æ

24. ORIGINS OF ORGANISMAL COMPLEXITY

11 January 2023

In the antecedent to this book, a case was made for the emergence of genome complexity as a passive response to mutational bias and the limited reach of natural selection in lineages experiencing elevated levels of genetic drift (Lynch 2007). This left open the question as to whether differences at the next highest level of organization – the cell – might also be a product of effectively neutral processes. One might anticipate that the higher the level of organization and the closer to outward phenotypic expression, the more likely it is that natural selection will be fully responsible for patterns of variation. With this reasoning, some would even argue that none of the internal molecular and cellular details matter. However, multiple observations summarized in the preceding chapters make clear that this is not the case.

What follows is a synthesis of results from comparative cell biology and evolutionary theory relevant to the matter of the emergence and diversification of cellular and organismal complexity. Evolutionary biology is more than the application of phylogenetic comparative analyses and the concoction of adaptive hypotheses for observed historical patterns. Understanding the mechanisms by which evolution proceeds requires an appreciation of the functional constraints on cells imposed by the historical roots of biology's invariants (e.g., the dependence on DNA-based genomes, RNA-based transcriptomes, protein- and lipid-based infrastructures, and the reliance on ATP as a supplier of energy). As the biochemical, bioenergetic, and biophysical properties of life's cellular components ultimately define the limited ways by which mutational processes can introduce variation into populations, a mechanistic understanding of evolution is inconceivable without information on its material basis. Likewise, conditional on the material endowment of life, the imposition of key population-genetic features dictate the paths of possible evolutionary trajectories, defining the types of changes that are open vs. closed to exploitation by natural selection and mutational biases.

At the time of Wallace and Darwin, essentially nothing was known about genetics or the molecular constituents of cells, removing all inhibitions to thinking that natural selection is essentially all powerful. Things might have been different had biology's great phase of molecular and cell emphasis come earlier. We now know that the need to expand evolutionary biology beyond a pan-adaptationist framework is deeply rooted in biological reality. Mutations are particulate in nature, with effects that are typically small and deleterious, and the transmission of alleles across generations is a stochastic process, with the role of chance in evolution varying by several orders of magnitude across phylogenetic lineages. Every genomic nucleotide site in every species is subject to mutational change and to the vagaries of genetic drift. However, genomes are also finite in size, and these matters along with issues such as recombination rates and nonadditive gene interactions dictate the kinds of

pathways that are open to evolutionary exploration.

Most notably, the capacity of natural selection to use genetic fuel to drive the expansion of biological diversity has significant bounds dictated by intrinsic and extrinsic factors influencing effective population sizes. Not all mutations are available for adaptive evolutionary utilization, as the effective size of a population dictates the granularities of mutational effects that are visible to the eyes of natural selection. In small populations, adaptive change can only proceed by use of beneficial mutations of relatively large effects, and mildly deleterious mutations cannot be purged. In contrast, in large populations, evolution is much more fine-tuned, as its arsenal expands to mutations with very small effects. Thus, even in the ideal case in which mutation and selection pressures operate in identical ways, phylogenetic lineages are expected to vary in their response to selection, simply as a consequence of variation in the power of random genetic drift. Drift not modifies the availability of existing variation, but through history defines the genomic architectures from which mutational changes can be built.

Given various assumed selection coefficients, mutation and recombination rates, and aspects of population demography, evolutionary-genetics theory has been very good at defining how changes are expected to proceed in a generic way, usually in terms of changes in allele frequencies without any explicit reference to phenotypes. The next step in the development of a more mature science of evolutionary biology is to elucidate from first principles the connections between molecular/cell-level features and their selective consequences – the so-called genotype-phenotype-fitness mapping. The critical importance of cell biology to evolutionary biology resides in the exacting details by which the functional links between genotypes and phenotypes can be defined. This being said, although more than one physicist has claimed that all of biology is physics, this final chapter highlights why biophysics, biochemistry, and cell biology alone are unlikely to ever be sufficient to understand evolutionary processes.

Deconstructing the Great Chain of Being

It is commonly asserted that eukaryotes are superior life forms to prokaryotes, that animals represent the pinnacle of the evolutionary process within eukaryotes, and that vertebrates occupy a higher rung on the imagined ladder of ascendancy than invertebrates, with humans highest of all. One extreme view is that prokaryotes are condemned to a pathetic and perpetual fate of simplicity in form and function owing only to the absence of a mitochondrion (Lane and Martin 2010).

There is, however, no justification for this view of life. As enticing as the idea might be, there is no evidence that the ultimate goal of natural selection is to build bigger, bonier, and more complex organisms. Humans are better at being human than are bacteria, and bacteria are better at carrying out microbial functions than vertebrates, but one cannot directly compare the genetic fitnesses of reproductively isolated organisms out of context with their ecologies. Evolutionary fitness is a function of the relative transmission rate of genomic material within a cohesive population of compatible individuals. Nonetheless, even where performances can be

compared in an absolute sense, simpler organisms generally come out on top.

For example, heterotrophic bacteria, particularly large-celled species, are capable of assimilating biomass at substantially higher rates than morphologically more complex unicellular eukaryotes of comparable size (Chapter 8). The highest rate of cell division in the smallest unicellular eukaryotes is nearly an order of magnitude less than the highest rate in bacteria, and the eukaryotic rate progressively declines with increasing cell size. In multicellular eukaryotes, the maximum mass-specific growth rate (at any life stage) continues to decline with increasing size at maturity, such that the highest growth rates achievable by animals and land plants are 10 to 100× below those for unicellular species (Figure 24.1). This ability of prokaryotes to produce daughter cells at higher rates than eukaryotes is accomplished without subsidizing a vast intracellular network of specialized organelles. It is also done with fewer and smaller proteins, which on average appear to have higher catalytic efficiencies and higher accuracy of substrate utilization (Chapters 4 and 12). Thus, by all accounts, given identical metabolic tasks, prokaryotes are just as good as, if not better than, their eukaryotic counterparts. The same appears to be true for motility and environmental sensing when properly scaled against organismal size (Chapters 16 and 22).

As in the case of reduced genome replication fidelity in organisms of increasing size (Chapter 4), this gradient of declining maximum growth rates in eukaryotes may be a consequence of the accumulation of mildly deleterious growth-reducing mutations in small-N_e species (Lynch 2020; Lynch et al. 2022). Particularly noteworthy in Figure 24.1 are the approximately -0.20 and -0.10 power-law scaling relationships with size for heterotrophs and autotrophs. As noted in Chapter 8, there has been an on-going debate as to whether mass-specific metabolic rate scales with the -0.25 or -0.33 power, but the observations for maximum growth rate are inconsistent with both hypotheses. One might imagine that biophysical constraints would set the baseline pattern, with deleterious-mutation accumulation then increasingly diminishing the capacity of larger organisms (with smaller N_e) relative to the biophysics barrier. However, this would lead to a steeper slope than the biophysical expectations, i.e., slopes < -0.25, contrary to what is actually seen.

If the drift-barrier hypothesis is the correct explanation for the patterns observed in Figure 24.1, the maintenance of a constant power-law relationship across a wide range of body sizes demands a particular distribution of growth-altering effects of mutations – for each proportional increase in organism size, there must be a constant proportional increase in the total fixed growth-reducing mutational load. As organism size increases, N_e declines (Figure 4.3), remarkably also with the -0.2 power of organism mass, and this subjects windows of more deleterious mutational effects to fixation by drift. However, to keep the increment in load constant with increasing organism size, a specific relationship is required – the number of genomic sites with a particular effect must be inversely proportional to the effect size, as the product of the two is the total load.

Although direct estimation of very small mutational effects is currently intractable, observational limits are not grounds for dismissing a hypothesis. Numerous areas of science, including bioenergetics and genetics, have advanced significantly by building around theoretical constructs that required decades of work by thousands of investigators and enormous funding to achieve validation with direct

observations. Not convinced? Consider the multiple billion-dollar projects currently funding team projects in particle physics and astronomy.

Genome complexity and organismal complexity. Three decades of whole-genome sequencing have revealed clear phylogenetic patterns of genome structure and organization (Lynch 2007). Rarely exceeding 10 Mb in length, prokaryotic genomes typically contain 1000 to 8000 protein-coding genes, and these generally comprise $> 95\%$ of the entire genome, with slightly less than 1 kb of genomic space allocated per gene (Figure 24.2). Prokaryotic gene regulation is often simplified to the point that multiple genes are coordinately expressed from the same operons. Transposons and retrotransposons are nearly entirely absent from prokaryotic genes. In contrast, eukaryotic genomes are rarely smaller than 8 Mb, with unicellular species commonly having genomes in the range of 20 to 100 Mb, and metazoan and land-plant genomes often expanding beyond 1 Gb. This vast expansion of genome size is largely a consequence of the increase of intergenic DNA caused by the proliferation of multiple classes of mobile-genetic elements (essentially parasitic DNAs) and by the colonization of protein-coding genes by spliceosomal introns. The latter are intergenic DNAs that must be spliced out of transcripts to achieve productive messenger RNAs, and are completely absent from prokaryotic genomes.

