcriptional regulatory networks. *BioEssays* 24:1118–1129
- Boyd R (2018) A different kind of animal. Princeton University Press, Princeton
- Calcott B, Sterelny K (eds) (2011) The major transitions in evolution revisited. MIT Press, Cambridge
- Conway Morris S (2010) Evolution: like any other science it is predictable. *Philos Trans R Soc Lond B Biol Sci* 365:133–145. <https://doi.org/10.1098/rstb.2009.0154>
- Cornforth DM et al (2014) Combinatorial quorum sensing allows bacteria to resolve their social and physical environment. *Proc Natl Acad Sci USA* 111:4280–4284. <https://doi.org/10.1073/pnas.1319175111>
- Couzin ID (2009) Collective cognition in animal groups. *Trends Cog Sci* 13:36–43. <https://doi.org/10.1016/j.tics.2008.10.002>
- Daniels BC, Ellison CJ, Krakauer DC, Flack JC (2016) Quantifying collectivity. *Curr Opin Neurobiol* 37:106–113. <https://doi.org/10.1016/j.conb.2016.01.012>
- Darwin C (1881) The formation of vegetable mould, through the action of worms. John Murray, London
- Davidson EH, Erwin DH (2010) Evolutionary innovation and stability in animal gene networks. *J Exp Zool (Mol Dev Evol)* 314B:182–186
- de Muizon C (1993) Walrus-like feeding adaptation in a new cetacean from the Pliocene of Peru. *Nature* 365:745–748
- Degnan BM, Vervoort M, Larroux C, Richards GS (2009) Early evolution of metazoan transcription factors. *Curr Opin Genet Dev* 19:591–599. <https://doi.org/10.1016/j.gde.2009.09.008>
- Delwiche CF (1999) Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* 154(Suppl):S164–S177
- Douglas AE (2014) Symbiosis as a general principle in eukaryotic evolution. *Cold Spring Harbor Perspect Biol* 6:a016113
- Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG (2017) Archaea and the origin of eukaryotes. *Nat Rev Microbiol* 15:711–723. <https://doi.org/10.1038/nrmicro.2017.133>
- Erkenbrack EM, Davidson EH (2015) Evolutionary rewiring of gene regulatory network linkages at divergence of the echinoid subclasses. *Proc Natl Acad Sci USA* 112:E4075–E4084. <https://doi.org/10.1073/pnas.1509845112>
- Erwin DH (2006) Dates and rates: temporal resolution in the deep time stratigraphic record. *Ann Rev Earth Planet Sci* 34:569–589
- Erwin DH (2008) Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol Evol* 23:304–310
- Erwin DH (2015) Early metazoan life: divergence, environment and ecology. *Philos Trans R Soc B*. <https://doi.org/10.1098/rstb.2015.0036>
- Erwin DH (2016) Wonderful life revisited: chance and contingency in the Ediacaran-Cambrian radiation. In: Ramsay G, Pence CH (eds) *Chance in evolution*. University of Chicago Press, Chicago, pp 277–298

- Erwin DH, Davidson EH (2009) The evolution of hierarchical gene regulatory networks. *Nat Rev Genet* 10:141–148. <https://doi.org/10.1038/nrg2499>
- Erwin DH, Tweedt SM (2011) Ecosystem engineering and the Ediacaran-Ordovician diversification of Metazoa. *Evol Ecol* 26:417–433
- Feinerman O, Korman A (2017) Individual versus collective cognition in social insects. *J Exp Biol* 220:73–82. <https://doi.org/10.1242/jeb.143891>
- Flack JC (2012) Multiple time-scales and the developmental dynamics of social systems. *Philos Trans R Soc B* 367:1802–1810. <https://doi.org/10.1098/rstb.2011.0214>
- Flack JC (2017a) Life's information hierarchy. In: Walker SI, Davies PCW, Ellis GFR (eds) *From matter to life*. Cambridge University Press, Cambridge, pp 283–302
- Flack JC (2017b) Coarse-graining as downward causation. *Philos Trans R Soc A* 375(2109):20160338
- Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433:481–487
- Grotzinger JP, Knoll AH (1999) Stromatolites in Precambrian carbonates: evolutionary mileposts or environmental dipsticks? *Ann Rev Earth Planet Sci* 27:313–358
- Hallgrímsson B et al (2019) The developmental-genetics of canalization. *Sem Cell Dev Biol* 88:69–79. <https://doi.org/10.1016/j.semcdb.2018.05.019>
- Hein AM, Rosenthal SB, Hagstrom GI, Berdahl A, Torney CJ, Couzin ID (2015) The evolution of distributed sensing and collective computation in animal populations. *eLife*. <https://doi.org/10.7554/eLife.10955>
