Received: 23 December 2020 | Revised: 14 October 2021 | Accepted: 26 October 2021

DOI: 10.1111/gbi.12481

gebiology WILEY

ORIGINAL ARTICLE

Molecular dating of the blood pigment hemocyanin provides new insight into the origin of animals

Elisa Maria Costa-Paiva[1,2] | Beatriz Mello[3] | Bruno Santos Bezerra[1] | Christopher J. Coates[4] | Kenneth M. Halanych[5] | Federico Brown[1] | Juliana de Moraes Leme[6] | Ricardo I. F. Trindade[2]

[1]Zoology Department, Institute of Biosciences, University of Sao Paulo, Sao Paulo, Brazil

[2]Geophysics and Atmospheric Sciences, Institute of Astronomy, University of Sao Paulo, Sao Paulo, Brazil

[3]Genetics Department, Biology Institute, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

[4]Department of Biosciences, Faculty of Science and Engineering, Swansea University, Swansea, UK

[5]Center for Marine Science, University of North Carolina Wilmington, Wilmington, NC, USA

[6]Geoscience Institute, University of Sao Paulo, Sao Paulo, Brazil

Correspondence
Elisa Maria Costa-Paiva, Zoology Department, Institute of Biosciences, University of Sao Paulo, Brazil.
Email: elisam.costapaiva@gmail.com

Funding information
National Science Foundation, Grant/Award Number: DEB-1036537 and OCE-1155188; Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Grant/Award Number: 2018/20268-1, 2016/06114-6 and 2015/50164-5

Abstract

The Neoproterozoic included changes in oceanic redox conditions, the configuration of continents and climate, extreme ice ages (Sturtian and Marinoan), and the rise of complex life forms. A much-debated topic in geobiology concerns the influence of atmospheric oxygenation on Earth and the origin and diversification of animal lineages, with the most widely popularized hypotheses relying on causal links between oxygen levels and the rise of animals. The vast majority of extant animals use aerobic metabolism for growth and homeostasis; hence, the binding and transportation of oxygen represent a vital physiological task. Considering the blood pigment hemocyanin (Hc) is present in sponges and ctenophores, and likely to be present in the common ancestor of animals, we investigated the evolution and date of Hc emergence using bioinformatics approaches on both transcriptomic and genomic data. Bayesian molecular dating suggested that the ancestral animal Hc gene arose approximately 881 Ma during the Tonian Period (1000–720 Ma), prior to the extreme glaciation events of the Cryogenian Period (720–635 Ma). This result is corroborated by a recently discovered fossil of a putative sponge ~890 Ma and modern molecular dating for the origin of metazoans of ~1,000–650 Ma (but does contradict previous inferences regarding the origin of Hc ~700–600 Ma). Our data reveal that crown-group animals already possessed hemocyanin-like blood pigments, which may have enhanced the oxygen-carrying capacity of these animals in hypoxic environments at that time or acted in the transport of hormones, detoxification of heavy metals, and immunity pathways.

KEYWORDS
metazoa emergence, multifunctional, Neoproterozoic, Oxygen-binding proteins

1 | INTRODUCTION

The emergence of animals and timing of divergences among early metazoan lineages are crucial to understanding the processes of biological evolution itself and the causative links between environmental changes and biological innovation (Mills et al., 2018; dos Reis et al., 2015b). The Earth's biosphere had experienced profound changes by the end of the Proterozoic Eon: from a unicellular world marked by deep-water anoxia arose a multicellular world with complex life forms accompanied by major changes in the environment, *for example*, the oxygenation of Earth's surface (Erwin et al., 2011; Knoll, 2011; Raff & Raff, 1970; Xiao et al., 2014). Although most extant animals use

aerobic metabolism for homeostasis and growth, much of the molecular toolkit was present in the closest relatives of animals, which may have originated during low-oxygen periods (Jabłońska & Tawfik, 2021; Sebé-Pedrós et al., 2011). The need for oxygen in animals is mistakenly associated with respiratory function exclusively—it is now well-known that oxygen is also required for collagen synthesis, wound healing, and some immune functions (Coates & Decker, 2017; Mills & Canfield, 2014; Schreml et al., 2010). Considering the close relationship between animals and oxygen, low-oxygen availability was hypothesized to have prevented the origin of animals until the late Neoproterozoic (Nursall, 1959). Increased dioxygen availability is frequently attributed as a main trigger for animal evolution coupled with aerobic metabolism (Mills & Canfield, 2014). Recent studies have challenged this canonical view that the origin of animals was controlled primarily by atmospheric oxygen levels (Mills, Francis, Vargas, et al., 2018; Mills et al., 2014; Sperling et al., 2013, 2015).

Early animals were likely small, soft-bodied, and collagen-limited (although not necessarily collagen-free) organisms that lived under low oxygen levels, restricting their use of oxygen to high-priority physiological functions (Mills & Canfield, 2014; Towe, 1970). Nevertheless, the diversification of lineages prompted increases in morphological, physiological, and ecological complexities. Simple oxygen diffusion became inefficient to sustain the animal's metabolic needs leading to the evolution of efficient circulatory systems and oxygen-binding/transport proteins that provided significant advantages (Burmester, 2001; Raff & Raff, 1970; Schmidt-Rhaesa, 2007). These carrier molecules are proteins that likely originated from enzymes whose primary function would be to protect the organism from dioxygen toxicity, having acquired the potential for molecule transport later (Terwilliger, 1998). The evolutionary history of oxygen-binding proteins and early metazoan metabolic demands have been intertwined over millions of years, and understanding their evolution may provide valuable information about the origin and diversification of animals.

Oxygen carrier proteins are biological macromolecules that can reversibly bind molecular dioxygen—often referred to as respiratory or blood pigments, since they tend to exhibit color when bound to oxygen (Coates & Nairn, 2014; Terwilliger, 1998). They are divided into three chemical categories: hemoglobins, hemerythrins, and hemocyanins (Terwilliger, 1998). Hemocyanins (Hc)—macromolecules of focus here—are large, extracellular glycoproteins found extensively among arthropods and mollusks (Burmester, 2002, 2015; Coates & Decker, 2017). Recent evidence has also demonstrated the presence of Hcs and Hc-like genes/proteins in hemichordates, tunicates, sponges, ctenophores, and annelids (Aguilera et al., 2013; Costa-Paiva et al., 2018; Immesberger & Burmester, 2004; Martín-Durán et al., 2013). Arthropod and mollusk Hcs are so named due to the presence of a conserved Type III dicupric active site, yet they emerged independently from an existing copper protein called tyrosinase (or phenoloxidase) (Burmester, 2001, 2015; van Holde et al., 2001; Terwilliger, 1998).

The genealogy of the Hc superfamily, as well as the evolutionary changes associated with the emergence of those different proteins, cannot be understood without considering animal phylogeny and divergence times among its lineages. Reconstruction of the last common ancestor of animals is problematic, partly because of the continued challenges in recovering early animal relationships (Giribet, 2016; Halanych, 2016; Pisani et al., 2015; Whelan et al., 2015) and because recovering early putative animal fossil records is challenging (Budd, 2008; Budd & Jensen, 2000). Fossil records of early animal life have generated considerable controversy over the years, especially when they conflict with timings based on molecular clock estimates (Budd & Mann, 2020a, 2020b). Current estimates for molecular origins for crown-group Metazoa range from 1,000 Ma to 615 Ma (Dohrmann & Wörheide, 2017; Peterson et al., 2004; Qun et al., 2007; dos Reis et al., 2015). From a biological perspective, the fossil record provides the only direct insight into evolutionary history (Wood et al., 2020); however, with recent advancements of molecular clock methodologies, estimates of divergence of major animal lineages are becoming more accurate, and the disparity between molecular dating and fossil evidence of clade age minima has reduced (dos Reis et al., 2015).

Molecular phylogenetic methods have revolutionized our knowledge about protein evolution and function, as well as the evolutionary history of taxa (Pagel et al., 1999; Perron et al., 2019; Swofford et al., 1996). Molecular dating, which is an age estimation of internal nodes based on molecular sequences, is now a standard approach and can be used successfully for deep time studies, helping to elucidate the diversification of major taxa and their association with Earth's history (e.g., Delsuc et al., 2018; Irisarri et al., 2017; Marin et al., 2016; Misof et al., 2014; Morris et al., 2018; Varga et al., 2019; Wolfe et al., 2019). Besides its application to infer the age of biological lineages, the inference of divergence times based on molecular data can be used to estimate the split times between homologous gene and protein sequences (Bezerra et al., 2021; Boden et al., 2021; Shih & Matzke, 2013; Yu & Li, 2014). Protein functions and adaptations at the molecular level cannot be understood without considering species phylogeny. In fact, the proteins of an organism frequently share its phylogenetic history, and physiological adaptations that have evolved in the organism can be recapitulated by changes in protein sequences (Burmester, 2002). Thus, dating specific genes has the potential to shed new light on pervasive issues, such as, the origin of animals.

Although deep divergence time studies can incorporate hundreds of genes to estimate divergence times of species lineages (Dohrmann & Wörheide, 2017; dos Reis et al., 2015), molecular dating of specific proteins can recover the evolutionary history of these proteins against a background of the evolution of the major taxa in which they are found/lost. Molecular dating of deep divergences may be challenging, mostly because of issues such as sequence saturation, which can affect analyses by biasing the estimated genetic distances (Magallón et al., 2013; Schwartz & Muller, 2010; Wilke et al., 2009; Zheng et al., 2011). However, estimated divergence times based on amino acid sequences that are more conserved compared with nucleotides sequences can alleviate the problem of saturation.

Considering the importance for animal physiology and deep divergence times of oxygen-binding proteins, the fact that only few studies have systematically addressed dating in the evolutionary history of these proteins is surprising (Burmester, 2001, 2002;

Prothmann et al., 2020). The recent discovery of Hc-like genes in early diverging lineages of Metazoa, including sponges and ctenophores (Costa-Paiva et al., 2018), suggests that these proteins were already present in the last common ancestor of animals. Here, we set-out to date the origin of the Hc superfamily using animal transcriptomic and genomic data. We have taken a comparative phylogenetic approach to access the evolutionary history of the Hc superfamily and Bayesian dating to infer its emergence. Our results are further contextualized with the major environmental changes that happened across the Neoproterozoic Era.