Much less pronounced is the expansion of gene number in eukaryotes, which except in extreme cases of recent polyploidy is typically $< 50,000$, and quite often $< 20,000$. The smallest eukaryotic genomes are quite similar in nature to those in prokaryotes, being devoid of mobile elements and generally highly depauperate of intronic DNA. In contrast, in the enormously bloated genomes of animals and land plants, $< 2\%$ of DNA is associated with protein coding, and in the extreme cases, neighboring genes are separated by ~ 100 or more kb.

The origins of these massive changes in gene and genome organization across the Tree of Life have little to do with adaptive evolutionary differentiation (Lynch 2007). For example, mobile-genetic elements are a burden to host cells for two reasons: 1) each such element typically spans a few hundred to a few thousand base-pairs, thereby imposing an energetic cost in terms of excess DNA synthesis; and 2) random insertions of offspring elements into novel genomic sites often have negative consequences for host-cell fitness as a consequence of gene disruption. Consistent with this view, despite having no innate immunity to mobile-element invasion, the genomes of microbial eukaryotes with relatively large effective population sizes are generally nearly devoid of such elements (as in prokaryotes); such insertions are simply kept rare in these species by purifying natural selection. Likewise, introns, which can exceed 10 kb in length in multicellular species, are an energetic burden at both the DNA and RNA levels, and also impose upon their host genes an elevated rate of mutation to defective alleles owing to the need to maintain specific splice-site recognition sequences.

The fact that the smallest eukaryotic genomes contain fewer genes than bacterial genomes in cells of comparable size (Figure 24.2) makes clear that an expanded gene set is not a requirement for the development of a eukaryotic cell plan. This upholds the hypothesis that most of the changes in eukaryotic genome organization are simple by-products of the shift in the population-genetic environment permissive to the expansion of noncoding DNA. Recall from Chapter 17 that based on the energetic

costs of DNA synthesis, species with small cells and large effective population sizes are able to selectively purge insertions of nonfunctional DNA as small as a few base pairs. In contrast, in large eukaryotic species, natural selection is essentially blind to the relative energetic costs of insertions as large as several kb.

A final point of caution in interpreting the features of contemporary biology in terms of adaptive establishment is that the current function of a cellular feature may have little to do with its origin. For example, introns enable multicellular species to engage in tissue-specific alternative splicing of precursor messenger RNAs, thereby increasing proteome diversity without adding more genes. However, introns were present in large numbers in LECA (Rogozin et al. 2012; Irimia and Roy 2014), well before the emergence of multicellularity. Similar arguments can be made with respect to the increase of intergenic DNA as a substrate for the development of novel mechanisms of gene regulation (Lynch 2007). Thus, although an expansion of genome size in the ancestral eukaryote was likely promoted by mutation pressure despite the intrinsic disadvantages, this set the stage for later adaptive exploitation. These observations again provide compelling support for the strong role of historical contingencies with nonadaptive roots in guiding the future paths of cellular evolution.

A shake-up of genomic organization in the ancestral eukaryote. Although part of the expansion of eukaryotic gene number was a consequence of mitochondrial-to-nuclear genome transfer (Chapter 23), other additions came from duplications within the nuclear genome itself (Lynch 2007; Vosseberg et al. 2021). Incremental single-gene duplication is an on-going process in all genomes, with rates per gene often of the same magnitude as base-substitution mutation rates per nucleotide site, i.e., in the range of 10^{-9} to 10^{-8}/gene copy/generation. However, such processes are typically balanced in the long run by an approximately equal rate of gene loss, and there is no evidence of an inexorable climb in gene number in any domain of life. There are, nonetheless, occasional episodes of massive genome expansion with more permanent effects. For example, many whole-genome duplication events have been recorded in lineages of plants, animals, yeasts, ciliates, and many other eukaryotes, although these are unknown within prokaryotes.

Of special note is a significant period of genome expansion on the road from FECA to LECA that led to the addition of several thousand genes. Given the low probability of preservation of duplicate genes (Chapter 6), the initial burst in gene number likely exceeded 10,000, with some eukaryotic lineages then losing more duplicates than others. One of the most pronounced lines of evidence for a basal expansion of eukaryotic gene number draws from the increased complexity of multimeric proteins in this lineage. Virtually every aspect of eukaryotic cell biology reveals the use of heteromeric protein complexes (with the component parts encoded by different genes) whose orthologs in prokaryotes are frequently homomeric (with all subunits encoded by the same genetic locus). A large fraction of these changes occurred in the stem eukaryote, with the added subcomponents arising by gene duplication. Already discussed in detail in prior chapters, just a few examples will be summarized here.

First, as outlined in Chapter 10, many of the homomeric proteins involved in prokaryotic maintenance of chromosome integrity and replication have eukaryotic

orthologs (involved in mitosis and meiosis) that assemble into heteromers consisting of subunits encoded by duplicated genes. The existence of numerous duplications in archaeal orthologs, and not present in bacteria, implies that the nuclear genome of the ancestral eukaryote (likely archaeal in origin) may have been endowed with such features from the outset (Makarova and Koonin 2013). However, the dimeric form of histones in archaea transitioned to a hetero-octomeric form in eukaryotes (Henneman et al. 2018). Second, numerous proteins involved in RNA processing exhibit a similar syndrome. Consider, for example, the Sm family of proteins, which are involved in the processing of single-stranded RNAs, including forming the core of the eukaryotic spliceosome that removes introns from transcripts. Whereas bacterial Sm proteins form 6- or 7-subunit homomeric rings, they are fully heteromeric in eukaryotes, with each subunit encoded by a different genetic locus (Scofield and Lynch 2008). Third, many of the key molecular machines involved in protein surveillance and processing, including chaperones, the proteasome, and the exosome, obtained their heteromeric structures in pre-LECA eukaryotes (Chapter 14). Fourth, the guardian of the nuclear environment, the nuclear-pore complex, consists of a layered series of duplications, and this is also true of nearly every aspect of the eukaryotic vesicle transport system (Chapter 15). Finally, the α and β subunits of structural tubulin filaments emerged prior to LECA, as did the δ and ϵ subunits deployed in the eukaryotic flagellum (Chapter 16).

Although many more examples could be given, these kinds of observations suffice to reveal that there was a substantial amount of gene duplication in the stem eukaryote, with an especially significant amount of duplicate preservation associated with the structural features of protein complexes. Given the frequency with which whole-genome duplications occur in modern-day lineages of eukaryotes, the possibility that one or more of such events precipitated this expansion in gene number in the ancestral eukaryote cannot be ruled out (Makarova et al. 2005; Zhou et al. 2010). Unfortunately, a formal evaluation of the matter by phylogenetic analysis is made difficult by secondary chromosome rearrangements and removals of large numbers of duplicates over the vast reach of time since the origin of eukaryotes.

Whatever the mechanism – one or more whole-genome duplications or cumulative duplications of smaller chromosomal regions, Marakova et al. (2005) suggest that the basal increase in eukaryotic gene number may have been precipitated by a cataclysmic event inducing a sudden and prolonged reduction in population size. Part of the rationale for this argument is the fact that the preservation of duplicate genes by subfunctionalization (as opposed to neofunctionalization) is facilitated in populations of small size. Recall from Chapter 6 that subfunctionalization is a process by which duplicate genes are preserved by the complementary loss of key subfunctions. Unless there is a resolution of an adaptive conflict, no fitness gain results from such a process. Rather, there are weak bioenergetic and mutational costs of relying on duplicated genes relative to single-copy genes with the same functions. Such costs will be invisible to the eyes of purifying selection in populations of sufficiently small size.

Was this expansion of the complexity of key molecular machines of any relevance to the establishment of the altered eukaryotic cell plan or to the subsequent diversification of eukaryotes into morphologically diverse descendent lineages? Evidence that increases in multimeric complexity endow organisms with superior molecular perfor-

mance or with a dramatic shift in function or diversity of functions has been elusive (Chapter 13). On the other hand, as outlined in Chapter 6, the differential loss of duplicate genes in parallel lineages can passively lead to genetic map changes that manifest as post-reproductive isolating mechanisms between sister taxa, thereby establishing novel and independent lineages. This again illustrates the point that cellular modifications arising entirely by nonadaptive mechanisms may have had a foundational role in channeling downstream evolutionary pathways, not just by the creation of novel molecular functions and structures, but by expanding the potential for novel lineage proliferation.

Multicellularity

Under the common belief that organismal complexity represents adaptive progress, the evolution of multicellularity is often viewed as the ultimate goal of natural selection. In this view, single-celled lineages (the vast majority of the Tree of Life) are condemned to a perpetual state of simplicity by the lack of one or more critical ingredients, such as the mitochondrion, rather than by a lack of selective incentive. For this reason, the evolution of multicellularity is often touted as a "major transition" in the history of life (Maynard Smith and Szathmáry 1995). The icons of multicellularity are metazoans and land plants, made not just of multiple cells but multiple cell types. It should be remembered, however, that complex multicellularity with multiple cell types has arisen many times across the eukaryotic Tree of Life (Bonner 2001; Grosberg and Strathmann 2007; Ruiz-Trillo and Nedelcu 2015; Sebé-Pedrós et al. 2017), including on multiple occasions in fungi, red and green algae, slime molds, ciliates, and as noted below, even in bacteria. Moreover, this broad phylogenetic distribution still understates the ease with which multicellularity can evolve.

