SYMPOSIUM INTRODUCTION

The Temporal and Environmental Context of Early Animal Evolution: Considering All the Ingredients of an “Explosion”

Erik A. Sperling¹ and Richard G. Stockey

Department of Geological Sciences, Stanford University, 450 Serra Mall, Building 320, Stanford, CA 94305, USA

From the symposium "From Small and Squishy to Big and Armored: Genomic, Ecological and Paleontological Insights into the Early Evolution of Animals" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2018 at San Francisco, California.

¹E-mail: esper@stanford.edu

Synopsis Animals originated and evolved during a unique time in Earth history—the Neoproterozoic Era. This paper aims to discuss (1) when landmark events in early animal evolution occurred, and (2) the environmental context of these evolutionary milestones, and how such factors may have affected ecosystems and body plans. With respect to timing, molecular clock studies—utilizing a diversity of methodologies—agree that animal multicellularity had arisen by ~800 million years ago (Ma) (Tonian period), the bilaterian body plan by ~650 Ma (Cryogenian), and divergences between sister phyla occurred ~560–540 Ma (late Ediacaran). Most purported Tonian and Cryogenian animal body fossils are unlikely to be correctly identified, but independent support for the presence of pre-Ediacaran animals is recorded by organic geochemical biomarkers produced by demosponges. This view of animal origins contrasts with data from the fossil record, and the taphonomic question of why animals were not preserved (if present) remains unresolved. Neoproterozoic environments demanding small, thin, body plans, and lower abundance/rarity in populations may have played a role. Considering environmental conditions, geochemical data suggest that animals evolved in a relatively low-oxygen ocean. Here, we present new analyses of sedimentary total organic carbon contents in shales suggesting that the Neoproterozoic ocean may also have had lower primary productivity—or at least lower quantities of organic carbon reaching the seafloor—compared with the Phanerozoic. Indeed, recent modeling efforts suggest that low primary productivity is an expected corollary of a low-O₂ world. Combined with an inability to inhabit productive regions in a low-O₂ ocean, earliest animal communities would likely have been more food limited than generally appreciated, impacting both ecosystem structure and organismal behavior. In light of this, we propose the “fire triangle” metaphor for environmental influences on early animal evolution. Moving toward consideration of all environmental aspects of the Cambrian radiation (fuel, heat, and oxidant) will ultimately lead to a more holistic view of the event.

Introduction

Questions regarding early animal evolution are a hot research topic—from determining the sister group to all other animals (compare Whelan et al. [2015] and Simion et al. [2017]), to understanding the genetic and developmental changes driving body plan evolution (Fernandez-Valverde et al. 2018; Paps 2018; Paps and Holland 2018), to determining how best to analyze the exponential increase in new genomic sequence data being generated (Laumer 2018). However, these debates can often take place in a sterile setting, between nodes on a cladogram. We gain a much deeper understanding of metazoan evolution by rooting our discussions in time and place. Here, we aim to review and synthesize recent advances in understanding the temporal and environmental context of early animal evolution.

Some of the most famous early animal fossils are exceptionally preserved specimens from the Cambrian Burgess Shale and Chengjiang lagerstätte. Their canonization as “weird wonders” (Gould 1990) stimulated decades of debate on the nature of early animal evolution, and they continue to yield remarkable insights into the nature of Cambrian animals and ecosystems (see for instance Sheppard et al. [2018]). However, phylogenetic research on these
Cambrian faunas has demonstrated that almost all of the “weird wonders” can convincingly be placed into the stem-groups of extant phyla, if not within the crown group (reviewed by Erwin and Valentine 2013; Briggs 2015). In other words, Cambrian faunas document evolution along the stem-lineages of individual phyla and classes. Thus these fossils represent the end of the first chapter in animal evolution. To start at the beginning of this chapter and the early “landmark” events (the origin of multicellularity, the nervous system, gut, mesoderm, and bilaterality) we must turn earlier in the geological record. But how much earlier? Darwin (1859) famously suggested animals had an extremely long pre-history, and this early history of animals was lost through geological processes of erosion and metamorphosism. This explained why animals appeared so suddenly on the scene at the base of the Cambrian: it was an artifact of an imperfect geological record. In the last three decades there has been a tremendous research effort focused on the time period preceding the Cambrian (the Neoproterozoic Era—comprised of the Tonian, Cryogenian, and Ediacaran periods; Fig. 1). Contrary to Darwin’s hypothesis, this interval is indeed present in the rock record and does record the roots of animal evolution. Geologists have established a robust chronology and stratigraphic framework for the interval (Narbonne et al. 2012; Shields-Zhou et al. 2016), and targeted geochemistry has firmly established ages and durations for the Neoproterozoic “Snowball Earth” glaciations (Fig. 1) (Macdonald et al. 2010; Rooney et al. 2015; Prave et al. 2016; Pu et al. 2016). Meanwhile, paleontologists have discovered new fossil localities, re-investigated known deposits with advanced imaging techniques, and harnessed information from new approaches such as molecular clocks and the organic geochemical (biomarker) record.

The first goal of this paper is to review these new advances regarding the timing of early animal evolution. Aside from putting ages on the evolutionary trees considered by neontologists, the question of timing is central to evaluating causal hypotheses for the Cambrian radiation. It is well agreed that animals are monophyletic—an evolutionary singularity—and thus the origins of animals and multicellularity likely have more to do with genetic challenges than an environmental “trigger” (Butterfield 2018). If crown-group animals (as a marker for complex multicellularity) evolved in the mid-to-late Ediacaran, just before the Cambrian, as has been proposed by some researchers (Antcliffe et al. 2014; Budd and Jensen 2017), the Cambrian radiation could simply be read as an adaptive radiation. Alternatively, if animals originated far earlier in the Proterozoic, with traits like large body size, skeletonization, and the range of new morphological features seen in Cambrian fossils appearing convergently and simultaneously, it is more likely an external ecological or environmental trigger was involved.

