CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Exceptionally preserved fossil assemblages through geologic time and space
Authors	Muscente, A. D.;Schiffbauer, James D.;Broce, Jesse;Laflamme, Marc;O'Donnell, Kenneth;Boag, Thomas H.;Meyer, Michael;Hawkins, Andrew D.;Huntley, John Warren;McNamara, Maria E.;MacKenzie, Lindsay A.;Stanley Jr, George D.;Hinman, Nancy W.;Hofmann, Michael H.;Xiao, Shuhai
Publication date	2017-04-26
Original Citation	Muscente, A. D., Schiffbauer, J. D., Broce, J., Laflamme, M., O'Donnell, K., Boag, T. H., Meyer, M., Hawkins, A. D., Huntley, J. W., McNamara, M. E., MacKenzie, L. A., Stanley Jr, G. D., Hinman, N. W., Hofmann, M. H. and Xiao, S. (2017) 'Exceptionally preserved fossil assemblages through geologic time and space', Gondwana Research, 48, pp. 164-188. doi:10.1016/j.gr.2017.04.020
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.gr.2017.04.020
Rights	© 2017, International Association for Gondwana Research. Published by Elsevier B.V. All rights reserved. This manuscript version is made available under the CC-BY-NC-ND 4.0 license. - http://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-11-07 07:15:56
Item downloaded from	https://hdl.handle.net/10468/4593

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Accepted Manuscript



Exceptionally preserved fossil assemblages through geologic time and space

A.D. Muscente, James D. Schiffbauer, Jesse Broce, Marc Laflamme, Kenneth O'Donnell, Thomas H. Boag, Michael Meyer, Andrew D. Hawkins, John Warren Huntley, Maria McNamara, Lindsay A. MacKenzie, George D. Stanley, Nancy W. Hinman, Michael H. Hofmann, Shuhai Xiao

PII: S1342-937X(17)30210-1
DOI: doi: 10.1016/j.gr.2017.04.020
Reference: GR 1792

To appear in:

Received date: 23 February 2017
Revised date: 13 April 2017
Accepted date: 24 April 2017

Please cite this article as: A.D. Muscente, James D. Schiffbauer, Jesse Broce, Marc Laflamme, Kenneth O'Donnell, Thomas H. Boag, Michael Meyer, Andrew D. Hawkins, John Warren Huntley, Maria McNamara, Lindsay A. MacKenzie, George D. Stanley, Nancy W. Hinman, Michael H. Hofmann, Shuhai Xiao , Exceptionally preserved fossil assemblages through geologic time and space. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Gr(2017), doi: 10.1016/j.gr.2017.04.020

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

GR Focus Review

Exceptionally preserved fossil assemblages through geologic time and space

A. D. Muscente[a*], James D. Schiffbauer[b], Jesse Broce[b], Marc Laflamme[c], Kenneth O'Donnell[d], Thomas H. Boag[e], Michael Meyer[f], Andrew D. Hawkins[g], John Warren Huntley[b], Maria McNamara[h], Lindsay, A. MacKenzie[i], George D. Stanley Jr.[j], Nancy W. Hinman[j], Michael. H. Hofmann[j], and Shuhai Xiao[g]

[a]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA

[b]Department of Geological Sciences, University of Missouri, Columbia, MO 65211, USA

[c]Department of Chemical and Physical Sciences, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada

[d]AECOM, Germantown, MD 20876, USA

[e]Department of Geological Sciences, Stanford University, Stanford, CA 94305, USA

[f]Geophysical Laboratory, Carnegie Institution for Science, Washington, DC 20015, USA

[g]Department of Geosciences, Virginia Tech, Blacksburg, VA 24061, USA

[h]School of Biological, Earth, and Environmental Sciences, University College Cork, Cork, T23 TK30, Ireland

[i]Department of Geological Sciences, SUNY Geneseo, Geneseo, NY 14454, USA

[j]Department of Geosciences, University of Montana, Missoula, MT 59812, USA

Corresponding author at: Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA. Email address: a.d.muscente@gmail.com (A. D. Muscente).

ACCEPTED MANUSCRIPT

ABSTRACT

Geologic deposits containing fossils with remains of non-biomineralized tissues (i.e. Konservat-Lagerstätten) provide key insights into ancient organisms and ecosystems. Such deposits are not evenly distributed through geologic time or space, suggesting that global phenomena play a key role in exceptional fossil preservation. Nonetheless, establishing the influence of global phenomena requires documenting temporal and spatial trends in occurrences of exceptionally preserved fossil assemblages. To this end, we compiled and analyzed a dataset of 694 globally distributed exceptional fossil assemblages spanning the history of complex eukaryotic life (~610 to 3 Ma). Our analyses demonstrate that assemblages with similar ages and depositional settings commonly occur in clusters, each signifying an ancient geographic region (up to hundreds of kilometers in scale), which repeatedly developed conditions conducive to soft tissue preservation. Using a novel hierarchical clustering approach, we show that these clusters decrease in number and shift from open marine to transitional and non-marine settings across the Cambrian-Ordovician interval. Conditions conducive to exceptional preservation declined worldwide during the early Paleozoic in response to transformations of near-surface environments that promoted degradation of tissues and curbed authigenic mineralization potential. We propose a holistic explanation relating these environmental transitions to ocean oxygenation and bioturbation, which affected virtually all taphonomic pathways, in addition to changes in seawater chemistry that disproportionately affected processes of soft tissue conservation. After these transitions, exceptional preservation rarely occurred in open marine settings, excepting times of widespread oceanic anoxia, when low oxygen levels set the stage. With these patterns, non-marine cluster count is correlated with non-marine rock quantity, and generally decreases with age. This result suggests that geologic processes, which progressively

ACCEPTED MANUSCRIPT

destroy terrestrial rocks over time, limit sampling of non-marine deposits on a global scale. Future efforts should aim to assess the impacts of such phenomena on evolutionary and ecological patterns in the fossil record.

Keywords: Taphonomy, Lagerstätten, big data, bioturbation, ocean oxygenation, rock record bias

Contents

1. Introduction
2. Review of taphonomic processes and pathways
   2.1. Taphonomic processes
   2.2. Taphonomic pathways: Combinations of processes
   2.3. The dichotomy of the exceptional fossil record
3. Data and methods
   3.1. Overview of dataset of exceptionally preserved fossil assemblages
   3.2. Binning by age and depositional setting
   3.3. Comparison with sampling patterns in the Paleobiology Database (PBDB)
   3.4. Nearest neighbor analysis and geospatial statistics
   3.5. Hierarchical clustering
   3.6. Estimations of rock quantity and past sea level
   3.7. Statistical testing of cluster counts and proportions
4. Results
   4.1. Comparison with global trends in skeletal fossil collection
   4.2. Nearest neighbor analysis and geospatial statistics
   4.3. Cluster count time series
   4.4. Proportion time series
5. Discussion
   5.1. Explanations for clustering of assemblages
   5.2. Trends in occurrences of clusters through time
   5.3. Relationships with quantities of rock available for sampling
   5.4. Trends in depositional settings of clusters
   5.5. Interpreting the cluster count and proportion time series results
      5.5.1. Exceptional preservation through time in marine settings
      5.5.2. Exceptional preservation through time in non-marine settings
      5.5.3. Other trends in exceptional preservation through time
6. Conclusions

Acknowledgements
Supplementary material
References

ACCEPTED MANUSCRIPT

1. Introduction

The bulk of the fossil record consists of skeletal materials (i.e. shells, bones, and teeth), biologically produced by organisms that control the formation of minerals within their tissues. Due to the abundance and wide availability of these fossils for study, they constitute the focus of most paleontological research, shaping scientific understanding of evolutionary and ecological patterns through geologic time. Problematically, however, organisms that produce mineralized body parts make up only a fraction of the total biodiversity in modern (Schopf, 1978; Valentine, 1989) and ancient ecosystems (Conway Morris, 1986), and those body parts convey limited information regarding non-biomineralized or "soft" anatomy. With these issues in mind, fossils of non-biomineralized tissues—or exceptionally preserved fossils (Fig. 1) —represent a critical resource for studying the ancient biosphere. Fossils of this type include some of the earliest organisms on Earth (Knoll, 1985) as well as remains of integuments (Manning et al., 2009; Navalón et al., 2015), exoskeletons (McNamara et al., 2012a), feathers (Colleary et al., 2015), eyes (Lee et al., 2011), muscles (Martill, 1990), "jellies" (Chen et al., 2007; Stanley and Sturmer, 1987), internal organs (Zhang et al., 2015), nerves (Yang et al., 2016), cells (Xiao et al., 1998), and sub-cellular structures (Muscente et al., 2015b), which offer many opportunities for work on anatomy, physiology, and systematics. Furthermore, geologic deposits containing exceptionally-preserved fossils—such as certain Konservat-Lagerstätten (Seilacher, 1970)—offer relatively complete snapshots for exploring the ecology and dynamics of ancient ecosystems, as they generally contain remains of both biomineralizing and non-mineralizing taxa (Conway Morris, 1986). For these reasons, exceptionally preserved fossil assemblages—such as the Weng'an (Xiao et al., 2014), Burgess Shale (Conway Morris, 1986), Hunsrück (Briggs et al., 1996), and Mazon Creek (Cotroneo et al., 2016) biotas—have received special attention among scientists

ACCEPTED MANUSCRIPT

studying Earth history. Work on such Konservat-Lagerstätten has led to innovative hypotheses regarding patterns in diversity and extinction of soft-bodied organisms through time (Darroch et al., 2015; Labandeira and Sepkoski, 1993; Van Roy et al., 2010).

Exceptionally preserved fossils occur less commonly than skeletal fossils because non-biomineralized body parts are rapidly destroyed in most environments. Preservation of soft tissues only takes place in rare circumstances where non-mineralized body parts survive degradational processes long enough to be transformed into recalcitrant carbonaceous materials or become secondarily replicated by minerals, which can persist over geologic timescales (Briggs, 2003). The taphonomic processes responsible for these transformations are modestly well understood. On the other hand, the environmental conditions that engender exceptional fossil preservation—and therefore, control the geographic and stratigraphic distributions of exceptionally preserved fossils—represent a subject of debate, as their relative importance remain unresolved. The preponderances of Konservat-Lagerstätten in particular facies and preserved via a limited number of taphonomic processes in the Ediacaran-lower Paleozoic suggest that the record of such assemblages reflects global controls on soft tissue conservation (Allison and Briggs, 1991; Butterfield, 2003). Whereas local and regional controls on preservational processes include climate, topography, (pore and bottom) water circulation, and sedimentation rate (Allison and Briggs, 1993), global controls are potentially related to diagenetic conditions (Butterfield, 1995; Muscente et al., 2015a) and seawater chemistry (Butterfield, 1995; Gaines et al., 2012b), including ocean oxygenation (Gaines and Droser, 2010). Geobiological agents, such as sediment-mixing animals and microbial mats, also influence exceptional preservation. These agents are geospatially variable in their local and regional distributions, and their evolution has shaped near-surface preservational environments

ACCEPTED MANUSCRIPT

on global scales (Muscente et al., 2015a; Schiffbauer et al., 2014b). Global controls, nonetheless, remain contentious (Butterfield, 2012; Pickerill, 1994) because, by and large, most studies have focused on particular assemblages (Schiffbauer et al., 2014b), regions (Muscente et al., 2015a), stratigraphic intervals (Butterfield, 2003), and preservational styles (Gaines et al., 2012b). As a result, efforts have not yielded a holistic understanding of soft tissue taphonomy.

One potential approach for addressing this knowledge gap involves studying the global distributions of exceptionally preserved fossil assemblages in geologic time and space in order to ascertain their relationships with geologic and environmental phenomena. Allison and Briggs (1993) conducted some of the pioneering work in this area by assembling a compilation of 44 exceptionally preserved faunas from the Phanerozoic and showing that the faunas occur most numerously in the Cambrian and Jurassic. Subsequent work by Schiffbauer and Laflamme (2012) brought the total number of Konservat-Lagerstätten in the compilation up to 53, revealing that they are actually most numerous in the Ediacaran and Cambrian. However, these compilations exclude numerous assemblages, which occur throughout the geologic record and have been described in numerous works (Baird et al., 1985b; Boag et al., 2016; Conway Morris, 1989; Muscente et al., 2015a). In addition, these efforts do not fully explore geologic overprints in the exceptional fossil record associated with variations in the quantities (areas and volumes) of sedimentary rocks, which affect overall geologic sampling (Allison and Briggs, 1993). Lastly, they fail to account for the geospatial clustering of assemblages (Plotnick, 2017), even though factors influencing fossil preservation operate throughout regions hundreds of kilometers in scale (Muscente et al., 2015a). In summary, by the modern standards of ‘big data’ paleobiology (Alroy, 2003), no comprehensive dataset has yet been compiled for rigorous analyses of temporal and spatial trends in Konservat-Lagerstätten.

ACCEPTED MANUSCRIPT

In this contribution, we examine the global, regional, and localized sedimentary conditions that allow for exceptional fossil preservation. We begin by reviewing the major taphonomic processes of soft tissue conservation. Through this review, we argue that a single unifying conceptual model—rooted in sedimentary geology and geomicrobiology—can account for virtually all preservational styles of soft tissue fossils, if their origins are considered in terms of taphonomic processes with unique and shared environmental controls. Using this model as basis for interpreting distributions of exceptionally preserved fossils in geologic time and space, we then present a quantitative meta-analysis of a new compilation of 694 exceptionally preserved fossil assemblages, which span the history of complex eukaryotic life (~635–0 Ma). These analyses reveal major patterns in fossil occurrence, which are pertinent to hypotheses regarding controls on exceptional preservation operating on regional and global levels. To conclude, we synthesize these results in light of our novel unifying conceptual model, and build a case that oxygen and bioturbation levels represent the principal global and regional controls on the exceptional preservation, owing to their influence over virtually all taphonomic pathways.

2. Environmental controls on taphonomic processes and pathways – a review

2.1. Taphonomic processes

Exceptional preservation of non-biomineralized tissues is, in essence, a race between fossil degradation and mineralization (Briggs, 2003), occurring in rare circumstances where tissues survive long enough to be transformed via *in situ* diagenetic polymerization into recalcitrant aliphatic components (Stankiewicz et al., 2000) and/or become replicated by authigenic/diagenetic minerals (Muscente et al., 2015a), which survive alteration over geologic time. The processes involved in fossil degradation and mineralization are not entirely

ACCEPTED MANUSCRIPT

independent, and microbes that utilize different metabolisms and are distributed in stratified zones within sediments play both constructive and destructive roles. Metabolisms that cause decay (e.g. sulfate reduction), for example, contribute to development of microenvironments in which minerals form in place of decomposing tissues. These preservational microenvironments, which geochemically differ in various ways from their immediate surroundings, may form internally or externally of carcasses (McNamara et al., 2009). Regardless, microenvironment development depends upon reductant and oxidant (i.e. electron donor and acceptor) availability, and ultimately, only leads to mineral formation if the chemical species that precipitate in response to microbial metabolisms are present in sufficient supply. As these geochemical variables hinge on global and regional factors (e.g. seawater chemistry and sedimentation rate), products of microenvironment development express patterns of broad environmental phenomena.

Aside from bioimmuration (Taylor, 1990) and preservation in amber (Martínez-Delclòs et al., 2004), the main taphonomic processes involved in conservation of soft tissues (Fig. 1) include silicification (Muscente et al., 2015a; Rice et al., 2002; Strang et al., 2016; Xiao et al., 2010), phosphatization (Briggs et al., 1993; Dornbos, 2011; McNamara et al., 2009; Muscente et al., 2015a; Schiffbauer et al., 2014a), pyritization (Briggs et al., 1991; Briggs et al., 1996; Guan et al., 2016; Schiffbauer et al., 2014b), aluminosilicification (Cai et al., 2012; Orr et al., 1998), and preservation of tissues with carbonate (i.e. calcite and siderite) minerals (Chen et al., 2014; Cotroneo et al., 2016). If these processes do not occur, soft tissues that survive early post-burial microbial degradation may be preserved as insoluble (i.e. kerogen) and/or soluble carbonaceous materials (Cai et al., 2012; Schiffbauer et al., 2014b; Stankiewicz et al., 2000), as observed in carbonaceous compressions (Xiao et al., 2002), small carbonaceous microfossils (Butterfield and Harvey, 2012), and organically preserved skeletal fossils (Muscente and Xiao, 2015a). Non-

ACCEPTED MANUSCRIPT

biomineralized tissues are also sometimes preserved in siliciclastic rocks as casts, molds, and impressions (Gutiérrez-Marco and García-Bellido, 2015; Muscente and Allmon, 2013), such as Ediacara-type fossils (Laflamme et al., 2011; Narbonne, 2005).

Notably, these taphonomic processes are exemplified by some well-known fossil assemblages (Butterfield, 2003). Such exemplary biotas show that the processes tend to conserve tissues of different types and degrees of recalcitrance, and capture anatomical information at varying levels of dimensionality (i.e. two versus three dimensions), scale (e.g. microscopic versus macroscopic features), and resolution (e.g. cellular detail versus bulk morphology; Xiao and Schiffbauer, 2009). Additionally, the exemplary biotas provide evidence that expressions of these processes vary through time and space. For example, depending on the environmental conditions and organisms, phosphatization may lead to preservation of "Doushantuo-type" or "Orsten-type" microfossils, distinguished by preservation of cellular structures in the former and recalcitrant tissues (e.g. cuticles) in the latter (Butterfield, 2003). Even so, Cambrian biotas with evidence of cellular and cuticular phosphatization bridge these end members (Xiao and Schiffbauer, 2009), and affirm that the different types of phosphatized fossils form under similar environmental conditions. Taking this observation into consideration, the processes described herein represent the fundamental mechanisms behind a variety of taphonomic pathways.

2.2. Taphonomic pathways: Combinations of processes

The preservation of a single fossil may involve one taphonomic process or many (Cai et al., 2012), and the processes involved in exceptional preservation may vary among tissues within a single specimen (Cai et al., 2012; McNamara et al., 2009), among specimens in a single assemblage (Guan et al., 2016; McNamara et al., 2012b; Schiffbauer et al., 2014b), or among

ACCEPTED MANUSCRIPT

assemblages located proximally in geologic time and space (Muscente et al., 2015a). Well-known fossils, which express evidence of multiple taphonomic processes, include the carbonaceous/pyritized/aluminosilicified fossils of the Ediacaran Gaojiashan Member (Dengying Formation) in South China (Cai et al., 2012); the carbonaceous/aluminosilicified fossils of the Burgess Shale (Orr et al., 1998); the phosphatized/calcified fossils of Libros in Spain (McNamara et al., 2012b; McNamara et al., 2009); the carbonaceous/phosphatized fossils of some Burgess Shale-type localities (Butterfield, 2002; Lerosey-Aubril et al., 2012); the phosphatized/pyritized/calcified fossils of the Eocene London Clay (Allison, 1988b); and the fossils of the Carboniferous Mazon Creek and Montceau-les-Mines biotas (Cotroneo et al., 2016; Perrier and Charbonnier, 2014). Other examples occur throughout the upper Neoproterozoic and Phanerozoic, such as phosphatized/silicified microfossils within chert nodules of the Ediacaran Doushantuo Formation in South China (Muscente et al., 2015a) and carbonaceous/pyritized/aluminosilcified fossils of the Cambrian Wheeler Formation in Utah (LoDuca et al., 2015; Fig. 2). Altogether, these fossils indicate that microenvironments conducive to different taphonomic processes develop under common global and regional circumstances. Considering that these microenvironments develop, in part, due to stratified microbial metabolisms (and availability of metabolites) within sediments, the processes themselves must proceed concurrently and/or sequentially as a fossil successively passes through each zone during its post-burial history (Schiffbauer et al., 2014b). Cross-cutting relationships provide a means to investigate these possibilities. Indeed, in some examples, the taphonomic processes evidently proceeded in sequence. In the phosphatized/silicified microfossils of the Doushantuo Formation, cell lumens were first phosphatized around the oxic-anoxic sedimentary boundary, where PO_4^{3-} concentrations were greatest, before the remaining tissues were silicified

ACCEPTED MANUSCRIPT

in deeper anoxic sediments with pH conditions that fostered calcite dissolution and chert precipitation (Muscente et al., 2015a). Similarly, in the case of carbonaceous/pyritized/aluminosilicified fossils, pyritization must have occurred prior to conversion of carbonaceous material into recalcitrant compounds via polymerization, as the microbial sulfate reduction (MSR) metabolisms that drive pyrite formation and pyritization require labile organic matter for fuel. Sulfate reduction, nevertheless, plays a significant role in all taphonomic processes. In addition to strongly influencing the overall rate of organic degradation (Allison, 1988a)—and therefore, the likelihood of organic matter preservation—MSR causes pyrite and carbonate precipitation (Briggs et al., 1996; Schiffbauer et al., 2014b), and represents the most efficient process of phosphorus remineralization (Arning et al., 2009), yielding PO_4^{3-} that may fuel phosphatization. In addition, MSR can cause pH changes, which may contribute to phosphatization (Briggs and Wilby, 1996), silicification (Xiao et al., 2010), and/or aluminosilicification (Gabbott et al., 2001). Consequently, under appropriate chemical conditions, a multitude of taphonomic processes may concurrently occur in the MSR zone (Briggs, 2003).