Under appropriate laboratory settings, it is not difficult to coax multicellularity out of cultures of unicellular organisms on short time scales (Boraas et al. 1998; Hammerschmidt et al. 2014; Fisher et al. 2016). This is most dramatically illustrated by an experiment that self-selected for aggregations of *S. cerevisiae* cells, so-called snowflake yeast colonies, that rapidly settle in test-tube environments (Ratcliff and Travisano 2014; Ratcliff et al. 2015). In parallel experiments, the initial transition to snowflake form results from a mutation in a single gene preventing mother-daughter cell separation. As discussed below, a frequently assumed challenge for the stable establishment of multicellularity is the emergence of cheater genotypes that harvest the benefits of group living without paying the price incurred by other group members (e.g., the release of group-beneficial products). However, in this particular system, the potential for genetic conflict is thwarted, as selection favors large colonies produced by fragmentation, eliminating the opportunities for transmission of cheaters that fail to engage in aggregation.

In situations like this, as long as the ecological conditions selecting for multicellularity remain, secondary genomic changes that refine such a state will be further promoted, provided the advantages of group living outweigh the interests of the selfish individual. However, it helps if the group members are close relatives.

With clonally reproducing snowflake yeast, a crude form of cellular differentiation emerges, including apoptosis (suicide) of central cells, which although disadvantageous to the individual cell is essential to offspring colony production. In effect, the colony (rather than the individual cell) becomes the unit of selection. That is, the colony becomes the individual.

Constructing a more contrived system, Wahl and Murray (2016) developed another form of multicellular yeast with somatic differentiation. In this case, a construct was engineered to enable the recurrent production of two cell types, one from the other. Upon chemical induction, a faster-growing "germ cell" gave rise to a slower growing "somatic" cell type that secreted invertase into the medium. This enzyme then hydrolyzed nonutilizable sucrose into fructose and glucose molecules essential for the growth of the germ cells. Multicellularity was maintained in this system because cheating cells (which don't invest in invertase secretion) rapidly segregate into cheating-only colonies, with reduced reproductive rates.

These two examples clearly show that the key issue with respect to the evolution of complex multicellularity is not the presence of an intrinsic barrier to initial emergence. Rather, multicellularity is typically held back by the limited ecological opportunities that encourage such body plans, by the investment costs imposed by larger somas, and by the presence of internal threats that can thwart persistence after initial establishment. Before addressing these matters in more detail, a brief consideration of bacteria will make clear that multicellularity is by no means restricted to the eukaryotic domain.

Multicellularity and cooperativity in bacteria. Although they have not gone to the extremes of metazoans and land plants, prokaryotes are not immune to evolving physical or behavioral consortia of mutually beneficial cells with specialized functions (Claessen et al. 2014). In effect, the individual cells of such collectives are similar to those in asexually propagating, multicellular eukaryotes, in that cell fitness is a function of the emergent properties of the group. Cell-cell communication is typically involved, and in some cases, subsets of cells terminally differentiate to the point of relinquishing the capacity to reproduce. This is reminiscent of the altruistic behaviors observed in social animals, where the sacrifice of individual fitness is advantageous provided sufficient benefits are accrued by close relatives, e.g., warning calls in prairie dogs (Hamilton 1964a,b; Diggle et al. 2007). The one major difference between multicellularity in eukaryotes and bacteria is that in the former case cells are usually aggregated from the outset (slime molds being exceptions), whereas in bacteria individual cells come together to form a unit (cyanobacteria and actinomycetes being exceptions). Just a few examples will be given here.

Numerous species of filamentous cyanobacteria are capable of producing specialized cells with anaerobic interiors for nitrogen fixation (heterocysts) as well as specialized spore cells (akinetes). The soil bacterium *Bacillus subtilis* undergoes periodic switching between motile unicellular and sessile filamentous states, with one transition being essentially random and the other involving a timer (Chapter 22). Multicellular magnetotactic bacteria have no known unicellular stage, and are linked together via structural filaments contained within a central acellular chamber (Keim et al. 2004; Shapiro et al. 2011). Mycobacteria harbor a family of excretion-system loci that coordinate intercellular communication and a form of sexual reproduction

known as conjugation (Gray et al. 2016). A cave-dwelling species, *Jeongeupia sacculi,* initially grows into well-organized sheets that then differentiate into mounds exceeding 1 mm in diameter, finally developing into closed volvano-like structures that explosively extrude single-celled propagules (Mizuno et al. 2022)

Particularly striking are the widespread quorum-sensing systems that bacteria use to induce intraspecific neighbors to cooperate in the production of biofilms, planktonic aggregates, or swarming motility (Ng and Bassler 2009). Such systems generally involve simple positive-feedback loops, wherein above a certain cell density, cells begin to release a chemical pheromone (a low molecular-weight chemical called an autoinducer). Pheromones are costly, and of little use at low population densities, but when the external autoinducer concentration exceeds a threshold level, cognate receptors bind them, triggering a signal-transduction cascade that coordinately turns on a group of response genes, including the gene that produces the pheromone.

The signaling molecules used in quorum sensing range widely among species. There is clear coevolution between these and their receptors (Wellington Miranda et al. 2021), although the extent to which this is driven by selection to avoid cross-species messaging or a simple consequence of drift in bivariate communication systems (Chapters 21 and 22) remains unclear. In some systems, the pheromone molecules are small enough to simply diffuse through the cell membrane and are picked up by cytoplasmic factors (e.g., lactones in *Pseudomonas;* Ng and Bassler 2009). In other cases, larger message molecules (e.g., modified peptides in Gram-positive bacteria) are bound by transporters located on the cell surface, which then transfer them to kinases used to phosphorylate response regulator proteins (Chapter 22). Some species have multiple quorum-sensing systems, providing a basis for complex environmental sensing mechanisms and combinatorial communication (Cornforth et al. 2014; Feng et al. 2015; Jemielita et al. 2018), behavioral attributes that are typically assumed to be restricted to metazoans.

Bacteria even have memories, in the sense that prior exposure to an environmental challenge can enhance the response at a later point in time, even extending into the next generation (Chapter 22). This occurs when the lifespans of key proteins exceed the time needed for cell division (Mitchell et al. 2009; Lambert and Kussell 2014; Mathis and Ackermann 2016). The benefits of collective behavior are not always clear, but include presumed advantages associated with biofilm formation, such as the aggregative behavior of food-degrading and/or tissue-invading enzymes and antibiotic production.

Given that signaling proteins are bioenergetically expensive, such aggregative systems are vulnerable to invasion by mutant cheaters that relinquish the costs of signaling while still profiting from the public goods produced by the remaining group members, a classical weak-link in social systems. A powerful defense against such invasions is made possible by kin recognition, wherein individuals can discriminate close from distant relatives, thereby providing opportunities for selectively dispensing benefits to individuals likely harboring similar genetic constitutions, including the underlying genes for helping (Diggle et al. 2007).

Such a recognition system is known to operate in the social soil bacterium *Myxococcus xanthus,* a striking example of prokaryotic multicellularity. In this species, individual cells aggregate into swarming masses that assemble into fruiting bodies, within which only a fraction of cells produce spores. Such aggregations are genet-

ically highly homogeneous, with kin detection and restricted cell adhesion being governed by a simple two-locus system sensitive to even single amino-acid changes (Cao and Wall 2017). Dozens to hundreds of mutually exclusive recognition groups, each operating in effect as a unit of selection, can coexist in single soil samples (Vos and Velicer 2009). Strikingly, kin discrimination in these myxobacteria involves a two-tiered system. Following the first coarse-grained recognition step, specific toxins are injected that kill recipients that do not carry the appropriate immunity genes, thereby virtually assuring that members of individual aggregates are clonemates (Vassallo and Wall 2019).

The deployment of kin-recognition groups extend beyond *Myxococcus.* For example, when grown on agar substrates, the soil bacterium *B. subtilis* forms distinct boundaries between incompatibility groups (Stefanic et al. 2015). Eukaryotic slime molds (*Dictyostelium*) are social amoebae that form stalks and fruiting bodies after solitary cells aggregate into kin-recognition groups with a simple genetic basis (Benabentos et al. 2009; Hirose et al. 2011). With collective benefits being dispensed only to close kin, these kinds of systems are organized in ways that are not much different than multicellular eukaryotes where all cells within individuals are clonemates. However, there is a bit of double-edged sword here – too little discrimination at the time of aggregation will lead to the loss of kin-group fitness, but too much discrimination will lead to aggregates that are physically too small to reap overall group-level benetifts (Márquez-Zacarías et al. 2021).