The nature of this trigger has been widely debated in the literature. Comprehensive review of all hypotheses is not possible here, but several recent papers have covered this topic (Erwin et al. 2011; Zhang et al. 2014; Schiffbauer et al. 2016). Proposed ecological triggers for the Cambrian radiation, or certain aspects of it, range from ecosystem engineering (Butterfield 2011, 2018; Erwin and Tweedt 2011), direct organism–organism interactions including predation and expansion of food webs (Peterson et al. 2005; Marshall 2006; Butterfield 2007; Schiffbauer et al. 2016), and changes in resource patchiness (Budd and Jensen 2017). Proposed environmental triggers include true polar wander (Mitchell et al. 2015), changes in the major ion composition of seawater (Peters and Gaines 2012), and most prominently in the geological literature, a substantial increase in oxygen levels in the ocean/atmosphere system (reviewed by Mills and Canfield 2014; Xiao 2014). The exact mechanistic relationship between increasing oxygen levels and the Cambrian radiation in the literature is varied. Many geochemical studies simply state that animals require oxygen, and assert that more oxygen must be better. Other studies have linked oxygen changes to ecological (Sperling et al. 2013a), developmental (Hammarlund et al. 2018), or evolutionary (Wood and Erwin 2018) mechanisms that were themselves proposed as the proximate cause of the Cambrian radiation. Indeed, as recognized in a recent paper by Schiffbauer et al. (2016), the search for a sole trigger may be simplistic given the myriad interconnected links between the biosphere and geosphere.

In this paper, we build on published geochemical modeling and present a new investigation of shale organic carbon to suggest that both oxygen and food supply to the benthos likely changed through the Neoproterozoic–Paleozoic transition. Looking across natural environmental gradients in the modern ocean, we then demonstrate that the ecological effects of increasing food or increasing oxygen are directionally similar. While it is evident that the Cambrian radiation was a multi-faceted event, we emphasize that directly assigning causal significance to individual forcing factors remains a challenging task, and demands consideration of a cascade of coupled genomic, environmental, and ecological changes.
Temporal context

Fossil record

The most obvious path to studying the early history of animals is by direct investigation of the fossil record. Moving back in time from the Cambrian, the Ediacara biota represents the first large, complex fossils in the fossil record (reviewed by Boag et al. 2016; Droser et al. 2017; Tarhan et al. 2018). The first Ediacaran organisms appear in the “Avalon” assemblage of Newfoundland and UK after ~571 million years ago (Ma) (Noble et al. 2015; Pu et al. 2016). These generally represent enigmatic sessile frond-like organisms with a “fractal” organization that are unlikely to be animals (Sperling et al. 2007)—or at least show no obvious characters relating them to animals. However, trace fossils possibly representing diploblastic organisms are also present (Liu et al. 2010), and some body fossils may represent cnidarians (Liu et al. 2014) or sponges (Sperling et al. 2011). The succeeding “White Sea” and “Nama” assemblages record a considerable increase in the diversity of form, movement, reproduction strategies, and biomineralization (Tarhan et al. 2018). The species making up these assemblages flourished only briefly, as possible Cambrian survivors are controversial, and certainly very rare even if genuine survivors (Laflamme et al. 2013; Budd and Jensen 2017). Whether these Ediacaran organisms suddenly became extinct at the end of the Ediacaran due to an environmental perturbation (Smith et al. 2016), or gradually disappeared due to biological interactions with evolving bilaterians (Darroch et al. 2015), remains a critical question in Ediacaran paleobiology (reviewed by Tarhan et al. [2018]). In any case, as noted by the title of a recent review paper on the Ediacara biota (Droser et al. 2017), these fossils in the latter half of the Ediacaran likely represent the “rise of animals” rather than their origin and early evolution, and consequently we must look earlier.
The early Ediacaran, Cryogenian, and Tonian fossil records have been relatively silent (or at least controversial) with respect to the origin of animals. As discussed below, a number of paleontologists believe this silence to be an accurate record. During these time periods a number of putative sponge fossils have been reported, but none can be fully accepted as representing evidence of early animals. A single phosphatized fossil named *Eocyathispongia gia-nia* from the Doushantuo Formation (~600 Ma) has sponge-like features and a body plan possibly consistent with a water-channel system (Yin et al. 2015). However, only a single specimen has been discovered, from a geological formation with a large variety of intriguing but clearly abiogenic phosphatic structures. For instance, *Vernanimalcula* from the Doushantuou was initially described as a bilaterian fossil (Chen et al. 2004), based on a few specimens discovered during an initial search of over 50,000 specimens (Bengtson et al. 2012). *Vernanimalcula* has now been convincingly demonstrated to be a pseudo-fossil, and examination of the full range of morphologies in the Doushantuo shows how *Vernanimalcula* is an end-member artifact of phosphatization (Bengtson et al. 2012). Putative siliceous demosponge spicules from the Doushantuo have also recently been invalidated (Muscente et al. 2015). Considering *Eocyathispongia* then, with only one specimen, and the possibility that its body plan is not consistent with a functioning water-channel system (Cavalier-Smith 2017), its status as a fossil sponge must remain uncertain (as acknowledged by the authors). Phosphatized embryos from the Doushantuo Formation (Xiao et al. 1998) still have promise for understanding evolutionary trends during the Ediacaran, but more plausibly represent single-celled organisms than animals (Huldtgren et al. 2011).

Moving farther back, sponge fossils have also been described from the Cryogenian of South Australia, but again they are not without issue. Maloof et al. (2010) applied a serial grinding technique to image specimens from the Trezona Formation, and described a population of millimeter- to centimeter-sized nonsymmetric organisms with a network of interior canals leading to openings on the fossil surface. However, these openings are recessed into the surface of the fossil, rather than elevated like most sponge oscula, and would lead to recycling of the feeding current. It is also unclear if the postulated fossils match expectations for flow in biological fluid transport systems (LaBarbera 1990), and it is possible these clasts represent ripped up microbial mats (Antcliffe et al. 2014). Other possible Neoproterozoic sponges, including even older reported fossils, were reviewed by Antcliffe et al. (2014) and found unconvincing.