- Hopfield JJ, Tank DW (1986) Computing with neural circuits: a model. *Science* 233:625–633. <https://doi.org/10.1126/science.3755256>
- Jekely G, Keijzer F, Godfrey-Smith P (2015) An option space for early neural evolution. *Philos Trans R Soc Lond B*. <https://doi.org/10.1098/rstb.2015.0181>
- Jensen S, Droser ML, Gehling JG (2005) Trace fossil preservation and the early evolution of animals. *Palaeogeogr Palaeoclimatol Palaeoecol* 220:19–29
- Jones CG, Lawton JH, Shachak M (1997) Positive and negative effects of organisms as physical ecosystem engineers. *Ecology* 78:1946–1957
- Kiessling W, Simpson C, Foote M (2010) Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. *Science* 327:196–198. <https://doi.org/10.1126/Science.1182241>
- Knoll AH (2011) The multiple origins of complex multicellularity. *Ann Rev Earth Planet Sci* 39:217–239
- Lapierrel P, Gogarten JP (2009) Estimating the size of the bacterial pan-genome. *Trends Genet* 25:107–110. <https://doi.org/10.1016/j.tig.2008.12.004>
- Lopez-Garcia P, Moreira D (2015) Open questions on the origin of eukaryotes. *Trends Ecol Evol* 30:697–708. <https://doi.org/10.1016/j.tree.2015.09.005>
- Maynard Smith J, Szathmáry E (1995) *The major transitions in evolution*. W. H. Freeman, New York
- McGhee GR Jr (2011) *Convergent evolution*. MIT Press, Cambridge
- McGhee GR Jr (2016) Can evolution be directional without being teleological? *Stud Hist Philos Biol Biomed Sci* 58:93–99. <https://doi.org/10.1016/j.shpsc.2015.12.006>
- McInerney J, Cummins C, Haggerty L (2011) Goods-thinking vs. tree-thinking: finding a place for mobile genetic elements. *Mob Genet Elem* 1:304–308. <https://doi.org/10.4161/mge.19153>
- McInerney JO, O'Connell MJ, Pisani D (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat Rev Microbiol* 12:449–455. <https://doi.org/10.1038/nrmicro3271>
- McInerney JO, McNally A, O'Connell MJ (2017) Why prokaryotes have pangenomes. *Nat Microbiol*. <https://doi.org/10.1038/nmicr.obiol.2017.40>
- McShea DW (1991) Complexity and evolution: what everybody knows. *Biol Philos* 6:303–324
- Melo D, Porto A, Cheverud JM, Marroig G (2016) Modularity: genes, development, and evolution. *Ann Rev Ecol Evol Syst* 47:463–486. <https://doi.org/10.1146/annurev-ecolsys-121415-032409>
- Michod RE (1999) *Darwinian dynamics*. Princeton University Press, Princeton
- Morris KV, Mattick JS (2014) The rise of regulatory RNA. *Nat Rev Genet* 15:423–437. <https://doi.org/10.1038/nrg3722>
- Nandell CD, Bucci V, Drescher K, Levin SL, Bassler BL, Xavier JB (2014) Cutting through the complexity of cell collectives. *Proc R Soc Lond B*. <https://doi.org/10.1098/rspb.2012.2770>
- Nordbotten JM, Levin SA, Szathmáry E, Stenseth NC (2018) Ecological and evolutionary dynamics of interconnectedness and modularity. *Proc Natl Acad Sci USA* 115:750–755. <https://doi.org/10.1073/pnas.1716078115>
- Odling-Smee FJ, Laland KN, Feldman MW (2003) *Niche construction: the neglected process in evolution*. Princeton University Press, Princeton
- Peak D, West JD, Messinger SM, Mott KA (2004) Evidence for complex, collective dynamics and emergent, distributed computation in plants. *Proc Natl Acad Sci USA* 101:918–922. <https://doi.org/10.1073/pnas.0307811100>
- Peter IS, Davidson EH (2011) A gene regulatory network controlling the embryonic specification of endoderm. *Nature* 474:635–639
- Peter IS, Davidson EH (2015) *Genomic control processes*. Development and evolution. Academic Press, London
- Peter IS, Davidson EH (2017) Assessing regulatory information in developmental gene regulatory networks. *Proc Natl Acad Sci USA* 114:5862–5869. <https://doi.org/10.1073/pnas.1610616114>
- Peter IS, Faure E, Davidson EH (2012) Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci USA* 109:16434–16442. <https://doi.org/10.1073/pnas.1207852109>
- Pigliucci M (2008) Is evolvability evolvable? *Nat Rev Genet* 9:75–82
- Plotnick RE (2003) Ecological and L-system based simulations of trace fossils. *Palaeogeogr Palaeoclimatol Palaeoecol* 192:45–58. [https://doi.org/10.1016/S0031-0182\(02\)00678-8](https://doi.org/10.1016/S0031-0182(02)00678-8)