How does a single bacterial species come to consist of multiple kin-recognition groups, as opposed to operating as a single species-wide unit? One possibility, demonstrated by experiments with *B. subtilis* (Pollak et al. 2016), is that the coexistence of incompatibility groups is facilitated by a form of frequency-dependent selection. In a mixed population, rare strains with one particular sensing mechanism are unable to activate their own quorum-sensing system because the concentration of their messenger molecule is too low, but such strains can still gain advantages elicited by common strains with activated quorum-sensing systems. However, as initially rare strains increase in abundance, and their own systems become active, they then become subject to exploitation by other rare strains. In principle, the entire system facilitates coexistence, with the fitnesses of strains increasing as their relative frequencies decrease, thereby bounding them away from extinction.

Finally, it is worth noting that just as the individual cells of multicellular eukaryotes exhibit division of labor, there are many examples of bacterial consortia in which different members provide complementary resources for metabolic cross-feeding. For example, Rosenthal et al. (2018) found that genetically uniform laboratory populations of *Bacillus subtilis* can subdivide into two subpopulations of metabolically differentiated cells, one producing harmful acetate, and another converting the latter into a benign storage molecule. Likewise, simple test-tube populations of *E. coli* often develop spatial structure and cross-feeding in ways that promote coexistence of different evolved clades (Good et al. 2017; Behringer et al. 2018). In an engineered example, Pande et al. (2014) constructed two strains of *E. coli* with complementary amino-acid requirements – each lacked the ability to synthesize one amino acid but released the other into the environment. Over a short time period, the strains evolved a division of metabolic labor that enhanced the productivity of the overall system, as the cost of producing excess an amino acid for the partner was less than

the advantage gained from being provisioned by the other.

All of these observations on bacteria highlight the importance of a key issue in evolutionary biology – the unit of selection. In principle, selection can operate at any level, discriminating among individuals within populations, among groups of individuals based on their emergent properties, or even among species on very long time scales. Moreover, a trait that is beneficial at one level need not be beneficial at another. Altruistic behavior is a prime case in point, as in this case the individual incurs a fitness cost while other members of the population experience a fitness gain. Generally, when conflicts like this arise, we expect selection at the individual level to prevail for the simple reason that the turnover rate of individuals exceeds that of groups (Williams 1966), but the distinction between individual- and group-selection becomes blurred when the interacting members of the population are close relatives.

What is an individual from an evolutionary perspective? In almost all organisms, individuality is clearly demarcated in a physical sense, e.g., by cell membranes in the case of unicellular species, and by distinct soma in the case of multicellulars. However, things are not so clear at the genetic level, as some individuals are more closely related than others, the extreme cases being monozygotic twins or members of an asexually propagating lineage. This matters because behavior that is harmful to an individual but sufficiently beneficial to close relatives can be promoted by kin selection (Wilson 1975). For example, a suicidal behavior that sufficiently elevates the fitness of multiple clone-mates (who also carry the genes for such behavior) will be advanced by selection. Given the clonal nature of microbial aggregates, it is then clear that the unit of selection at the genetic level extends beyond individual cells to the local kin group, not greatly different from the situation in multicellular eukaryotes in which the soma is derived by clonal expansion. Thus, many microbes are multicellular in an evolutionary sense even in the absence of structural connections between cells. The key point is that eukaryogenesis was not a pre-requisite for admission into the multicellular world.

The costs of multicellularity. Certain forms of multicellularity can open up access to novel resource pools not physically possible for single cells, e.g., the ability of animals to overwhelm smaller prey items and land plants to experience the full spectrum of ambient sunlight. Multicellularity can also provide added benefits via improvements in extracellular metabolism, protection against physical environmental challenges, stress resistance, and avoiding predation, although increased size may invite the attention of still larger consumers (Tong et al. 2022). However, to be promoted by natural selection, transitions to novel ecological niches require that any benefits accrued outweigh the costs, at least in the early stages of establishment. For example, a mutation that expands dietary breadth needs to do so in a way that magnifies the net rate of acquisition of resources beyond what was possible in the ancestral state, and needs to do so for a long enough time period to move to fixation (with the number of required generations being on the order of the effective population size in the simplest of cases; Chapter 4).

With their structural support systems, animals and land plants were predisposed to proliferate across uncolonized land masses, provided the advantages outweighed the costs of cellular cooperation. However, once enough multicellular lineages had evolved diverse and refined mechanisms for occupying most niches afforded by large

size, there would be no continuing evolutionary incentive for the mass movement of microbes towards multicellularity, any more than established multicellular species would be expected to undergo regressive evolution to unicellularity.

Taken in the broader picture of evolution across the Tree of Life, the often-stated beauty and wondrous nature of land plants and animals may have artistic appeal, but belies the underlying costs of multicellularity. The prices to be paid are not simple tradeoffs involving allocations to different functions, but are more fundamental consequences of the shift in the population-genetic environment. Once propelled into a new ecological niche, the reduction in N_e incurred by larger, multicellular organisms would have had downstream effects, including subjection of the genome and associated cellular features to drift and mutational pressures towards undesirable states.

First, as discussed above, whereas large complex organisms may have access to novel resources, they also ultimately succumb to a reduction in bioenergetic capacity (Figure 24.1). Specifically, maximum growth rates decline with organism size, such that even with an unlimited food supply, large eukaryotes are unable to convert resources into biomass at the rates of smaller organisms. Thus, the long-term benefits of multicellularity do not reside in a greater capacity for assimilating biomass essential for growth and reproduction, at least not at high resource levels. Are there energetic advantages at low resource levels? Given the increased maintenance costs of larger organisms per unit time and their longer generation times (Chapter 8), this too seems unlikely.

Second, the idea that selection relentlessly promotes increased complexity to confer robustness in the face of environmental and mutational changes also comes up short, as it ignores both the physical costs of complexity and the constraints on evolutionary processes. As noted above, more complex gene structures magnify both the bioenergetic costs and the mutational vulnerability of alleles. There is no evidence that genes endowed with introns, proteins of greater length, or expansions of mobile-element families were promoted by advantages conferred upon their host genomes. Quite the contrary, the initial establishment of all of these genomic features appears to have been made possible only by the reduction in the efficiency of negative selection.

Third, as outlined in Chapter 20, while added layers of intracellular surveillance, such as DNA polymerase proof-readers, constitute an increase in complexity by any definition of the term and superficially lead to the impression of an elevation in robustness, this too is an illusion. Although the initial establishment of a new layer of protection may be promoted by positive selection, in the long run, the overall advantage is expected to dissipate as a consequence of the further accumulation of mildly deleterious mutations to the multiple-component system. This then leads to a more complex, seemingly robust system with no better capacity than the originally simpler version, but now with added costs of energetic investment and mutational vulnerability. A rough analogy from human behavior is Parkinson's (1957) law, which states that the work required to complete a task expands as the time necessary for completion increases.

Fourth, multicellularity imposes the necessity of costly mechanisms for the suppression of renegade cells. As noted above for species with aggregative groups, the constant threat from emerging cheater cells is arguably one of the greatest challenges

to the maintenance of multicellularity (Frank 1995; Michod 1999). This is also true for organisms with complex development, particularly for metazoans, where many cells are not permanently bonded together. A recurrent problem for species with somatic tissues is the emergence of cells that proliferate at the expense of the overall organism, cancer being one of the most obvious manifestations of this kind of problem. Indeed, because somatic mutations are not inherited across generations, selection for genomic-repair mechanisms appears to be substantially reduced relative to the situation in the germline, with mutation rates in somatic cells typically being elevated at least ten-fold (Lynch 2010; Blokzijl et al. 2016; Abascal et al. 2021; Ueda et al. 2022). As an additional side consequence, as intrasomatic selection progressively selects for clonal variants that expand at the expense of their neighboring cells, somatic mutations inevitably lead to senescence (Nelson and Masel 2017; Rozhok and DeGregori 2019).

The claim here is not that complex multicellularity emerged in the face of short-term disadvantages, but rather that once set in motion, the transition to such body plans created unavoidable downstream consequences. During the evolutionary diversification of eukaryotes, short-term opportunities guided particular lineages down trajectories leading to larger body plans with structural support only possible with multicellularity. However, once established, such body plans not only resulted in lineages with reduced bioenergetic capacity, but also altered the population-genetic environment in ways that encourage the establishment of genomic and subcellular features with added expenses. In particular, multicellularity set up a scenario by which effectively neutral (but absolutely deleterious) changes could accumulate through time by mutation pressure, perhaps in some cases even being driven by positive selection for mechanisms that enhance competitive ability while reducing individual productivity (e.g., physical dominance of larger over smaller individuals). Combined with the partitioning of the full repertoire of functions essential in unicellular organisms to specialized cell types (discussed in the following section), complex multicellularity then becomes an evolutionary trap in the sense that reversion to unicellularity becomes nearly impossible.