2 | MATERIAL AND METHODS

Arthropod Hcs are members of a protein superfamily that also includes (a) arthropod phenoloxidases (POs) whose functions include sclerotization of the cuticle, wound healing, and innate immunity (Whitten & Coates, 2017); (b) hexamerins (HEX), proteins present in insects that do not bind oxygen but are considered storage proteins associated with molting or nutrition (Burmester, 1999a); (c) decapod pseudo-hemocyanins or cryptocyanins (pHc) that are similar to Hcs but appear to act as storage proteins in the hemolymph (Burmester, 1999b); and (d) hexamerin receptors that are present in dipteran insects and are related to their own ligands (Burmester & Schellen, 1996). Although these proteins form a functionally diverse superfamily, their sequences present highly conserved core elements that allow their evolutionary history to be traced (Burmester, 2001; Costa-Paiva et al., 2018).

2.1 | Dataset assembly and alignment

The Hc dataset was formed using 108 previously published Hc superfamily sequences distributed as: (a) 60 Hc sequences (including Hc-like); (b) 34 POs; (c) 11 HEXs; and (d) 3 pHc (Aguilera et al., 2013; Burmester, 2001; Costa-Paiva et al., 2018; Martín-Durán et al., 2013) (Figure 1). Protein sequences with their respective accession numbers from NCBI are presented in Table 1.

In order to infer homology between amino acid positions, datasets were compiled and aligned with MAFFT using the accurate "E-INS-i" algorithm (Katoh & Standley, 2013). The completed alignment was trimmed using trimAl (Capella-Gutiérrez et al., 2009) with a 50% gap threshold to eliminate poorly aligned regions and used for all subsequent analyses (File S1–S2).

2.2 | Phylogenetic reconstructions

The LG+C40+F+Γ4 mixture model, the best-fit model of protein evolution for the dataset, was selected using ModelFinder, a software implemented in the IQ-TREE software (Kalyaanamoorthy et al., 2017), which uses Akaike and Bayesian Information Criteria methods (AIC and BIC, respectively). IQ-TREE was also used to perform a maximum likelihood inference (Nguyen et al., 2015), with branch supports obtained by the ultrafast bootstrap approximation with 1,000 replicates (Hoang et al., 2018). The tree was rooted using two amoebozoan homologue sequences (File S3).

2.3 | Molecular dating

Molecular dating was performed in PhyloBayes (Lartillot & Philippe, 2004) using a mixture model, and the phylogenetic tree was inferred by IQ-TREE. Estimation of divergence times was performed with the LG+C40 model using a gamma distribution (Γ4) of site-rate heterogeneity and a birth-death prior on divergence times (File S4). We inferred divergence times using both the log-normal autocorrelated relaxed clock (-ln) and the uncorrelated gamma relaxed clock (-ugam). MCMC (Markov Chain Monte Carlo) was run for 36,000 cycles and a burn-in period of 10%. Convergence of chains was accessed by running two independent MCMC runs. In both runs, ESS (effective sample sizes) values were higher than 200, after discarding the burn-in period.

To calibrate divergence times, we first identified duplication and speciation nodes with the gene duplication wizard tool in MEGA 7 (Kumar et al., 2016). This was performed because calibration information derived from fossil data provides information regarding the split times between biological lineages (*i.e.*, speciation events). So, divergences classified as speciation nodes that reflected robust biological clades and were free of duplication events were chosen for calibration. This search for gene duplications implemented the algorithm described in Zmasek and Eddy (2001) to infer gene duplications and speciation events for all internal nodes in the gene tree. The algorithm assumed that the gene tree and species tree are both properly rooted and biologically correct.

We used four calibration nodes according to best practice recommendations (Parham et al., 2012): Annelida, Arthropoda, Pancrustacea, and Lobopodia (Arthropoda + Onychophora). As PhyloBayes requires a root calibration, we assigned flexible boundaries to the divergence between amoebozoans and the ingroup, which followed the maximum dates for Eukarya MRCA (2,400 Ma) reported on the TimeTree database (Kumar et al., 2017) and the minimum date was based on the oldest fossil remains of acritarchs that can be ascribed with certainty to total-group Eukaryota (1,619 Ma) from the Changcheng Formation, North China (Lamb et al., 2009). The fossil structures do not indicate membership of any specific crown eukaryote clade, only to use these records to minimally constrain the timing of divergence between the Eukaryota and their archaebacterial sister lineage, Asgardarchaeota (Betts et al., 2018).

To estimate the tMRCA (time to the most recent common ancestor) of the Annelida, clade was assigned boundaries of 476 Ma and 636 Ma (Benton et al., 2015). This constraint was based on the maximum age interpretation of the Lantian Biota (Yuan et al., 2011). The tMRCA of crown arthropods was calibrated with a minimum value of 514 Ma and a maximum value of 636 Ma (Benton et al., 2015) based on the fossil *Yicaris dianensis* (Zhang et al., 2007). Thus, a minimum

TABLE 1 List of taxa and genes analyzed with their respective NCBI accession numbers

Taxon	Gene identification	Accession number
METAZOA		
Porifera		
Amphimedon queenslandica	PPO	XP_003390261.1
Kirkpatrickia variolosa	Hc	MF998096
Latrunculia apicalis	Hc	MF998097
Ctenophora		
Coeloplana astericola	Hc	MF998091
Mnemiopsis leidyi	Hc	MF998101
	Hc	ML0447
	Hc	ML0463
	Hc	ML0910
Pleurobrachia bachei	Hc	MF998107
Hemichordata		
Harrimaniidae gen sp. (from Iceland)	Hc	MF998095
Saccoglossus kowalevski	Hc	ACY92544
	Hc	XP002734027.1
Chordata		
Ciona intestinales	Tyr	XP002128449.1
	Tyr	XP002119145.1
	Tyr	NP001029009.1
	Hc	XP002119145
	Hc	XP002128449
Annelida		
Paramphinome jeffreysii	Hc	MF998102
Pista macrolobata	Hc	MF998106
Streblosoma hartmanae	Hc	MF998109
Terebellides stroemii	Hc	MF998112
Thelepus crispus	Hc	MF998113
Onychophora		
Epiperipatus sp.	Hc	Q8MPM7
	Hc	CAD12808.1
Arthropoda		
Chelicerata		
Androctonus australis	Hc	P80476
Aphonopelma sp.	HcA	P14750
	HcB	Q9NFH9
	HcG	Q9NFL4
	HcF	Q9NFL5
	HcC	Q9NFL6
	HcD	P02241
	HcE	P02242
Cupiennius salei	Hc	CAC44749
	Hc	CAC44750
	Hc	CAC44751
	Hc	CAC44752
	Hc	CAC44753
	Hc	CAC44755
Euphrynichus bacillifer	HcA	CCA94920.1
	HcB	CCA94921.1
Limulus polyphemus	HcA	CAJ91099.1
	HcB	CAJ91100.1
	Hc	P04253
Mastigoproctus giganteus	HcA	CCA94927.1
	HcB	CCA94928.1
Myriapoda		
Scutigera coleoptrata	HcA	CAC69246
	HcB	CAD55132
	HcC	CAD24086
	HcD	CAC69247
	HcX	CAD24085
Spirostreptus sp.	Hc	CAC33894
Pancrustacea		
Anopheles gambiae	Tyr	XP307623.1
	Tyr	XP312089.2
	Tyr	XP315073.2
	Tyr	XP315074.1
	Tyr	XP315075.1
	Tyr	XP315076.1
	Tyr	XP315083.1
	Tyr	XP315084.2
	Tyr	XP316323.2
Apriona germani	Hex	AAM44045
Blaberus discoidalis	Hex	AAA74579
Bombyx mori	PPO	BAA08368
	Tyr	BAA08368.1
	Tyr	BAA08369.1
Callinectes sapidus	Hc	AAF64305
Cancer magister	Hc	AAA96966
Daphnia pulex	Tyr	GW183100.1
Drosophila melanogaster	Hex1	NP476624
	Hex1	NP511138
	Hex1	NP523868
	Hex2	NP524816
	PPO	NP476812
	PPO	NP524760
	PPO	NP610443
	Tyr	NP476812.1
	Tyr	NP524760.1
	Tyr	NP610443.1
Galleria mellonella	PPO	AAK64363
Homarus americanus	Hc	CAB75960
	pHc	CAB38042
	pHc	CAB38043
Homarus gammarus	PPO	Q70GP3
Manduca sexta	PPO1	AAC05796
	PPO2	AAC37243
Marsupenaeus japonicus	PPO	BAB83773
Metacarcinus magister	pHc	AAD09762

(Continues)

TABLE 1 (Continued)

Taxon	Gene identification	Accession number
Neobellieria bullata	PPO	AAD45526
	PPO	AAD45527
Pacifastacus leniusculus	Hc	AAM81357
	PPO	CAA58471
Palinurus vulgaris	Hc	CAC69243
	Hc	CAC69244
	Hc	CAC69245
Panaeus vannamei	Hc	CAA57880
Penaeus monodon	PPO	AAD45201
Penaeus semisulcatus	PPO	AAM77690
Periplaneta americana	Hex	AAB09629
Plodia interpunctella	Hex	AAK71136
Schistocerca americana	Hc	AAC16760
Spodoptera litura	HexA	CAB55603
	HexB	CAB55602
Tenebrio molitor	Hex	AAK77560
	PPO	BAA75470
Tribolium castaneum	Tyr	NP001034493.1

constraint was established on the age of the top of the Nangaoan Stage of the Qiandongian Series of the Cambrian of China, which has been dated to 514 Ma (Peng & Babcock, 2008; Peng et al., 2012) and a soft maximum constraint was based on the maximum age interpretation of the Lantian Biota (Yuan et al., 2011).