Fossils preserved via multiple taphonomic processes provide evidence for a single unifying conceptual model, rooted in sedimentary geology and geomicrobiology, which accounts for all variations in preservational style among fossils that are not preserved via bioimmuration or in amber (Fig. 3). Overall, in this model, taphonomic pathways differ in several aspects of post-burial history, including the durations of the fossils in the various microbial zones of the sediment column and the availabilities at each level of geochemical species (e.g. O_2, organic matter, reactive iron, SO_4^{2-}, H_2S, PO_4^{3-}, SiO_2, etc.) essential for fossil degradation and mineralization. In general, these variables hinge on sedimentation rates, which control the rate of

ACCEPTED MANUSCRIPT

fossil burial through each zone (Schiffbauer et al., 2014b), and on geomicrobiological agents (e.g. burrowing animals, microbial mats, etc.) that modulate geochemical gradients, influence microbial zone thicknesses, and play destructive and constructive roles in soft tissue preservation (Callow and Brasier, 2009). Bottom water chemistry, which affects pore water geochemical gradients and the location of the redox boundary, also influences post-burial history (Guan et al., 2016; Muscente et al., 2015a). In this context, an organism may be preserved as a fossil typifying a combination of processes (Fig. 2) or as an end-member epitomizing just one (Fig. 1), depending on its pathway through preservational zones of the sedimentary column.

2.3. The dichotomy of the exceptional fossil record

In exploring the stratigraphic and geographic distributions of exceptionally preserved fossils, a conspicuous dichotomy presents itself. On the one hand, each taphonomic process has its own unique requirements (Briggs et al., 1996; Gaines et al., 2012b; Guan et al., 2016; Muscente et al., 2015a; Schiffbauer et al., 2014b; Tarhan et al., 2016), and the conditions conducive to the taphonomic processes can develop in isolation. In this light, the temporal and spatial distributions of the fossils exemplifying each of the major processes are independent of each other. Indeed, separate and generally unrelated aspects of seawater chemistry may account for the different albeit overlapping stratigraphic distributions of silicified (Knoll, 1985), phosphatized (Butterfield, 2003), and Burgess Shale-type fossils (Gaines et al., 2012b). On the other hand, certain factors—oxidant availability, microbial mats, animal scavenging, and sediment mixing—influence virtually all taphonomic pathways, owing to their effects on the durations of fossils in the various microbial zones of the sedimentary column and on the relative degrees of soft tissue decay and mineralization at each level. Consequently, fossils of all

ACCEPTED MANUSCRIPT

preservational styles may track those sedimentary factors through space and time, occurring most commonly in facies deposited when circumstances favored mineralization by any means over total degradation. Taking this dichotomy into consideration, phenomena affecting all taphonomic pathways should manifest in the stratigraphic and geographic distributions of Konservat-Lagerstätten, even though fossils of distinct preservational styles are conserved in partly different environmental conditions. In the remainder of this contribution, we examine the fossil record for overarching trends through space and time in occurrences of exceptionally preserved fossils, regardless of their taphonomic pathways and preservational styles. We recognize a number of trends, which in light of our unifying conceptual model, illuminate the environmental factors that influence virtually all taphonomic pathways and exert primary (first-order) control on the temporal and spatial distributions of Konservat-Lagerstätten.

3. Data and methods

If global phenomena control exceptional fossil preservation and sampling, we predict that the numbers, geospatial spreads, and depositional settings of Konservat-Lagerstätten follow paleoenvironmental and geologic changes through stratigraphy. To test this prediction, we compiled time series data on the frequency of exceptional preservation, and compared the time series to trends in the rock record. Accordingly, we devised methods for examining the geospatial and chronostratigraphic distributions of Konservat-Lagerstätten, which record patterns in development of exceptional preservational conditions through space and time. Besides studying these distributions, we assessed for biases—for instance, whether Konservat-Lagerstätten of similar ages and facies tend to cluster together in stratigraphy and geography, as expected if regional factors control exceptional preservation—and investigated for possible

ACCEPTED MANUSCRIPT

relationships with sampling, rock availability, and sea level. Additional information on the methods, as well as supplementary data, can be found with the online version of the report.

In devising our methods, we aimed to take advantage of the rich literature available pertaining to Konservat-Lagerstätten. Of course, there is no assurance that any two Konservat-Lagerstätten are directly comparable. Because there are no standards for defining and sampling these deposits, different workers have incorporated varying amounts of faunal, facial, geologic, and taphonomic variation in lumping and splitting them. We have attempted to rectify this issue by (1) sampling stratigraphically and geographically distinct points of fossil collection from the literature (see Section 3.1), and (2) lumping these fossil "assemblages" into equivalent groups based on their ages, geographic proximities, and depositional settings (see Section 3.5). By following this approach, we have generated a standardized dataset that provides an overview of the exceptional fossil record and captures the broadest and most fundamental patterns in fossil occurrence and sampling. Data processing yielded units with comparable temporal, environmental, and geospatial coverages. Using the standardized units as the building blocks of numerous time series, we conducted statistical analyses (see Section 3.7) to test our main predictions. We did not take variations in preservational style into consideration in identifying the standardized units because local and microenvironmental conditions strongly influence processes of soft tissue mineralization (McNamara et al., 2009). In addition, fossils of distinct styles often occur in geographic and stratigraphic proximity (Cai et al., 2012; Muscente et al., 2015a; Wang et al., 2012), indicating that regional and perhaps global phenomena allow for soft tissue conservation via numerous taphonomic pathways. On the whole, our work aims to identify patterns in the fossil record produced by such broadly influential phenomena, as they likely exert the greatest effects on the temporal and spatial distributions of exceptional preserved fossils.

ACCEPTED MANUSCRIPT

3.1. Overview of dataset of exceptionally preserved fossil assemblages

Assemblages—representing beds, facies, localities, sections, formations, and regions—were sampled from 650 primary and secondary literature references (see Supplementary Database and references therein) and included in the dataset if they have been sources of exceptionally-preserved animal, algal, and/or Ediacara-type fossils. The dataset includes the best age estimate, geographic location (lat/long and Cartesian coordinates), paleolatitude, and litho- and chronostratigraphic units of each assemblage (see Supplementary Information). The smallest units, given available information, were generally selected for inclusion in the dataset. As a result, in some cases, deposits often recognized as singular examples of Konservat-Lagerstätten (Seilacher, 1970) were entered in the dataset as multiple points of fossil collection. This strategy helps to account for the broad geospatial coverages of some Konservat-Lagerstätten. For instance, the Green River Formation Lagerstätte—which encompasses numerous localities in Wyoming, Colorado, and Utah (Bradley, 1964)—is included in the dataset as representative assemblages, which cover the geographic range of the formation. Similar examples include the Burgess Shale (Collins et al., 1983), Chengjiang (Hou and Bergström, 2003), and Jehol biotas (Pan et al., 2013). Due to limitations in published data, other Lagerstätten—known from multiple beds, localities, and sections—were entered in the dataset as singular units (i.e. formations or regions). For example, although fossils have been collected from >200 collection areas in a range of approximately 50 km around Morris, Illinois (Baird et al., 1985a), the geographic coordinates and fossils of each collection area are not reported in the literature. Consequently, the collection areas are included in the dataset as a single Mazon Creek assemblage. Similar examples include the Permian KwaZulu (van Dijk, 1997), Triassic Buntsandstein (Bashkuev et al., 2012), and

ACCEPTED MANUSCRIPT

Jurassic Hartford Basin (Huber and McDonald, 2003) assemblages. Future work will undoubtedly allow for improvement of the dataset over time.

The dataset does not include assemblages of fossils preserved via bioimmuration or in amber because the records of these assemblages reflect evolutionary trends among biomineralizing animals (Taylor, 1990) and embryophytes (Martínez-Delclòs et al., 2004), respectively. In addition, the dataset does not include assemblages made up exclusively of weakly biomineralized animal exoskeletons, embryophyte plants, hemichordates, chitinozoans, and/or putative hydroids for the following reasons. Because biominerals encapsulate and protect organic matter (Briggs, 1999), weakly biomineralized exoskeletons have higher preservation potentials than non-biomineralized tissues. Embryophytes do not occur until the middle Paleozoic and are mostly restricted to non-marine settings (Taylor et al., 2009). Hemichordate and chitinozoan fossils occur only in the Paleozoic (Clarkson, 1998), and include clades with planktonic and cosmopolitan lifestyles that influence the preservation and geospatial distributions of their assemblages (Elles, 1939; Finney and Berry, 1997; Rombouts, 1982). Many putative hydroids from the Paleozoic may be hemichordates, and those from the Mesozoic and Cenozoic are chiefly preserved through bioimmuration (Muscente et al., 2016).

3.2. Binning by age and depositional setting

Assemblages were assigned—based on descriptions of their paleoenvironments in the literature—to mutually exclusive marine, transitional, and non-marine depositional setting categories, as well as an indeterminate category for assemblages that are not well constrained. Additionally, assemblages from marine and transitional environments were assigned to a combined marine/transitional category. Based on their ages, assemblages were assigned to time

ACCEPTED MANUSCRIPT

bins comprising two time series: a time series of 30.5 (n=20) million year (my) equal duration time bins and another of 20.3 (n=30) my equal duration time bins. We selected these durations because their corresponding time series identically span the shortest geochronological interval (613–3 Ma) that encompasses all assemblages in the dataset albeit with different numbers of time bins (n=20 for 30.5 my; n=30 for 20.3 my). Based on their time bins, assemblages were additionally assigned to time domains, which were compared via statistical difference testing (see Section 3.7).

3.3. Comparison with sampling patterns in the PBDB

The latitudinal, longitudinal, and paleogeographic distributions of the assemblages in the dataset were compared to the corresponding distributions of (principally skeletal) collections in the PBDB in order to assess for relative geographic sampling biases (see Supplementary Information, Figs. S1, S2). Whereas each assemblage signifies a stratigraphic and geographic point of fossil collection, each PBDB collection represent a sample acquired from a similar collection point. Hence, while the units are not strictly equivalent, they are broadly analogous. Collections were sampled from the PBDB in a controlled manner, such that, their distribution among geologic systems and depositional settings identically matches the corresponding distribution of assemblages in our new dataset. Assemblages and collections were assigned—based on their geographic locations and estimated ages and paleolatitudes—to latitudinal (n=12), longitudinal (n=24), and age/paleolatitudinal (n=144) bins. Latitudinal and longitudinal bins each correspond to 15 degrees of geographic range (e.g. 0–15°, 15–30°, etc.), and the 144 age/paleolatitudinal bins include 12 groups—corresponding to the 12 geologic systems—that consist of 12 bins, each representing 15 degrees of paleolatitude range. For each bin type, the

ACCEPTED MANUSCRIPT

mean counts of collections were compared to the corresponding counts of exceptionally preserved assemblages through correlation analysis.

3.4. Nearest neighbor analysis and geospatial statistics

We employed nearest neighbor and geospatial statistical analyses to explore for patterns in the geographic distribution of exceptionally preserved fossil assemblages. For the nearest neighbor analysis, we calculated the distances among all pairs of assemblages, as measured between their Cartesian coordinates, and identified all unique assemblage-nearest neighbor pairs as well as the distances, chronostratigraphic ages, and depositional settings of the assemblages. To assess if assemblages occur heterogeneously in their geographic distribution and/or tend to occur near others of similar age and depositional setting, we calculated several descriptive geospatial statistics (see Supplementary Information). Ripley's L(d) was calculated for various subsets of the dataset in order to assess whether the assemblages are uniform, random, or clumped in their geographic distribution, relative to the total dataset or some subset. Ignoring sample size effects, L(d) is generally equal to zero for randomly distributed points, greater than zero for uniformly spaced points; and less than zero for clumped points. To assess significance, we performed randomization tests and produced confidence intervals (CIs). Observed values above and below the CI represent significant evidence of assemblage regularity and clustering, respectively. Moran's I, which provides a measure of spatial autocorrelation (i.e. correlation in signal among proximally-located points), was calculated to assess spatial autocorrelation in geologic age amongst nearest neighbors and amongst points located at various relative distances between 0 and 2000 km. As with traditional correlation coefficient (R) calculations, values

ACCEPTED MANUSCRIPT

greater than 1 and less than 1 signify positive and negative correlations, respectively. To assess the significance of these values, we determined Z values and their corresponding probabilities.

3.5. Hierarchical clustering

Geospatial statistics indicate that the assemblages commonly occur in groups (or clusters) with similar geographic locations, ages, and environmental settings (see Section 4.2). To standardize the data so that clusters of comparable temporal, geospatial, and environmental coverages are the sampling units in the analyses, we hierarchically clustered the assemblages within each time bin in three-dimensional space based on their geographic locations (i.e. the pairwise Euclidean distances among their Cartesian coordinates), and counted in each time bin the number of geospatially distinct clusters, including "clusters" signifying solitary collection points (see methods below, Fig. S3). From the counts, we determined proportions of marine, transitional, non-marine, and marine/transitional clusters, and altogether, generated assorted time series of counts and proportions of clusters for hypothesis testing. Herein, we primarily focus on the results of clustering assemblages within equal duration time bins, as opposed to within periods, epochs, or ages. Everything else being equal, it is expected that there is a positive between assemblage count and geochronological bin duration. Thus, by clustering within equal duration time bins, we have corrected for the influence of bin duration on the results. Of course, assigning data points with age uncertainties to time bins, which do not correspond exactly to correlative geologic units, can introduce errors into datasets. To confirm that these potential errors did not significantly affect the results, we clustered assemblages within geologically defined chronostratigraphic intervals (systems, series, and sets of series and biozones) and comparatively assessed the resulting time series (see Supplementary Information).

ACCEPTED MANUSCRIPT

Overall, this hierarchical clustering work takes into consideration the regionalized nature of exceptional preservational conditions (Muscente et al., 2015a), and mitigates effects of incomplete and geospatially uneven sampling, as assemblages variably corresponding to beds, facies, localities, sections, and formations are all combined into clusters (geospatially equivalent to regions), which to a degree, encompass missing data (i.e. assemblages that have similar ages and locations to those in the dataset but have not yet been discovered or reported). Large clusters more comprehensively account for missing data than small ones, as they encompass broader areas. Cluster size, however, trades off with total cluster count (i.e. sample size), and thus, can affect results of hypothesis testing. In addition, clusters of sizes larger than typical depositional and tectonic systems (e.g. basins) may include genetically-distinct deposits. If so, their counts may not accurately reflect natural patterns of exceptional preservation. To account for these issues, we compiled counts of clusters of various sizes (i.e. geographic ranges) for each time bin, and assessed changes in the results (i.e. test statistics and *P* values) as a function of cluster size. Using this approach, we searched for trends in time series, which do not vary with cluster sizes in the ranges of typical depositional and tectonic systems (10s to 100s of kilometers in scale). Trends apparent in all time series, regardless of cluster sizes in this range, represent robust patterns in fossil occurrence and sampling of potential significance.

In this contribution, we quantify cluster size using a metric—maximum potential linkage distance (MPLD)—corresponding to the greatest possible Euclidean distance that may separate points or groups within a cluster, given the method used in hierarchical clustering. This value equals the length of each branch of equal measure (or "height") cut from a dendrogram produced via hierarchical analysis. Because various potential criteria exist for combining groups into larger clusters, the relationship between MPLD and actual geographic range slightly varies with

ACCEPTED MANUSCRIPT

method. For example, in clusters identified using the complete-linkage distance method, clusters represent non-overlapping circular regions with diameters equal to their MPLDs, as groups are combined via this method based on the distances among their most distant points (Milligan, 1980; Milligan and Cooper, 1987). Conversely, for the centroid method, the MPLD of a cluster may be shorter than its geographic range because the distances among groups are measured between their means (i.e. central positions) and points on opposite ends of a cluster may be separated by greater distances than the centroids of their local groups (Milligan, 1980; Milligan and Cooper, 1987). Nonetheless, for any given method, two groups are considered distinct clusters if they are separated by a Euclidean distance greater than a particular MPLD benchmark, and the geographic ranges of clusters generally increase with their MPLDs. As the clusters at a given MPLD represent broadly equivalent geospatial areas, their count represents an approximation of geospatial spread (i.e. a high number corresponding to widespread occurrence).

Hierarchical clustering analyses were performed in JMP (Sall et al., 2012). For most analyses, we determined counts of clusters with specific (n=15) MPLDs ranging from 0 to 4000 km. Analyses were repeated using procedures varying with regard to the dataset fractions, time bin durations, and depositional setting categories used in cluster counting. Independent analyses were conducted on the entire dataset (i.e. the global dataset) and the North American subset. In analyses of the global dataset, assemblages within 30.5 my and 20.3 my duration time bins were hierarchically clustered in several ways, such that, assemblages were clustered (1) independently of depositional setting category; (2) according to three (marine, transitional, and non-marine) depositional setting categories; and (3) according to two (marine/transitional and non-marine) depositional setting categories. Analyses in which assemblages were clustered independently of depositional setting category and according to marine, transitional, and non-marine categories

ACCEPTED MANUSCRIPT

were repeated using five clustering methods: complete-linkage, average, centroid, single-linkage, and Ward's methods (Milligan, 1980; Milligan and Cooper, 1987). Assemblages were clustered according to the marine/transitional and non-marine categories using the complete-linkage method. In analyses of the North America subset, assemblages within 30.5 and 20.3 my duration time bins were clustered using the complete-linkage method independently of depositional setting category; according to the marine, transitional, and non-marine categories; and according to the marine/transitional and non-marine categories.

3.6. Estimation of rock quantity and past sea level

We compiled data regarding rock quantity and past sea level for comparison with the time series of cluster counts. The global sedimentary rock outcrop areas for the time bins in this study were estimated from geologic maps in the Geological Atlas of the World (Choubert and Faure-Muret, 1976). Time bins' global marine and non-marine sedimentary rock volumes were estimated from published data on the major lithologic associations of the Phanerozoic (Ronov, 1982). The surface/subsurface areas of North American marine/transitional and non-marine rocks in the time bins were estimated using data accessed from the Macrostrat database (https://macrostrat.org/classic/about.php). Lastly, for each time bin used in this study, we determined the minimum and maximum global sea levels (measured in meters relative to present day sea level) along with the levels at the bins early and late boundaries, as reconstructed in published sea level curves (Haq and Schutter, 2008; Hardenbol et al., 1999). From these values, we estimated mean global sea levels (see Supplementary Information).