This being said, additional considerations remain to be explored from the standpoint of population-genetic mechanisms. In particular, in the initial stages of a transition, mutations to multicellularity would need to proliferate in a background population of unicellularity. Assuming a life cycle with a sexual phase, this leaves many questions unanswered with respect to genetic compatibility of unicellular and multicellular variants. These are key issues given that, owing to initial rarity in the critical phase of establishment, all mutations in sexual populations face the challenge of the fitness consequences of mixing with foreign types, effectively being tested only on heterotypic backgrounds. Thus, long-term phases of clonal propagation, haploidy, and/or local inbreeding would seem to facilitate the emergence of multicellularity. If so, how is inbreeding depression and mutational meltdown avoided in the initially small isolates of multicellulars? Are multicellular variants immediately isolated from their unicellular ancestors? Can multicellularity drive itself through the unicellular subpopulation by contagious spread through outcrossing?

The emergence of cell-type specialization. As noted above, the average land-plant and animal genome harbors two- to three-fold more genes than in their uni-

cellular relatives. However, little of this increase seems to be related to necessities associated with multicellularity, as opposed to being indirect by-products of alterations in the population-genetic environment. Consider, for example, that in the early days of genome sequencing, the observation that numerous genes in mammalian genomes were absent from *Drosophila* and *Caenorhabditis* genomes led to the idea that such genes play key roles in the development of vertebrates. As a broader phylogenetic survey of genome contents began to emerge, it became clear that these apparent "vertebrate-specific" genes were actually cases of gene loss in invertebrate lineages.

Similarly, many of the genes originally thought to have been unique to metazoans have since been found in basal unicellular lineages (e.g., choanoflagellates and ichthyosporeans), and in some cases even more deeply (Ocaña-Pallarès et al. 2022). For example, the integrin proteins used for cell-cell adhesion in metazoans are present in the apusozoan lineage (basal to both animals and fungi) but absent from fungi and choanoflagellates (basal lineages outside of metazoans) (Sebé-Pedrós et al. 2010). Likewise, many proteins initially thought to be uniquely involved in cell signaling, immune response, and development in metazoans were subsequently found to be present in choanoflagellates (King et al. 2008; Richter et al. 2018). While there are many gene gains on the branch subtending the metazoan lineage, the number of gene losses appears to be just as great, and included among them are the genes for the biosynthesis of nine amino acids (Richter et al. 2018). Thus, the evolution of complex multicellularity did not involve a major influx of new genes. Rather, many of the genes deployed in the unique features of complex multicellular organisms are modified descendants of those used in related functions in unicellular ancestors. Notably, the genesis of a number of novel cell functions in multicellular species seem to involve the co-option and repurposing of preexisting stress-response pathways (Love and Wagner 2022).

Under the assumption that division of labor leads to the whole being more than the sum of its parts, it has been suggested that cell-type specialization is critical to the evolution of complex multicellularity (Maynard Smith and Szathmáry 1995; Michod 1999). However, although metazoans and land plants exhibit dozens to hundreds of cell types, the increase in complexity of cellular functions is relatively small. Moreover, although some metazoan cell types have evolved new sets of tasks (e.g., nerve and bone cells), most cell types in multicellular species have simply lost a range of features found in ancestral unicellular species. A plausible explanation for such outcomes is that whereas single-celled organisms must be capable of multitasking with respect to nutrient acquisition, avoidance of predators, dealing with unfavorable environments, etc., multicellularity offers the possibility of subfunctionalization at the level of cell-type specialization. Indeed, drawing from such patterns of partitioning, some have suggested that cell types can be classified based on shared and divergent patterns of gene expression (Arendt 2008; Arendt et al. 2016; Kishi and Parker 2021). The key point is that, as with many aspects of genome and cell biology, the emergence of multicellularity relies much more heavily on the subdivision of ancestral functions than on neofunctionalization.

A common way in which cell specialization emerges from single-celled life appears to be the conversion of a prior temporal pattern of cell life-cycle differentiation into a developmentally regulated spatial pattern (Mikhailov et al. 2009; Sebé-Pedrós

et al. 2017). Examples of preexisting phenotypic variants include bimodal phases in bacteria (Chapter 22) and alterations between asexual and sexual generations in unicellular eukaryotes. The expected scenario here is one in which cell differentiation based on external environmental cues is progressively eliminated and replaced by internal signaling mechanisms based on cell-cell communication and developmental regulation. This potential for partitioning preexisting gene functions combined with regulatory rewiring clarifies why the evolution of multicellularity does not require a massive investment in new genomic real estate.

Just as the loss of complementary functions by subfunctionalization leads to the preservation of duplicate genes, division of labor among cell types plays a key role in ensuring the stability of complex body plans. That is, the establishment of cell-type specific features that enhance overall organismal performance but come at an expense to free-living cells (e.g., functional partitioning) reduces the likelihood of reversion to unicellularity even if this were to be beneficial. In this sense, the division of labor in multicellular organisms, whether intrinsically beneficial or not, has a ratchet-like effect on the stability of multicellularity – as cell types become more and more specialized, the likelihood of reversion of any single cell type to a form containing all ancestral-cell features necessary for independent living becomes less and less likely (Libby and Ratcliff 2014; Cooper and West 2018).

Separation of germline and somatic cells represents an extreme form of division of labor, as the former become increasingly specialized for the single function of propagation, releasing the latter from any special requirements associated with meiosis and sexual reproduction of offspring. However, to see that such an extreme form of partitioning does not depend on the prior establishment of multicellularity, one need only look to the ciliated protozoa (Chapter 10). These highly diverse and globally distributed protists are binucleate, with the diploid micronucleus serving as a transcriptionally silent germline, which recombines during cell fusion and sexual reproduction. The macronucleus, a highly polyploid and edited version of the micronucleus, serves as the locale of somatic gene expression, and is transmitted without recombination during asexual cell division, but after numerous cell divisions is disposed of and replaced following sexual reproduction.

Studies on the Volvocales, an order of green algae that includes the unicellular *Chlamydomonas,* provide further insight into many of the above points. The order includes species with 2-, 4-, 8-celled body plans, etc., extending up to large spherical, multicellular forms known as *Volvox,* containing thousands of cells. Notably, the increase in colony size is accompanied by a substantial expansion in noncoding DNA in both the mitochondrial and plastid genomes, consistent with a shift in the population-genetic environment given that these organelles are uninvolved in cell-type differentiation (Smith et al. 2013). Although it has been argued that collective aggregates are more efficient at responding to weak chemotactic signals (Colizzi et al. 2020), and that flagellar stirring of boundary layers can magnify advective transport beyond the limits of diffusion (Solari et al. 2006), this leaves unexplained the predominance of the simplest body plans, and the advantages of multicellularity vs. unicellularity in the group remain unclear.

According to the phylogeny, the first steps in the transition to multicellularity in the Volvocales involve the evolution of colonies of undifferentiated cells, with division of labor arising secondarily (Kirk 2005; Herron et al. 2009; Herron 2016; Featherston

et al. 2018; Matt and Umen 2018). Notably, the genus *Volvox* is polyphyletic, as the volvocine form has actually evolved on several independent occasions, and these changes are accompanied by significant reductions in gene number (Lindsey et al. 2021; Jiménez-Marín and Olson 2022). The most extreme form is *Volvox carteri,* which harbors 16 large nonmobile germ cells embedded in an extracellular matrix surrounded by a single-layered sphere of ~ 2000 terminally differentiated flagellated cells. The germ cells exhibit the greatest breadth of gene expression, relative to the more specialized transcriptomes of the somatic cells, a pattern similar to that seen in pluripotent stem cells in metazoans (Matt and Umen 2018).

Finally, we return to the common assertion that the establishment of the mitochondrion was central to the emergence of multicellularity (e.g., Bendich 2010; Lane and Martin 2010; Medini et al. 2020). The reasoning behind this speculation is diverse, ranging from the supposed bioenergetic advantages of mitochondria (shown in preceding chapters to be incorrect), to the sequestration of germ cells from mutagenic by-products of organelle respiration, to the involvement of mitochondria in various aspects of developmental control. All of these arguments ignore the existence of multicellular prokaryotes and fail to make the distinction between mechanisms of origin of organismal features and their secondary, downstream modifications. As just one example, cellular apoptosis is often viewed as a unique innovation that emerged after the evolution of multicellularity. The process is triggered by events associated with the mitochondrion, and can be essential to development and cancer suppression. However, the apoptotic machinery appears to predate not only the evolution of complex multicellularity but even the origin of eukaryotes, as it is present in multiple lineages of unicellular eukaryotes as well as in bacteria (Koonin and Aravind 2002; Klim et al. 2018).

Closing Comments

We get excited by things that we see day to day. As a consequence, the majority of current research in evolutionary biology has focused on things like butterflies, vertebrates, and flowering plants. The bulk of the work done in the remaining areas in biology is concentrated on a few model organisms, such as yeast and *E. coli,* often pursued under the guise of better understanding of human biology and promoting biomedical applications.

From work on metazoans and land plants, we have learned a lot about agents of natural and sexual selection, at least for multicellular species. What remains to be understood are the molecular/cellular mechanisms underlying phenotypic divergence and the degree to which these vary among phylogenetic lineages. Although the issues are common to all organisms, unicellular organisms provide a logical starting point for such investigation. The root of the Tree of Life as well as most of the branches is unicellular, and each cell in all of today's organisms is a product of a continuous cell lineage tracing back to the beginning of biological time.