In short, while many fossils have not been conclusively falsified as sponges—including the two fossils discussed above—neither can there be considered strong fossil evidence for sponges (or any other animal group) in the Cryogenian or Tonian. The key question is whether the fossil record, with an absence of Tonian to mid-Ediacaran animals, can be read literally with respect to the timing of animal origins. Many paleontologists believe it can (e.g., Budd and Jensen 2017). Alternative sources of evidence, such as the organic geochemical record and molecular clocks, suggest otherwise; below we discuss both of these records in turn.

**Organic geochemistry**

While most paleontologists search for body fossils, an alternative approach to interrogate the early history of clades is to search for molecular fossils: diagnostically organic molecules produced by organisms (Briggs and Summons 2014). These “biomarkers” are the hydrocarbon remains of lipids such as sterols in eukaryotes and hopanoids in bacteria. Compared with microbes, the biomarker complement of animals is relatively “boring” (i.e., not taxonomically specific)—with the exception of demosponges. One sterol in particular (a C₃₀ sterol, 24-isopropylcholesterol, with its fossilized sterane complement known as 24-isopropylcholestanol, or 24-ipc) is exclusively made as a major sterol by living demosponges.

Love et al. (2009) published a record of 24-ipc steranes from the South Oman Salt Basin, where it was found in Cryogenian strata—beneath the Marinoan glaciation—and throughout the Ediacaran into the Cambrian (Fig. 1). Critically, 24-ipc was found in both conventional solvent extracts and in the insoluble kerogen phase, helping rule out secondary contamination (which has been a recent concern in previous ancient biomarker studies; see French et al. [2015]). 24-ipc has also been found in “relatively high” abundance in other rocks worldwide of this age (McCaffrey 1994; Kelly et al. 2011). Note that “high” really is relative, as total C₃₀ steranes (of which 24-ipc is just one component) represent a minor component of total steranes in these Neoproterozoic rocks—typically only 1–4% of the total C₂₇–C₃₀ steranes (Love et al. 2009). Based on the phylogenetic pattern of 24-isopropylcholesterol synthesis in living organisms, Love et al. (2009) proposed that these Cryogenian and Ediacaran 24-ipc steranes were likely produced by total-group demosponges.
The inference that 24-ipc in Neoproterozoic rocks represent the preserved organic remains of Neoproterozoic demosponges has been controversial. Brocks and Butterfield (2009) suggested that unicellular relatives of animals could have produced these ancient biomarkers. Antcliffe (2013) suggested that these steranes could have been produced by pelagophyte algae (some of which produce trace quantities of 24-isopropylcholesterol), could have been a contaminant introduced by drilling fluid, or could have been produced by diagenetic alteration of other sterol skeletons. These latter two points are based on fundamental misunderstandings of organic geochemistry (Love and Summons 2015), but concerns about 24-ipc as a phylogenetically unique marker for demosponges remain. To address this, Gold et al. (2016) conducted sterol assays of the living unicellular relatives to animals, and demonstrated they do not make 24-ipc sterol precursors (or indeed any C_{30} sterols at all; see also Kodner et al. [2008]). Gold and colleagues then investigated the molecular underpinnings of 24-isopropylcholesterol synthesis, and demonstrated that demosponges and pelagophyte algae independently evolved the ability to methylate C_{30} sterols. Given the timing of gene duplications for the known extant sterol producers, only demosponges, and not pelagophytes, could have been the source of Neoproterozoic 24-ipc. Most recently, Zumberge et al. (2017) discovered a new C_{30} sterane biomarker, named 26-methylstigmastane (26-mes), whose sterol precursors are only produced by demosponges among the living biota. Thus, 26-mes does not suffer from some ambiguities present in interpreting the 24-ipc record. 26-mes was found in the same Cryogenian through Cambrian rock and oil samples as 24-ipc, providing convincing evidence of the pattern described by Love et al. (2009). Cryostane, a C_{28} sterane, appears in even earlier pre-Sturtian rocks of the Tonian period and may have also been made by demosponges, but this is speculative because unlike 24-ipc or 26-mes the natural sterol precursors have not been identified (Brocks et al. 2016).

A subtle but important distinction is that, as far as currently known, 24-ipc and 26-mes are not sponge biomarkers but rather demosponge biomarkers. Certainly, more sterol assays and genomic studies are needed on other sponge groups and unicellular relatives of animals, but no other sponge or holozoan protist group has been found to have any unconventional sterol structure with respect to side-chain chemistry. Only conventional sterols with common straight-chain alkylated side-chains are found. The fact that these two unconventional biomarkers are only made as a major sterol by demosponges (and not all sponges) indicates the presence of a derived group of sponges in the Cryogenian. Consequently, the Porifera crown group must have originated earlier. In other words, no matter the phylogenetic relationship of sponges and ctenophores to other animals, the biomarkers provide a minimum age constraint for crown-group sponges and suggest multicellular animals with a water-canal system were living prior to 635 Ma.