3.7. Statistical testing of cluster counts and proportions

ACCEPTED MANUSCRIPT

To test for differences among the time bins and time domains, data were compared through Pearson's chi-square (equality and homogeneity) testing, randomization testing, and Fisher's exact testing. The tests were performed in JMP (Sall et al., 2012) and Microsoft Excel. Chi-square equality tests were performed to compare all the time bins in each time series and to assess whether, in terms of cluster count, they are significantly unequal. Randomization tests were performed in Excel to corroborate chi-square equality testing results and to assess if observed trends in cluster count through geologic time could be explained as random variations or sampling artefacts (Huntley et al., 2006). Additional chi-square tests were performed on pairs of time domains to assess whether, in terms of their proportions of clusters from various settings, their differences are significant. As some of the chi-square homogeneity tests failed the standard sample size rule (i.e. 80% of the expected counts are >5 and all expected counts >1 in contingency tables used in calculation of test statistics), Fisher's exact tests were also performed on pairs of time domains to assess whether differences among their proportions are significant.

Linear polynomial regression models were fit to the time bins' proportions of clusters (y) and ages (x) in order to assess whether the proportions of clusters from marine, transitional, marine/transitional, and non-marine depositional settings significantly change through time. To assess the fits of the models, *P* values were tabulated through analysis of variance (ANOVA) *F*-testing. The fit of any regression model to the data (except for linear model with slope = 0) represents significant evidence that the proportion of clusters changes through time. To assess the accuracy of the *F*-tests, which assume that residuals are normally-distributed, Shapiro-Wilk and Jarque-Bera tests were performed on the residuals in the regression analyses.

The time bins' cluster counts were plotted against their corresponding rock areas, rock volumes, and sea levels to assess for potential correlations among these variables. Analyses

ACCEPTED MANUSCRIPT

involving non-marine cluster counts did not include time bins older than 430 Ma, which pre-date the earliest robust fossil evidence of widespread terrestrial plant (Kenrick and Crane, 1997) and animal (Shear and Selden, 2001) life. In addition, analyses involving global rock volume and sea level did not include time bins older than 531.7 Ma and Ma 552 Ma, respectively, due to data limitations. The analyses were repeated for time series based on various clustering methods, time bin durations, MPLDs, and data categories used in cluster counting. To verify that the statistically significant results of this work are robust to autocorrelation, the analyses were repeated with data subjected to first differencing (see Supplementary Information).

4. Results

The dataset provides evidence for several striking patterns in the distribution of Konservat-Lagerstätten in geologic time and space. First, according to the raw data but regardless of time bin duration, assemblages are most common in the Ediacaran, Cambrian, lower-middle Cretaceous, Paleogene, and Neogene, and relatively rare in the Devonian, upper Triassic, and uppermost Cretaceous (Fig. 4a). Second, the data suggest that the depositional settings of the deposits changed through time with marine assemblages occurring most abundantly in the Ediacaran-lower Paleozoic interval, transitional assemblages occurring most abundantly in the middle-upper Paleozoic, and non-marine assemblages occurring most abundantly in the Mesozoic and Cenozoic (Fig. 4b, c). In addition, the data indicate that marine and transitional assemblages increase in frequency in the Jurassic and lower-upper Cretaceous. Lastly, the assemblages in the dataset show a marked change in paleolatitudinal range through the Phanerozoic (Fig. 4d, e). The assemblages occur at high and low (northern and southern)

ACCEPTED MANUSCRIPT

latitudes in the Ediacaran-Ordovician, but concentrate at equatorial latitudes in the middle-late Paleozoic and at high northern latitudes in the Mesozoic and Cenozoic.

4.1. Comparison with global trends in skeletal fossil collection

The latitudinal, longitudinal, and age/paleolatitudinal distributions of the assemblages in the dataset broadly resemble those of principally skeletal fossil collections randomly sampled from the PBDB (Fig. 5, see Supplementary Information Fig. S2). Most assemblages and collections have been sampled from the northern hemisphere at latitudes between 30 and 60 degrees (Fig. 5a, b), and analyses indicate a strong positive correlation (Fig. 5e; *F*-test *P*: 0.001; R: 0.930; R^2: 0.860) between the assemblage and collection counts in the latitudinal bins. Likewise, the assemblages and collections follow similar tri-modal longitudinal distributions (Fig. 5a, b)—with peaks in the -120 to -75 degree, -15 to 30 degree, and 105 to 120 degree ranges—that broadly correspond to the longitudinal ranges associated with maximum landmass (i.e. North and South America; Europe and Africa; and eastern Asia and Australia). Their counts in the longitudinal bins are also positively correlated (Fig. 5f; *F*-test *P*: 0.001; R: 0.830; R^2: 0.680). In conjunction with these trends, the assemblages and collection have similar distributions in terms of their age and paleolatitude (Fig. 5c, d). Correlation analyses (Fig. 5g; *F*-test *P*: 0.001; R: 0.760; R^2: 0.580) indicate that the occurrences of the assemblages and collection are broadly congruent in terms of age and paleolatitude. Therefore, no broad scale differences exist among the geospatial distributions of exceptionally preserved and skeletal fossils.

4.2. Nearest neighbor analysis and geospatial statistics

ACCEPTED MANUSCRIPT

The assemblages in the dataset commonly occur in clusters. These clusters are evident in maps of assemblage localities (see Supplementary Information, Fig. S1, and Supplementary Animations). The majority (61%) of assemblages occur within 50 km of their nearest neighbors (Fig. 6a), and half of all assemblage-nearest neighbor pairs include points from matching chronostratigraphic series and depositional settings (Fig. 6b-d). Spatial autocorrelation analyses provide significant evidence of autocorrelation in age for nearest neighbors, and shows that similarity in age between assemblages generally increases with geographic proximity between 0 and 800 km distance (Fig. 6e). In addition, Ripley's L(d) function analyses provide significant evidence that the assemblages occur in clusters (Fig. 6f). Although all subsets of the dataset exhibit evidence of clustering, regardless of data category, the analyses indicate that the assemblages most commonly cluster with those from matching depositional settings with the transitional assemblages exhibiting the greatest degree of intraspecific categorical association, particularly at great distances.

4.3. Cluster count time series

The number of clusters of exceptionally preserved fossil assemblages varies through the Ediacaran-Neogene interval. A general trend persists in all time-series of counts of clusters with MPLDs between 50 and 1000 km (Fig. 7) irrespective of clustering method (see Supplementary Information, Figs. S4, S5). This trend is evident, regardless of whether the assemblages are clustered within equal duration time bins (Fig. 7a, b, i–l) or chronostratigraphic interval (Fig. S6a–d). In all cases, the total number of clusters increases from the lower Ediacaran to the Cambrian, declines from the lower to middle Paleozoic, and generally rises from the Devonian to the Neogene with notable drops in total cluster count observed in the Upper Triassic and

ACCEPTED MANUSCRIPT

uppermost Cretaceous. These trends reflect underlying patterns in occurrences of marine, transitional, and non-marine assemblages (Fig. 7c–h). In general, marine clusters increase in number from the Ediacaran to Cambrian before declining in number through the lower Paleozoic; clusters signifying transitional depositional settings occur sporadically throughout the Phanerozoic, but peak in the Silurian, Carboniferous, and upper Mesozoic; and non-marine clusters first appear in the Silurian and rise in number through the Phanerozoic. Assemblages of all categories are rare in the Upper Triassic and late Cretaceous. Chi-square equality testing shows that the time bins are significantly unequal with respect to cluster count ($P < 0.05$ for analyses with MPLDs < 1000 km, Fig. 8), and randomization tests indicate that the general trend in total cluster counts cannot be explained by random variations or sampling artefacts (Fig. 9).

Correlation analyses of the time series data—comparing marine, non-marine, and marine/transitional cluster counts to rock areas and volumes (Fig. 10)—provide statistically significant evidence for correlations between some pairs of these variables (Fig. 11). Correlations supported by *F*-test *P* values < 0.05 for all MPLDs and time bin durations indicate that non-marine cluster counts are positively correlated (generally R > 0.65) with the areas and volumes of non-marine rocks (Fig. 11a–c, j–l). Global outcrop area accounts for more variation (R^2 generally > 0.7 and > 0.55 for the 30.5 and 20.3 my duration time bins, respectively) in non-marine cluster count than global rock volume (R^2 generally > 0.6 and > 0.45 for the 30.5 and 20.3 my duration time bins, respectively). North American area of non-marine rock accounts for comparatively less variation (R^2 generally > 0.45 and > 0.45 for the 30.5 and 20.3 my duration time bins, respectively) in North American non-marine cluster count (Fig. 11a–c, j–l), but these low R^2 values may reflect the smaller size of the North American dataset. Regardless, these data indicate a positive correlation exists between the quantity of non-marine rock and the number of

ACCEPTED MANUSCRIPT

exceptionally preserved non-marine assemblages. First difference cross-plots of these variables for 30.5 my time bins corroborate this correlation (see Supplementary Information, Fig. S7), and the same significant correlations are observed, even when assemblages are clustered within geologically defined chronostratigraphic intervals (Fig. S8).

Correlations between marine/transitional cluster count and global marine rock volume—which are supported by *P* values < 0.05 and generally have R values > 0.5 for all time bin durations and MPLDs between 300 and 1000 km—provide some statistical evidence for a relationship between those variables (Fig. 11 g–i, p–r). Yet, the R^2 values of these correlations indicate that global marine rock volume accounts for < 40% of the variation in marine/transitional cluster count. The corresponding values from correlations between marine cluster count and global marine rock volume as well as from correlation analyses comparing North American marine and marine/transitional cluster counts to North America marine/transitional rock area also provide evidence for positive correlations among the variables (Fig. 11d–i, m–r), but in all cases, rock quantity accounts for only a minor fraction ($R^2 < 0.35$) of the variation in cluster count. As opposed to these results, those from correlation analyses comparing global marine and marine/transitional cluster counts to global marine rock outcrop area did not yield significant results (Fig. 11d–i, m–r; *F*-test *P* > 0.05 for all MPLDs, regardless of clustering method and time bin duration).

Correlation analyses comparing marine, non-marine, and marine/transitional cluster counts (based on the complete-linkage method and various MPLDs and time bin durations) to mean global sea level indicate that no relationship exists between cluster count and sea level (*F*-test *P* > 0.05 for all correlations and MPLDs; Fig. 11). Sea level does not account for variation in cluster count at any MPLD. Results of correlation analyses comparing marine and

ACCEPTED MANUSCRIPT

marine/transitional cluster counts to mean sea level have R^2 values < 0.05 (R: -0.2–0) for all MPLDs, regardless of time bin duration. In comparison, correlation analyses comparing non-marine cluster counts to mean sea level have greater R^2 values (0.11–0.16 and 0.05–0.07 for the 30.5 and 20.3 my duration time bins, respectively, and all MPLD between 300 and 1000 km), but the corresponding *F*-test *P* values > 0.05 suggest these results are not significant.

4.4. Proportion time series

Proportion time series indicate that, like the assemblages, the clusters shift from predominantly marine facies in the Ediacaran-Ordovician interval (Fig. 12a, b) to transitional facies in the Silurian-Permian interval (Fig. 12c, d) and non-marine facies in Triassic-Neogene interval (Fig. 12e, f). This result does not vary with clustering method (see Supplementary Information, Figs. S6, S9–S11). Pairwise homogeneity tests show that the three time domains are significantly different in terms of proportions of clusters in each depositional setting category (Chi-square and Fisher's Exact tests, $P < 0.05$ for all analyses, Fig. 13a–c, f–h, k–m, r–t). Furthermore, regression modeling shows that the marine (Fig. 12a, b) and non-marine (Fig. 12e, f) proportions decrease and increase through geologic time, respectively, while the transitional proportion (Fig. 12c, d) peaks in the upper Paleozoic (*F*-test $P < 0.05$ for all analyses). Although exceptionally preserved fossil assemblages primarily occur in non-marine facies in the Triassic-Neogene interval, stratigraphic intervals in the Jurassic (~186–145 Ma) and Cretaceous (~104.7–84.3 Ma) include greater marine/transitional and lesser non-marine proportions than predicted by regression models (Fig. 14a, b). Pairwise homogeneity tests also suggest these intervals include significantly greater marine/transitional and lesser non-marine proportions than other Triassic-

ACCEPTED MANUSCRIPT

Neogene time domains (Chi-square and Fisher's Exact tests, $P < 0.05$ for MPLDs < 800 km, Fig. 13d, e, i, j, n–q, u–x).

5. Discussion

Our dataset of exceptionally preserved fossil assemblages represents a major expansion of previous compilations in terms of size and number of data fields (Allison and Briggs, 1993; Hendy, 2011; Schiffbauer and Laflamme, 2012). The global geographic and paleogeographic distributions of the assemblages in the dataset broadly resemble those of skeletal fossils, suggesting that the dataset does not contain significant geospatial biases relative to common trends in fossil sampling. Like skeletal fossils (Plotnick, 2017), the assemblages themselves commonly occur in clusters with others of similar age and matching depositional setting. Well-known examples of these clusters occur in the Ediacaran of South China (i.e. the Doushantuo and Lantian formations, (Muscente et al., 2015a), Cambrian of British Columbia (i.e. the Burgess Shale and related deposits, (Conway Morris, 1989), Cambrian of the western US (i.e. the Wheeler, Marjum, and related deposits, (Conway Morris, 1989), Carboniferous of the mid-continental US (i.e. the Mazon Creek and related deposits, (Baird et al., 1985b), Jurassic of Germany (i.e. the Solnhofen and related deposits, (Ebert et al., 2015), and the Eocene of the western US (i.e. the Green River Formation, (Bradley, 1964). Of course, numerous other clusters occur around the world and throughout the geologic record of complex life. The results of geospatial statistical analyses (Fig. 6) as well as maps showing occurrences of assemblages through time (see Supplementary Information, Fig. S1, and Animation S1 and S2) indicate that these clusters are generally tens to hundreds of kilometers in scale, though their full extents remain ambiguous, given incomplete exposure of geologic units and the challenges of evenly and

ACCEPTED MANUSCRIPT

completely sampling assemblages from primary literature sources, which often do include exact geospatial coordinates. Geospatial statistics, nonetheless, provide some constraints on the maximum sizes of the clusters. At distances greater than 800 km, assemblages demonstrate negligible evidence of spatial autocorrelation in age (Fig. 6e), and at distances greater than 1000 km, marine assemblages do not exhibit significant evidence of clustering (Fig. 6f). Thus, the geospatial statistics suggest that clusters of related assemblages only rarely exceed 800 km, and indicate that efforts to quantify and interpret changes in exceptional preservational conditions through time should focus on counts and proportions of clusters of sizes no more than 1000 km.

Our hierarchical clustering represents a "lumping approach," providing relatively conservative estimates of cluster counts and proportions. The work notably does not factor in differences in preservation, instead grouping assemblages into clusters, regardless of their taphonomic characteristics. Our review of environmental controls on preservational processes shows that assemblages preserved via different taphonomic processes do not necessarily represent unrelated environmental factors, as certain conditions can contribute to exceptional preservation via several preservational pathways. Indeed, stylistic differences among proximally arranged assemblages may actually represent alternative expressions of regional or global phenomena, which have been modulated by local and microenvironmental conditions. In any case, consideration of preservational styles would not have affected the raw assemblage counts and proportions, but it would have increased cluster counts and shifted proportions in favor of time bins and depositional settings with the greatest varieties of taphonomic pathways. We predict that, if we took preservational styles into account in clustering, the main differences in the results would pertain to the upper Neoproterozoic-lower Paleozoic interval. Assemblages in this interval constitute a number of fully marine "taphonomic windows" (Allison and Briggs,

ACCEPTED MANUSCRIPT

1991), each corresponding to a particular style of preservation (Butterfield, 2003). These windows generally closed between the lower and middle Paleozoic (Butterfield, 2003; Xiao and Schiffbauer, 2009). In the upper Paleozoic-Cenozoic interval, assemblages sometimes reflect multiple preservational processes (Allison, 1988b; Wang et al., 2012), but infrequently exemplify taphonomic processes other than carbonaceous compression (Muscente and Xiao, 2015b) or carbonate mineralization (Baird et al., 1985b; Cotroneo et al., 2016). Thus, had preservational styles been considered in hierarchical clustering, it is likely that the upper Neoproterozoic-lower Paleozoic interval would have been disproportionately affected. Future work should address these issues. In the meantime, it seems that consideration of taphonomic pathways would not have affected the overarching trends observed in cluster counts and proportions through geologic time and space (Figs. 7, 12, 14), as those trends broadly follow the raw data (Fig. 4).

5.1. Explanations for clustering of assemblages

Three non-exclusive explanations can account for the clustering of assemblages in our dataset. First, at the broadest scale (1000s of kilometers), the clustering likely signifies global-scale patterns in geologic sampling (Plotnick, 2017). Because assemblages have predominantly been sampled from surface outcrops in the northern hemisphere, they may appear to cluster together. This explanation, nonetheless, does not account for clustering of assemblages at local to regional scales (10s to 100s of kilometers). Second, because the assemblages represent outcrops of lithological packages with shared tectonic and erosional histories, the clustering of assemblages may reflect patterns in the exposures of rocks available for geologic sampling (Plotnick, 2017). Accordingly, the number of assemblages in each cluster may correspond, in

ACCEPTED MANUSCRIPT

part, to the amount of exposure and geospatial size of its hosting lithological package. However, the availability of rock alone does not guarantee occurrence of exceptionally preserved fossils. Thus, the clustering of assemblages lastly indicates that, in the past, exceptional preservational conditions sometimes arose repeatedly over millions of years throughout laterally extensive regions (e.g. basins). In these cases, regional factors influencing taphonomic processes most likely allowed for conservation of soft tissues, as the recurrent conditions are suggestive of common causes. Of course, the relative significance of each of these explanations for explaining patterns in occurrences of assemblages may vary from cluster to cluster or time interval to time interval. If changes in the Earth system through time affected the potential for development of exceptional preservational conditions, the likelihood of finding exceptional fossils per unit of rock quantity may follow a similar trend.

5.2. Trends in occurrences of clusters through time

The hierarchical clustering work in this study takes into account the clustering of assemblages, and mitigates the consequences of uneven sampling of assemblages with similar geographic locations, ages, and depositional settings. Time series illustrate the effects of this clustering work, showing for instance, that the raw unclustered data (Fig. 4a; MPLD=0) include overabundances of Ediacaran (ca. 560 Ma) and Lower Cretaceous assemblages relative to geographic expansiveness, as illustrated in the clustered data (Fig. 7a, b, i–l; MPLDs 50–1000 km). For the Ediacaran, this result is a reflection of the numerous albeit geospatially concentrated collection points of Ediacara-type fossils provided at bed level in primary literature (Hofmann et al., 2008). Conversely, for the Cretaceous, it may reflect geographic concentrations of localities in China and the UK, where the Jehol (Wang et al., 2012) and Purbeck Limestone (Coram, 2003)

ACCEPTED MANUSCRIPT

groups respectively contain exceptional insects at numerous sites. In any case, counts of clusters versus time (Fig. 7a, b, i–l) generally corroborate the major trends observed in the raw data. Clusters increase in number (geospatial spread) from the lower Ediacaran into the Phanerozoic, peak in the Cambrian, and decline through the lower-middle Paleozoic before rising through the remainder of the Phanerozoic, excluding intervals in the late Triassic and latest Cretaceous (Fig. 7b, j, l), which include relatively few clusters of assemblages. These differences in cluster count among time bin are statistically significant (Fig. 8, 9) for MPLDs < 1000 km. Analyses involving clusters of MPLDs > 1000 km did not yield statistically significant evidence of cluster count variation (Fig. 8), primarily due to sample size constraints (i.e. too few clusters in some time bins) but also low magnitude differences among time bins (Fig. 7). However, given that clusters of MPLDs < 1000 km best capture the clustering patterns of interest (Fig. 6), the results confirm that the frequency of exceptional preservation has significantly changed through Earth history.