For Pancrustacea, the aforementioned requirements were met twice (*i.e.*, two speciation nodes that included only Pancrustacea sequences were recovered twice in the estimated phylogeny). Because of that, four speciation nodes were calibrated with uniform distributions and lower/upper boundaries based on dos Reis et al. (2015). The time range used to calibrate the tMRCA (of pancrustaceans was between 514 and 531 Ma (dos Reis et al., 2015). Lastly, the tMRCA of lobopodians was a minimum of 528 Ma and a maximum of 636 Ma (Benton et al., 2015) based on Rusophycus trace fossils that are widely accepted to have been produced by arthropod-grade organisms (Budd & Jensen, 2000). Rusophycus occurs well below the first animal body fossils in Cambrian sections around the world (Crimes & Jiang, 1986; MacNaughton & Narbonne, 1999; Weber & Zhu, 2003). A soft maximum constraint is based on the maximum age interpretation of the Lantian Biota (Yuan et al., 2011).

3 | RESULTS & DISCUSSION

Molecular dating with PhyloBayes inferred the existence of a last common ancestral hemocyanin sequence with a median at 881 Ma, indicating an emergence of animals in the Tonian, prior to the extreme glaciation events .

3.1 | Non-arthropod hemocyanin and hemocyanin-like proteins

Maximum likelihood and Bayesian inference analysis revealed two highly supported clades: 1) a clade formed by hemichordate, tunicate, and sponge Hcs, Hc-like, and PO sequences; 2) a clade formed by ctenophore, annelid, and panarthropod sequences (Figure 2). The last common ancestor (LCA) of the first clade was originated at approximately 633 Ma (range 846–492 Ma; Tonian - Cambrian Period) (Figure 3), while the LCA of the second had its inferred ages centered at 737 Ma (831 – 679 Ma; Tonian - Cryogenian Period).

A monophyletic sponge clade of Hcs (Figure 2, dark yellow clade, bs = 100%) was the sister taxon to a deuterostome clade of PO and Hcs (Figure 2, purple clade, bs = 100%). The deuterostome clade presented a clear distinction between hemichordate Hcs and tunicate sequences, which included Hcs and PO sequences. The tMRCA estimates for sponges and deuterostome Hcs were 420 Ma (607–252 Ma) and 362 Ma (541–215 Ma), respectively (Figure 3, dark yellow and purple clade, respectively). The temporal mismatch observed for sponges and for deuterostomes may have occurred due to a limited representation of Hcs available for these specific lineages. It is clear that sponges are much older than deuterostomes (dos Reis et al., 2015), evidence that include a recently discovered putative keratose sponge about 890 Ma (Turner, 2021).

Regarding the ctenophore, annelid, and panarthropod clade, our results demonstrated the presence of a monophyletic ctenophore Hc clade (Figure 2, dark blue, bs = 100%), a annelid Hc sequences (Figure 2, light yellow clade, bs >95%), and one composed by panarthropod sequences. The first clade contains ctenophore representatives exclusively with an estimated time of the emergence of ctenophore Hcs about 558 Ma (633 – 430 Ma), which average is centered in the Ediacaran Period. As ctenophores are soft-bodied animals, they are sparsely represented in the rock record with mostly species restricted to Cambrian Burgess Shale-type deposits (Parry et al., 2021).

The date of the annelid/panarthropod split (a.k.a. the lophotrochozoan/ecdysozoan split) was estimated to have occurred about 674 Ma (713–646 Ma) at the Cryogenian period, with the origin of annelid Hcs at approximately 574 Ma (619–498 Ma) during the Ediacaran Period (Figure 3). Although the dating estimates of early evolution of annelids remains obscure or controversial—mostly due to a discordance between molecular phylogenies and fossils (Chen et al., 2020; Eibye-Jacobsen & Vinther, 2012; Parry et al., 2015)—our results agree with the oldest annelid fossil record, *that is*, a bristle worm that unambiguously belongs to crown annelids from the Canglangpu formation, which was dated to the early Cambrian (Chen et al., 2020).

3.2 | Panarthropod hemocyanin superfamily

Hcs, pHcs, PPOs, and HEXs together form a functionally diverse protein superfamily, where most sequences and core structural elements are strikingly conserved. These core elements allow tracing the evolutionary history of this protein superfamily (Burmester, 2001). The

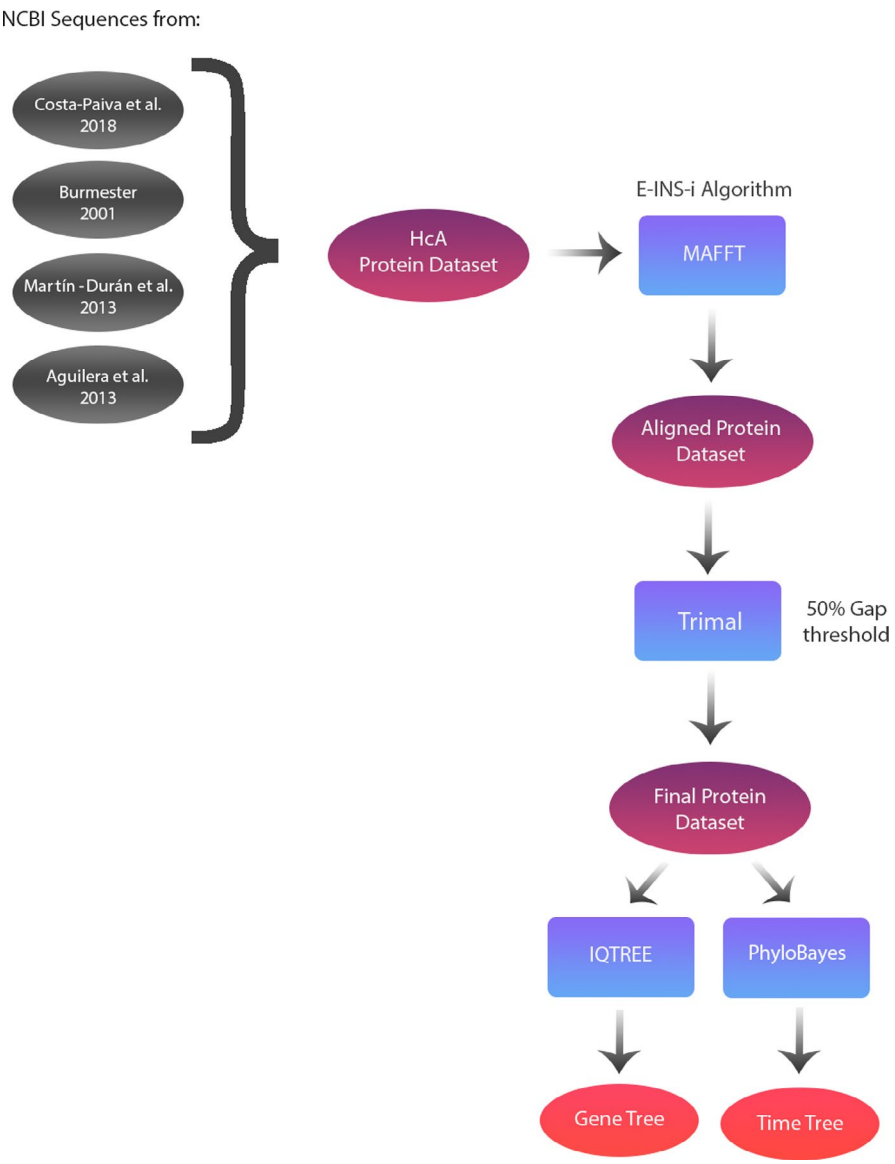

FIGURE 1 Bioinformatics pipeline. Rounded rectangles represent software or scripts and ovals represent input/output files. The dataset was formed using 108 previously published hemocyanin superfamily sequences distributed as: 60 hemocyanin sequences; 34 phenoloxidases; 11 hexamerins; 3 pseudo-hemocyanin

emergence of a panarthropod Hc superfamily was dated to 633 Ma (636–625 Ma) during the Ediacaran Period (Figure 3)—suggesting that this blood pigment was present in the most recent common ancestor of extant arthropods (Burmester, 2001, 2015; van Holde & Miller, 1995; Markl & Decker, 1992). Our findings corroborate previous inferences about arthropod Hc origins, around 700 – 600 Ma, based also on molecular dating (Burmester, 2001, 2002).

Within the panarthropod clade, we recovered an onychophoran Hc clade (Figure 2, red clade, bs =100%) and a clade formed by arthropod sequences with two main gene lineages. One of the lineages comprised pancrustacean, chelicerate, and myriapod Hcs Hc, pHcs, and HEXs (Figure 2, gray, orange, light green, and light blue clades), while the other is composed of pancrustacean POs (Figure 2, pink clade). Divergence between these two gene lineages was dated at approximately 621 (630 – 609 Ma), thereby indicating a likely origin during the Ediacaran Period (Figure 3). The origin of the well-supported clade formed by pancrustacean POs (Figure 2, pink clade, bs = 100%) was centered at 586 Ma (604–567 Ma) in the Ediacaran Period (Figure 3, pink clade).