**Molecular clocks**

Molecular clocks calculate the rate of molecular evolution through time in groups with a presumably robust fossil record, and then in simplest form apply these rates across the phylogenetic tree to infer divergence times. This practice as applied to animal evolution has a somewhat sordid history (Graur and Martin 2004). However, beginning with the use of multiple calibration points spread phylogenetically across invertebrates (Peterson et al. 2004), and most importantly through the introduction of “relaxed clock” methods, analyses have reached broad agreement on the timing of animal origins (Fig. 1) (Aris-Brosou and Yang 2003; Peterson et al. 2008; Erwin et al. 2011; Parfrey et al. 2011; dos Reis et al. 2015; Dohrmann and Wörheide 2017; Lozano-Fernandez et al. 2017). Significant uncertainties do still remain regarding the validity of specific paleontological calibration points, how uncertainty around calibration points should be modeled, and whether uncorrelated or autocorrelated rate models should be used. In light of these uncertainties both dos Reis et al. (2015) and Cunningham et al. (2017) have provided useful and timely discussion of the precision afforded by molecular clocks. To address this uncertainty, the full range of possible parameter space in terms of gene sequences, partitions, possible paleontological calibration points, soft- versus hard-bounds on calibration points, calibration probability densities, root priors, phylogenetic topologies, rate models, and software choices have now been explored in sensitivity analyses—filling pages of supplementary tables and figures (see references above for details). Two robust messages have emerged.

First, and consistent with demosponge biomarker data, animal multicellularity emerged prior to the Ediacaran, possibly in the Cryogenian but more likely in the Tonian. Although there is considerable uncertainty on these deeper nodes, the mean divergence between sponges and all other animals is in the Tonian in all “most likely” analyses (e.g., the
main summary figure of a paper). Only in one analysis does the 95% confidence interval (CI) for crown-group animals extend into the Cryogenian (dos Reis et al. [2015]; note that we use updated definitions for the Cryogenian: 720–635 Ma, see Shields-Zhou et al. [2016]). Molecular divergence estimations using topologies where cnidarians rather than sponges are the sister group of all other animals have not been extensively examined, but analyses by Dohrmann and Wörheide (2017) demonstrated that the 95% CI for crown-group animals does not extend into the Cryogenian using this topology. Indeed, across the dozens of different sensitivity analyses conducted by multiple research groups, the mean estimate for crown group animals is never in the Cryogenian, and 95% CIs generally exclude this period. Likewise, in all “most likely” analyses the 95% CIs for crown-group Metazoa exclude the Mesoproterozoic (>1000 Ma), and only rarely do 95% CIs extend to the Mesoproterozoic in sensitivity analyses. The animal crown node of course represents the minimum time for the origin of animal multicellularity, which must have evolved along the metazoan stem-lineage. Most estimates place the animal–choanoflagellate divergence at ~900–1000 Ma (Fig. 1), suggesting that the dawn of animal multicellularity was likely in the Tonian even accounting for all sources of uncertainty (specifically between the white and gray nodes in Fig. 1).

The origin of eumetazoan and bilaterian body plans generally falls in the Cryogenian but there is considerable uncertainty. In contrast to the relative cross-study consensus for the animal crown node, for Eumetazoa and Bilateria the 95% CIs of some studies exclude the mean estimate of other analyses. The disparate ages between studies for Eumetazoa and Bilateria remain an area for investigation. In any case, the second robust finding from molecular clock studies is that the divergences between individual sister phyla likely occurred in the late Ediacaran (Fig. 1, black nodes). This pattern is best illustrated by the taxon-rich study of Erwin et al. (2011), where many more sister phyla were analyzed than could be visually included in Fig. 1. Molecular clock evidence often places the crown nodes of these phyla within the Cambrian when uncertainty is included, and/or identifiable fossil representatives of these phyla occur in the early-middle Cambrian (even if not having all the characteristics of the crown group; Budd and Jensen [2000]). Consequently, it is clear that the pulse of body plan innovation, skeletonization, and size increase that characterizes the Cambrian radiation took place as a polyphyletic radiation across multiple lineages around the Precambrian-Cambrian boundary (Peterson et al. 2005). In essence, while animals have a cryptic Tonian origin according to molecular clocks and biomarker data, the explosion of new body plans recorded in the Cambrian fossil record is a real evolutionary event. As discussed in the “Introduction” section, this polyphyletic radiation implies an external forcing.

We recognize that this reading of the animal evolutionary record leans heavily on molecular clock and biomarker data over a literal reading of the fossil record. Most persuasive to us, in this regard, is that there is now strong concordance between molecular clocks and the fossil record in the Phanerozoic (Erwin et al. 2011; Cunningham et al. 2017; Daley et al. 2018). Put simply, if clades such as crown-group Arthropoda or Brachiopoda diverged in the early Cambrian, as indicated by both clocks and rocks, there is no way to telescope all the molecular evolution between these divergences and the last common ancestor of animals into the ~30 million years demanded by proponents of a literal reading of the fossil record (Antcliffe et al. 2014; Budd and Jensen 2017). This is true even if accelerated (but plausible) rates of molecular evolution existed in some lineages (Lee et al. 2013).

The view presented here then leads to the eternal question of why early animals are not fossilized if they arose ~800 Ma. Some of this tension relates to time of clade origin (reflected by molecular clocks) versus time of ecological dominance (recorded by the fossil record), and also to some mis-understandings of various records. For instance, the presence of Cryogenian demosponges as supported by biomarkers evokes a seafloor covered in sponges. But as discussed above, the 24-ipc biomarker only occurs at the ppm level in sterol extracts: only minute inputs are required. A long preservational gap is also not an anomaly in the deep-time fossil record. As an example, oxygenic photosynthesis (the major source of oxygen on Earth) must have appeared prior to or coincident with the Great Oxygenation Event at ~2300–2400 Ma, but the first cyanobacterial fossils do not appear until hundreds of millions of years later, at ~2000 Ma (Fischer et al. 2016). Below we discuss some factors (mainly rarity and small body size) that may have played into low preservational potential for early Neoproterozoic animals. However, like most paleontologists, do not find them fully convincing. We view this taphonomic debate, as well as the basic question of temporal origins, as still-open scientific questions. For the purposes of discussing causes, though, we must take this current best-
Environmental context