5.3. Relationships with quantities of rock available for sampling

Differences in cluster count among time bins may, in part, follow variation in the quantity of sedimentary rock. The estimates of the global outcrop areas, global volumes, and North American surface/subsurface areas of the non-marine rocks in the time bins in this study follow similar trends through geologic time (Fig. 10). Like non-marine cluster count, the quantity of non-marine rock increases from the lower-middle Paleozoic to the Pliocene (Quaternary rocks were not included in our estimates). Indeed, correlation analyses indicate that non-marine sedimentary rock quantity (outcrop area or volume) accounts for the majority of non-marine cluster count variation observed among time bins younger than 430 Ma (Fig. 11a–c; $R^2 > 0.6$ and *F*-test $P < 0.05$ for most MPLDs ≤ 1000 km), and first difference cross-plots of these variables

ACCEPTED MANUSCRIPT

for 30.5 my time bins corroborate these correlations (see Supplementary Information, Fig. S7). Thus, on a global scale, the number of non-marine assemblages and clusters may depend on the amount of rock available for geologic sampling of exceptionally preserved fossils.

The estimates regarding marine rock areas and volumes follow markedly different patterns through the Phanerozoic. Whereas global marine rock outcrop area remains more-or-less constant from the Cambrian to the Cretaceous, the global volumes and North American surface/subsurface areas of marine/transitional rocks fluctuate across several orders of magnitude with minimum volumes/areas in the Permian-Triassic interval and maximum volumes/areas recorded in the Ordovician and upper Mesozoic. These differences may reflect the inclusion of subsurface deposits in the estimates of global volumes and North American areas. Because the exceptionally preserved assemblages in the database—excluding the Cambrian Mount Cap Formation (Butterfield, 1994), Carboniferous Castlecomer (Orr et al., 2008; Orr et al., 1996), and Permian Pechora Basin (Rasnitsyn et al., 2005) assemblages—were discovered in surface outcrops, global rock outcrop area may represent the best proxy for considering biases in sampling. Global rock volume and North American surface/subsurface rock area, in contrast, may represent the best proxies for investigating preservational biases caused by tectonically-controlled patterns in sedimentation. In any case, the correlation analyses indicate that neither rock quantity (Fig. 11e–i, m–r; $R^2 < 0.33$ and F-test $P < 0.05$ for most MPLDs ≤ 1000 km) nor global relative sea level (Haq and Schutter, 2008; Hardenbol et al., 1999; Figs. 10b, 11e–i, m–r; F-test $P > 0.05$ for all MPLDs) account for the variations observed in marine and marine/transitional cluster counts. This result does not preclude a relationship between rock quantity and assemblage/cluster count in marine and transitional settings, or between count and sea level. The rise in marine cluster count from the lower Ediacaran to the Cambrian, for

ACCEPTED MANUSCRIPT

example, may signify an increase in the amount of marine rock available for exploration (Fig. 10). Regardless, given a lack of correlations, the trends in marine assemblages and clusters point to other factors of influence.

5.4. Trends in depositional settings of clusters

Because total cluster count significantly varies through the geologic history of complex life, regardless of its method of calculation (Fig. 7a, b, i–l), efforts to quantify the predominance of exceptional preservation in various depositional settings must account for changes in sample size from time bin to time bin. For this reason, in this study, we normalized counts of marine, transitional, non-marine, and marine/transitional clusters relative to total cluster counts (i.e. marine+transitional+non-marine and marine/transitional+non-marine counts), and for all methods of clustering the data, determined proportions of clusters in each depositional setting category for each time bin (Figs. 12, 14). Statistical testing shows that the proportions of clusters in the categories change significantly throughout the stratigraphic record (Figs. 12–14). Evidently, exceptional preservational conditions shifted from predominantly marine depositional settings in the Ediacaran-early Paleozoic to transitional and non-marine settings in the Silurian-Carboniferous and Permian-Neogene intervals, respectively (Fig. 12; Allison and Briggs, 1991). In addition, the data indicates that exceptional preservation rose in frequency within marine and transitional settings in the Jurassic and early-late Cretaceous (Fig. 14a, b).

Several observations rule out the interpretation of the major proportion pattern as a product resulting from inversed ‘wedges’ in marine and non-marine cluster counts. If decreasing marine cluster counts and increasing non-marine cluster counts did cause the proportion pattern, then their trends would overlap, and transitional clusters would not predominate in any time

ACCEPTED MANUSCRIPT

domain. The decline in marine cluster count, however, began in the Cambrian (Fig. 7c, d), prior to the rise of terrestrial animal life (Shear and Selden, 2001), and although the oldest unambiguous non-marine clusters first appear in the Devonian (Trewin, 1985; Trewin and Davidson, 1995), they remain infrequent until the Carboniferous (Fig. 7g, h). Moreover, transitional clusters are predominant in the interim between the Cambrian and Carboniferous systems (Fig. 12c, d). Thus, in part, the pattern must reflect independent trends through time in occurrences of marine and non-marine fossil assemblages.

5.5. Interpreting the cluster count and proportion time series results

Stratigraphic and geographic distributions of fossils reflect evolutionary and ecological trends, but also patterns in fossil preservation and sampling associated with rock record biases and paleoenvironmental gradients. In this context, the major trends in the exceptional fossil record through Earth history may, to varying extents, signify geologic and taphonomic overprints. Both types of overprints affect the fossil record on a global scale, but as we argue based on our dataset, they disproportionately affect the records of marine, non-marine, and transitional assemblages. In this section, we explore the trends in Konservat-Lagerstätten through geologic time in each of these depositional settings, and overall, provide a comprehensive account of the environmental and geological phenomena responsible for their distributions.

5.5.1. Exceptional preservation through time in marine settings

Secular environmental changes of global scale can affect occurrences of fossils by influencing preservational processes and creating taphonomic biases. In regards to exceptional fossil preservation, such taphonomic overprints result from environmental changes that affect the

ACCEPTED MANUSCRIPT

degradation and mineralization of soft tissues. Such phenomena best explain the variation through geologic time in marine cluster count, which overall, does not significantly co-vary with marine rock quantity. In particular, the decline in the number of marine clusters through the lower Paleozoic interval likely reflects environmental changes in open marine settings, which promoted destruction of soft tissues and/or reduced the potential for authigenic mineralization. Along these same lines, the rise in marine cluster count from the lower Ediacaran to Cambrian may reflect global environmental changes that limited degradation and/or fostered tissue mineralization. This trend, however, may signify a sampling bias caused by geologic overprinting (rock quantity increases from the Ediacaran to the Cambrian, Fig. 10), or by the Cambrian radiation of sclerotized animals, which had recalcitrant tissues with greater preservational potentials than those of Ediacaran eukaryotes (Butterfield, 2003). Thus, we focus on the decline in marine cluster count through the lower-middle Paleozoic, which cannot be explained by such alternatives.

In light of our unifying conceptual model (Fig. 3) and previous hypotheses regarding the opening and closing of taphonomic windows through geologic time (Butterfield, 2003; Gaines et al., 2012b; Muscente et al., 2015a; Tarhan et al., 2016), two non-mutually exclusive (and holistically complementary) explanations emerge for the overarching trend in the marine record. The first explanation attributes the abundance of exceptionally preserved fossil assemblages in the upper Neoproterozoic-lower Paleozoic to seawater chemistry, which given its unique make-up in that interval, supported a multitude of preservational processes and opened all taphonomic windows. High silica allowed for silicification (Knoll, 1985; Maliva et al., 1989); influx of phosphate promoted phosphatization (Muscente et al., 2015a); and aspects that inhibited degradative chemistry resulted in organic preservation (Butterfield, 1990; Gaines et al., 2012a;

ACCEPTED MANUSCRIPT

Gaines et al., 2012b; Orr et al., 1998; Petrovich, 2001). Yet, as our conceptual model shows, the major taphonomic processes do not occur under entirely unique environmental conditions. Focusing on shared conditions, the second explanation ascribes the decline in exceptional preservation in the upper Neoproterozoic-lower Paleozoic to secular changes in near-surface marine environments (Allison and Briggs, 1993), which broadly affected all taphonomic processes and lastingly restricted soft tissue conservation in marine settings. Overall, the trend broadly follows the prolonged oxygenation of the ocean-atmosphere system (Gill et al., 2011; Sperling et al., 2015) and the protracted development of the sediment mixed layer (Tarhan et al., 2015), the zone of sediment homogenized and fluidized via bioturbation by burrowing animals, in the early Paleozoic. These changes in marine environments affected exceptional preservation by promoting scavenging of buried carcasses (Allison and Briggs, 1993); enhancing the seawater concentrations of O_2 and SO_4^{2-} used in the main microbial metabolisms of decay (Canfield and Farquhar, 2009; Tarhan et al., 2015); deepening the sedimentary aerobic and sulfate reduction zones in which soft tissues are most aggressively degraded (Muscente et al., 2015a; Schiffbauer et al., 2014b); and reducing the prevalence of microbial mats, which facilitate authigenic/diagenetic mineralization (Wilby et al., 1996) by sealing fossils off from oxic or suboxic bottom waters (Gehling, 1999; Laflamme et al., 2011) and preventing efflux of precipitating geochemical species (Callow and Brasier, 2009; Muscente et al., 2015a). Collectively, these changes impacted the likelihood of organic remains surviving long enough to undergo authigenic/diagenetic mineralization prior to their destruction. As a result, exceptional preservation environments in the late Paleozoic, Mesozoic, and Cenozoic were largely confined to transitional and non-marine depositional settings, where conditions that facilitate soft tissue

ACCEPTED MANUSCRIPT

conservation sometimes develop within restricted local and regional systems (Allison and Briggs, 1991; Allison and Briggs, 1993).

Global preservational biases caused by changing oxygen levels may not be unique to the Ediacaran-Ordovician interval. The prevalence of exceptional assemblages within marine/transitional facies associated with the Toarcian (Ansorge, 2003; Martindale et al., 2017) and Cenomanian/Turonian (Martill et al., 2011) global oceanic anoxic events (Takashima et al., 2006; Figs. 12 a–d, 14a, b) indicates such events favored exceptional preservation. Low oxygen levels during the events may have deterred scavengers, limited soft tissue decomposition, and fostered sedimentary microbial mats. Future investigations of assemblages in these intervals may further illuminate the role of oxygen in exceptional preservation.

5.5.2. Exceptional preservation through time in non-marine settings

The exposure and quantity (i.e. availability) of sedimentary rock represents a potential geologic control on fossil sampling. The correlation between non-marine cluster count and non-marine rock outcrop area and volume suggests that rock quantity largely determines the number of exceptional fossil deposits sampled from terrestrial facies in each time bin. Given that non-marine rock quantity increases from the Paleozoic to present, the rise in non-marine cluster count through that interval most likely signifies a sampling bias caused by underlying geologic phenomena. In general, the quality of the fossil record declines with age (Kowalewski and Flessa, 1996), as rocks are progressively modified and destroyed by tectonic, diagenetic, and erosional processes over time. As a direct consequence of such phenomena, the availability of non-marine rock improves toward the present (Fig. 10; Wall et al., 2009), and more non-marine

ACCEPTED MANUSCRIPT

assemblages have been sampled from younger than older geologic intervals. Whereas older deposits were destroyed, younger ones have more readily survived to the present.

Overall, the linear relationship between cluster count and rock availability suggests that, on a global scale, the probability of exceptional preservational conditions developing in non-marine settings did not significantly fluctuate through Earth history; if it did, then the relationship between rock quantity and cluster count would not be resolvable. In light of this relationship, the results indicate that exceptional preservation occurs within non-marine settings in response to local and provincial conditions, largely immune to global secular environmental changes. As non-marine assemblages occur in clusters throughout the middle Devonian-Neogene interval (Bradley, 1964; Huber and McDonald, 2003; Trewin, 1985; Trewin and Davidson, 1995; Wang et al., 2012), these phenomena must include regional scale factors, like topography, sedimentation rate, climate, primary productivity, salinity, and paleogeography (Allison and Briggs, 1993). In general, exceptional preservation may be favored in regional settings with bottom water stagnation, episodic burial, pore water anoxia, and early-diagenetic sealing of sediments in addition to low levels of animal scavenging, bioturbation, and sediment reworking (Allison and Briggs, 1991; Seilacher et al., 1985). Collectively, all these conditions influence soft tissue degradation. The specific preservational styles in a region, however, depend upon factors affecting mineralization, which may be highly localized to fossils (McNamara et al., 2009), such as the availabilities of reactive iron, silica, and phosphate (Guan et al., 2016; Muscente et al., 2015a) and their durations within the various microbial zones within the sediment column (Schiffbauer et al., 2014b). Preservational style can also be influenced by the localized perturbations caused by carcasses entering microbial zones as well as the composition of the microbiota driving decay and mineralization of soft tissues (Sagemann et al., 1999).

ACCEPTED MANUSCRIPT

5.5.3. Other trends in exceptional preservation through time

Exceptionally preserved fossil assemblages were predominantly preserved in transitional settings in the middle and upper Paleozoic (Fig. 12c, d). The origin of this trend is ambiguous, in part, due to limited data regarding the exposure, area, and volume of transitional facies. The abundance of transitional assemblages may reflect widespread coastlines bordering epicontinental seas, or specifically in the Carboniferous, broad coastal delta plains at low latitudes (Allison and Briggs, 1993). If the trend does not signify geologic overprinting, it may simply reflect the absence of exceptional preservation in marine settings and the paucity of non-marine sedimentary rocks in that interval. Total cluster counts are low throughout the interval (Fig. 7a, b), outside of the Pennsylvanian, which contains a plethora of assemblages (Baird et al., 1985b). Alternatively, the trend may correspond to globally low bioturbation levels, as sediment mixing in transitional environments did not begin approaching modern intensities and depths until the Permian or Triassic (Buatois et al., 2005; Mángano and Buatois, 2015). Future work should investigate this possible relationship between bioturbation and exceptional preservation.

The dearth of assemblages in the upper Triassic (Norian) and uppermost Cretaceous (Campanian and Maastrichtian), likewise, remains problematic. The absence of assemblages in these intervals may signify paucity of non-marine sedimentary rock at a level that cannot be resolved in the available rock area and volume data. Alternatively, perhaps global climatic conditions limited fossil preservation. None of the cluster time series directly or inversely follow the Phanerozoic paleotemperature curve, which has highs in the lower Paleozoic, upper Paleozoic, and upper Cretaceous as well as lows in the Ordovician, Silurian, Carboniferous, and Cenozoic. However, this lack of co-variation through the record does not rule out the possibility

ACCEPTED MANUSCRIPT

of climatic influence. Although there is generally a lack paleoclimatic information for the Norian (Preto et al., 2010), oxygen isotopes suggest that both the late Triassic and late Cretaceous witnessed climatic cooling of hothouse conditions (Veizer and Prokoph, 2015). As the solubility of oxygen in water is inversely related to temperature, this cooling trend may have reduced the frequency of dysoxic environments conducive to soft tissue preservation. Of course, even as the climate cooled, arid conditions may have also prevailed in parts of the continents. Such arid conditions could have limited intracontinental precipitation, and thereby, curbed development of lakes and intermontane basins, which might produce Konservat-Lagerstätten. Thus, the paucity of assemblages in the upper Triassic and uppermost Cretaceous may, in part, signify aspects of Earth's climate. Future work on Triassic and Cretaceous paleoclimatology may illuminate the trend. In any case, the general dearth of exceptionally preserved fossils in these intervals presents a challenge for paleontological research on the history of the terrestrial biosphere. Research efforts should target non-marine facies in these intervals in order to potentially fill the gaps.

6. Conclusions

In summation, with the goal of holistic understanding of soft tissue taphonomy, our review puts forth a unifying conceptual model—rooted in sedimentary geology and geomicrobiology—that accounts for all variations in fossil preservational pathway and style by considering taphonomic processes with shared and unique environmental controls. Taking this conceptual model into consideration, we interpret the results of a meta-analysis of a new and comprehensive compilation of exceptionally preserved fossil assemblages, which markedly expands on previous work. Konservat-Lagerstätten commonly cluster in geographic and stratigraphic space with others of similar age and matching depositional setting. Such clusters

ACCEPTED MANUSCRIPT

signify exceptional preservational conditions that arose repeatedly over millions of years throughout laterally extensive regions (10s to 100s of kilometers in scale). Trends in the counts and proportions of these clusters through geologic time illuminate the environmental controls on exceptional preservation operating on interrelated local, regional, and global scales. In addition, the results indicate that the marine and non-marine records of exceptionally preserved fossils reflect different principal 'mega-biases' (Kowalewski and Flessa, 1996), signifying environmental and geologic phenomena affecting fossil preservation and sampling on global levels. Whereas preservational biases prevail in the marine record, sampling biases predominate in the non-marine record. Research on the history of life in the fossil record should account for such megabiases by integrating geospatially-minded approaches (Plotnick, 2017) and paleoenvironmental data with metrics of diversity, disparity, and sampling effort.

Remaining questions pertain to the consequences of global megabiases for ecological and evolutionary patterns. Our results provide some support for the possible mass extinction of life around the Precambrian-Cambrian boundary (Laflamme et al., 2013), which is purportedly captured in the fossil record of soft-bodied Ediacara-type taxa (Darroch et al., 2015). Preservation of these organisms evidently ceased during a time when global marine conditions were broadly conducive to soft tissue conservation. Accordingly, the simplest explanation for the absence of Ediacara-type fossils beyond the Ediacaran is extinction. However, the pattern could alternatively reflect a preservational bias superimposed on evolutionary and ecological trends. For instance, perhaps some Ediacara-type taxa persisted into the Phanerozoic (Conway Morris, 1993; Hagadorn et al., 2000; Jensen et al., 1998), but due to ecological pressures (Schiffbauer et al., 2016), became rare and/or restricted to environments, which did not allow for soft tissue conservation. Under such circumstances, preservation of Ediacara-type taxa may have been

ACCEPTED MANUSCRIPT

uncommon. Ultimately, to test for the effects of such biases, future efforts should aim to improve upon our work, documenting patterns in exceptional fossil occurrence and sampling at higher temporal and spatial resolutions and integrating further taphonomic information.

Notably, the megabiases described in this study support alternative hypotheses for other purported evolutionary and ecological patterns among ancient soft-bodied organisms. The overall rise in insect family richness through the Carboniferous-Neogene interval (Labandeira and Sepkoski, 1993), for example, could be explained as a sampling bias caused by the declining quality of the non-marine record with age. In another case, preservational bias caused by environmental changes affecting soft tissue conservation in the early Paleozoic could account for the purported extinction of Burgess shale-type organisms in the Ordovician (Van Roy et al., 2010). These examples affirm that an improved understanding of phenomena affecting exceptional fossil preservation and sampling and their temporal and geospatial heterogeneity will be critical for developing, testing, and improving hypotheses regarding the history of life.