Within the monophyletic clade composed by pancrustacean, chelicerate, and myriapod Hcs, pHcs and HEX, significant differentiation was recovered between chelicerate and myriapod Hcs (Figure 2, gray and orange, bs > 95%) and pancrustacean, Hcs, pHcs, and HEXs (Figure 2, light blue and light green clade, bs = 100%). The well-supported clade formed by chelicerate and myriapods Hcs was divided into two maximally supported clades: (1) chelicerate Hc sequences (Figure 2, gray clade, bs =100%) and (2) myriapod Hc sequences (Figure 2, orange clade, bs = 100%). The date of the pancrustacean Hc/myriapod and chelicerate Hc split was estimated to have occurred about 612 Ma (624–599 Ma), in the Ediacaran Period. This estimate agrees with a previous estimate for Hcs that dated it around 600 Ma (Burmester, 2001) and corroborates previous studies that suggest this divergence happened before the radiation of arthropod subphyla, which occurred no later than in the Cambrian period (Burmester, 2002; Conway-Morris, 1993; Gu, 1998; Valentine et al., 1999). Nevertheless, new evidence suggests a more recent slip between pancrustaceans and myriapods (Lozano-Fernandez et al., 2016). Our time estimate

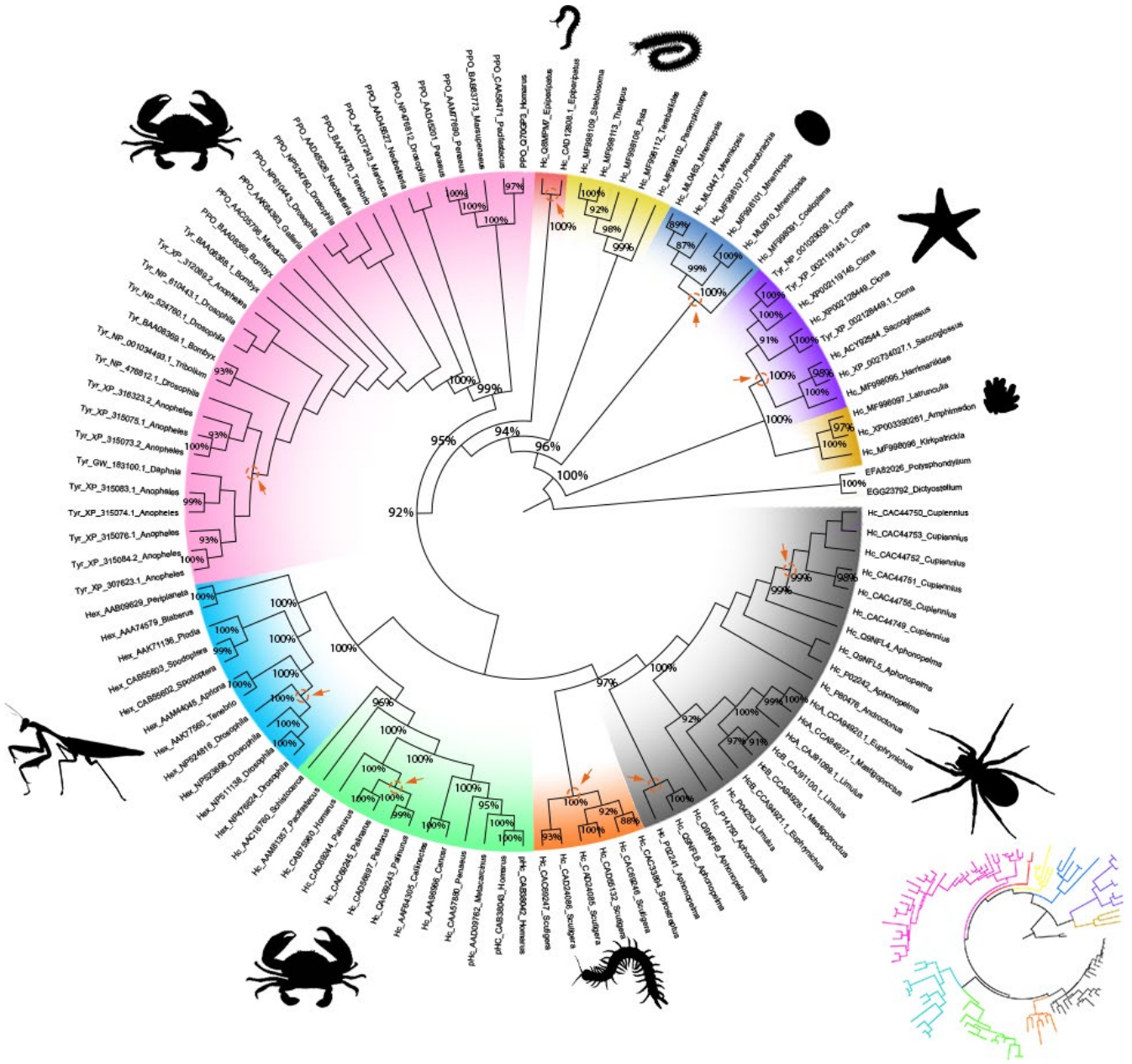

FIGURE 2 ML tree for the hemocyanin gene superfamily rooted with two amoebozoan Hc sequences. (A) Purple clade is formed by deuterostome Hcs and POs; (B) Dark yellow clade is sponge Hc-like; (C) Dark blue clade is ctenophore Hc-like; (D) Light yellow clade is formed by annelid Hc-like; (E) Red clade is onychophoran Hcs; (F) Pink clade is panarthropod phenoloxidases; (G) Light blue clade is hexapod hexamerins; (H) Light green clade is pancrustacea Hcs and pHcs; (I) Orange clade is myriapod Hcs; (J) Gray clade is chelicerate Hcs; (K) Colorless clade is the outgroup. Doted circles and arrows indicate gene duplication events. The number after the protein abbreviation in each sequence indicates the GenBank accession number for each gene. Only bootstrap support values over 80% are indicated

for the myriapod/chelicerate Hc divergence was dated at 584 Ma (604–559 Ma) (Figure 3, orange and gray clades).

Pancrustacean, Hcs, pHcs, and HEXs clade slip in two highly supported clades: (1) pancrustacea Hcs and pHcs sequences (Figure 2; light green clade, bs >95%) and (2) insect HEXs (Figure 2, light blue clade, bs =100%). Our results suggested an origin for both insect HEXs and pancrustacean Hcs and pHcs during the Cambrian Period at 516 Ma (524–514 Ma) and 511 (530 – 485 Ma), respectively (Figure 3). Our results corroborated previous findings by Burmester (2002), showing a close relationship between crustacean and insect Hc genes. The phylogenetic position of insect HEXs suggests that this copper-less storage proteins evolved within the insect stem lineage around 516 Ma ago, contradicting previous findings, which proposed an origin around 400 Ma (Burmester, 2002). Hexapods are derived from aquatic crustaceans, yet the timing of this event remains controversial (Burmester, 2015). The first terrestrial hexapod fossils were dated from the early Devonian period approximately 400 Ma (Kenrick et al., 2012; Misof et al., 2014), suggesting a Silurian origin around 450 Ma.

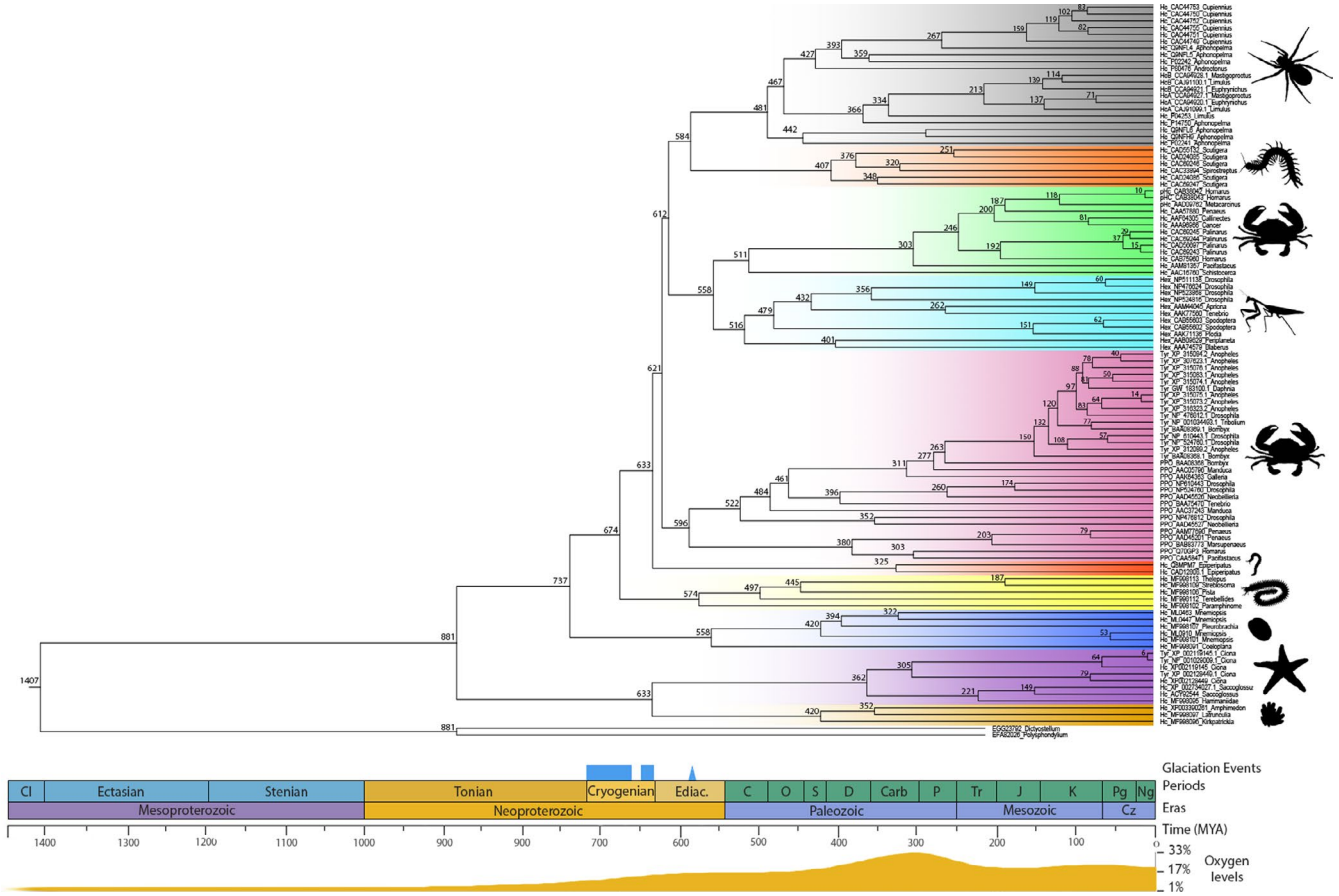

FIGURE 3 Hemocyanin gene superfamily tree with time estimates. (A) Purple clade is formed by deuterostome Hcs and POs; (B) Orange clade is sponge Hc-like; (C) Dark blue clade is ctenophore Hc-like; (D) Yellow clade is formed by annelid Hc-like; (E) Red clade is onychophoran Hcs; (F) Dark green clade is myriapod Hcs; (G) Gray clade is chelicerate Hcs; (H) Pink clade is panarthropod phenoloxidases; (I) Light green clade is pancrustacean Hcs and pHcs; (J) Light blue clade is hexapod hexamerins sequences. The number after the protein abbreviation in each sequence indicates the GenBank accession number for each gene. Average node ages are plotted. C, Cambrian; O, Ordovician; S, Silurian; D, Devonian; Carb, Carboniferous, P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; Pg, Paleogene; Ng, Neogene; Cz, Cenozoic; Ma, Million years ago. Glaciation events and atmospheric oxygen levels estimates are indicated

3.3 | Emergence of metazoan hemocyanins

A much-debated topic in geobiology refers to the influence of atmospheric oxygenation on Earth and the origin and diversification of animal lineages. All extant animals require oxygen for at least a fraction of their life cycle, suggesting that life cycle completion in total anoxia is either incompatible with metazoan ecology and physiology, or an extremely rare and derived metazoan trait (Cole et al., 2020). However, some extant animals can tolerate, and live in, low-oxygen environments (Sperling, Halverson, et al., 2013; Sperling et al., 2015). Although oxygen requirements of early animals are not fully understood, there are multiple theoretical estimates of oxygen consumption in animals, which depend primarily on the organism's length, width, and possession of a vascular system with oxygen-carrying proteins, such as, hemocyanins (Mills et al., 2014; Sperling, Halverson, et al., 2013).