Animals evolved in a low-oxygen ocean

It has long been hypothesized that animals evolved in a less oxygenated ocean than exists today (e.g., Nursall 1959). Some of this initial reasoning was circular—for instance some evidence for lower oxygen levels hinged on the fact that there were no animals (reviewed by Mills and Canfield 2014)—but it was also clear from straightforward clues in the rock record, like widespread Archean and Proterozoic Banded Iron Formations, that the early Earth was less oxygenated than the modern Earth. In the past decade, more sophisticated geochemical redox studies have generally supported this narrative, providing both local and global evidence for widespread Proterozoic subsurface anoxia (Partin et al. 2013; Reinhard et al. 2013; Liu et al. 2016; Sperling et al. 2015b; Hardisty et al. 2017). The presumption is that widespread Proterozoic marine anoxia was a direct result of lower atmospheric oxygen concentrations, although there are arguments against this (Butterfield 2018). Oxygen levels are generally believed to have increased substantially at some point between the mid-Ediacaran to early Cambrian (the hypothesized Neoproterozoic Oxygenation Event; Shields-Zhou and Och 2011). It is this rise in oxygen that is commonly invoked as the geologic trigger for the Cambrian “explosion.” Whether atmospheric oxygen levels rose to modern in the Ediacaran–Cambrian, or the increase was smaller and oxygen levels remained relatively muted through the early Paleozoic, is a current topic of debate (Dahl et al. 2010; Chen et al. 2015; Sperling et al. 2015b; Wallace et al. 2017; Stolper and Keller 2018), as are the implications of a muted rise for the Cambrian explosion (Sperling et al. 2015a). However, the prevalence of subsurface marine anoxia is a clear message of the Neoproterozoic geochemical record.

The implications for early animals of a poorly oxygenated ocean have been extensively explored in the literature. Based on theoretical calculations, animals would be limited to relatively small and thin body plans (Raff and Raff 1970; Runnegar 1982a; Payne et al. 2010). This is matched by empirical evidence from transects across modern oxygen minimum zones (OMZs), which show that animals living in naturally very low-O₂ settings are small, thin, generally unmineralized, and have low-energy (non-carnivorous) feeding strategies (Rhoads and Morse 1971; Levin 2003; Gooday et al. 2009; Sperling et al. 2013b, 2015a). At the very lowest oxygen levels, macrofaunal abundance (organisms retained on a 300 μm sieve) drops dramatically, and the dominant animals are meiofauna such as nematode worms (Gooday et al. 2009). This has important implications for fossil taphonomy and reconciliation between the fossil record and molecular/biomarker data. As noted by Sperling (2013), nematodes occur by the 100–1000 s in every square meter of modern marine seafloor, and probably have for quite some time (the nematode crown group is likely at least 500 million years old; Erwin et al. 2011). Yet, paleontologists have only rarely, if ever, found a marine nematode body fossil (most putative fossil marine nematodes are at least an order of magnitude larger than living nematodes and show few characteristic features). Nematodes are preserved under exceptional conditions such as in terrestrial amber, but in the marine realm such small, thin animals simply do not preserve (reviewed by Poinar 2011). Budd and Jensen (2000) argued that ancestral bilaterians must have been larger than nematodes, at least several millimeters long, but such organisms also lack an extensive fossil record. Consequently, if oxygen levels were low enough to keep marine “worms” small and thin, there is little discrepancy between molecular clock ages for bilaterians in the Cryogenian and the absence of an unequivocal bilaterian fossil record until the late Ediacaran.

Animals evolved in a food-limited ocean

While most studies of the Neoproterozoic environment have focused on oxygen and redox, emerging evidence suggests there were also important differences in primary productivity and food delivery to the seafloor compared with the Phanerozoic. Analysis of a large compilation of sedimentary phosphorous data supports modeling results indicating that anoxic, ferruginous Proterozoic ocean waters may have effectively scavenged bioavailable P, and that oceans during that time could have been considerably P-limited (Derry 2015; Reinhard et al. 2017). Phosphorous limitation appears to have eased at some point in the Neoproterozoic, although the exact global timing is obscured by a dearth of compiled samples from outside South China (Reinhard et al. 2017). Whether linked to these changes in nutrients or not, the Neoproterozoic also witnessed a change in the dominant primary producers, with
organic geochemical data demonstrating that algae broke the incumbency of cyanobacteria between the Sturtian and Marinoan “Snowball Earth” glaciations (Brocks et al. 2017). Most discussion in these and other papers have related changes in primary productivity back to oxygen production in the Earth system, and a more detailed examination of how changes in food supply might have affected early animal ecosystems is required.

To further understand how food supply to the seafloor changed through time, we investigate here the direct record of organic carbon in sedimentary rocks. Specifically, we have conducted the largest compilation of total organic carbon (TOC) contents in shale (fine-grained clastic rocks) to date. This lithology was chosen because (1) it provides a consistent and geographically widespread lithology through time, (2) there are abundant shale TOC measurements in the literature, and (3) shale depositional environments represent the greatest sink for organic carbon in the geological record (Berner 1982). Our analyses combine previous compilations of TOC assembled from US Geological Survey (USGS) data (compiled by Husson and Peters [2017]; 3345 samples), and geochemical compilations focused on other elemental systems but with associated TOC data: specifically studies by Reinhard et al. (2017; 5207 samples) and Partin et al. (2013; 711 samples). To these existing compilations we added new data from the literature (3437 samples). Finally, 992 new shale TOC analyses conducted at Stanford University are reported here. The complete dataset consisted of 13,650 shale TOC measurements through Earth history. Full analytical methods, details of the compilation, and additional caveats are located in the Supplementary Information.