Acknowledgements

We thank B. Gill, S. Nesbitt, K. Eriksson, and A. Knoll for discussions of this manuscript, and acknowledge the Macrostrat and Paleobiology Databases as the sources of some of the data in this report. This research was supported by the Keck Foundation-funded project "The Co-Evolution of the Geo- and Biospheres: An Integrated Program for Data-Driven Abductive Discovery in the Earth Sciences" (A. D. M. and M. M.). Additional funding to support this research was provided by NSF EAR-1528553 (S. X.), NASA Exobiology NNX15AL27G (S. X.), and NSERC Discovery Grant (M. L.).

ACCEPTED MANUSCRIPT

Supplementary material

Supplementary information (discussion, figures, and tables), animations (Animations S1, S2), and data (Dataset S1) can be found online at *Insert web address here*.

References

Allison, P.A., 1988a. The role of anoxia in the decay and mineralization of proteinaceous macro-fossils. Paleobiology 14, 139–154.

Allison, P.A., 1988b. Taphonomy of the Eocene London Clay biota. Palaeontology 31, 1079–1100.

Allison, P.A., Briggs, D.E.G., 1991. Taphonomy of nonmineralized tissues, In: Allison, P.A., Briggs, D.E.G. (Eds.), Taphonomy: Releasing the data locked in the fossil record. Plenum Press, New York, NY, United States, pp. 25–70.

Allison, P.A., Briggs, D.E.G., 1993. Exceptional fossils record: Distribution of soft-tissue preservation through the Phanerozoic. Geology 21, 527–530.

Alroy, J., 2003. Global database will yield reliable measures of global biodiversity. Paleobiology 29, 26–29.

Ansorge, J., 2003. Insects from the lower Toarcian of middle Europe and England. Acta Zoologica Cracoviensia 46, 291–310.

Arning, E.T., Birgel, D., Brunner, B., Peckmann, J., 2009. Bacterial formation of phosphatic laminites off Peru. Geobiology 7, 295–307.

Baird, G.C., Shabica, C.W., Anderson, J.L., Richardson, E.S., 1985a. Biota of a Pennsylvanian muddy coast: Habitats within the Mazonian delta complex, northeast Illinois. Journal of Paleontology 59, 253–281.

ACCEPTED MANUSCRIPT

Baird, G.C., Sroka, S.D., Shabica, C.W., Beard, T.L., Scott, A.C., Broadhurst, F.M., 1985b. Mazon Creek-Type fossil assemblages in the U.S. midcontinent Pennsylvanian: Their recurrent character and palaeoenvironmental significance [and discussion]. Philosophical Transactions of the Royal Society B: Biological Sciences 311, 87–99.

Bashkuev, A., Sell, J., Aristov, D., Ponomarenko, A., Sinitshenkova, N., Mahler, H., 2012. Insects from the Buntsandstein of lower Franconia and Thuringia. Paläontologische Zeitschrift 86, 175–185.

Boag, T.H., Darroch, S.A.F., Laflamme, M., 2016. Ediacaran distributions in space and time: Testing assemblage concepts of earliest macroscopic body fossils. Paleobiology 42, 574–594.

Bradley, W.H., 1964. Geology of the Green River Formation and associated Eocene rocks in south-western Wyoming and adjacent parts of Colorado and Utah. Geological Survey Professional Paper 496–A, A1–A86.

Briggs, D.E.G., 1999. Molecular taphonomy of animal and plant cuticles: Selective preservation and diagenesis. Philosophical Transactions of the Royal Society B: Biological Sciences 354, 7–17.

Briggs, D.E.G., 2003. The role of decay and mineralization in the preservation of soft-bodied fossils. Annual Review of Earth and Planetary Sciences 31, 275–301

Briggs, D.E.G., Bottrell, S.H., Raiswell, R., 1991. Pyritization of soft-bodied fossils: Beecher's Trilobite Bed, Upper Ordovician, New York State. Geology 19, 1221–1224.

Briggs, D.E.G., Kear, A.J., Martill, D.M., Wilby, P.R., 1993. Phosphatization of soft-tissue in experiments and fossils. Journal of the Geological Society, London 150, 1035–1038.

ACCEPTED MANUSCRIPT

Briggs, D.E.G., Raiswell, R., Bottrell, S.H., Hatfield, D., Bartels, C., 1996. Controls on the pyritization of exceptionally preserved fossils: An analysis of the Lower Devonian Hunsrück Slate of Germany. American Journal of Science 296, 633–663.

Briggs, D.E.G., Wilby, P.R., 1996. The role of calcium carbonate-calcium phosphate switch in the mineralization of soft-bodied fossils. Journal of the Geological Society, London 153, 665–668.

Buatois, L.A., Gingras, M.K., Maceachern, J., Mangano, M.G., Zonneveld, J.-P., Pemberton, S.G., Netto, R.G., Martin, A., 2005. Colonization of brackish-water systems through time: Evidence from the trace-fossil record. PALAIOS 20, 321–347.

Butterfield, N.J., 1990. Organic preservation of non-mineralizing organisms and the taphonomy of the Burgess Shale. Paleobiology 16, 272–286.

Butterfield, N.J., 1994. Burgess Shale-type fossils from a Lower Cambrian shallow-shelf sequence in northwestern Canada. Nature 369, 477–479.

Butterfield, N.J., 1995. Secular distribution of Burgess Shale-type preservation. Lethaia 28, 1–13.

Butterfield, N.J., 2002. *Leanchoilia* guts and the interpretation of three-dimensional structures in Burgess Shale-type fossils. Paleobiology 28, 155–171.

Butterfield, N.J., 2003. Exceptional fossil preservation and the Cambrian Explosion. Integrative and Comparative Biology 43, 166–177.

Butterfield, N.J., 2012. Does cement-induced sulfate limitation account for Burgess Shale-type preservation? Proceedings of the National Academy of Sciences 109, E1901–E1901.

Butterfield, N.J., Harvey, T.H.P., 2012. Small carbonaceous fossils (SCFs): A new measure of early Paleozoic paleobiology. Geology 40, 71–74.

ACCEPTED MANUSCRIPT

Cai, Y., Schiffbauer, J.D., Hua, H., Xiao, S., 2012. Preservational modes in the Ediacaran Gaojiashan Lagerstätte: Pyritization, aluminosilicification, and carbonaceous compression. Palaeogeography Palaeoclimatology Palaeoecology 326–328, 109–117.

Callow, R.H.T., Brasier, M.D., 2009. Remarkable preservation of microbial mats in Neoproterozoic siliciclastic settings: Implications for Ediacaran taphonomic models. Earth-Science Reviews 96, 207–219.

Canfield, D.E., Farquhar, J., 2009. Animal evolution, bioturbation, and the sulfate concentration of the oceans. Proceedings of the National Academy of Sciences 106, 8123–8127.

Chen, J.-Y., Schopf, J.W., Bottjer, D.J., Zhang, C.-Y., Kudryavtsev, A.B., Tripathi, A.B., Wang, X.-Q., Yang, Y.-H., Gao, X., Yang, Y., 2007. Raman spectra of a lower Cambrian ctenophore embryo from southwestern Shaanxi, China. Proceedings of the National Academy of Sciences 104, 6289–6292.

Chen, Z., Zhou, C., Xiao, S., Wang, W., Guan, C., Hua, H., Yuan, X., 2014. New Ediacara fossils preserved in marine limestone and their ecological implications. Scientific Reports 4 (4180): 1–10.

Choubert, G., Faure-Muret, A., 1976. Geological atlas of the world. 1:10,000,000. 22 sheets with explanations. UNESCO, Paris, France.

Clarkson, E., 1998. Invertebrate Palaeontology and Evolution. Blackwell Science Ltd, Malden, Massachusetts.

Colleary, C., Dolocan, A., Gardner, J., Singh, S., Wuttke, M., Rabenstein, R., Habersetzer, J., Schaal, S., Feseha, M., Clemens, M., Jacobs, B.F., Currano, E.D., Jacobs, L.L., Sylvestersen, R.L., Gabbott, S.E., Vinther, J., 2015. Chemical, experimental, and

ACCEPTED MANUSCRIPT

morphological evidence for diagenetically altered melanin in exceptionally preserved fossils. Proceedings of the National Academy of Sciences 112, 12592–12597.

Collins, D., Briggs, D.E.G., Conway Morris, S., 1983. New Burgess Shale fossil sites reveal middle Cambrian faunal complex. Science 222, 163–167.

Conway Morris, S., 1986. The community structure of the Middle Cambrian Phyllopod bed (Burgess Shale). Palaeontology 29, 423–467.

Conway Morris, S., 1989. The persistence of Burgess Shale-type faunas: Implications for the evolution of deeper-water faunas. Transactions of the Royal Society of Edinburgh: Earth Sciences 80, 271–283.

Conway Morris, S., 1993. Ediacaran-like fossils in Cambrian Burgess Shale-type faunas of North America. Palaeontology 36, 593–635.

Coram, R.A., 2003. Taphonomy and ecology of Purbeck fossil insects. Acta Zoologica Cracoviensia 46, 311–318.

Cotroneo, S., Schiffbauer, J.D., McCoy, V.E., Wortmann, U.G., Darroch, S.A.F., Peng, Y., Laflamme, M., 2016. A new model of the formation of Pennsylvanian iron carbonate concretions hosting exceptional soft-bodied fossils in Mazon Creek, Illinois. Geobiology 14, 543–555.

Darroch, S.A.F., Sperling, E.A., Boag, T.H., Racicot, R.A., Mason, S.J., Morgan, A.S., Tweedt, S., Myrow, P., Johnston, D.T., Erwin, D.H., Laflamme, M., 2015. Biotic replacement and mass extinction of the Ediacara biota. Proceedings of the Royal Society B: Biological Sciences 282, 1–10.

Dornbos, S., 2011. Phosphatization Through the Phanerozoic, In: Allison, P., Bottjer, D. (Eds.), Taphonomy. Springer Netherlands, pp. 435–456.

ACCEPTED MANUSCRIPT

Ebert, M., Kölbl-Ebert, M., Lane, J.A., 2015. Fauna and predator-prey relationships of Ettling, an actinopterygian fish-dominated Konservat-Lagerstätte from the late Jurassic of southern Germany. PLoS ONE 10, e0116140.

Elles, G.L., 1939. Factors controlling graptolite successions and assemblages. Geological Magazine 76, 181–187.

Finney, S.C., Berry, W.B.N., 1997. New perspectives on graptolite distributions and their use as indicators of platform margin dynamics. Geology 25, 919–922.

Gabbott, S.E., Norry, M.J., Aldridge, R.J., Theron, J.N., 2001. Preservation of fossils in clay minerals; a unique example from the Upper Ordovician Soom Shale, South Africa. Proceedings of the Yorkshire Geological Society 53, 237–244.

Gaines, R.R., Droser, M.L., 2010. The paleoredox setting of Burgess Shale-type deposits. Palaeogeography, Palaeoclimatology, Palaeoecology 297, 649–661.

Gaines, R.R., Droser, M.L., Orr, P.J., Garson, D., Hammarlund, E., Qi, C., Canfield, D.E., 2012a. Burgess shale−type biotas were not entirely burrowed away. Geology 40, 283–286.

Gaines, R.R., Hammarlund, E.U., Hou, X., Qi, C., Gabbott, S.E., Zhao, Y., Peng, J., Canfield, D.E., 2012b. Mechanism for Burgess Shale-type preservation. Proceedings of the National Academy of Sciences 109, 5180–5184.

Gehling, J.G., 1999. Microbial mats in terminal Proterozoic siliciclastics: Ediacaran death masks. PALAIOS 14, 40–57.

Gill, B.C., Lyons, T.W., Young, S.A., Kump, L.R., Knoll, A.H., Saltzman, M.R., 2011. Geochemical evidence for widespread euxinia in the later Cambrian ocean. Nature 469, 80–83.

ACCEPTED MANUSCRIPT

Guan, C., Wang, W., Zhou, C., Muscente, A.D., Wan, B., Chen, X., Yuan, X., Chen, Z., Ouyang, Q., 2016. Controls on fossil pyritization: Redox conditions, sedimentary organic matter content, and *Chuaria* preservation in the Ediacaran Lantian Biota. Palaeogeography, Palaeoclimatology, Palaeoecology 474, 26–35.

Gutiérrez-Marco, J.C., García-Bellido, D.C., 2015. Micrometric detail in palaeoscolecid worms from Late Ordovician sandstones of the Tafilalt Konservat-Lagerstätte, Morocco. Gondwana Research 28, 875–881.

Hagadorn, J.W., Fedo, C.M., Waggoner, B.M., 2000. Early Cambrian Ediacaran-type fossils from California. Journal of Paleontology 74, 731–740.

Haq, B.U., Schutter, S.R., 2008. A chronology of Paleozoic sea-level changes. Science 322, 64–68.

Hardenbol, J., Thierry, J., Farley, M.B., Jacquin, T., Graciansky, P.-C.d., Vail, P.R., 1999. Mesozoic and Cenozoic Sequence Chronostratigraphic Framework of European Basins, Mesozoic and Cenozoic Sequence Stratigraphy of European Basins. SEPM Society for Sedimentary Geology, pp. 3–13.

Hendy, A.J.W., 2011. Taphonomic Overprints on Phanerozoic Trends in Biodiversity: Lithification and Other Secular Megabiases, In: Allison, P.A., Bottjer, D.J. (Eds.), Taphonomy: Process and Bias Through Time. Springer Netherlands, Dordrecht, pp. 19–77.

Hofmann, H.J., O'Brien, S.J., King, A.F., 2008. Ediacaran biota on Bonavista Peninsula, Newfoundland, Canada. Journal of Paleontology 82, 1–36.

Hou, X., Bergström, J., 2003. The Chengjiang fauna — the oldest preserved animal community. Paleontological Research 7, 55–70.

ACCEPTED MANUSCRIPT

Huber, P., McDonald, N.G., 2003. Early Jurassic insects from the Newark Supergroup, northeastern United States, In: LeTourneau, P.M., Olsen, P.E. (Eds.), The Great Rift Valleys of Pangea in Eastern North America, vol. 2: Sedimentology, Stratigraphy, and Paleontology. Columbia University Press, New York City, New York, pp. 206–223.

Huntley, J.W., Xiao, S., Kowalewski, M., 2006. On the morphological history of Proterozoic and Cambrian acritarchs, In: Xiao, S., Kaufman, A.J. (Eds.), Neoproterozoic Geobiology and Paleobiology. Springer, Dordrecht, the Netherlands, pp. 23–56.

Jensen, S., Gehling, J.G., Droser, M.L., 1998. Ediacara-type fossils in Cambrian sediments. Nature 393, 567–569.

Kenrick, P., Crane, P.R., 1997. The origin and early evolution of plants on land. Nature 389, 33–39.

Knoll, A.H., 1985. Exceptional preservation of photosynthetic organisms in silicified carbonates and silicified peats. Philosophical Transactions of the Royal Society B: Biological Sciences 311, 111–122.

Kowalewski, M., Flessa, K.W., 1996. Improving with age: The fossil record of lingulide brachiopods and the nature of taphonomic megabiases. Geology 24, 977–980.

Labandeira, C., Sepkoski, J., 1993. Insect diversity in the fossil record. Science 261, 310–315.

Laflamme, M., Darroch, S.A.F., Tweedt, S.M., Peterson, K.J., Erwin, D.H., 2013. The end of the Ediacara biota: Extinction, biotic replacement, or Cheshire Cat? Gondwana Research 23, 558–573.

Laflamme, M., Schiffbauer, J.D., Narbonne, G.M., Briggs, D.E.G., 2011. Microbial biofilms and the preservation of the Ediacara biota. Lethaia 44, 203–213.

ACCEPTED MANUSCRIPT

Lee, M.S.Y., Jago, J.B., Garcia-Bellido, D.C., Edgecombe, G.D., Gehling, J.G., Paterson, J.R., 2011. Modern optics in exceptionally preserved eyes of Early Cambrian arthropods from Australia. Nature 474, 631–634.

Lerosey-Aubril, R., Hegna, T.A., Kier, C., Bonino, E., Habersetzer, J., Carré, M., 2012. Controls on gut phosphatisation: The trilobites from the Weeks Formation Lagerstätte (Cambrian; Utah). PLoS ONE 7, e32934.

LoDuca, S.T., Caron, J.-B., Schiffbauer, J.D., Xiao, S., Kramer, A., 2015. A reexamination of *Yuknessia* from the Cambrian of British Columbia and Utah. Journal of Paleontology 89, 82–95.

Maliva, R.G., Knoll, A.H., Siever, R., 1989. Secular change in chert distribution: A reflection of evolving biological particpation in the silica cycle. PALAIOS 4, 519–532.

Mángano, M.G., Buatois, L.A., 2015. The trace-fossil record of tidals flats through the Phanerozoic: Evolutionary innovations and faunal turnover, In: McIlroy, D. (Ed.), Ichnology: Papers from ICHNIA III. Geological Association of Canada, St. John's, Canada, pp. 157–177.

Manning, P.L., Morris, P.M., McMahon, A., Jones, E., Gize, A., Macquaker, J.H.S., Wolff, G., Thompson, A., Marshall, J., Taylor, K.G., Lyson, T., Gaskell, S., Reamtong, O., Sellers, W.I., van Dongen, B.E., Buckley, M., Wogelius, R.A., 2009. Mineralized soft-tissue structure and chemistry in a mummified hadrosaur from the Hell Creek Formation, North Dakota (USA). Proceedings of the Royal Society B: Biological Sciences 276, 3429–3437.

Martill, D.M., 1990. Macromolecular resolution of fossilized muscle tissue from an elopomorph fish. Nature 346, 171–172.

ACCEPTED MANUSCRIPT

Martill, D.M., Ibrahim, N., Brito, P.M., Baider, L., Zhouri, S., Loveridge, R., Naish, D., Hing, R., 2011. A new Plattenkalk Konservat Lagerstätte in the Upper Cretaceous of Gara Sbaa, south-eastern Morocco. Cretaceous Research 32, 433–446.

Martindale, R.C., Them, T.R., Gill, B.C., Marroquín, S.M., Knoll, A.H., 2017. A new Early Jurassic (ca. 183 Ma) fossil Lagerstätte from Ya Ha Tinda, Alberta, Canada. Geology.

Martínez-Delclòs, X., Briggs, D.E.G., Peñalver, E., 2004. Taphonomy of insects in carbonates and amber. Palaeogeography Palaeoclimatology Palaeoecology 203, 19–64.

Martínez-Delclòs, X., Briggs, D.E.G., Peñalver, E., 2004. Taphonomy of insects in carbonates and amber. Palaeogeography, Palaeoclimatology, Palaeoecology 203, 19–64.

McNamara, M.E., Briggs, D.E.G., Orr, P.J., Noh, H., Cao, H., 2012a. The original colours of fossil beetles. Proceedings of the Royal Society B: Biological Sciences 279, 1114–1121.

McNamara, M.E., Orr, P.J., Alcala, L., Anadon, P., Penalver Molla, E., 2012b. What controls the taphonomy of exceptionally preserved taxa—Environment or biology? A case study using frogs from the Miocene Libros Konservat-Lagerstatte (Teruel, Spain). PALAIOS 27, 63–77.

McNamara, M.E., Orr, P.J., Kearns, S.L., Alcala, L., Anadon, P., Penalver Molla, E., 2009. Soft-tissue preservation in Miocene frogs from Libros, Spain: Insights into the genesis of decay microenvironments. PALAIOS 24, 104–117.

Milligan, G., 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika 45, 325–342.

Milligan, G.W., Cooper, M.C., 1987. Methodology review: Clustering methods. Applied Psychological Measurement 11, 329–354.