The estimated age for the origin of animal Hc superfamily was approximately 881 Ma (1117–756 Ma) in the Tonian Period during the Neoproterozoic Era (Figure 3). The first members of this family were likely to have emerged before metazoans, as there are incomplete Hc homologues in amoebozoans—only two of the three protein domains are present (Martín-Durán et al., 2013). In animals, the first Hcs were likely derived from a phenoloxidase-like enzyme, as previously proposed by Burmester (2002) for the arthropod stemline. Phenoloxidases play an important role in the initial stages of the melanization process, where they catalyze the hydroxylation of monophenols (e.g., L-tyrosine) to *ortho*-diphenols and the oxidation of *ortho*-diphenols to *ortho*-quinones, which eventually go on to form melanins (Burmester, 2002). Phenoloxidases and melanins act in the front line of innate immunity of many aquatic and terrestrial invertebrates, contribute to clot sealing during wound healing, and participate in the sclerotization of the arthropod cuticle (Ashida & Yoshida, 1988; Àspan & Söderhäll, 1991; Söderhäll & Cerenius, 1998; Whitten & Coates, 2017). Thus, it is likely that the emergence of POs was directly related to a detoxification or defense response in early animals in the Neoproterozoic. In arthropods, it is conceivable that

this enzyme was also linked to the evolution of hardened exoskeletons in the late Precambrian period (Burmester, 2002). Moreover, the high alkalinity of seawater at the end of Ediacaran (Xiao et al., 2016) could have triggered biomineralization (Cui et al., 2016, 2019; Wood et al., 2017).

The emergence of the Hc superfamily during the Neoproterozoic Era corroborates modern molecular dating of an age between 1,000 and 650 Ma for the origin of metazoans (Dohrmann & Wörheide, 2017; Erwin et al., 2011; Lozano-Fernandez et al., 2017; Peterson et al., 2004; Qun et al., 2007; dos Reis et al., 2015). In addition, these results suggest that early-branching animals may have already possessed blood pigments (Hc-like), which may have enhanced their respiratory capacity in a hypoxic environment at that time. In animals without circulatory systems, Hcs might act in other cellular processes beyond oxygen loading. Early Hc and Hc-like proteins may have also acted in the transport of hormones, detoxification of heavy metals, and innate immunity (Coates & Costa-Paiva, 2020; Coates et al., 2011, 2013; Coates & Nairn, 2014; Coates & Talbot, 2018).

The Neoproterozoic Era (1000–541 Ma) was characterized by significant modifications on the planet dynamics, including massive lithosphere alteration (Trindade et al., 2006); extreme glaciation events (Rooney et al., 2015; Spence et al., 2016); and the rise of atmospheric oxygen levels (Sperling et al., 2015; Tostevin & Mills, 2020). At least three great Neoproterozoic glaciations occurred, the extreme Sturtian (720–660 Ma) and Marinoan (650–636 Ma) ice ages and the more restricted, Gaskiers glaciation (582 Ma) (Cordani et al., 2020; Rooney et al., 2015; Spence et al., 2016). Our results indicate that the emergence of animal Hcs occurred prior to these extreme glaciation events that could suggest the presence of animals by this time. This assumption is corroborated by molecular estimates (Dohrmann & Wörheide, 2017; dos Reis et al., 2015), the presence of a putative sponge fossil from 890 Ma (Turner, 2021) and diagnostic of pre-Marinoan (<650–635 Ma) demosponges (Love et al., 2009; Love & Summons, 2015), and possible cryostane demosponges biomarkers indigenous to bitumens and oils compatible with pre-Sturtian metazoans (<800–740 Ma) (Brocks et al., 2016). However, new evidence calls into question the veracity of biomarkers from the first metazoans, drawing attention to the fact that they may be algal biomarkers, which are common during the Neoproterozoic (Bobrovskiy et al., 2021; van Maldegem et al., 2021).

During extreme glaciations, animals could live on the underside of the ice or on the sediment under the ice sheet, although a variety of biotic refugia during the Sturtian and Marinoan have been identified, including marine and terrestrial hydrothermal vents (Costas et al., 2008; Fraser et al., 2014), sea-ice brine channels within ice grounding-line crack systems (Thomas & Dieckmann, 2002), and cryoconite (Christner et al., 2003; Hoffman, 2016). As suggested by the date of the arthropod Hcs origin calculated here (around 584 Ma), arthropods may have evolved immediately before Gaskiers glaciation (Figure 3). This result is consistent with the evolution of novel and more metabolically demanding traits, such as sclerotization and higher motility during the emergence of more complex food webs (Sperling, Frieder, et al., 2013; Wood et al., 2019).

Regarding oxygenation events on Earth, as early as 3.0 Gya ago, a dynamic rising and falling oxygen levels in the ocean and atmosphere took place, superimposed on a first-order trend from generally low to intermediate to high concentrations over a period of perhaps two and half billion years (Lyons et al, 2014). A recent review of atmospheric oxygen and marine redox state(s) through the Neoproterozoic–Palaeozoic demonstrated that oxygen fluctuated by about an order of magnitude, suggesting that instead of a single Neoproterozoic oxygenation event, there were multiple ocean oxygenation events during this period (Tostevin & Mills, 2020). Interestingly, our results suggest a pattern of Hc diversification and expansion that could be related to these multiple oxygenation events. More studies tracing the co-correlation of these multiple oxygenation events and genomic expansion of Hc would offer further insight into the effects, and selective pressure, of oxygen-binding proteins in early metazoans when oxygen availability was a key limiting factor.

4 | CONCLUSIONS

Our results suggest major evolutionary steps occurred before the extreme glacial events of the Neoproterozoic as marked by the emergence of the metazoan Hc superfamily at around 1117–756 Ma (average 881 Ma). They imply that crown-group animals were likely to possess blood pigments (Hc-like), which may have enhanced their respiratory capacity under the predicted low-oxygen conditions of that time. Moreover, Hcs might also have worked as a means for the transport of hormones, detoxification of heavy metals, and innate immunity pathways in animals without circulatory systems. Obtaining functional and experimental data on Hcs at different oxygen levels is still needed to evaluate the significance of their widespread occurrence in metazoans in the context of the metazoan dawn.

ACKNOWLEDGMENTS

A fellowship to E.M.C-P. was provided by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil—2018/20268–1). The study was supported by FAPESP, by thematic project (Proc. 2016/06114-6), coordinated by R.I.T. Use of SkyNet computational resources at Auburn University is acknowledged. This work was also funded in part by National Science Foundation (grants DEB–1036537 to K.M.H. and Scott R. Santos and OCE–1155188 to K.M.H.), and FAPESP JP grant 2015/50164-5 to F.D.B. C.J.C.'s contributions are facilitated by start-up funds from the College of Science, Swansea University.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Figshare at https://doi.org/10.6084/m9.figshare.c.5662558.v1

ORCID

Elisa Maria Costa-Paiva https://orcid.org/0000-0002-8470-1271
Beatriz Mello https://orcid.org/0000-0002-4511-2757
Christopher J. Coates https://orcid.org/0000-0002-4471-4369
Kenneth M. Halanych https://orcid.org/0000-0002-8658-9674
Federico Brown https://orcid.org/0000-0001-9250-5011
Juliana de Moraes Leme https://orcid.org/0000-0002-5833-4885
Ricardo I. F. Trindade https://orcid.org/0000-0001-9848-9550

REFERENCES

Aguilera, F., McDougall, C., & Degnan, B. M. (2013). Origin, evolution and classification of type-3 copper proteins: lineage-specific gene expansions and losses across the Metazoa. *BMC Evolutionary Biology*, *13*(1), 96. https://doi.org/10.1186/1471-2148-13-96.

Ashida, M., & Yoshida, H. (1988). Limited proteolysis of prophenoloxidase during activation by microbial products in insect plasma and effect of phenoloxidase on electrophoretic mobility of plasma proteins. *Insect Biochemistry*, *18*, 11–19. https://doi.org/10.1016/0020-1790(88)90031-5.

Àspan, A., & Söderhäll, K. (1991). Purification of prophenoloxidase from crayfish blood cells, and its activation by an endogenous serine proteinase. *Insect Biochemistry*, *21*, 363–373. https://doi.org/10.1016/0020-1790(91)90002-V.

Benton, M. J., Donoghue, P. C. J., Asher, R. J., Friedman, F., Near, T. J., & Vinther, J. (2015). Constraints on the timescale of animal evolutionary history. *Palaeontologia Electronica*, *18.1.1FC*, 1–107. https://doi.org/10.26879/424.

Betts, H. C., Puttick, M. N., Clark, J. W., Williams, T. A., Donoghue, P. C., & Pisani, D. (2018). Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nature Ecology & Evolution*, *2*(10), 1556–1562. https://doi.org/10.1038/s41559-018-064.