We also gravitate towards simple explanations for biodiversification. However, although it is relatively easy to understand the superficial features of the process of natural selection and to concoct adaptive stories as to how biodiversity arose, evo-

lution is not a simple matter of natural selection pushing around mean phenotypes. Rather, the paths open to exploitation by selection are governed by the nonadaptive processes of mutation, recombination, and random genetic drift, all of which have been universal genetic forces since the origin of life. Although we do not know their relative strengths at the earliest stages in evolution, they vary by orders of magnitude among today's phylogenetic lineages, often in ways that scale with organism size, and there is no reason to think that this has not been the case for the past three billion or so years. Such variation has significant consequences for the emergence and utilization of the genetic material upon which natural selection acts.

Thus, whereas the technical field of population genetics is often conveniently viewed as being marginal with respect to questions concerning deep phylogenetic divergence, it is actually front and center. Likewise, while the details of molecular and cellular biology are often viewed as largely irrelevant to the ways in which populations respond to natural selection, it is precisely here that structural and functional modifications must be made to yield new phenotypes, so these details also matter.

Formally demonstrating that selection operates in an effectively deterministic manner, with the underlying details being immaterial, would be a great achievement for evolutionary biology. Indeed, such a demonstration would effectively confine the future of the field to the monotonous cataloging of selective forces operating on different organisms. However, the preceding pages challenge this view. Evolutionary processes are inherently stochastic, with the potential paths open to evolutionary exploitation depending on historical contingencies and the granularity of mutational effects relative to the power of random genetic drift. This makes the development of evolutionary theory more challenging than desirable for those confined to a Darwinian mode of thinking, but a desire for simplicity is no substitute for the scientific goal of describing reality.

Although no serious scientist any longer argues against the central role played by natural selection in evolutionary change, it is now clear that organisms exposed to identical selection pressures will respond in qualitatively different ways depending on their population and cellular environments, in some cases being completely impervious to selection and largely driven by mutation pressure. Once the subject of a long debate confined to molecular evolution, effectively neutral processes now appear to have played a key role in the evolution of genome architecture, cellular features, and by extension, whole-organism biology. If this view is correct, it offers a unifying and mechanistic approach to thinking about the history and dynamics of evolutionary change across the Tree of Life and across levels of organization, liberating us from the century-old tradition of assuming that all of biodiversity reflects an optimization process dictated by a supreme designer called natural selection.

Summary

- Were life given the chance to start anew, cells bounded by external envelopes and harboring polymeric genomes would likely evolve. However, the underlying elemental features of life might be entirely different from the particular hand

dealt by the early Earth. With different baseline genomic features and ecological settings, population-genetic environments might differ radically as well. It is interesting to contemplate how such altered beginnings might modify downstream paths of biodiversification. However, a big enough problem for now is explaining the emergence and maintenance of organismal simplicity vs. complexity across this particular planet's biosphere.

- Contrary to popular belief, evolution is not on an inexorable path to build more complex organisms. There no evidence for an intrinsic advantage of complexity at the molecular, cellular, or organismal levels, and empirical observations show that simple bacteria often carry out the most basic functions of life more efficiently than do the more complex cells of eukaryotes.

- The emergence of eukaryotes from prokaryotic ancestors was accompanied by a substantial increase in genome size, particularly in multicellular lineages. However, the vast majority of this expansion is a consequence of the colonization by noncoding and nonfunctional DNA. Rather than being driven by positive selection, such changes appear to have arisen as passive by-products of a reduction in effective population size, which diminishes the ability of natural selection to oppose insertions of excess DNA.

- There is no evidence that an increase in gene number played a primary role in the establishment of the eukaryotic cell plan. Instead, it appears that the reorganization of the ancestral eukaryotic genome, possibly spawned in part by one or two whole-genome duplication events followed by gene loss and subfunctionalization, led to the emergence of a new form of genomic architecture, which in turn opened up new paths for evolutionary change. One of the most striking sets of such changes involved the transition of homomeric molecular complexes in prokaryotes to heteromeric forms in eukaryotes.

- Complex multicellularity, involving large numbers of cell types as embodied in land plants and metazoans, is restricted to eukaryotes. With just two instances of such evolution across the Tree of Life, there is no statistical basis for arguing that such evolution was dependent on the prior emergence of the eukaryotic cell plan.

- Bacterial species exhibit a wide range of features that enable collectives of individuals to operate in ways that are more than the sum of their parts, including division of labor among different cell types. Such traits include morphological differentiation at the cellular level, quorum-sensing mechanisms that promote coordinated behavioral changes in response to changes in cell density, and cross-feeding of complementary resources. Many of these mutual benefits are specifically dispensed to close relatives (kin-recognition groups), an essential behavior

for thwarting the emergence of cheater cells.

- Ecological opportunity played a central role in the emergence of land plants and animals, but such establishment induced downstream costs, some of which were inevitable consequences of a reduction in long-term effective population sizes. These include a reduced ability to assimilate biomass, a substantial increase in the energetic cost and mutational vulnerability of genomes, gratuitous investment in overly complex cellular features, and the constant threat of mutant somatic cheater cells.

- Arguments about the causes and consequences of multicellularity may profit from a broader consideration of the unicellular branches on the Tree of Life. For example, although multicellular organisms often have sequestered germlines, multicellularity is not a prerequisite for this condition, as all ciliates have separate germline and somatic nuclei. Likewise, although mitochondria are involved in apoptotic cell death in multicellular species, apoptosis is not a unique feature of multicellular organisms, as its antecedents can be found in unicellular lineages.

- The emergence of multicellularity relies much less on the evolution of novel gene functions than on the partitioning and/or loss of ancestral cell functions leading to the division of labor among cell types. As multicellular organisms become more complex, this leads to a situation in which reversion to unicellularity becomes effectively impossible. As a consequence, multicellular lineages become terminal branches in the Tree of Life, whereas unicellular lineages retain the capacity to become multicellular should the ecological opportunity emerge.

- Avoid left by the complete extinction of all animals and land plants would likely be entered quickly by multicellular descendants of unicellular species. Although the specific phenotypes to emerge under such a scenario would likely deviate from the situation today, the same syndrome of genomic changes would be expected to unfold.

Literature Cited

Abascal, F., L. M. R. Harvey, E. Mitchell, A. R. J. Lawson, S. V. Lensing, P. Ellis, A. J. C. Russell, R. E. Alcantara, A. Baez-Ortega, Y. Wang, et al. 2021. Somatic mutation landscapes at single-molecule resolution. Nature 593: 405-410.

Arendt, D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. Nat. Rev. Genet. 9: 868-882.

Arendt, D., J. M. Musser,C. V. H. Baker,A. Bergman,C. Cepko,D. H. Erwi,M. Pavlicev,G. Schlosser,S. Widder,M. D. Laubichler,and G. P. Wagner. 2016. The origin and evolution of cell types. Nat. Rev. Genet. 17: 744-757.

Behringer, M. G., B. I. Choi, S. F. Miller, T. G. Doak, J. A. Karty, W. Guo, and M. Lynch. 2018. *Escherichia coli* cultures maintain stable subpopulation structure during long-term evolution. Proc. Natl. Acad. Sci. USA 115: E4642-E4650.

Benabentos, R., S. Hirose, R. Sucgang, T. Curk, M. Katoh, E. A. Ostrowski, J. E. Strassmann, D. C. Queller, B. Zupan, G. Shaulsky, et al. 2009. Polymorphic members of the lag gene family mediate kin discrimination in *Dictyostelium*. Curr. Biol. 19: 567-572.

Bendich, A. J. 2010. Mitochondrial DNA, chloroplast DNA and the origins of development in eukaryotic organisms. Biol. Direct 5: 42.

Blokzijl, F., J. de Ligt, M. Jager, V. Sasselli, S. Roerink, N. Sasaki, M. Huch, S. Boymans, E. Kuijk, P. Prins, et al. 2016. Tissue-specific mutation accumulation in human adult stem cells during life. Nature 538: 260-264.

Bonner, J. T. 2001. First Signals: The Evolution of Multicellular Development. Princeton Univ. Press, Princeton, NJ.

Boraas, M. E., D. B. Seale, and J. E. Boxhorn. 1998. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. Evol. Ecol. 12: 153-164.

Cao, P., and D. Wall. 2017. Self-identity reprogrammed by a single residue switch in a cell surface receptor of a social bacterium. Proc. Natl. Acad. Sci. USA 114: 3732-3737.

Claessen, D, D. E. Rozen, O. P. Kuipers, L. Søgaard-Andersen, and G. P. van Wezel. 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. Nat. Rev. Microbiol. 12: 115-124.

Colizzi, E. S., R. M. Vroomans, and R. M. Merks. 2020. Evolution of multicellularity by collective integration of spatial information. eLife 9: e56349.

Cooper, G. A., and S. A. West. 2018. Division of labour and the evolution of extreme specialization. Nat. Ecol. Evol. 2: 1161-1167.