- Queller DC (1997) Cooperators since life began. *Quart Rev Biol* 72:184–188
- Rebeiz M, Patel NH, Hinman VF (2015) Unraveling the tangled skein: the evolution of transcriptional regulatory networks in development. *Ann Rev Genom Human Genet* 16:103–131. <https://doi.org/10.1146/annurev-genom-091212-153423>
- Reid P, Dupraz CD, Visscher PT, Sumner DY (2003) Microbial processes forming marine stromatolites. In: Krumbein WE, Pateron DM, Zavarzin GA (eds) *Fossil and recent biofilms*. Springer, Dordrecht, pp 103–118
- Schmitz JF, Zimmer F, Bornberg-Bauer E (2016) Mechanisms of transcription factor evolution in Metazoa. *Nucl Acids Res* 44:6287–6297. <https://doi.org/10.1093/nar/gkw492>
- Sebe-Pedros A, de Mendoza A (2016) Transcription factors and the origin of animal multicellularity. In: Ruiz-Trillo I, Nedelcu AM (eds) *Evolutionary transitions to multicellular life*, vol 2. Springer, Dordrecht, pp 379–394
- Sebe-Pedros A, Ruiz-Trillo I, de Mendoza A, Lang BF, Degnan BM (2011) Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzakii*. *Mol Biol Evol* 28:1241–1254. <https://doi.org/10.1093/molbev/msq309>
- Sebe-Pedros A et al (2016) The dynamic regulatory genome of *Capsaspora* and the origin of animal multicellularity. *Cell* 165:1224–1237. <https://doi.org/10.1016/j.cell.2016.03.034>

- Seilacher A (1986) Evolution of behavior as expressed in marine trace fossils. In: Nitecki MH, Kitchel JA (eds) *Evolution of animal behavior*. Oxford University Press, Oxford, pp 62–87
- Shapiro JA (1988) Thinking about bacterial populations as multicellular organisms. *Ann Rev Microbiol* 52:81–104
- Smaldino PE, Richerson PJ (2013) Human cumulative cultural evolution as a form of distributed computation. In: Mechelucci P (ed) *Handbook of human computation*. Springer, New York, pp 979–992
- Szabo A, Mayor R (2016) Modelling collective cell migration of neural crest. *Curr Opin Cell Biol* 42:22–28. <https://doi.org/10.1016/j.ceb.2016.03.023>
- Takahashi KH (2018) Multiple modes of canalization: links between genetic, environmental canalizations and developmental stability, and their trait-specificity. *Sem Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.05.018>
- Tosches MA (2017) Developmental and genetic mechanisms of neural circuit evolution. *Dev Biol* 431:16–25. <https://doi.org/10.1016/j.ydbio.2017.06.016>
- Tweedt SM, Erwin DH (2015) Origin of metazoan developmental toolkits and their expression in the fossil record. In: Ruiz-Trillo I, Nedelcu AM (eds) *Evolution of multicellularity*. Academic Press, London, pp 47–77
- Valiant L (2013) *Probably approximately correct*. Basic Books, New York
- Van Belleghem SM et al (2017) Complex modular architecture around a simple toolkit of wing pattern genes. *Nat Ecol Evol* 1:52. <https://doi.org/10.1038/s41559-016-0052>
- Wagner GP, Altenberg L (1996) Complex adaptations and the evolution of evolvability. *Evolution* 50:967–976
- Wagner PJ, Kosnik MA, Lidgard S (2006) Abundance distributions imply elevated complexity of post-Paleozoic marine ecosystems. *Science* 314:1289–1292
- Wagner GP, Pavlicev M, Cheverud JM (2007) The road to modularity. *Nat Rev Genet* 8:921–931
- Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. *Ann Rev Cell Dev Biol* 21:319–346. <https://doi.org/10.1146/annurev.cellbio.21.012704.131001>
- Watson RA, Szathmary E (2016) How can evolution learn? *Trends Ecol Evol* 31:147–157. <https://doi.org/10.1016/j.tree.2015.11.009>
- Watson RA et al (2016) Evolutionary connectionism: algorithmic principles underlying the evolution of biological organisation in evo-devo, evo-eco and evolutionary transitions. *Evol Biol* 43:553–581. <https://doi.org/10.1007/s11692-015-9358-z>
- Webster M (2018) Morphological homeostasis in the fossil record. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.05.016>
- Wood R (1999) *Reef evolution*. Oxford University Press, Oxford

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.