Bezerra, B. S., Belato, F. A., Mello, B., Brown, F., Coates, C. J., de Moraes Leme, J., Trindade, R. I. F., & Costa-Paiva, E. M. (2021). Evolution of a key enzyme of aerobic metabolism reveals Proterozoic functional subunit duplication events and an ancient origin of animals. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-95094-4.

Bobrovskiy, I., Hope, J. M., Nettersheim, B. J., Volkman, J. K., Hallmann, C., & Brocks, J. J. (2021). Algal origin of sponge sterane biomarkers negates the oldest evidence for animals in the rock record. *Nature Ecology and Evolution*, *5*, 165–168. https://doi.org/10.1038/s41559.

Boden, J. S., Konhauser, K. O., Robbins, L. J., & Sánchez-Baracaldo, P. (2021). Timing the evolution of antioxidant enzymes in cyanobacteria. *Nature Communications*, *12*(1), 1–12. https://doi.org/10.1038/s41467-021-24396-y.

Brocks, J. J., Jarrett, A. J. M., Sirantoine, E., Kenig, F., Moczydłowska, M., Porter, S., & Hope, J. (2016). Early sponges and toxic protists: possible sources of cryostane, an age diagnostic biomarker antedating Sturtian Snowball Earth. *Geobiology*, *14*, 129–149. https://doi.org/10.1111/gbi.12165.

Budd, G. E. (2008). The earliest fossil record of the animals and its significance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 1425–1434. https://doi.org/10.1098/rstb.2007.2232.

Budd, G. E., & Jensen, S. (2000). A critical reappraisal of the fossil record of the bilaterian phyla. *Biological Reviews*, *75*(2), 253–295. https://doi.org/10.1017/S000632310000548X.

Budd, G. E., & Mann, R. P. (2020). The dynamics of stem and crown groups. Science. *Advances*, *6*(8), eaaz1626. https://doi.org/10.1126/sciadv.aaz1626.

Budd, G. E., & Mann, R. P. (2020). Survival and selection biases in early animal evolution and a source of systematic overestimation in molecular clocks. *Interface Focus*, *10*, 20190110. https://doi.org/10.1098/rsfs.2019.0110.

Burmester, T. (1999a). Evolution and function of the insect hexamerins. *European Journal of Entomology*, 213–226.

Burmester, T. (1999b). Identification, molecular cloning, and phylogenetic analysis of a non-respiratory pseudo-hemocyanin of *Homarus americanus*. *Journal of Biological Chemistry*, *274*(19), 13217–13222. https://doi.org/10.1074/jbc.274.19.13217.

Burmester, T. (2001). Molecular evolution of the arthropod hemocyanin superfamily. *Molecular Biology and Evolution*, *18*(2), 184–195. https://doi.org/10.1093/oxfordjournals.molbev.a003792.

Burmester, T. (2002). Origin and evolution of arthropod hemocyanins and related proteins. *Journal of Comparative Physiology B*, *172*(2), 95–107. https://doi.org/10.1007/s00360-001-0247-7.

Burmester, T. (2015). Evolution of respiratory proteins across the Pancrustacea. *Integrative and Comparative Biology*, *55*(5), 792–801. https://doi.org/10.1093/icb/icv079.

Burmester, T., & Schellen, K. (1996). Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and dipteran arylphorin receptor. *Journal of Molecular Evolution*, *42*, 713–728. https://doi.org/10.1007/BF02338804.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Chen, H., Parry, L. A., Vinther, J., Zhai, D., Hou, X., & Ma, X. (2020). A Cambrian crown annelid reconciles phylogenomics and the fossil record. *Nature*, *583*(7815), 249–252, https://doi.org/10.1038/s41586-020-2384-8.

Christner, B. C., Kvitko, B. H., & Reeve, J. N. (2003). Molecular identification of bacteria and eukarya inhabiting an Antarctic cryoconite hole. *Extremophiles*, *7*(3), 177–183. https://doi.org/10.1007/s00792-002-0309-0.

Coates, C. J., & Costa-Paiva, E. M. (2020). Multifunctional roles of hemocyanins. In U. Hoeger, & J. R. Harris (Eds.). *Vertebrate and invertebrate respiratory proteins, lipoproteins and other body fluid proteins* (pp. 233-250). Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-41769-7_9

Coates, C. J., & Decker, H. (2017). Immunological properties of oxygen-transport proteins: hemoglobin, hemocyanin and hemerythrin. *Cellular and Molecular Life Sciences*, *74*(2), 293–317. https://doi.org/10.1007/s00018-016-2326-7.

Coates, C. J., Kelly, S. M., & Nairn, J. (2011). Possible role of phosphatidylserine–hemocyanin interaction in the innate immune response of Limulus polyphemus. *Developmental & Comparative Immunology*, *35*(2), 155–163. https://doi.org/10.1016/j.dci.2010.08.015.

Coates, C. J., & Nairn, J. (2014). Diverse immune functions of hemocyanins. *Developmental & Comparative Immunology*, *45*(1), 43–55. https://doi.org/10.1016/j.dci.2014.01.021.

Coates, C. J., & Talbot, J. (2018). Hemocyanin-derived phenoloxidase reaction products display anti-infective properties. *Developmental & Comparative Immunology*, *86*, 47–51. https://doi.org/10.1016/j.dci.2018.04.017.

Coates, C. J., Whalley, T., Wyman, M., & Nairn, J. (2013). A putative link between phagocytosis-induced apoptosis and hemocyanin-derived phenoloxidase activation. *Apoptosis*, *18*(11), 1319–1331. https://doi.org/10.1007/s10495-013-0891-x.

Cole, D. B., Mills, D. B., Erwin, D. H., Sperling, E. A., Porter, S. M., Reinhard, C. T., & Planavsky, N. J. (2020). On the co-evolution of surface oxygen levels and animals. *Geobiology*, *18*(3), 260–281. https://doi.org/10.1111/gbi.12382.

Conway-Morris, S. (1993). The fossil record and the early evolution of the Metazoa. *Nature*, *361*, 219–225. https://doi.org/10.1038/361219a0.

Cordani, U. G., Fairchild, T. R., Ganade, C. E., Babinski, M., & Leme, J. D. M. (2020). Dawn of metazoans: to what extent was this influenced

by the onset of "modern-type plate tectonics"? Brazilian. *Journal of Geology*, *50*(2), https://doi.org/10.1590/2317-4889202020190095.

Costa-Paiva, E. M., Schrago, C. G., Coates, C. J., & Halanych, K. M. (2018). Discovery of novel hemocyanin-like genes in Metazoans. *Biological Bulletin*, *235*(3), 134–151. https://doi.org/10.1086/700181.

Costas, E., Flores-Moya, A., & López-Rodas, V. (2008). Rapid adaptation of phytoplankters to geothermal waters is achieved by single mutations: were extreme environments 'Noah's Arks' for photosynthesizers during the Neoproterozoic 'snowball Earth'? *New Phytologist*, *180*(4), 922–932. https://doi.org/10.1111/j.1469-8137.2008.02620.x.

Crimes, T. P., & Jiang, Z. W. (1986). Trace fossils from the Precambrian-Cambrian boundary candidate at Meishucun, Jinning, Yunnan. *China. Geological Magazine*, *123*(6), 641–649. https://doi.org/10.1017/S0016756800024158

Cui, H., Kaufman, A. J., Xiao, S., Peek, S., Cao, H., Min, X., Cai, Y., Siegel, Z., Liu, X. M., Peng, Y., Schiffbauer, J. d, & Martin, A. J. (2016). Environmental context for the terminal Ediacaran biomineralization of animals. *Geobiology*, *14*, 344–363. https://doi.org/10.1111/gbi.12178

Cui, H., Xiao, S., Yaoping, C., Peek, S., Plummer, R. E., & Kaufman, A. J. (2019). *Sedimentology and chemostratigraphy of the terminal Ediacaran Dengying Formation at the Gaojiashan section*. Geol. Mag. https://doi.org/10.1017/S0016756819000293

Delsuc, F., Philippe, H., Tsagkogeorga, G., Simion, P., Tilak, M. K., Turon, X., López-Legentil, S., Piette, J., Lemaire, P., & Douzery, E. J. (2018). A phylogenomic framework and timescale for comparative studies of tunicates. *BMC Biology*, *16*(1), 1–14. https://doi.org/10.1186/s12915-018-0499-2.

Dohrmann, M., & Wörheide, G. (2017). Dating early animal evolution using phylogenomic data. *Scientific Reports*, *7*, 3599. https://doi.org/10.1038/s41598-017-03791-w.

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M. J., Donoghue, P. C., & Yang, Z. (2015). Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Current Biology*, *25*(22), 2939–2950. https://doi.org/10.1016/j.cub.2015.09.066.

Eibye-Jacobsen, D., & Vinther, J. (2012). Reconstructing the ancestral annelid. *Journal of Zoological Systematics and Evolutionary Research*, *50*(1), 85–87. https://doi.org/10.1111/j.1439-0469.2011.00651.x.

Erwin, D., Laflamme, M., Tweedt, S., Sperling, E., Pisani, D., & Peterson, K. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*, *334*(6059), 1091–1097. https://doi.org/10.1126/science.1206375.

Fraser, C. I., Terauds, A., Smellie, J., Convey, P., & Chown, S. L. (2014). Geothermal activity helps life survive glacial cycles. *Proceedings of the National Academy of Sciences*, *111*(15), 5634–5639. https://doi.org/10.1073/pnas.1321437111.

Giribet, G. (2016). New animal phylogeny: future challenges for animal phylogeny in the age of phylogenomics. *Organisms Diversity & Evolution*, *16*(2), 419–426. https://doi.org/10.1007/s13127-015-0236-4.

Gu, X. (1998). Early metazoan divergence was about 830 million years ago. *Journal of Molecular Evolution*, *47*, 369–371. https://doi.org/10.1007/PL00013150.