Cornforth, D. M., R. Popat, L. McNally, J. Gurney, T. C. Scott-Phillips, A. Ivens, S. P. Diggle, and S. P. Brown. 2014. Combinatorial quorum sensing allows bacteria to resolve their social and physical environment. Proc. Natl. Acad. Sci. USA 111: 4280-4284.

Diggle, S. P., A. S. Griffin, G. S. Campbell, and S. A. West. 2007. Cooperation and conflict in quorum-sensing bacterial populations. Nature 450: 411-414.

Featherston, J., Y. Arakaki, E. R. Hanschen, P. J. Ferris, R. E. Michod, B. J. S. C. Olson, H. Nozaki, and P. M. Durand. 2018. The 4-celled *Tetrabaena socialis* muclear genome reveals the

essential components for genetic control of cell number at the origin of multicellularity in the volvocine lineage. Mol. Biol. Evol. 35: 855-870.

Feng, L., S. T. Rutherford, K. Papenfort, J. D. Bagert, J. C. van Kessel, D. A. Tirrell, N. S. Wingreen, and B. L. Bassler. 2015. A qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. Cell 160: 228-240.

Fisher, R. M., T. Bell, and S. A. West. 2016. Multicellular group formation in response to predators in the alga *Chlorella vulgaris*. J. Evol. Biol. 29: 551-559.

Frank, S. A. 1995. Mutual policing and repression of competition in the evolution of cooperative groups. Nature 377: 520-522.

Good, B. H., M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai. 2017. The dynamics of molecular evolution over 60,000 generations. Nature 551: 45-50.

Gray, T. A., R. R. Clark, N. Boucher, P. Lapierre, C. Smith, and K. M. Derbyshire. 2016. Intercellular communication and conjugation are mediated by ESX secretion systems in mycobacteria. Science 354: 347-350.

Grosberg, R. K., and R. R. Strathmann. 2007. The evolution of multicellularity: a minor major transition? Ann. Rev. Ecol. Evol. Syst. 38: 621-654.

Hamilton, W. D. 1964a. The genetical evolution of social behaviour. I. J. Theor. Biol. 7: 1-16.

Hamilton, W. D. 1964b. The genetical evolution of social behaviour. II. J. Theor. Biol. 7: 17-52.

Hammerschmidt, K., C. J. Rose, B. Kerr, and P. B. Rainey. 2014. Life cycles, fitness decoupling and the evolution of multicellularity. Nature 515: 75-79.

Henneman, B., C. van Emmerik, H. van Ingen, and R. T. Dame. 2018. Structure and function of archaeal histones. PLoS Genet. 14: e1007582.

Herron, M. D, J. D. Hackett, F. O. Aylward, and R. E. Michod. 2009. Triassic origin and early radiation of multicellular volvocine algae. Proc. Natl. Acad. Sci. USA 106: 3254-3258.

Herron, M. D. 2016. Origins of multicellular complexity: *Volvox* and the volvocine algae. Mol. Ecol. 25: 1213-1223.

Hirose, S., R. Benabentos, H. I. Ho, A. Kuspa, and G. Shaulsky. 2011. Self-recognition in social amoebae is mediated by allelic pairs of tiger genes. Science 333: 467-70.

Irimia, M., and S. W. Roy. 2014. Origin of spliceosomal introns and alternative splicing. Cold Spring Harb. Perspect. Biol. 6: a016071.

Jemielita, M., N. S. Wingreen, and B. L. Bassler. 2018. Quorum sensing controls *Vibrio cholerae* multicellular aggregate formation. eLife 7: e42057.

Jiménez-Marín, B., and B. J. S. C. Olson. 2022. The curious case of multicellularity in the volvocine algae. Front. Genet. 13: 787665.

Keim, C. N., F. Abreu, U. Lins, H. Lins de Barros, and M. Farina. 2004. Cell organization and ultrastructure of a magnetotactic multicellular organism. J. Struct. Biol. 145: 254-262.

King, N., M. J. Westbrook, S. L. Young, A. Kuo, M. Abedin, J. Chapman, S. Fairclough, U. Hellsten, Y. Isogai, I. Letunic, et al. 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature 451: 783-788.

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

Kishi, Y., and J. Parker. 2021. Cell type innovation at the tips of the animal tree. Curr. Opin. Genet. Dev. 69: 112-121.

Klim, J., A. Gładki, R. Kucharczyk, U. Zielenkiewicz, and S. Kaczanowski. 2018. Ancestral state reconstruction of the apoptosis machinery in the common ancestor of eukaryotes. G3 (Bethesda) 8: 2121-2134.

Koonin, E. V., and L. Aravind. 2002. Origin and evolution of eukaryotic apoptosis: the bacterial connection. Cell Death Differ. 9: 394-404.

Lambert, G., and E. Kussell. 2014. Memory and fitness optimization of bacteria under fluctuating environments. PLoS Genet. 10: e1004556.

Lane, N., and W. Martin. 2010. The energetics of genome complexity. Nature 467: 929-934.

Libby, E., and W. C. Ratcliff. 2014. Ratcheting the evolution of multicellularity. Science 346: 426-427.

Lindsey, C. R., F. Rosenzweig, and M. D. Herron. 2021. Phylotranscriptomics points to multiple independent origins of multicellularity and cellular differentiation in the volvocine algae. BMC Biol. 19: 182.

Love, A. C., and G. P. Wagner. 2022. Co-option of stress mechanisms in the origin of evolutionary novelties. Evolution 76: 394-413.

Lynch, M. 2007. The Origins of Genome Architecture. Sinauer Assocs., Inc., Sunderland, MA.

Lynch, M. 2010. Evolution of the mutation rate. Trends Genetics 26: 345-352.

Lynch, M. 2020. The evolutionary scaling of cellular traits imposed by the drift barrier. Proc. Natl. Acad. Sci. USA 117: 10435-10444.

Lynch, M., B. Trickovic, and C. P. Kempes. 2022. Evolutionary scaling of maximum growth rate with organism size. Sci. Reports 12: 22586.

Makarova, K. S., and E. V. Koonin. 2013. Archaeology of eukaryotic DNA replication. Cold Spring Harb. Perspect. Biol. 5: a012963.

Makarova, K. S., Y. I. Wolf, S. L. Mekhedov, B. G. Mirkin, and E. V. Koonin. 2005. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. Nucleic Acids Res. 33: 4626-4638.

Márquez-Zacarías, P., P. L. Conlin, K. Tong, J. T. Pentz, and W. C. Ratcliff. 2021. Why have aggregative multicellular organisms stayed simple? Curr. Genet. 67: 871-876.

Mathis, R., and M. Ackermann. 2016. Response of single bacterial cells to stress gives rise to complex history dependence at the population level. Proc. Natl. Acad. Sci. USA 113: 4224-4229.

Matt, G. Y., and J. G. Umen. 2018. Cell-type transcriptomes of the multicellular green alga *Volvox carteri* yield insights into the evolutionary origins of germ and somatic differentiation programs. G3 (Bethesda) 8: 531-550.

Maynard Smith, J., and E. Szathmáry. 1995. The Major Transitions in Evolution. Oxford Univ. Press, Oxford, UK.

Medini, H., T. Cohen, and D. Mishmar. 2020. Mitochondria are fundamental for the emergence of metazoans: on metabolism, genomic regulation, and the birth of complex organisms. Annu. Rev. Genet. 54: 151-166.

Michod, R. E. 1999. Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality. Princeton Univ. Press, Princeton, NJ.

Mikhailov, K. V., A. V. Konstantinova, M. A. Nikitin, P. V. Troshin, L. Y. Rusin, V. A. Lyubetsky, Y. V. Panchin, A. P. Mylnikov, L. L. Moroz, S. Kumar, et al. 2009. The origin of Metazoa: a transition from temporal to spatial cell differentiation. Bioessays 31: 758-768.

Mitchell, A., G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel. 2009. Adaptive prediction of environmental changes by microorganisms. Nature 460: 220-224.

Mizuno, K., M. Maree, T. Nagamura, A. Koga, S. Hirayama, S. Furukawa, K. Tanaka, and K. Morikawa. 2022. Novel multicellular prokaryote discovered next to an underground stream. eLife 11: e71920.

Nelson, P., and J. Masel. 2017. Intercellular competition and the inevitability of multicellular aging. Proc. Natl. Acad. Sci. USA 114: 12982-12987.

Ng, W. L., and B. L. Bassler. 2009. Bacterial quorum-sensing network architectures. Annu. Rev. Genet. 43: 197-222.

Ocaña-Pallarès, E., T. A. Williams, D. López-Escardó, A. S. Arroyo, J. S. Pathmanathan, E. Bapteste, D. V. Tikhonenkov, P. J. Keeling, G. J. Szöllõsi, and I. Ruiz-Trillo. 2022. Divergent genomic trajectories predate the origin of animals and fungi. Nature 609: 747-753.

Pande, S., H. Merker, K. Bohl, M. Reichelt, S. Schuster, L. F. de Figueiredo, C. Kaleta, and C. Kost. 2014. Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. ISME J. 8: 953-962.

Parkinson, C. N. 1958. Parkinson's Law, or The Pursuit of Progress. John Murray, London, UK.