Our results show a distinct change in shale TOC contents at the base of the Phanerozoic (Fig. 2). Specifically, Neoproterozoic shales have a mean organic carbon content of 0.8 wt% (median, 0.2%), compared with 4.3 wt% (median, 2.6%) for the Phanerozoic. These differences are highly significantly different (e.g., $P < 2.2 \times 10^{-16}$ for a two-group Mann–Whitney $U$-test), with our statistical subdivisions supported by a large TOC increase evident from the first time bin fully in the Cambrian (525–500 Ma). When considering these results, it is worth noting there are a large number of depositional and post-depositional factors that influence the TOC value of a rock. A non-exhaustive list includes sedimentation rate, bottom-water $O_2$ content, and type of organic carbon (Canfield 1994), organic carbon loss during thermal maturation (and eventually, metamorphism) (Raiswell and Berner 1987), and oxidative weathering in outcrop versus core (Petsch et al. 2000). We have not attempted to correct for these factors, or for spatial-temporal sampling density, and thus these results should be viewed as a “first pass.” However, we do not anticipate secular changes in either biology or geology that would inhibit interpretation of these data as a first-order summary of benthic food supply.

Given available data, then, the Neoproterozoic is a clear outlier. In fact, the main difference in shale TOC contents through time is not between the Proterozoic and Phanerozoic, but rather between the Neoproterozoic and all other time periods (see Summary Distributions, Fig. 2). Considering the relative lack of data from the Paleoproterozoic and Mesoproterozoic, and the sampling dominance of a few formations, we are agnostic as to whether these Era-level differences are real. The post-850 Ma interval of the Neoproterozoic, though, is much more consistently sampled in space and time. The substantial difference between this interval—when cryptic animals evolved—and the Phanerozoic is likely to be real. It is also pertinent that the most obvious geochemical signals are in high TOC shales, and consequently, geochemists tend to preferentially sample the most TOC-rich horizons. If abundant, high TOC Neoproterozoic shale formations existed, they would have been sampled. These data contrast with the relatively enriched background carbonate carbon isotope record of the Neoproterozoic (Macdonald et al. 2010), which is often taken to represent high fractional rates of organic carbon burial. Aspects of the Neoproterozoic carbon isotope record in general are difficult to explain using conventional frameworks (discussed in Schrag et al. 2013), and these shale TOC data are another line of evidence that must be reconciled.

Other anecdotal lines of evidence can be found to support the shale TOC record. For instance, the rate of organic carbon supply to investigated Neoproterozoic manganese deposits was lower (Johnson et al. 2016), and paleontological evidence suggests Avalonian Ediacaran ecosystems may have been more food-limited than the Cambrian (Vinther et al. 2014; Hoyal-Cuthill and Conway Morris 2017). Many Neoproterozoic carbonates also have low TOC (Johnston et al. 2012a). Finally, given the feedback between organic carbon burial and oxygen production, recent modeling results suggest that a low-$O_2$ ocean/atmosphere system is difficult (or impossible) to achieve without low primary productivity (Laakso and Schrag 2014, 2017; Derry 2015). Accepting a low $O_2$ Neoproterozoic world...
likely requires accepting a low-food Neoproterozoic world.

The shale TOC results in Fig. 2 do not necessarily speak to overall rates of primary productivity, oceanic nutrient inventories, or the dominance of primary producer types, which likely began shifting earlier in the Neoproterozoic (Brocks et al. 2017; Reinhard et al. 2017). Higher primary productivity may have instead been balanced by more respiration of sinking organic matter. However, what the TOC results do estimate is the flux of organic carbon to seafloor ecosystems (export productivity), which is what would have mattered to early benthic animals. And this is essentially all early animals, as aside from possibly medusozoon cnidarians, Neoproterozoic animal evolution was a benthic affair (reviewed by Gold 2018). Further, if oxygen levels were low in the Proterozoic, $O_2$ would have interacted synergistically with low organic carbon supply in two ways to limit early animals’ access to food. First, some early animals such as sponges (which are constrained to be present by the Cryogenian, see above) are likely to have had very high energetic costs of feeding (Leys and Kahn 2018). High relative metabolic cost of feeding combined with low food supply would obviously limit scope for growth, and perhaps the geographic extent of metabolically viable habitat. Second, in a low-$O_2$ world, any area of higher primary productivity (such as upwelling zones) would have quickly lost all oxygen in the water column due to aerobic remineralization of relatively abundant organic carbon. In the modern ocean animals can survive and even thrive in low oxygen, but not in the complete absence of oxygen (Sperling et al. 2013b), and early skeletal metazoans were unable to inhabit low-oxygen areas (Tostevin et al. 2016). This is not to argue that Neoproterozoic animals would have been excluded from all possible food-rich areas, but basic biogeochemistry does suggest they would have been pushed toward primary productivity “deserts.” Hammarlund et al. (2017) noted that exclusion from higher productivity areas may have played a role in structuring Cambrian ecosystems. We argue this effect likely played an even bigger role in the Neoproterozoic, when oxygen levels are canonically believed to be far lower.

**Fig. 2** Boxplot analysis of 13,650 shale total organic carbon (TOC) weight percent measurements through Earth history. Left panel depicts the last 1000 million years of Earth history with data binned into 25 million year intervals; the Cambrian “explosion” is at ~525 Ma, the base of the Paleozoic. In boxplots, the thick horizontal lines represent median values, boxes represent first quartile and third quartile, and dashed lines extend to 1.5 times inter-quartile range. Outliers removed for figure clarity. Right panel depicts summary distributions for the Paleoproterozoic (2500–1600 Ma), Mesoproterozoic (1600–1000 Ma), Neoproterozoic (1000–541 Ma), and Phanerozoic (541–0 Ma) eras. In summary boxes, thick white lines represent means, dark colored boxes represent first and third quartiles, and light colored boxes represent 5th and 95th percentiles.
Implications of a food-limited Neoproterozoic ocean

If Neoproterozoic food supply was lower, and synergistic environmental interactions kept animals from productive regions, how did this affect early animals? In terms of body plan evolution, many morphological characteristics of early animals may have been adaptations to food-poor environments (Leys and Kahn 2018). To answer such questions at the ecosystem level, paleontologists have long looked to ecological patterns along natural environmental gradients in the modern ocean (e.g., Rhoads and Morse [1971] for oxygen). The sharpest natural gradients for almost every environmental parameter occur along continental margins (Levin and Sibuet 2012), with the change from very low to higher oxygen levels (re-capitulating the presumed Neoproterozoic–Paleozoic environmental transition) occurring at relatively shallow bathyal depths in OMZs. In contrast, the transition from very low to higher food supply (again re-capitulating the presumed ancient trend) occurs in the deep abyss, toward the center of ocean gyres. This is due to low primary productivity in the overlying waters, and very long transport distance to the seafloor, across which organic carbon rain is attenuated by progressive remineralization.