Halanych, K. M. (2016). How our view of animal phylogeny was reshaped by molecular approaches: lessons learned. *Organisms Diversity and Evolution*, *16*, 319–328. https://doi.org/10.1007/s13127-016-0264-8.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Hoffman, P. F. (2016). Cryoconite pans on Snowball Earth: supraglacial oases for Cryogenian eukaryotes? *Geobiology*, *14*(6), 531–542. https://doi.org/10.1111/gbi.12191.

Immesberger, A., & Burmester, T. (2004). Putative phenoloxidases in the tunicate Ciona intestinalis and the origin of the arthropod hemocyanin superfamily. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *174*(2), 169–180. https://doi.org/10.1007/s00360-003-0402-4.

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J. Y., Kupfer, A., Petersen, J., Jarek, M., Meyer, A., Vences, M., & Philippe, H. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nature Ecology & Evolution*, *1*(9), 1370–1378. https://doi.org/10.1038/s41559-017-0240-5.

Jabłońska, J., & Tawfik, D. S. (2021). The evolution of oxygen-utilizing enzymes suggests early biosphere oxygenation. *Nature Ecology & Evolution*, *5*(4), 442–448. https://doi.org/10.1038/s41559-020-01386-9.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

Kenrick, P., Wellman, C. H., Schneider, H., & Edgecombe, G. D. (2012). A timeline for terrestrialization: Consequences for the carbon cycle in the Palaeozoic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1588), 519–536. https://doi.org/10.1098/rstb.2011.0271.

Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences*, *39*, 217–239. https://doi.org/10.1146/annurev.earth.031208.100209.

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*, 1812–1819. https://doi.org/10.1093/molbev/msx116.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lamb, D. M., Awramik, S. M., Chapman, D. J., & Zhu, S. (2009). Evidence for eukaryotic diversification in the ~1800 million-year-old Changzhougou Formation, North China. *Precambrian Research*, *173*, 93–104. https://doi.org/10.1016/j.precamres.2009.05.005.

Lartillot, N., & Philippe, H. A. (2004). Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution*, *21*, 1095–1109. https://doi.org/10.1093/molbev/msh112.

Love, G. D., Grosjean, E., Stalvies, C., Fike, D. A., Grotzinger, J. P., Bradley, A. S., Kelly, A. E., Bhatia, M., Meredith, W., Snape, C. E., Bowring, S. A., Condon, D. J., & Summons, R. E. (2009). Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature*, *457*(7230), 718–721. https://doi.org/10.1038/nature07673.

Love, G. D., & Summons, R. E. (2015). The molecular record of Cryogenian sponges–a response to Antcliffe (2013). *Palaeontology*, *58*(6), 1131–1136. https://doi.org/10.1111/pala.12196.

Lozano-Fernandez, J., Carton, R., Tanner, A. R., Puttick, M. N., Blaxter, M., Vinther, J., Olesen, J., Giribet, G., Edgecombe, G. D., & Pisani, D. (2016). A molecular palaeobiological exploration of arthropod terrestrialization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1699), 20150133. https://doi.org/10.1098/rstb.2015.0133.

Lozano-Fernandez, J., dos Reis, M., Donoghue, P. C., & Pisani, D. (2017). RelTime rates collapse to a strict clock when estimating the timeline of animal diversification. *Genome Biology and Evolution*, *9*(5), 1320–1328. https://doi.org/10.1093/gbe/evx079.

MacNaughton, R. B., & Narbonne, G. M. (1999). Evolution and ecology of Neoproterozoic lower Cambrian trace fossils, NW Canada. *Palaios*, *14*, 97–115. https://doi.org/10.2307/3515367

Magallón, S., Hilu, K. W., & Quandt, D. (2013). Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American Journal of Botany*, *100*, 556–573. https://doi.org/10.3732/ajb.1200416

Marin, J., Battistuzzi, F. U., Brown, A. C., & Hedges, S. B. (2016). The timetree of prokaryotes: new insights into their evolution and speciation. *Molecular Biology and Evolution*, *34*(2), 437–446. https://doi.org/10.1093/molbev/msw245

Markl, J., & Decker, H. (1992). Molecular structure of the arthropod hemocyanins. In C. P. Mangum (Ed). *Blood and tissue oxygen carriers. advances in comparative and environmental physiology*, (vol 13, pp 325–376) Springer. https://doi.org/10.1007/978-3-642-76418-9_12

Martín-Durán, J. M., de Mendoza, A., Sebé-Pedrós, A., Ruiz-Trillo, I., & Hejnol, A. (2013). A broad genomic survey reveals multiple origins and frequent losses in the evolution of respiratory hemerythrins and hemocyanins. *Genome Biology and Evolution*, *5*, 1435–1442. https://doi.org/10.1093/gbe/evt102.

Mills, D. B., & Canfield, D. E. (2014). Oxygen and animal evolution: Did a rise of atmospheric oxygen "trigger" the origin of animals? *BioEssays*, *36*, 1145–1155. https://doi.org/10.1002/bies.201400101.

Mills, D. B., Francis, W. R., & Canfield, D. E. (2018). Animal origins and the Tonian Earth system. *Emerging Topics in Life Sciences*, *2*(2), 289–298. https://doi.org/10.1042/ETLS20170160.

Mills, D. B., Francis, W. R., Vargas, S., Larsen, M., Elemans, C. P. H., Canfield, D. E., & Wörheide, G. (2018). The last common ancestor of animals lacked the HIF pathway and respired in low-oxygen environments, *eLife*, *7*, e31176. https://doi.org/10.7554/eLife.31176.001.

Mills, D. B., Ward, L. M., Jones, C., Sweeten, B., Forth, M., Treusch, A. H., & Canfield, D. E. (2014). Oxygen requirements of the earliest animals. *Proceedings of the National Academy of Sciences USA*, *111*(11), 4168–4172. https://doi.org/10.1073/pnas.1400547111.

Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., & Niehuis, O. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*(6210), 763–767. https://doi.org/10.1126/science.1257570.

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C. H., Yang, Z., Schneider, H., & Donoghue, P. C. (2018). The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences*, *115*(10), E2274–E2283. https://doi.org/10.1073/pnas.1719588115.

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300.

Nursall, J. R. (1959). Oxygen as a prerequisite to the origin of the Metazoa. *Nature*, *183*(4669), 1170–1172. https://doi.org/10.1038/1831170b0.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*, 877–884. https://doi.org/10.1038/44766.

Parham, J. F., Donoghue, P. C., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., Inoue, J. G., Irmis, R. B., Joyce, W. G., Ksepka, D. T., Patané, J. S. L., Smith, N. D., Tarver, J. E., van Tuinen, M., Yang, Z., Angielczyk, K. D., Greenwood, J. M., Hipsley, C. A., Jacobs, L., … Benton, M. J. (2012). Best practices for justifying fossil calibrations. *Systematic Biology*, *61*(2), 346–359. https://doi.org/10.1093/sysbio/syr107.

Parry, L. A., Lerosey-Aubril, R., Weaver, J. C., & Ortega-Hernández, J. (2021). Cambrian comb jellies from Utah illuminate the early evolution of nervous and sensory systems in ctenophores, *iScience*, 102943. https://doi.org/10.1016/j.isci.2021.102943

Parry, L., Vinther, J., & Edgecombe, G. D. (2015). Cambrian stem-group annelids and a metameric origin of the annelid head. *Biology Letters*, *11*(10), 20150763. https://doi.org/10.1098/rsbl.2015.0763.

Peng, S., & Babcock, L. (2008). Cambrian Period. In J. G. Ogg, G. Ogg, & F. M. Gradstein (Eds.), *The concise geologic time scale* (pp. 37–46). Cambridge University Press.

Peng, S., Babcock, L. E., & Cooper, R. A. (2012). The Cambrian Period. In F. M. Gradstein, J. G. Ogg, M. Schmitz, & G. Ogg (Eds.), *The Geologic Timescale 2012* (pp. 437–488). Elsevier.

Perron, U., Moal, I. H., Thorne, J. L., & Goldman, N. (2019). Probabilistic models for the study of protein evolution. In D. J. Balding, I. Moltke, & J. Marioni (Eds.), *Handbook of Statistical Genomics* (pp. 347–430). John Wiley & Sons Ltd. https://doi.org/10.1002/9781119487845.ch12.

Peterson, K. J., Lyons, J. B., Nowak, K. S., Takacs, C. M., Wargo, M. J., & McPeek, M. A. (2004). Estimating metazoan divergence times with a molecular clock. *Proceedings of the National Academy of Sciences USA*, *101*(17), 6536–6541. https://doi.org/10.1073/pnas.0401670101.

Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., Lartillot, N., & Wörheide, G. (2015). Genomic data do not support comb jellies as the sister group to all other animals. *Proceedings of the National Academy of Sciences USA*, *112*(50), 15402–15407. https://doi.org/10.1073/pnas.1518127112.

Prothmann, A., Hoffmann, F. G., Opazo, J. C., Herbener, P., Storz, J. F., Burmester, T., & Hankeln, T. (2020). The globin gene family in arthropods: evolution and functional diversity. *Frontiers in Genetics*, *11*, 858. https://doi.org/10.3389/fgene.2020.00858.

Qun, Y., Junye, M., Xiaoyan, S., & Peiyun, C. (2007). Phylochronology of early metazoans: combined evidence from molecular and fossil data. *Geological Journal*, *42*(3–4), 281–295. https://doi.org/10.1002/gj.1074.

Raff, R., & Raff, E. (1970). Respiratory mechanisms and the metazoan fossil record. *Nature*, *228*(5275), 1003–1005. https://doi.org/10.1038/2281003a0.

Rooney, A. D., Strauss, J. V., Brandon, A. D., & Macdonald, F. A. (2015). A Cryogenian chronology: Two long-lasting synchronous Neoproterozoic glaciations. *Geology*, *43*(5), 459–462. https://doi.org/10.1130/G36511.1.

Schmidt-Rhaesa, A. (2007). *The Evolution of Organs Systems*. Oxford University Press.

Schreml, S., Szeimies, R. M., Prantl, L., Karrer, S., Landthaler, M., & Babilas, P. (2010). Oxygen in acute and chronic wound healing. *British Journal of Dermatology*, *163*(2), 257–268. https://doi.org/10.1111/j.1365-2133.2010.09804.x.