Pollak, S., S. Omer-Bendori, E. Even-Tov, V. Lipsman, T. Bareia, I. Ben-Zion, and A. Eldar. 2016. Facultative cheating supports the coexistence of diverse quorum-sensing alleles. Proc. Natl. Acad. Sci. USA 113: 2152-2157.

Ratcliff, W. C., J. D. Fankhauser, D. W. Rogers, D. Greig, and M. Travisano. 2015. Origins of multicellular evolvability in snowflake yeast. Nat. Commun. 6: 6102.

Ratcliff, W. C., and M. Travisano. 2014. Experimental evolution of multicellular complexity in *Saccharomyces cerevisiae*. BioScience 64: 383-393.

Richter, D. J., P. Fozouni, M. B. Eisen, and N. King. 2018. Gene family innovation, conservation and loss on the animal stem lineage. eLife 7: e34226.

Rogozin, I. B., L. Carmel, M. Csuros, and E. V. Koonin. 2012. Origin and evolution of spliceosomal introns. Biol. Direct. 7: 11.

Rosenthal, A. Z., Y. Qi, S. Hormoz, J. Park, S. H. Li, and M. B. Elowitz. 2018. Metabolic interactions between dynamic bacterial subpopulations. eLife 7: e33099.

Rozhok, A., and J. DeGregori. 2019. Somatic maintenance impacts the evolution of mutation rate. BMC Evol. Biol. 19: 172.

Ruiz-Trillo, I., and A. M. Nedelcu. (eds.) 2015. Evolutionary Transitions to Multicellular Life:

Principles and Mechanisms. Springer-Verlag, New York, NY.

Scofield, D. G., and M. Lynch. 2008. Evolutionary diversification of the Sm family of RNA-associated proteins. Mol. Biol. Evol. 25: 2255-2267.

Sebé-Pedrós, A., B. M. Degnan, and I. Ruiz-Trillo. 2017. The origin of Metazoa: a unicellular perspective. Nat. Rev. Genet. 18: 498-512.

Sebé-Pedrós, A., A. J. Roger, F. B. Lang, N. King, and I. Ruiz-Trillo. 2010. Ancient origin of the integrin-mediated adhesion and signaling machinery. Proc. Natl. Acad. Sci. USA 107: 10142-10147.

Shapiro, O. H., R. Hatzenpichler, D. H. Buckley, S. H. Zinder, and V. J. Orphan. 2011. Multicellular photo-magnetotactic bacteria. Environ. Microbiol. Rep. 3: 233-238.

Smith, D. R., T. Hamaji, B. J. Olson, P. M. Durand, P. Ferris, R. E. Michod, J. Featherston, H. Nozaki, and P. J. Keeling. 2013. Organelle genome complexity scales positively with organism size in volvocine green algae. Mol. Biol. Evol. 30: 793-797.

Solari, C. A., S. Ganguly, J. O. Kessler, R. E. Michod, and R. E. Goldstein. 2006. Multicellularity and the functional interdependence of motility and molecular transport. Proc. Natl. Acad. Sci. USA 103: 1353-1358.

Stefanic, P., B. Kraigher, N. A. Lyons, R. Kolter, and I. Mandic-Mulec. 2015. Kin discrimination between sympatric *Bacillus subtilis* isolates. Proc. Natl. Acad. Sci. USA 112: 14042-14047.

Tong, K., G. O. Bozdag, and W. C. Ratcliff. 2022. Selective drivers of simple multicellularity. Curr. Opin. Microbiol. 67: 102141.

Ueda, S., S. Yamashita, M. Nakajima, T. Kumamoto, C. Ogawa, Y. Y. Liu, H. Yamada, E. Kubo, N. Hattori, H. Takeshima, et al. 2022. A quantification method of somatic mutations in normal tissues and their accumulation in pediatric patients with chemotherapy. Proc. Natl. Acad. Sci. USA 119: e2123241119.

Vassallo, C. N., and D. Wall. 2019. Self-identity barcodes encoded by six expansive polymorphic toxin families discriminate kin in myxobacteria. Proc. Natl. Acad. Sci. USA 116: 24808-24818.

Vos, M., and G. J. Velicer. 2009. Social conflict in centimeter-and global-scale populations of the bacterium *Myxococcus xanthus.* Curr. Biol. 19: 1763-1767.

Vosseberg, J., J. J. E. van Hooff, M. Marcet-Houben, A. van Vlimmeren, L. M. van Wijk, T. Gabaldón, and B. Snel. 2021. Timing the origin of eukaryotic cellular complexity with ancient duplications. Nat. Ecol. Evol. 5: 92-100.

Wahl, M. E., and A. W. Murray. 2016. Multicellularity makes somatic differentiation evolutionarily stable. Proc. Natl. Acad. Sci. USA 113: 8362-8367.

Wellington Miranda, S., Q. Cong, A. L. Schaefer, E. K. MacLeod, A. Zimenko, D. Baker, and E. P. Greenberg. 2021. A covariation analysis reveals elements of selectivity in quorum sensing systems. eLife 10: e69169.

Williams, G. C. 1966. Adaptation and Natural Selection. Princeton Univ. Press, Princeton, NJ.

Wilson, E. O. 1975. Sociobiology: the New Synthesis. Harvard Univ. Press, Cambridge, MA.

Zhou, X., Z. Lin, and H. Ma. 2010. Phylogenetic detection of numerous gene duplications shared by animals, fungi and plants. Genome Biol. 11: R38.

Figure 24.1. Scaling of maximum interval-specific growth rates with size at maturity. For unicellular species, the growth rates are equivalent to cell-division rates, whereas for multicellular species, they are the maximum rates among all age classes during individual development. All data are standardized to 20°C. The upper dashed line is the approximate position of the upper limit to evolved maximum growth rates (not including larval fishes, with maternal provisioning); solid lines are fitted regressions for individual phylogenetic groups. Note that this figure is an extension of Figure 8.5 to multicellular species. From Lynch et al. (2022).

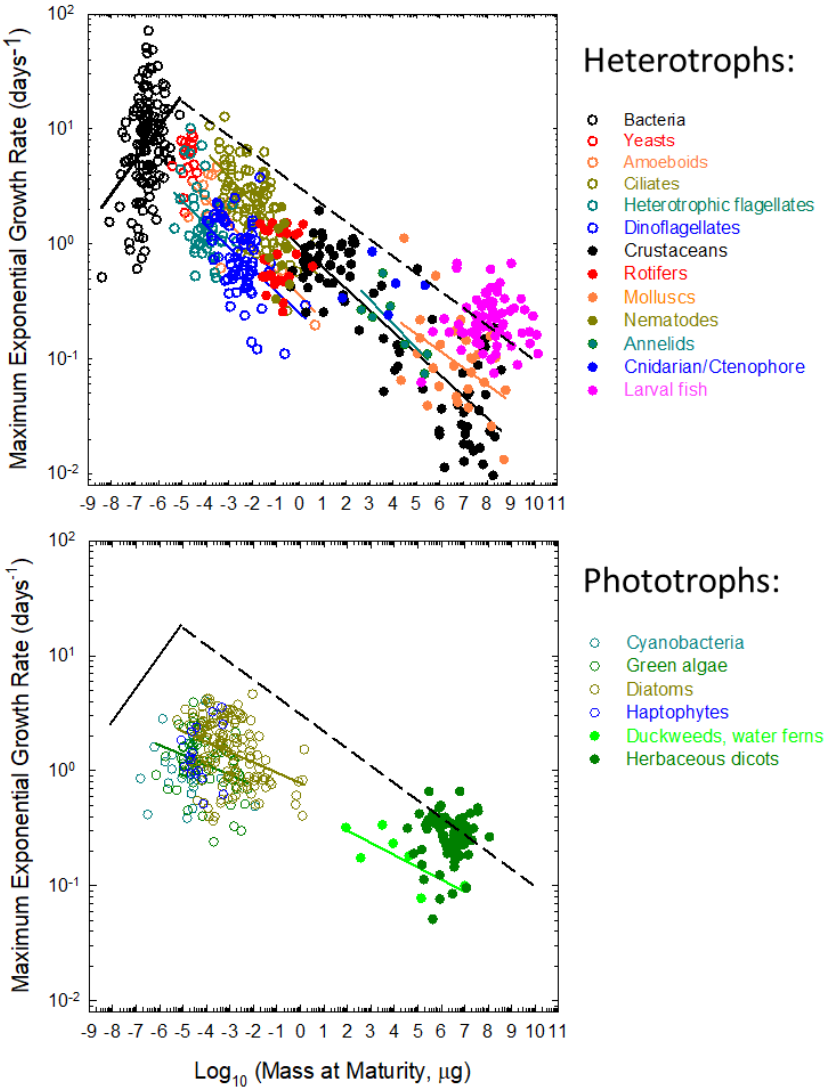

Figure 24.2. The scaling of the number of protein-coding genes per genome vs. total genome size (in millions of bases) across the Tree of Life. The numbers on the dashed isoclines denote the average kb of genomic DNA per gene for points located on the lines. Data obtained from the NCBI genome summaries.