Ecological changes across these abyssal food supply gradients are interesting to consider. The frigid, stenothermal, continuously dark, and high-pressure environments of the abyss inhibit a perfect analogy for shelfal/bathyal evolution in the Neoproterozoic and Cambrian. However, the abyss is attractive as a model because so many environmental factors are invariant, and variations in chemical energy (food) can therefore be studied as a relatively isolated system (in contrast to OMZs, where many environmental parameters co-vary, Sperling et al. [2016]). Recent mega-analyses of the deep-sea have established firm links between organismal abundance, body size, animal biomass, and food supply (Smith et al. 2008; Wei et al. 2010; McClain et al. 2012). The most recent large-scale analysis of food supply gradients in the global ocean also demonstrated a positive relationship between species-level diversity in the deepsea and chemical energy (Woolley et al. 2016). This contrasts with shallow-water environments, where kinetic energy (temperature) is a better predictor of diversity (Tittensor et al. 2010; Woolley et al. 2016). The overall pattern in the deep-sea seems clear, but note questions remain regarding the exact relationship between diversity and chemical energy (for instance, linear versus quadratic [McClain et al. 2012], and whether the relationship exists in all areas and for all taxonomic groups, e.g., Glover et al. [2002]). The sediment mixed layer depth (bioturbation) is also generally shallower in areas with lower food supply or nutritional quality (Smith and Rabouille 2002; Smith et al. 2008), although more study is clearly needed (Teal et al. 2008). Unfortunately for consideration of how feeding strategies change with food supply, quantitative analysis of feeding strategies, or indeed even direct evidence of feeding mode for any given species, is relatively rare for deep-sea organisms (Gage and Tyler 1991). Qualitative analyses, though, suggest that low-food abyssal regions are (1) dominated by deposit feeders, (2) have higher occurrences of omnivory, to better utilize any available food source, and (3) have relatively few specialist carnivores, likely because prey simply is not available (Gage and Tyler 1991).

A final and intriguing lesson regarding food supply in the modern ocean concerns organismal responses to exogenous environmental stresses. Recent experimental studies have demonstrated that animals can offset the negative effects of stressors such as ocean acidification through increased food supply (Melzner et al. 2011; Hettinger et al. 2013; Sanders et al. 2013). It is more energetically costly to calcify under acidified conditions, but increased food provides energy to pump ions more vigorously and maintain the correct carbonate chemistry at the site of calcification. Growth will even increase with more food in spite of adverse conditions: over a range of external carbonate chemistry conditions, this is simply an energetics problem. Likewise, heat-shock proteins and other molecular chaperones involved in stress responses require energy for protein synthesis. Environmental instability in Neoproterozoic oceans at a variety of temporal and spatial scales is increasingly being recognized (Johnston et al. 2012b; Reinhard et al. 2016; Lu et al. 2017; Wood and Erwin 2018), and given the relationship between food supply and stress tolerance seen in modern experiments, it can be predicted that food-limited Neoproterozoic animals would have been relatively more susceptible to short-term environmental fluctuations.

Oxygen, food, and the fossil record

Studies of natural environment gradients in the modern ocean suggest that under very low oxygen levels animal communities are dominated by small size classes, are comprised of unskeletonized organisms with low-energy, non-carnivorous feeding strategies, have shallower (but often intense) bioturbation levels, and have low species-level diversity (Levin
isms that can tolerate very low O₂—can be incredibly rare. Abundances in OMZs—for the small, thin organisms—animal body sizes are depressed, and abyssal plain: early animals would be small, thin, rare. This hinges of course on the veracity and calibration of the geochemical record; it has been proposed that geochemists have fundamentally mis-understood the record (Butterfield 2018). However, if true, this environmental landscape would provide an even more plausible explanation for the discrepancy between molecular clock evidence for bilaterians in the early Ediacaran or before and the first fossil evidence for bilaterians at the end of the Ediacaran (Droser et al. 2017; Parry et al. 2017) than oxygen-related hypotheses alone.

From this (likely) small, thin, rare, and unfossilized Neoproterozoic animal biota, the end of the Ediacaran and early Cambrian record increases in body size, diversity, skeletonization, abundance, biomass, and depth of bioturbation. The early Cambrian also witnessed the rise of bilaterian predators (Sperling et al. 2013a). When environmental factors are considered as possible drivers of these Cambrian biological changes, oxygen is most often invoked. But again, when using the modern ocean space-for-time translation most of these biological changes could alternatively be explained through changing food supply (Fig. 3). Looking across natural gradients in the modern ocean, perhaps the only Ediacaran–Cambrian biological change that appears influenced by oxygen gradients but not food gradients is the appearance of biomineralized forms, as calcified echinoderms and molluscs are relatively common in the abyss (Smith et al. 2008). Note though that we are not aware of any actual studies of biomineralization in the deep sea, and the relative lack of calcified forms in the cores of OMZs could alternatively be because modern O₂ minima are CO₂ maxima (i.e., acidic) due to the remineralization of organic carbon (Paulmier et al. 2011).