ACCEPTED MANUSCRIPT

Muscente, A.D., Allmon, W.D., 2013. Revision of the hydroid *Plumalina* Hall, 1858 in the Silurian and Devonian of New York. Journal of Paleontology 87, 710–725.

Muscente, A.D., Allmon, W.D., Xiao, S., 2016. The hydroid fossil record and analytical techniques for assessing the affinities of putative hydrozoans and possible hemichordates. Palaeontology 59, 71–87.

Muscente, A.D., Hawkins, A.D., Xiao, S., 2015a. Fossil preservation through phosphatization and silicification in the Ediacaran Doushantuo Formation (South China): A comparative synthesis. Palaeogeography, Palaeoclimatology, Palaeoecology 434, 46–62.

Muscente, A.D., Michel, F.M., Dale, J.G., Xiao, S., 2015b. Assessing the veracity of Precambrian 'sponge' fossils using *in situ* nanoscale analytical techniques. Precambrian Research 263, 142–156.

Muscente, A.D., Xiao, S., 2015a. New occurrences of *Sphenothallus* in the lower Cambrian of South China: Implications for its affinities and taphonomic demineralization of shelly fossils. Palaeogeography, Palaeoclimatology, Palaeoecology 437, 141–146.

Muscente, A.D., Xiao, S., 2015b. Resolving three-dimensional and subsurficial features of carbonaceous compressions and shelly fossils using backscattered electron scanning electron microscopy (BSE-SEM). PALAIOS 30, 462–481

Narbonne, G.M., 2005. The Ediacara Biota: Neoproterozoic origin of animals and their ecosystems. Annual Review of Earth and Planetary Sciences 33, 421–442.

Navalón, G., Marugán-Lobón, J., Chiappe, L.M., Luis Sanz, J., Buscalioni, Á.D., 2015. Soft-tissue and dermal arrangement in the wing of an Early Cretaceous bird: Implications for the evolution of avian flight. Scientific Reports 5 (14864), 1–7.

ACCEPTED MANUSCRIPT

Orr, P.J., Briggs, D.E.G., Kearns, S.L., 1998. Cambrian Burgess Shale animals replicated in clay minerals. Science 281, 1173–1175.

Orr, P.J., Briggs, D.E.G., Kearns, S.L., 2008. Taphonomy of exceptionally preserved crustaceans from the upper Carboniferous of southeastern Ireland. PALAIOS 23, 298–312.

Orr, P.J., Briggs, D.E.G., Parkes, M.A., 1996. The 'Castlecomer Fauna': A new Konservat-Lagerstätte from the upper Carboniferous of Ireland. Irish Journal of Earth Sciences 15, 93–106.

Pan, Y., Sha, J., Zhou, Z., Fürsich, F.T., 2013. The Jehol biota: Definition and distribution of exceptionally preserved relicts of a continental Early Cretaceous ecosystem. Cretaceous Research 44, 30–38.

Perrier, V., Charbonnier, S., 2014. The Montceau-les-Mines Lagerstätte (Late Carboniferous, France). Comptes Rendus Palevol 13, 353–367.

Petrovich, R., 2001. Mechanisms of fossilization of the soft-bodied and lightly armored faunas of the Burgess Shale and of some other classical localities. American Journal of Science 301, 683–726.

Pickerill, R.K., 1994. Exceptional fossil record: Distribution of soft-tissue preservation through the Phanerozoic: discussion. Geology 22, 183–184.

Plotnick, R.E., 2017. Recurrent hierarchical patterns and the fractal distribution of fossil localities. Geology 45, 295–298.

Preto, N., Kustatscher, E., Wignall, P.B., 2010. Triassic climates — State of the art and perspectives. Palaeogeography, Palaeoclimatology, Palaeoecology 290, 1–10.

Rasnitsyn, A.P., Sukacheva, I.D., Aristov, D., 2005. Permian insects of the Vorkuta Group in the Pechora Basin, and their stratigraphic implications. Paleontological Journal 39, 404–416.

ACCEPTED MANUSCRIPT

Rice, C.M., Trewin, N.H., Anderson, L.I., 2002. Geological setting of the Early Devonian Rhynie cherts, Aberdeenshire, Scotland: An early terrestrial hot spring system. Journal of the Geological Society 159, 203–214.

Rombouts, L.P., 1982. Factors controlling the distribution of Chitinozoa in the Gleedon Chronozone (Wenlockian) of Northern Europe. Annales de la Société géologique de Belgique 105, 259–271.

Ronov, A.B., 1982. The Earth's sedimentary shell (quantitative patterns of its structures, compositions, and evolution): The 20th V.I. Vernadskly lecture, March 12, 1978. International Geology Review 24, 1365–1388.

Sagemann, J., Bale, S.J., Briggs, D.E.G., Parkes, R.J., 1999. Controls on the formation of authigenic minerals in association with decaying organic matter: An experimental approach. Geochimica et Cosmochimica Acta 63, 1083–1095.

Sall, J., Lehman, A., Stephens, M.L., Creighton, L., 2012. JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP, Fifth Edition. SAS Institute.

Schiffbauer, J.D., Huntley, J.W., O'Neil, G.R., Darroch, S.A.F., Laflamme, M., Cai, Y., 2016. The latest Ediacaran wormworld fauna: Setting the ecological stage for the Cambrian explosion. GSA Today 26, 4–11.

Schiffbauer, J.D., Laflamme, M., 2012. Lagerstätten through time: A collection of exceptional preservational pathways from the terminal Neoproterozoic through today. PALAIOS 27, 275–278.

Schiffbauer, J.D., Wallace, A.F., Broce, J., Xiao, S., 2014a. Exceptional fossil conservation through phosphatization, In: Laflamme, M., Schiffbauer, J.D., Darroch, S.A.F. (Eds.), The Paleontological Society Papers, pp. 59–82.

ACCEPTED MANUSCRIPT

Schiffbauer, J.D., Xiao, S., Cai, Y., Wallace, A.F., Hua, H., Hunter, J., Xu, H., Peng, Y., Kaufman, A.J., 2014b. A unifying model for Neoproterozoic–Palaeozoic exceptional fossil preservation through pyritization and carbonaceous compression. Nature Communications 5, 1–12.

Schopf, T.J.M., 1978. Fossilization potential of an intertidal fauna: Friday Harbor, Washington. Paleobiology 4, 261–270.

Seilacher, A., 1970. Begriff und bedeutung der Fossil-Lagerstätten. Neues Jahrbuch für Geologie und Paläontologie, Monatshefte 1970, 34–39.

Seilacher, A., Reif, W.-E., Westphal, F., 1985. Sedimentological, ecological and temporal patterns of fossil Lagerstätten. Philosophical Transactions of the Royal Society B: Biological Sciences 311, 5–24.

Shear, W.A., Selden, P.A., 2001. Rustling in the undergrowth: Animals in early terrestrial ecosystems, In: Gensel, P.G., Edwards, D. (Eds.), Plants Invade the Land: Evolutionary and Environmental Perspectives. Columbia Uniersity Press, New York, pp. 29–51.

Sperling, E.A., Wolock, C.J., Morgan, A.S., Gill, B.C., Kunzmann, M., Halverson, G.P., Macdonald, F.A., Knoll, A.H., Johnston, D.T., 2015. Statistical analysis of iron geochemical data suggests limited late Proterozoic oxygenation. Nature 523, 451–454.

Stankiewicz, B.A., Briggs, D.E.G., Michels, R., Collinson, M.E., Flannery, M.B., Evershed, R.P., 2000. Alternative origin of aliphatic polymer in kerogen. Geology 28, 559–562.

Stanley, G.D., Sturmer, W., 1987. A new fossil ctenophore discovered by X-rays. Nature 328, 61–63.

ACCEPTED MANUSCRIPT

Strang, K.M., Armstrong, H.A., Harper, D.A.T., Trabucho-Alexandre, J.P., 2016. The Sirius Passet Lagerstätte: Silica death masking opens the window on the earliest matground community of the Cambrian explosion. Lethaia 49, 631–643.

Takashima, R., Nishi, H., Huber, B.T., Leckie, R.M., 2006. Greenhouse world and the Mesozoic ocean. Oceanography 19, 82–92.

Tarhan, L.G., Droser, M.L., Planavsky, N.J., Johnston, D.T., 2015. Protracted development of bioturbation through the early Palaeozoic Era. Nature Geoscience 8, 865–869.

Tarhan, L.G., Hood, A.v.S., Droser, M.L., Gehling, J.G., Briggs, D.E.G., 2016. Exceptional preservation of soft-bodied Ediacara Biota promoted by silica-rich oceans. Geology 44, 951–954.

Taylor, P.D., 1990. Preservation of soft-bodied and other organisms by bioimmuration — A review. Palaeontology 33, 1–17.

Taylor, T.N., Taylor, E.L., Krings, M., 2009. Paleobotany (Second Edition). Academic Press, London.

Trewin, N.H., 1985. Mass mortalities of Devonian fish — The Achanarras Fish Bed, Caithness. Geology Today 1, 45–49.

Trewin, N.H., Davidson, R.G., 1995. An Early Devonian lake and its associated biota in the Midland Valley of Scotland. Transactions of the Royal Society of Edinburgh: Earth Sciences 86, 233–246.

Valentine, J.W., 1989. How good was the fossil record? Clues from the Californian Pleistocene. Paleobiology 15, 83–94.

van Dijk, D.E., 1997. Insect faunas of South Africa from the upper Permian and the Permian-Triassic boundary. Palaeontologica Africana 34, 43–48.

ACCEPTED MANUSCRIPT

Van Roy, P., Orr, P.J., Botting, J.P., Muir, L.A., Vinther, J., Lefebvre, B., el Hariri, K., Briggs, D.E.G., 2010. Ordovician faunas of Burgess Shale type. Nature 465, 215–218.

Veizer, J., Prokoph, A., 2015. Temperatures and oxygen isotopic composition of Phanerozoic oceans. Earth-Science Reviews 146, 92–104.

Wall, P.D., Ivany, L.C., Wilkinson, B.H., 2009. Revisiting Raup: Exploring the influence of outcrop area on diversity in light of modern sample-standardization techniques. Paleobiology 35, 146–167.

Wang, B.O., Zhao, F., Zhang, H., Fang, Y.A.N., Zheng, D., 2012. Widespread pyritization of insects in the Early Cretaceous Jehol Biota. PALAIOS 27, 708–712.

Wilby, P.R., Briggs, D.E.G., Bernier, P., Gaillard, C., 1996. Role of microbial mats in the fossilization of soft tissues. Geology 24, 787–790.

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, 498–520.

Xiao, S., Schiffbauer, J.D., 2009. Microfossil Phosphatization and Its Astrobiological Implications, In: Seckbach, J., Walsh, M. (Eds.), From Fossils to Astrobiology: Record of Life on Earth and Search for Extraterrestrial Biosignatures. Springer Netherlands, pp. 89–117.

Xiao, S., Schiffbauer, J.D., McFadden, K.A., Hunter, J., 2010. Petrographic and SIMS pyrite sulfur isotope analyses of Ediacaran chert nodules: Implications for microbial processes in pyrite rim formation, silicification, and exceptional fossil preservation. Earth and Planetary Science Letters 297, 481–495.

ACCEPTED MANUSCRIPT

Xiao, S., Yuan, X., Steiner, M., Knoll, A.H., 2002. Macroscopic carbonaceous compressions in a terminal Proterozoic shale: A systematic reassessment of the Miaohe biota, South China. Journal of Paleontology 76, 347–376.

Xiao, S., Zhang, Y., Knoll, A.H., 1998. Three-dimensional preservation of algae and animal embryos in a Neoproterozoic phosphorite. Nature 391, 553–558.

Yang, J., Ortega-Hernández, J., Butterfield, N.J., Liu, Y., Boyan, G.S., Hou, J.-b., Lan, T., Zhang, X.-g., 2016. Fuxianhuiid ventral nerve cord and early nervous system evolution in Panarthropoda. Proceedings of the National Academy of Sciences 113, 2988–2993.

Zhang, H., Xiao, S., Liu, Y., Yuan, X., Wan, B., Muscente, A.D., Shao, T., Gong, H., Cao, G., 2015. Armored kinorhynch-like animals from the early Cambrian. Scientific Reports 5 (16521), 1–10.



Drew Muscente is a Postdoctoral Fellow at Harvard University. He received a B.S. in Biological Sciences from Cornell University and a Ph.D. in Geosciences from Virginia Tech. His work in paleobiology, taphonomy, and sedimentary geology is largely focused on the late Neoproterozoic-early Paleozoic interval, which signifies a transformative time in Earth history for life and near-surface environments. To improve understanding of biosphere-geosphere coevolution during this time, Drew studies exceptionally preserved fossils using electron and X-ray microbeam techniques; reconstructs paleoenvironmental conditions and processes from sedimentologic, petrographic, and geochemical data; and investigates for novel paleontologic and geobiologic patterns by compiling and analyzing ‘big data’ geoscientific resources.

ACCEPTED MANUSCRIPT



James D. (Jim) Schiffbauer is an Assistant Professor in the Department of Geological Sciences and Director of the MizzoμX Core Facility for X-ray microanalysis at the University of Missouri. He received his Ph.D. in Geosciences from Virginia Tech. An expert in electron and X-ray microbeam techniques, his research draws primarily from paleobiology and geochemistry to focus on dissecting pathways of exceptional preservation at the Ediacaran–Cambrian transition. He has additionally contributed to our understanding of the tubular fauna in the last ten million years of the Ediacaran Period.



Jesse S. Broce is a Ph.D. student at University of Missouri, supervised by James Schiffbauer. He received a M.S. in Geosciences from Virginia Tech, where he studied phosphatized fossil embryos. His current research focus is taphonomy of soft-tissue preservation, particularly Burgess Shale-type preservation, which he investigates through a combination of fossil microanalysis and experimental decay/mineralization.



Marc Laflamme is an Assistant Professor in the Department of Chemical and Physical Sciences at the University of Toronto Mississauga. He received his PhD at Queen's University. His interests lie in the classification and preservation of the Ediacara biota, the oldest large and complex organisms in the rock record. His research focuses on the evolutionary hierarchy and relationships among the Ediacara biota, their unique mode of preservation, and their relationship to animals.

ACCEPTED MANUSCRIPT



Ken O'Donnell is a geologist at AECOM. He received his B.S. in Geology at Beloit College and completed a master's degree in Geology at Virginia Tech. His research interests include sequence and biostratigraphy, paleoenvironmental analysis, conservation paleobiology, and shale petrography. His current work includes providing geological support for construction and infrastructure projects in the USA and abroad, and the application of sequence stratigraphy to optimize environmental remediation strategies.



Thomas Boag is a graduate student in Historical Geobiology at Stanford University, supervised by Erik Sperling. He received his undergraduate training in Geological Sciences at Queen's University (Kingston, Ontario) in 2014 and completed a masters degree in Paleobiology at the University of Toronto in 2015. His research focuses on the origins and early evolution of animal life and the changes to both the geosphere and biosphere during the Neoproterozoic to Paleozoic transition, specifically the Ediacaran. He integrates stratigraphic fieldwork with laboratory analyses of sedimentary geochemistry and animal physiology to better understand the Ediacaran biostratigraphic record.



Michael Meyer is a researcher at the Geophysical Laboratory at the Carnegie Institution for Science and the Project Science Manager of the Keck Foundation funded project "The Co-evolution of the Geo- and Bio-spheres." He received a B.S. in Geology and Anthropology from Beloit College, a M.S. in Geology

ACCEPTED MANUSCRIPT

from the University of South Florida, and a Ph.D. in Geosciences from Virginia Tech. The primary focus of his research is the evolutionary history of complex life, which he studies through paleontological, sedimentological, and biogeochemical investigations. His work often focuses on the Ediacaran (635–542 million years ago); a time when large perturbations to the atmosphere and ocean were ongoing. Due to these factors, his research is inherently multidisciplinary, and takes advantage of numerous analytical methodologies to answer novel questions about the development and evolution of early Earth systems while easily collaborating across fields.



Andrew Hawkins is a Ph.D. candidate at Virginia Tech, supervised by Shuhai Xiao. He received a B.S. in Geology from James Madison University and an M.S. in Geology from SUNY Buffalo. His research interests are diverse and include identifying and quantifying sources of bias in the fossil record, the Paleozoic fossil record of bromalites (a class of trace fossils that includes coprolites and cololites, i.e. mineralized gut contents), and the biostratigraphy of Ediacaran acritarchs. His work involves modeling approaches to understanding how factors such as lithification and collecting methodology can bias the diversity of fossil samples, Ordovician bromalites from North America, and acritarch biostratigraphy in the Doushantuo Formation of South China.



John Warren Huntley is an Assistant Professor in the Department of Geological Sciences at the University of Missouri. He received a B.S. in Geology at Appalachian State University, an M.S. in Geology at the University

ACCEPTED MANUSCRIPT

of North Carolina Wilmington, and a Ph.D. in Geosciences at Virginia Tech. His work focuses on the relationships between biotic interactions, environmental change, and macroevolutionary patterns and how to disentangle all of these signals from the various geological and preservational biases inherent in the fossil record.



Maria McNamara is a Senior Lecturer at University College Cork (Ireland). She received a Ph.D. in paleontology from University College Dublin. Her primary research interests relate to soft tissue preservation in fossil animals, in particular, the preservation of colour in fossil insects and vertebrates and the preservation of fossil skin and feathers. Her approach combines chemical and ultrastructural analyses of fossil tissues with fossilization experiments in order to better constrain interpretations of fossil data.



Lindsay A. MacKenzie is a Visiting Assistant Professor in the Department of Geological Sciences at SUNY Geneseo. She received a B.Sc. in Solid Earth Geology from Purdue University, a M.Sc. in Biological Sciences from the University of Alberta and a Ph.D in Geosciences from the University of Montana. She researches the taphonomy of fossil-*Lagerstättten*, with focuses on the Early Cambrian Chengjiang biota and the Devonian Man on the Hill assemblage. To fully understand the processes responsible for exceptional preservation, Lindsay utilizes a multidisciplinary approach combining sedimentological, paleobiological and geochemical analyses in addition to experimental taphonomy. She is working on expanding her focus to find novel ways to compare

ACCEPTED MANUSCRIPT

these different deposits to reveal specific taphonomic pathways responsible for exceptional preservation throughout the geologic record.



George D. Stanley is a Professor and Director of the Paleontology Center at the University of Montana. He received his Ph.D. from the University of Kansas. His research interests include soft-bodied cnidarians and gelatinous fossils of the Cambrian. His other research focuses on the evolution of reefs, mass extinctions and recoveries and the evolution of corals.



Nancy Hinman is a Professor in the Geosciences Department at the University of Montana. She received her undergraduate degree in chemistry from Reed College and her Ph.D. in oceanography from University of California, San Diego. Overall, she studies the effects of local environmental conditions on the preservation of biosignatures. Her approach involves unraveling complex organic-mineral interactions in extreme environments with high-resolution techniques.



Michael H. Hofmann is a Research Assistant Professor in the Department of Geosciences at the University of Montana and CEO of AIM GeoAnalytics. He received his Dipl. Geol. (M.S. equivalent) in Geology from the University of Erlangen-Nürnberg, Germany, and his Ph.D. in Geology from the University of Montana. His research interest includes the effects that primary sedimentary processes and stratigraphic architecture have on early diagenesis and the preservation of organic matter. He uses a multitude of analytical techniques in the field and laboratory to unravel the rock record.

ACCEPTED MANUSCRIPT



Shuhai Xiao is a Professor of Geobiology in the Department of Geosciences at Virginia Tech. He received a B.S. in Geology from Peking University (China) and a Ph.D. in Biology from Harvard University. Prior to joining the faculty at Virginia Tech, he taught for several years at Tulane University. His primary research interests include Proterozoic and Cambrian paleobiology, stratigraphy, paleoenvironments, geobiology, and taphonomy.