Schwartz, R. S., & Mueller, R. L. (2010). Branch length estimation and divergence dating: estimates of error in Bayesian and maximum likelihood frameworks. *BMC Evolutionary Biology*, *10*, 5. https://doi.org/10.1186/1471-2148-10-5.

Sebé-Pedrós, A., de Mendoza, A., Lang, B. F., Degnan, B. M., & Ruiz-Trillo, I. (2011). Unexpected repertoire of metazoan transcription factors in the unicellular holozoan Capsaspora owczarzaki. *Molecular Biology and Evolution*, *28*(3), 1241–1254. https://doi.org/10.1093/molbev/msq309.

Shih, P. M., & Matzke, N. J. (2013). Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *Proceedings of the National Academy of Sciences*, *110*(30), 12355–12360. https://doi.org/10.1073/pnas.1305813110.

Söderhäll, K., & Cerenius, L. (1998). Role of the prophenoloxidase-activating system in invertebrate immunity. *Current Opinion in Immunology*, *10*(1), 23–28. https://doi.org/10.1016/S0952-7915(98)80026-5.

Spence, G. H., Le Heron, D. P., & Fairchild, I. J. (2016). Sedimentological perspectives on climatic, atmospheric and environmental change in the Neoproterozoic Era. *Sedimentology*, *63*(2), 253–306. https://doi.org/10.1111/sed.12261.

Sperling, E. A., Frieder, C. A., Raman, A. V., Girguis, P. R., Levin, L. A., & Knoll, A. H. (2013). Oxygen, ecology, and the Cambrian radiation of animals. *Proceedings of the National Academy of Sciences USA*, *110*(33), 13446–13451. https://doi.org/10.1073/pnas.1312778110.

Sperling, E. A., Halverson, G. P., Knoll, A. H., Macdonald, F. A., & Johnston, D. T. (2013). A basin redox transect at the dawn of animal life. *Earth and Planetary Science Letters*, *371–372*, 143–155. https://doi.org/10.1016/j.epsl.2013.04.003.

Sperling, E. A., Knoll, A. H., & Girguis, P. R. (2015). The ecological physiology of Earth's second oxygen revolution. *Annual Review of Ecology Evolution and Systematics*, *46*, 215–235. https://doi.org/10.1146/annurev-ecolsys-110512-135808.

Swofford, D. L., Olsen, G. J., Waddell, P. J., & Hillis, D. M. (1996). Phylogenetic inference. In D. M. Hillis, C. Moritz, & B. K. Mable (Eds) *Molecular systematics* (pp 407–514). Sinauer Associates, Mass.

Terwilliger, N. B. (1998). Functional adaptations of oxygen-transport proteins. *Journal of Experimental Biology*, *201*, 1085–1098. https://doi.org/10.1242/jeb.201.8.1085.

Thomas, D. N., & Dieckmann, G. S. (2002). Antarctic sea ice–a habitat for extremophiles. *Science*, *295*(5555), 641–644. https://doi.org/10.1126/science.1063391.

Tostevin, R., & Mills, B. J. W. (2020). Reconciling proxy records and models of Earth's oxygenation during the Neoproterozoic and Palaeozoic. *Interface Focus*, *10*, 20190137. https://doi.org/10.1098/rsfs.2019.0137.

Towe, K. M. (1970). Oxygen-collagen priority and the early metazoan fossil record. *Proceedings of the National Academy of Sciences USA*, *65*(4), 781–788. https://doi.org/10.1073/pnas.65.4.781.

Trindade, R., D'Agrella Filho, M., Epof, I., & Brito Neves, B. (2006). Paleomagnetism of Early Cambrian Itabaiana mafic dikes (NE Brazil) and the final assembly of Gondwana. *Earth and Planetary Science Letters*, *244*(1–2), 361–377. https://doi.org/10.1016/j.epsl.2005.12.039.

Turner, E. C. (2021). Possible poriferan body fossils in early Neoproterozoic microbial reefs. *Nature*, *596*, 87–91. https://doi.org/10.1038/s41586-021-03773-z.

Valentine, J. W., Jablonski, D., & Erwin, D. E. (1999). Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development*, *126*(5), 851–859.

van Holde, K. E., & Miller, K. I. (1995). Hemocyanins. *Advances in Protein Chemistry*, *47*, 1–81. https://doi.org/10.1016/S0065-3233(08)60545-8.

van Holde, K. E., Miller, K. I., & Decker, H. (2001). Hemocyanins and invertebrate evolution. *Journal of Biological Chemistry*, *276*, 15563–15566. https://doi.org/10.1074/jbc.R100010200.

van Maldegem, L. M., Nettersheim, B. J., Leider, A., Brocks, J. J., Adam, P., Schaeffer, P., & Hallmann, C. (2021). Geological alteration of Precambrian steroids mimics early animal signatures. *Nature Ecololy and Evolution*, *5*, 169–173. https://doi.org/10.1038/s41559.

Varga, T., Krizsán, K., Földi, C., Dima, B., Sánchez-García, M., Sánchez-Ramírez, S., Szöllősi, G. J., Szarkándi, J. G., Papp, V., Albert, L., Andreopoulos, W., Angelini, C., Antonín, V., Barry, K. W., Bougher, N. L., Buchanan, P., Buyck, B., Bense, V., Catcheside, P., … Nagy, L. G. (2019). Megaphylogeny resolves global patterns of mushroom evolution. *Nature Ecology & Evolution*, *3*(4), 668–678. https://doi.org/10.1038/s41559-019-0834-1.

Weber, B., & Zhu, M. Y. (2003). Arthropod trace fossils from the Zhujiaqing Formation (Meishucunian, Yunnan) and their palaeobiological implications. *Progress in Natural Science*, *13*, 795–800. https://doi.org/10.1080/10020070312331344450.

Whelan, N. V., Kocot, K. M., Moroz, L. L., & Halanych, K. M. (2015). Error, signal, and the placement of Ctenophora sister to all other animals. *Proceedings of the National Academy of Sciences USA*, *112*, 5773–5778. https://doi.org/10.1073/pnas.1503453112.

Whitten, M., & Coates, C. J. (2017). Re-evaluation of insect melanogenesis research: views from the dark side. *Pigment Cell & Melanoma Research*, *30*, 386–401. https://doi.org/10.1111/pcmr.12590.

Wilke, T., Schultheiß, R., & Albrecht, C. (2009). As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. *American Malacoogical Bulletin*, *27*, 25–45. https://doi.org/10.4003/006.027.0203.

Wolfe, J. M., Breinholt, J. W., Crandall, K. A., Lemmon, A. R., Lemmon, E. M., Timm, L. E., Siddall, M. E., & Bracken-Grissom, H. D. (2019). A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans. *Proceedings of the Royal Society B*, *286*(1901), 20190079. https://doi.org/10.1098/rspb.2019.0079.

Wood, R., Donoghue, P. C. J., Lenton, T. M., Liu, A. G., & Poulton, S. W. (2020). The origin and rise of complex life: progress requires interdisciplinary integration and hypothesis testing. *Interface Focus*, *10*, 20200024. https://doi.org/10.1098/rsfs.2020.0024.

Wood, R., Ivantsov, A. Y., & Zhuralev, A. Y. (2017). First macrobiota biomineralization was environmentally triggered. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20170059. https://doi.org/10.1098/rspb.2017.0059.

Wood, R., Liu, A. G., Bowyer, F., Wilby, P. R., Dunn, F. S., Kenchington, C. G., Cuthill, J. F. H., Mitchell, E., & Penny, A. (2019). Integrated records of environmental change and evolution challenge the Cambrian Explosion. *Nature Ecology & Evolution*, *3*(4), 528–538. https://doi.org/10.1038/s41559-019-0821-6.

Xiao, S., Muscente, A. D., Chen, L., Zhou, C., Schiffbauer, J. D., Wood, A. D., Polys, N. F., & Yuan, X. (2014). The Weng'an biota and the Ediacaran radiation of multicellular eukaryotes. *National Science Review*, *1*(4), 498–520. https://doi.org/10.1093/nsr/nwu061.

Xiao, S., Narbonne, G. M., Zhou, C., Lafl Amme, M., Grazhdankin, D. V., Moczydlowska-Vidal, M., & Cui, H. (2016). Toward an Ediacaran time scale: problems, protocols, and prospects. *Episodes 39*, 540–555. https://doi.org/10.18814/epiiugs/2016/v39i4/103886

Yu, H., & Li, L. (2014). Phylogeny and molecular dating of the cerato-platanin-encoding genes. *Genetics and Molecular Biology*, *37*(2), 423–427. https://doi.org/10.1590/S1415-47572014005000003.

Yuan, X., Chen, Z., Xiao, S., Zhou, C., & Hua, H. (2011). An early Ediacaran assemblage of macroscopic and morphologically differentiated eukaryotes. *Nature*, *470*, 390–393. https://doi.org/10.1038/nature09810.

Zhang, X. G., Siveter, D. J., Waloszek, D., & Maas, A. (2007). An epipodite-bearing crown-group crustacean from the Lower Cambrian. *Nature*, *449*, 595–598. https://doi.org/10.1038/nature06138.

Zheng, Y., Peng, R., Kuro-o, M., & Zeng, X. (2011). Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of salamanders (Order Caudata). *Molecular Biology and Evolution*, *28*, 2521–2535. https://doi.org/10.1093/molbev/msr072.

Zmasek, C. M., & Eddy, S. R. (2001). A simple algorithm to infer gene duplication and speciation events on a gene tree. *Bioinformatics*, *17*(9), 821–828. https://doi.org/10.1093/bioinformatics/17.9.821.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Maria Costa-Paiva, E., Mello, B., Santos Bezerra, B., Coates, C. J., Halanych, K. M., Brown, F., de Moraes Leme, J., & Trindade, R. I. F. (2021). Molecular dating of the blood pigment hemocyanin provides new insight into the origin of animals. *Geobiology*, 00, 1–13. https://doi.org/10.1111/gbi.12481