An increase in nutrients and food has been proposed as a driver of marine ecosystem change in the Phanerozoic (Bambach 1999), and during the 1980s and 90s was considered as a possible player in the Cambrian radiation (Cook and Shergold 1984; Brasier 1992; Cook 1992). But these Cambrian studies were mainly concerned with trends in phosphatic skeletonization and preservation, and in any case, such hypotheses have fallen from favor. Given new observations on food supply through the Neoproterozoic–Paleozoic, and the inherent linkage between oxygen and primary productivity (Laakso and Schrag 2014; Derry 2015), they should clearly be revisited. Discussion of the ultimate cause(s) of such oxygen and/or primary productivity changes—from disparate geological triggers to ecological feedbacks—are beyond the scope of this article (see recent papers by Lenton et al. 2014; Husson and Peters 2017; Laakso and Schrag 2017; Reinhard et al. 2017; Butterfield 2018). We note though that specific causal forcings of oxygen and organic carbon dynamics are irrelevant to early animal ecosystems, aside from any additional environmental side effects.

Rather, here we simply raise the alternative hypothesis that if oxygen and food supply changed in tandem through the Ediacaran–Cambrian interval, many of the biological responses commonly ascribed to oxygen changes could equally well have been driven by food.
Considering all ingredients of an explosion

In an insightful paper on early animal evolution and the Cambrian “explosion,” Runnegar (1982b) evocatively closed by writing “the other ingredient, as in most explosives, may have been a strong oxidizing agent.” The idea that oxygen change was the key factor in the geologically sudden appearance of large, complex animals in the fossil record has increasingly dominated geological thinking and research in the subsequent decades. We do not disagree with Runnegar’s conclusion or this general line of research. Animals are clearly limited by low oxygen, and if oxygen levels rose past critical thresholds, this is an obvious trigger to explain many aspects of the Cambrian radiation. [Note such thresholds are lower than previously believed, and also that animal ecosystems respond non-linearly to oxygen changes (Sperling et al. 2015a).] But we also believe it is worthwhile to step back and consider Runnegar’s closing metaphor more broadly. Students of (literal) explosions know that oxygen is one key ingredient of an explosion, but insufficient on its own. For instance, discussing explosive forest fire events, Maclean (1992) wrote “An old-timer knows that, when a ground fire explodes into a crown fire with nothing he can see to cause it, he has not witnessed spontaneous combustion but the outer appearance of the invisible pressure of a ‘fire triangle’ suddenly in proper proportions for an explosion.”

For future study of environmental influences on early animal evolution and particularly the Cambrian
“explosion,” we suggest this “fire triangle” might be an appropriate metaphor. The classic fire triangle relates heat (flame or spark), an oxidizing agent (oxygen) and fuel (buildings, plants, trees, etc.), in the correct mixture (Fig. 4). For environmental influences on the Cambrian radiation, we might consider the first leg of the triangle to be global and regional temperature dynamics (heat). Temperature interacts synergistically with oxygen (through effects on metabolic rate) to limit aerobic habitability (Portner 2010; Deutsch et al. 2015), plays important and sometimes non-intuitive roles in oxygen solubility and diffusion (Verberk et al. 2011), and is the source of kinetic energy for ecosystems (Tittensor et al. 2010). The second leg would be oxygen, the focus of extensive geobiological research. The third leg, fuel, would represent primary productivity and labile organic carbon delivery to the seafloor. We note the “fire triangle” metaphor is not a perfect one—fires/explosions are destructive, whereas the Cambrian radiation is about the construction of animal body plans and ecosystems. Nonetheless, we believe it has conceptual utility as a current reading of the geological literature might suggest it was only the second leg (oxygen) that was important. Considering there were very large temperature changes in the Neoproterozoic (Hoffman et al. 2017), and clear changes to organic carbon abundance, type, and patchiness (Fig. 2; Brocks et al. 2017; Budd and Jensen 2017; Reinhard et al. 2017), this is unlikely to be the case. In other words, even if oxygen is ultimately demonstrated to be the most important environmental factor, we propose that consideration of a complete “fire triangle” of metabolically important environmental factors for the Cambrian “explosion” will help prevent tunnel vision that may be creeping in regarding the role of oxygen. Recognizing this broader biogeochemical framework is critical as environmental factors themselves are just one leg of an even broader causal triangle composed of the evolution of genomic/developmental factors, ecological, and environmental influences (Fig. 4). Given the fundamental linkages between biology and geology within the Earth system, the Cambrian radiation increasingly appears driven by a cascade of factors rather than a single, isolated trigger.

**Acknowledgments**

We thank Kevin Kocot, the co-organizer of this SICB symposium, and all symposium participants for a wide-ranging and enlightening discussion in San Francisco. We thank Gordon Love, Alex Zumberge, Tom Laakso, Doug Erwin, and members of the Historical Geobiology Laboratory for helpful discussion regarding this manuscript, and three anonymous reviewers for helpful comments. For the TOC analysis through time, we thank Jon Husson for providing a copy of compiled USGS data published in Husson and Peters (2017). We thank Francis Macdonald, Justin Strauss, Emily Smith, Tiffani Fraser, Samantha Ritzier, Phil Wilby, Don Canfield and Guy Narbonne for help collecting samples, Sabrina Tecklenburg, Austin Miller, Una Farrell, Thomas Boag, Stephanie Plaza-Torres, Paulo Linarde, Liam Bhajan, Austin Patch, Junyao Kang, Will Thompson-Butler, and Samantha Ritzier for help running samples, and David Mucciarone and Doug Turner for technical assistance.

**Funding**

This work was supported by the Alfred P. Sloan Foundation and National Science Foundation [grant number DEB-1747731].

**Supplementary data**

Supplementary data available at ICB online.

**References**


Leys SP, Kahn A. 2018. Oxygen and the energetic requirements of the first multicellular animals. Integr Comp Biol published online (doi: 10.1093/icb/icy051).


Sanders MB, Bean TP, Hutchinson TH, Le Quesne WJF. 2013. Juvenile King Scallop, Pecten maximus, is potentially tolerant to low levels of ocean acidification when food is unrestricted. PLoS One 8:e74118.