Fig. 1. Exceptionally preserved fossils. (a) Phosphatized embryo *Megasphaera* from the Ediacaran Doushantuo Formation, China. (b) Silicified multicellular alga *Wengania* from the from Ediacaran Doushantuo Formation, China. (c) Pyritized tubular metazoan *Conotubus* from the Ediacaran Dengying Formation, China. (d) Ediacara-type fossil *Swartpuntia* from the Ediacaran Nama Group, Namibia. (e) Aluminosilicified carbonaceous compression of arthropod *Marrella* from the Cambrian Burgess Shale, Canada. (f) Oxidized compression of a ctenophore ("comb jelly") from the Cambrian Qiongzhusi Formation (Chengjiang Biota), China (Yunnan University specimen RCCBYU 10217). (g) Carbonaceous compression of eurypterid from the Silurian Bertie Waterlime (Fiddlers Green Formation), US. (h) Silicified mayfly from the Miocene Barstow Formation, US. (i) Carbonaceous compression of insect *Fulgora* from the Eocene Green River Formation, US. (j) Fish from the Eocene Green River Formation, US.

Fig. 2. Fossil evidence of interconnected taphonomic pathways. (a–h) Phosphatized/silicified acritarch in chert nodule from the Jiulongwan section of the Ediacaran Doushantuo Formation of Hubei Province, South China. (a) Transmitted-light image. (b–d) Scanning electron microscopy

ACCEPTED MANUSCRIPT

(SEM) images of specimen in (a), showing calcium phosphate (granular white material) surrounded by chert (gray amorphous material). (e–h) Energy dispersive X-ray spectroscopy (EDS) elemental maps of (c). (i–p) Pyritized/aluminosilicified/carbonaceous hemichordate *Yuknessia simplex* from the Cambrian Wheeler Formation in the Drum Mountains, Utah, US. (i) Transmitted-light image. (j–l) SEM images of specimen in (a) showing black carbonaceous material (j), light-gray clay minerals (k), and white pyrite framboids (k, l). (m–p) EDS elemental maps of (j).

Fig. 3. Unifying model of exceptional taphonomic pathways with reactions, geochemical gradients, and microbial zones (not to scale) associated with preservational processes distributed in an ideal sediment profile. Model assumes soft tissue mineralization occurs exclusively within sediment, as suggested by studies of various Konservat-Lagerstätten (Briggs et al., 1991; Briggs et al., 1996; Cai et al., 2012; Guan et al., 2016; Muscente et al., 2015a). Naturally, sedimentary environments often vary from this ideal profile, with some environments disproportionately favoring certain preservational processes over others, owing to the processes' different prerequisites. Such natural variation in sedimentary environments sets the stage for all taphonomic pathways involving combinations of processes. Geochemical perturbations caused by carcasses entering microbial zones (Sagemann et al., 1999) as well as decay and mineralization of soft tissues within microenvironments of carcasses (McNamara et al., 2009) may cause preservational styles to deviate from those predicted by this simple model.

Fig. 4. Overview of dataset. For all plots, the *x* axis represents geologic time. (a) Number of exceptionally preserved assemblages. (b, c) Proportions of assemblages from marine,

ACCEPTED MANUSCRIPT

transitional, and non-marine depositional settings in 30.5 my (b) and 20.3 my (c) duration time bins. (d, e) Paleolatitudes. (d) Paleolatitudes of all assemblages (n=694) in dataset, estimated using G-Plates. (e) Paleolatitudes of subset (n=548) of dataset, estimated from collections in Paleobiology Database (PBDB). Cryo, Cryogenian; E, Ediacaran; Є, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian, Tr, Triassic; J, Jurassic; K, Cretaceous; Pg, Paleogene; N, Neogene; Q, Quaternary.

Fig. 5. Latitudinal, longitudinal, and age/paleolatitudinal distributions of exceptionally preserved fossil assemblages and randomly sampled PBDB fossil collections. (a, b) Latitudinal (red) and longitudinal (blue) distributions of PBDB collections (a) and exceptionally preserved assemblages (b) with map of present day locations of exceptionally preserved assemblages included in (b). (c, d) Age/paleolatitudinal distributions of PBDB collections (c) and exceptionally preserved assemblages (d) for each of the geologic systems (Ediacaran-Neogene). (e–g) Plots of counts of exceptionally preserved assemblages and mean counts of PBDB collections for latitudinal (e), longitudinal (f), and age/paleolatitudinal bins (g) with fitted linear regression models and correlation analysis results.

Fig. 6. Nearest neighbor analysis and geospatial statistics. (a) Plot showing number of unique assemblage pairs (n=583) versus the distance between the points in the couples. Each pair is a unique combination of an assemblage and its nearest neighbor, but some assemblages occur in multiple pairs because some nearest neighbor relationships are non-mutual and some assemblages are equidistant from two or more points. (b–d) Pie charts showing percentages of unique nearest neighbor pairings of assemblages from matching and different chronostratigraphic

ACCEPTED MANUSCRIPT

series and depositional settings. (e) Moran's I values of spatial autocorrelation in age with respect to distance between assemblages. Histogram shows counts of point pairs used in calculating the values. Crosses indicate values that fail significance testing ($\alpha = 0.01$), and do not represent evidence of spatial autocorrelation. Moran's I value ($I = 0.79$) for nearest neighbors indicated by dashed line (significant at $\alpha = 0.01$). (f) Ripley's L(d) function values. As indicated in the key, black line values are based on counts of all assemblages (regardless of their depositional settings); the blue, red, and green lines include values calculated from the marine, transitional, or non-marine subsets of the dataset, respectively; and the values in remaining lines are based on counts of assemblages from different depositional settings. Results of randomization testing of assemblage geospatial homogeneity and categorical (i.e. depositional setting) association are included as 95% CI envelopes. For each specific distance, L(d) values less than this CI represent significant evidence of clustering, and those greater than the interval represent evidence of uniformity.

Fig. 7. Cluster counts derived using the complete-linkage method. For all plots, the *x* axis is geologic time and *y* axis is the number of clusters. Assemblages in left and right columns were hierarchically clustered in 30.5 my and 20.3 my duration time bins, respectively. (a, b) Total cluster counts from hierarchical clustering of assemblages independently of depositional setting categories. (c–j) Cluster counts from hierarchical clustering of assemblages according to depositional setting categories. (c, d) Marine cluster counts. (e, f) Transitional cluster counts. (g, h) Non-marine cluster counts. (i) Total cluster counts (c+e+g). (j) Total cluster counts (d+f+h). (k, l) Total cluster counts from hierarchical clustering of assemblages according to marine/transitional and non-marine categories. Note that the general trends through geologic

ACCEPTED MANUSCRIPT

time are similar regardless of the time bin durations (a, c, e, g, i, k versus b, d, f, h, j, l) and data categories (a, b versus i, j versus k, l) used in cluster counting. See Supplementary Information for cluster counts from analyses employing other clustering methods.

Fig. 8. Pearson's chi-square testing of cluster count equality among time bins. Each plot contains *P* values (y-axis) and corresponding MPLDs (x-axis) determined from various tests (n=75 in plots of left and middle columns; n=15 in plots of right column) of total cluster count equality among time bins in time series, as shown in the example on the left. Different line styles correspond to different clustering methods, columns correspond to different data categories, and rows correspond to different time bin durations used in clustering. Diagonally hatched areas indicate *P* values determined from data tables where fewer than 5 clusters occur in >80% of expected value cells, and Pearson's chi-square test has low statistical power. Note that the results for MPLDs < 1000 km are similar regardless of clustering method (lines), time bin duration (rows), and data categories (columns) used in cluster counting.

Fig. 9. Dataset randomization testing of cluster count equality among time bins. For all plots, the y axis is cluster count, and the x axis is geologic time. Black lines indicate observed values; green lines indicate expected values, assuming time bins generally have equal cluster counts; and red lines indicate 99% CI, based on 1000 replicates. Different line styles correspond to different clustering methods, columns correspond to different time bin durations, and rows correspond to different data categories (clustering independently of depositional setting categories or dependent upon marine, transitional, and non-marine categories) and MPLDs (300 and 1000 km) used in cluster counting. Note that the results corroborate chi-square equality testing results.

ACCEPTED MANUSCRIPT

Fig. 10. Rock quantity estimates and Phanerozoic sea level curve. For all plots, the *x* axis is geologic time. Left and right columns provide rock quantity estimates for 30.5 my and 20.3 my duration time bins, respectively. Estimates of both marine and non-marine rocks are provided. (a, b) Global rock outcrop areas. (c, d) Global rock volumes. (e, f) North American surface/subsurface rock areas. A relative sea level curve is provided in (b).

Fig. 11. Correlation analyses. For all plots, x-axis is MPLD. Cluster counts based on various MPLDs were determined for 30.5 my and 20.3 my duration (t) time bins using the complete-linkage method. The *y* axes in the left, middle, and right columns correspond to P (a, d, g, j, m, p), R (b, e, h, k, n, q), and R^2 (c, f, i, l, o, r) values, respectively, for the correlations of the x and y variables listed on the left. Whereas the cluster counts compared to global rock outcrop areas and global rock volumes are based on the entire database of exceptionally preserved assemblages, the cluster counts compared to North American rock areas are based only on the North American subset of the database. The global sea level value of each bin is the mean of the levels at the bins' early and late boundaries. See Supplementary Information for first-difference testing of correlation results.

Fig. 12. Marine, transitional, and non-marine cluster proportions. For all rectangular plots, the *x* axis is geologic time and *y* axis is the proportion of clusters. Assemblages in left and right columns were hierarchically clustered in 30.5 my (a, c, e) and 20.3 my (b, d, f) duration time bins, respectively, according to marine (a, b) , transitional (c, d), and non-marine (e, f) depositional setting categories using the complete-linkage method (see Supplementary

ACCEPTED MANUSCRIPT

Information for results from analyses using other clustering methods). Vertical lines delineate time domains compared in chi-square homogeneity testing. Linear (1^{st}-degree for marine and non-marine clusters; 2^{nd}-degree for transitional clusters) polynomial regression models fit to the time series are also included in the rectangular plots. For all square plots (g–l), the *x* axis is MPLD and *y* axis is *P* value. The *P* values were compiled from F (g, j), Shapiro-Wilk (h, k), and Jarque-Bera (i, l) testing of regression models. Note that the results indicate that the proportions of clusters in each of the categories have significantly changed through time.

Fig. 13. Pearson's chi-square (a–e, k–q) and Fisher's exact (f – j, r–x) testing of time domain heterogeneity. For all plots, x axis is MPLD and y axis is *P* value. Compilation plots show *P* values from testing of pairwise time domain homogeneity using contingency table in key. Each plot includes *P* values determined from numerous tests of data compiled based on various clustering methods and MPLDs. The time series in each specific plot were compiled via hierarchical clustering of assemblages in time bins of duration (t) according to specific depositional setting categories. Based on the ages of their assemblages, the resulting clusters were assigned to time domains. As shown in the key, the values (*a*, *b*, *x*, *y*) in each plot indicate time domains and depositional setting categories used in determination of *P* values. Diagonally hatched areas indicate *P* values determined from contingency tables with fewer than 5 clusters in one or more cells of expected values, and Pearson's chi-square test has low statistical power.

Fig. 14. Marine/transitional and non-marine cluster proportions. For all rectangular plots, the *x* axis is geologic time and *y* axis is the proportion of clusters. Assemblages in left and right columns were hierarchically clustered in 30.5 my (a, c) and 20.3 my (b, d) duration time bins,

ACCEPTED MANUSCRIPT

respectively, according to marine/transitional (a, b) and non-marine (c, d) depositional setting categories using the complete-linkage method. Vertical lines delineate time domains compared in chi-square homogeneity testing. Linear 1st-degree polynomial regression models fit to the time series are also included in the rectangular plots. For all square plots (e–j), the *x* axis is MPLD and *y* axis is *P* value. The *P* values were compiled from F (e, h), Shapiro-Wilk (f, i), and Jarque-Bera (g, j) testing of regression models. Note that the results indicate that the proportions of clusters in each of the categories have significantly changed through time.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



Figure 1

ACCEPTED MANUSCRIPT

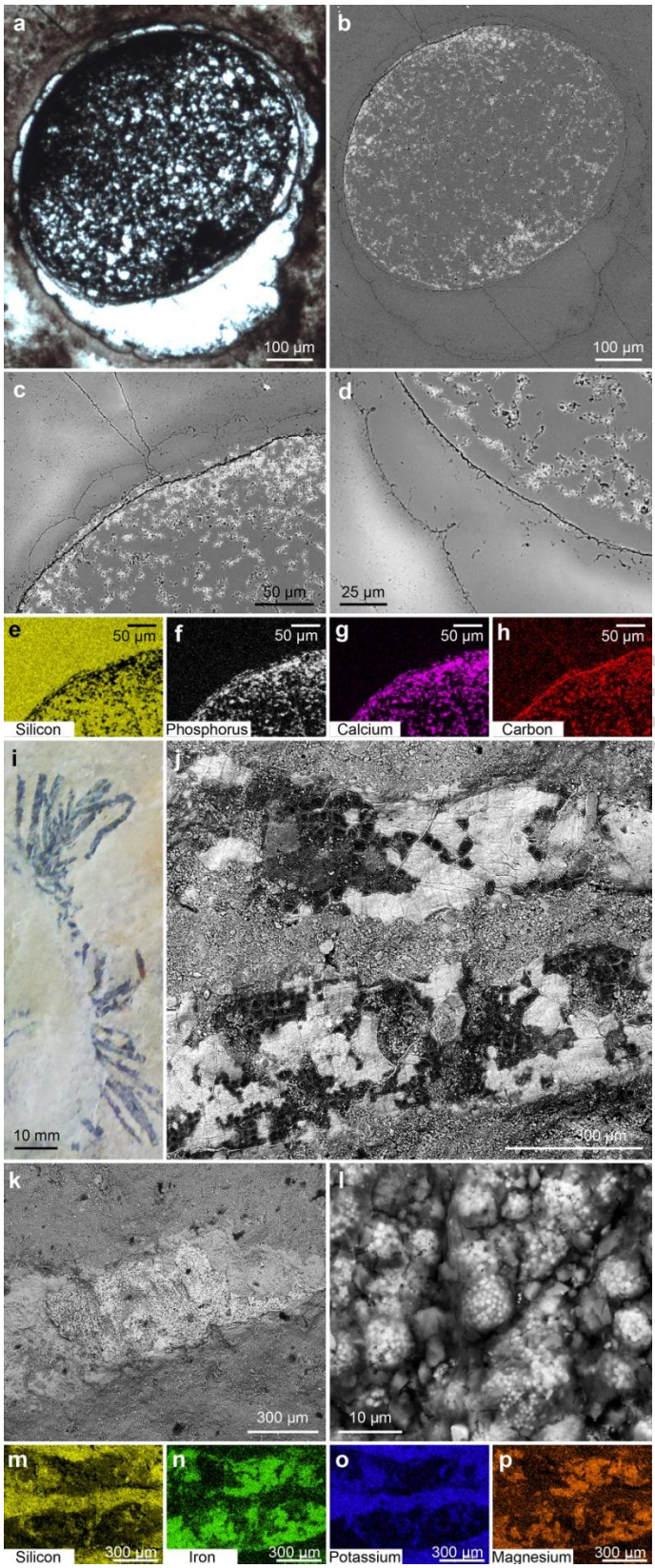


Figure 2

ACCEPTED MANUSCRIPT

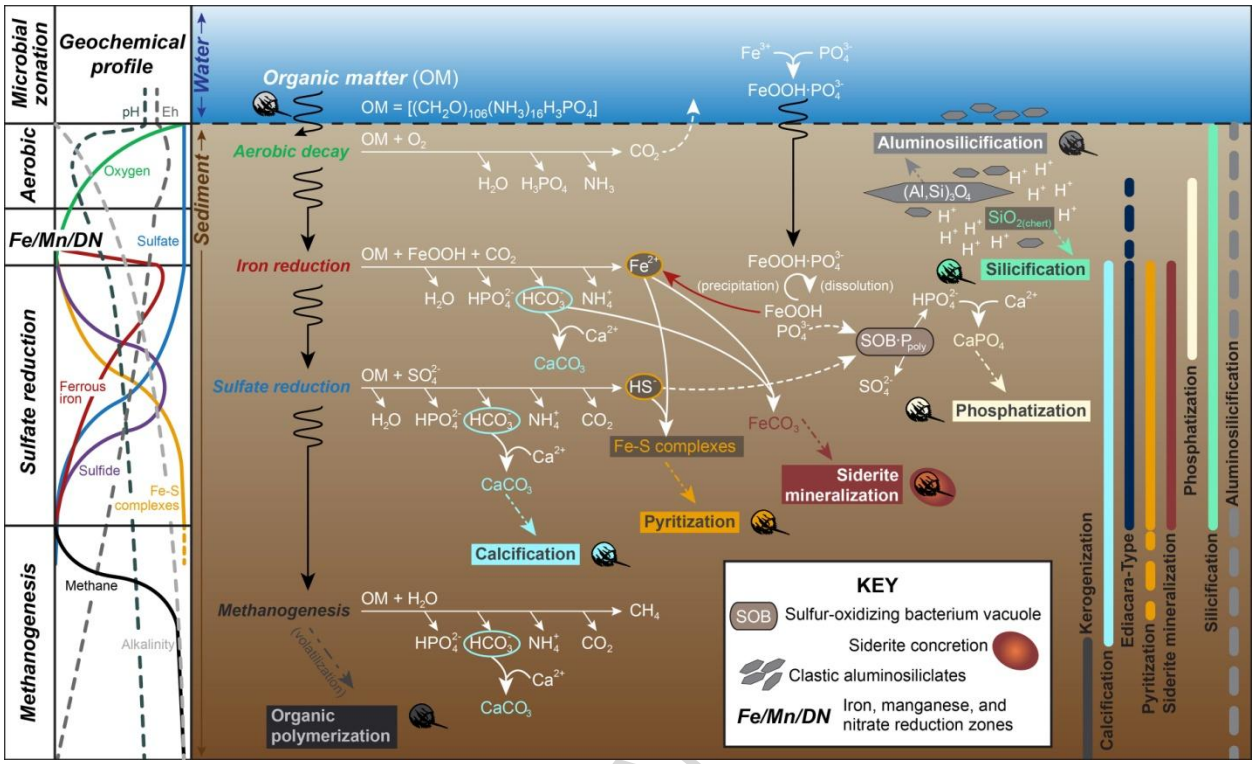


Figure 3

ACCEPTED MANUSCRIPT

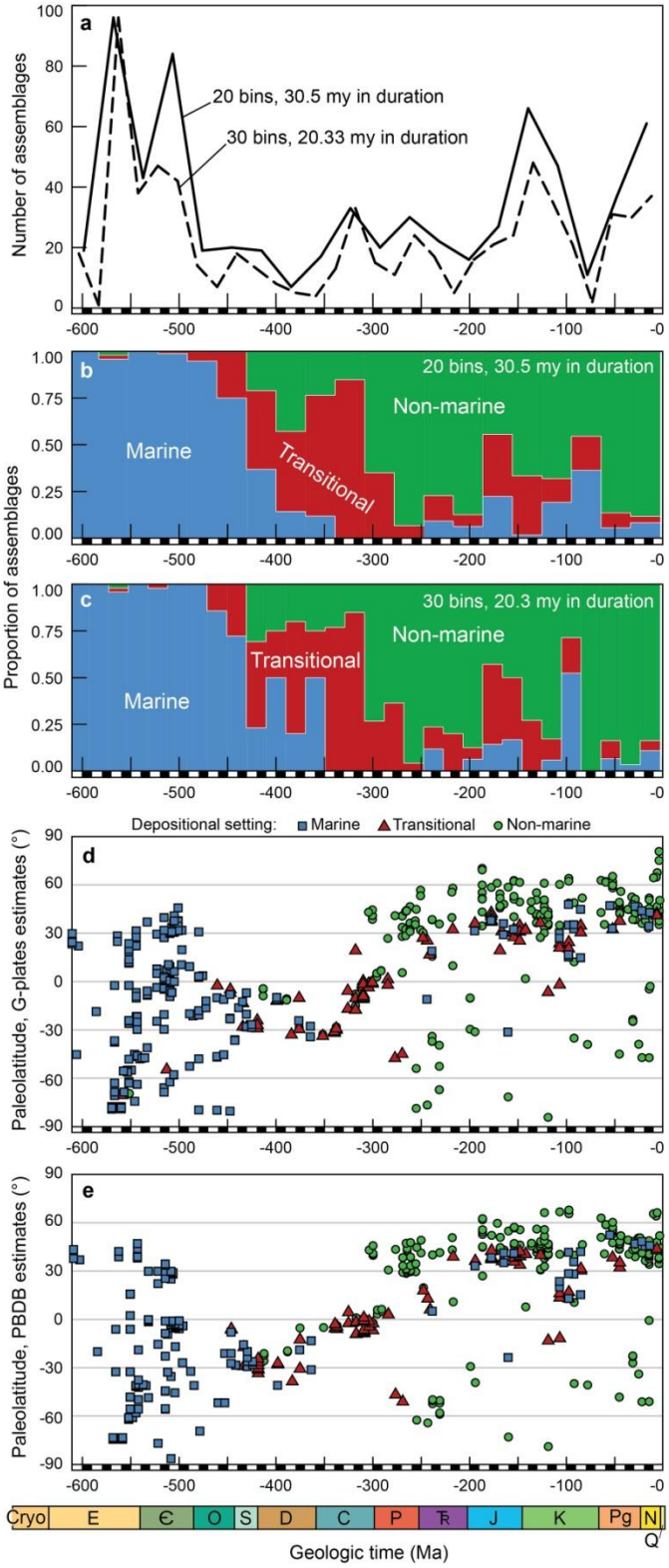


Figure 4

ACCEPTED MANUSCRIPT

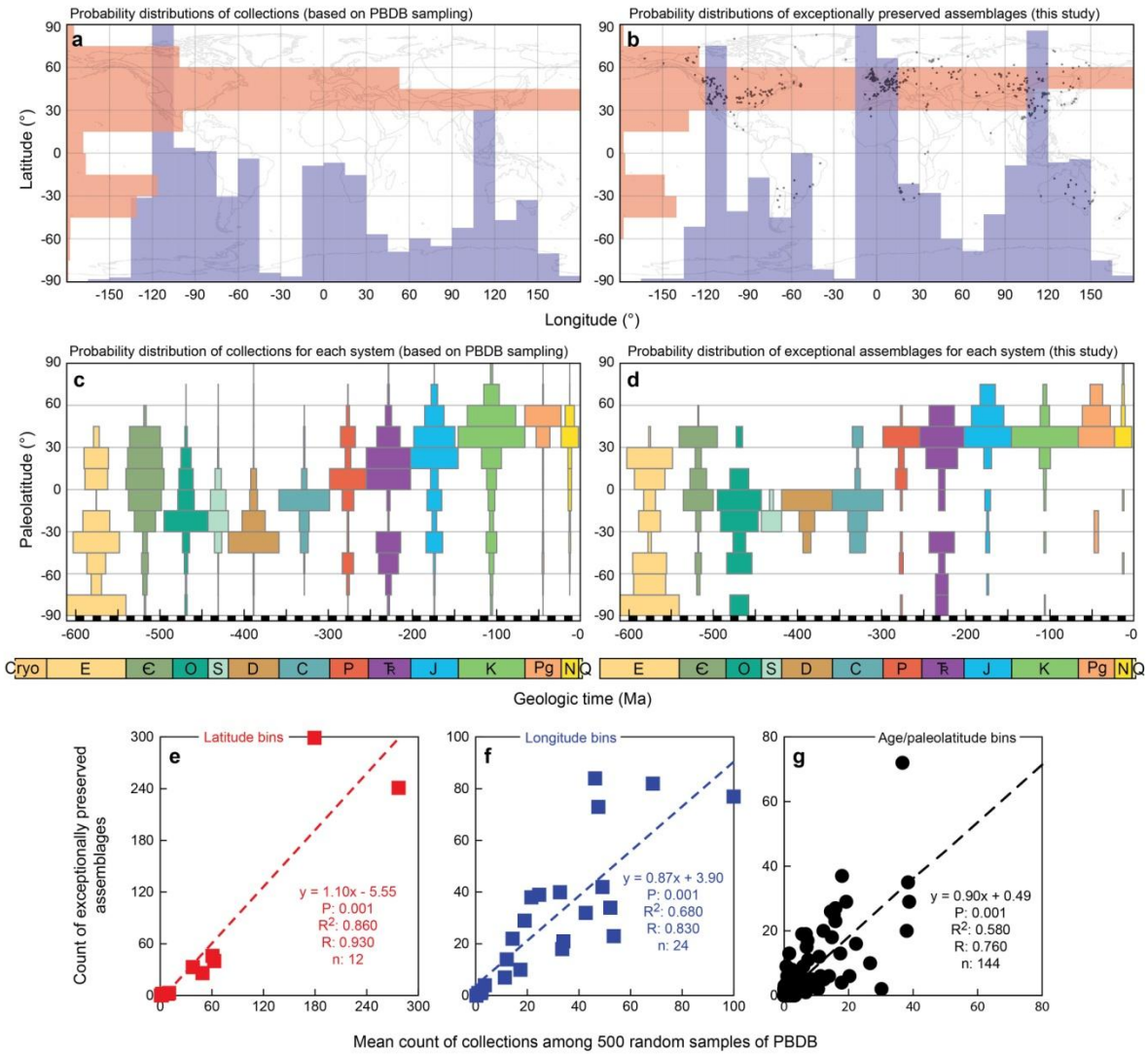


Figure 5

ACCEPTED MANUSCRIPT

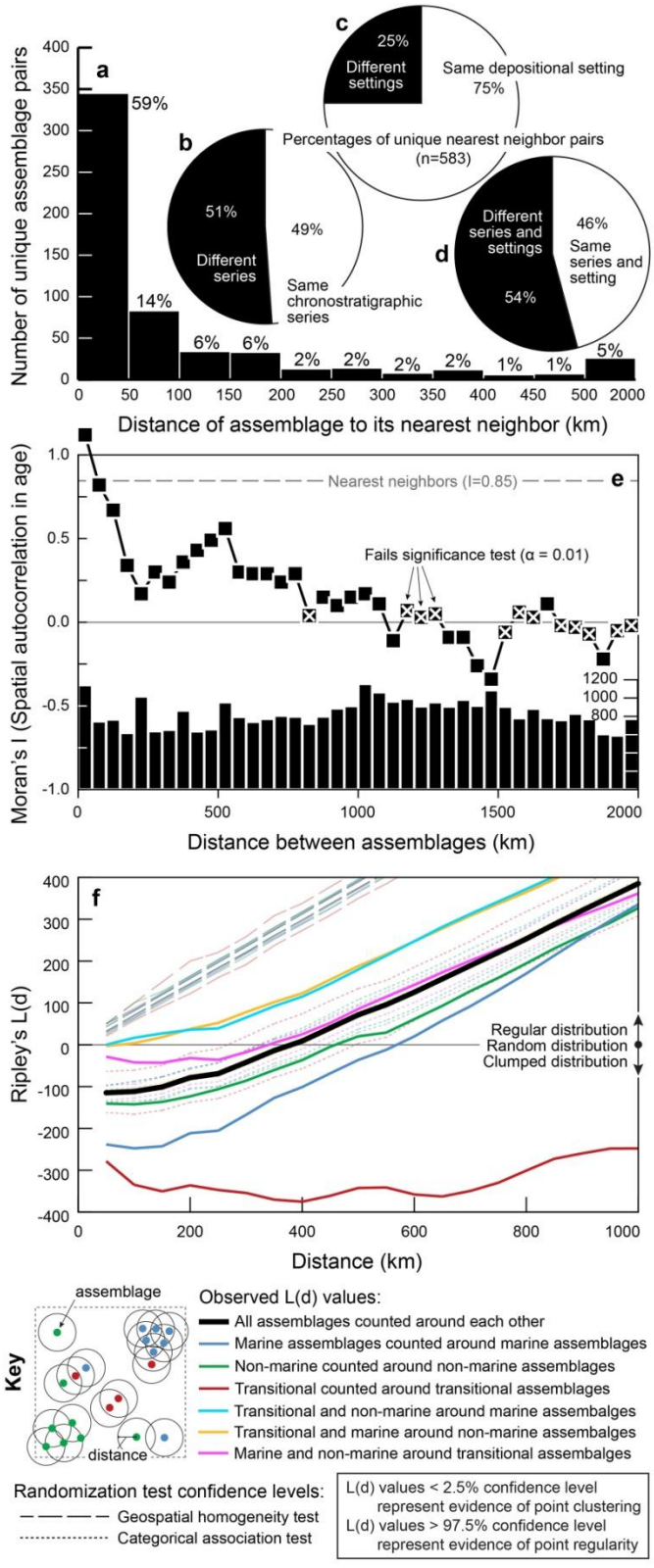


Figure 6

ACCEPTED MANUSCRIPT

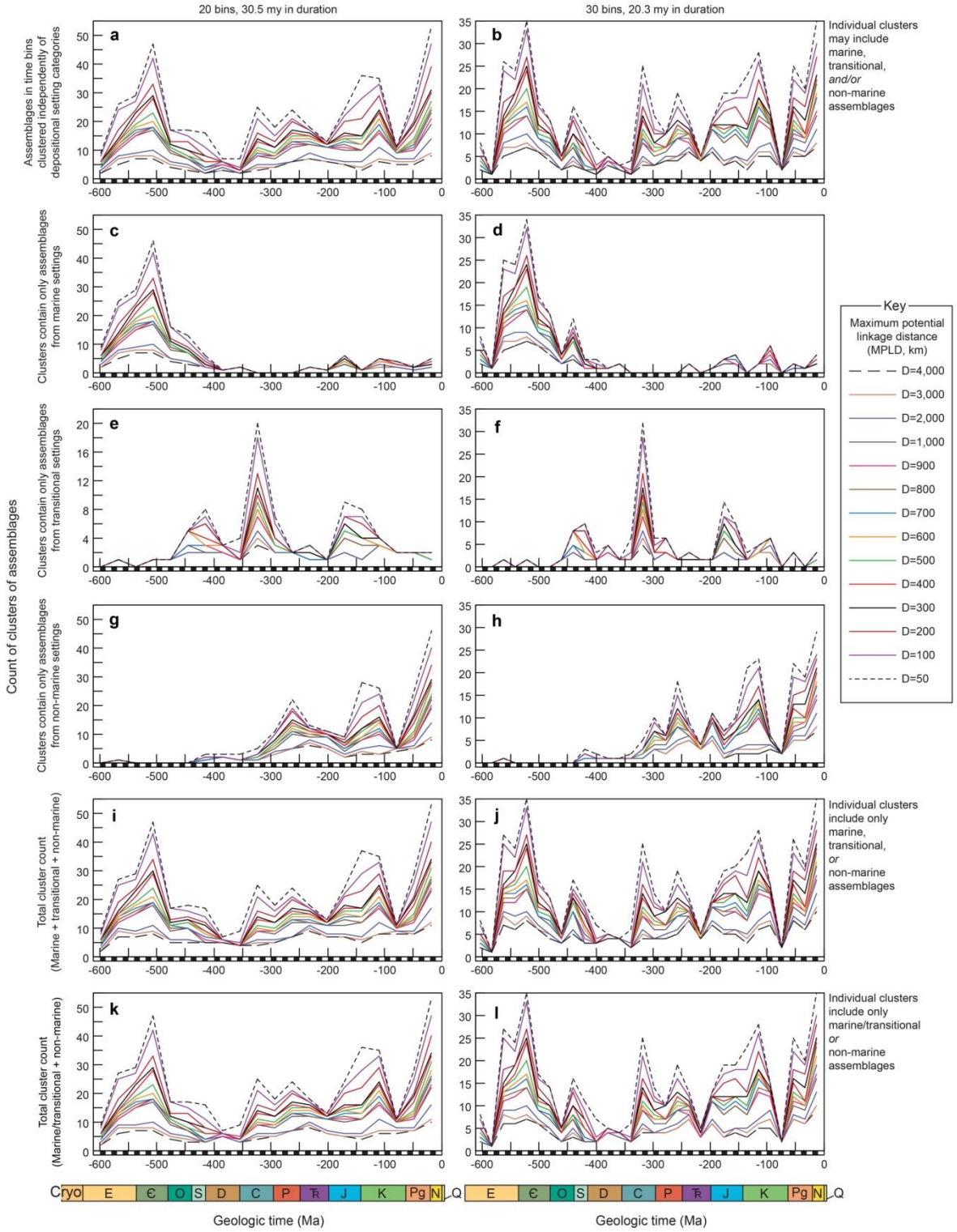


Figure 7

ACCEPTED MANUSCRIPT

Pearson's chi-square test:

H_0 = Time bins equal with respect to assemblage/cluster counts

H_A = Time bins are not equal with respect to assemblage/cluster counts

Data table (n = 20 or n = 30)

Bin	Observed	Expected
1 2 3 4 5 ... n	For each time bin in each time series: = Total count observed within time bin	= Mean total count among n time bins in time series

Example (time series, n = 7)

Bin	Observed	Expected
1	45	60
2	50	60
3	55	60
4	60	60
5	65	60
6	70	60
7	75	60
	P-value = 0.07 Because P>α, fail to reject H_0	

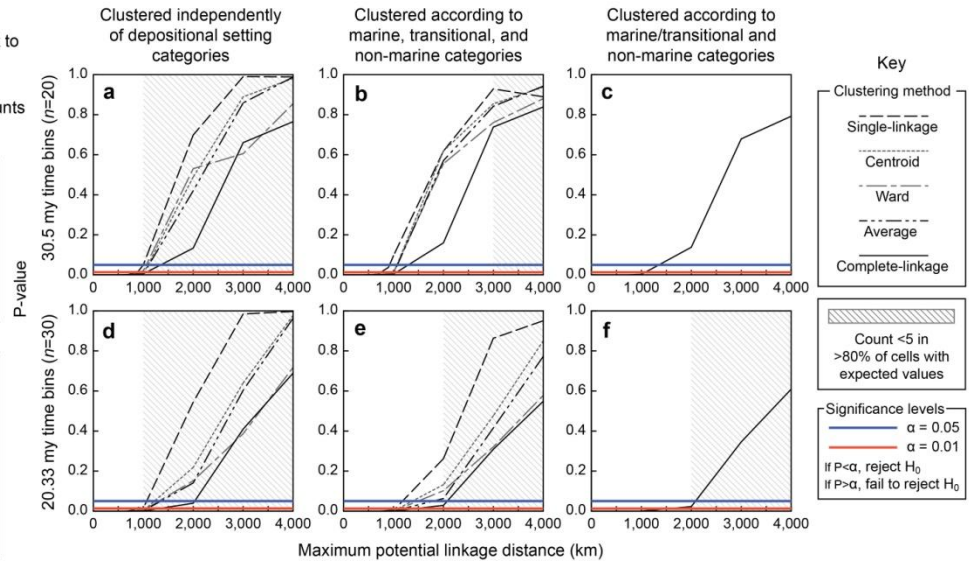


Figure 8

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

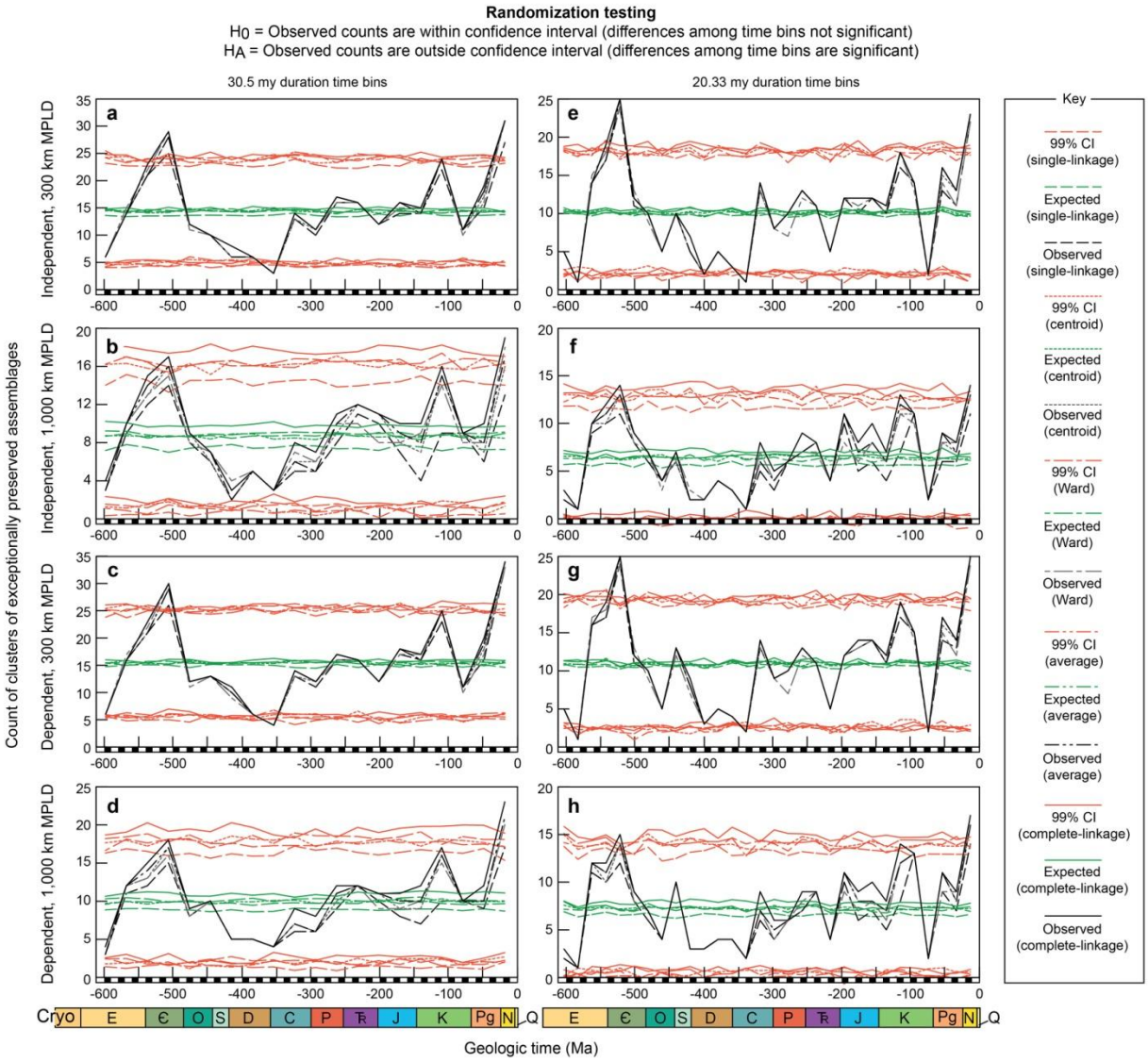


Figure 9

ACCE

ACCEPTED MANUSCRIPT

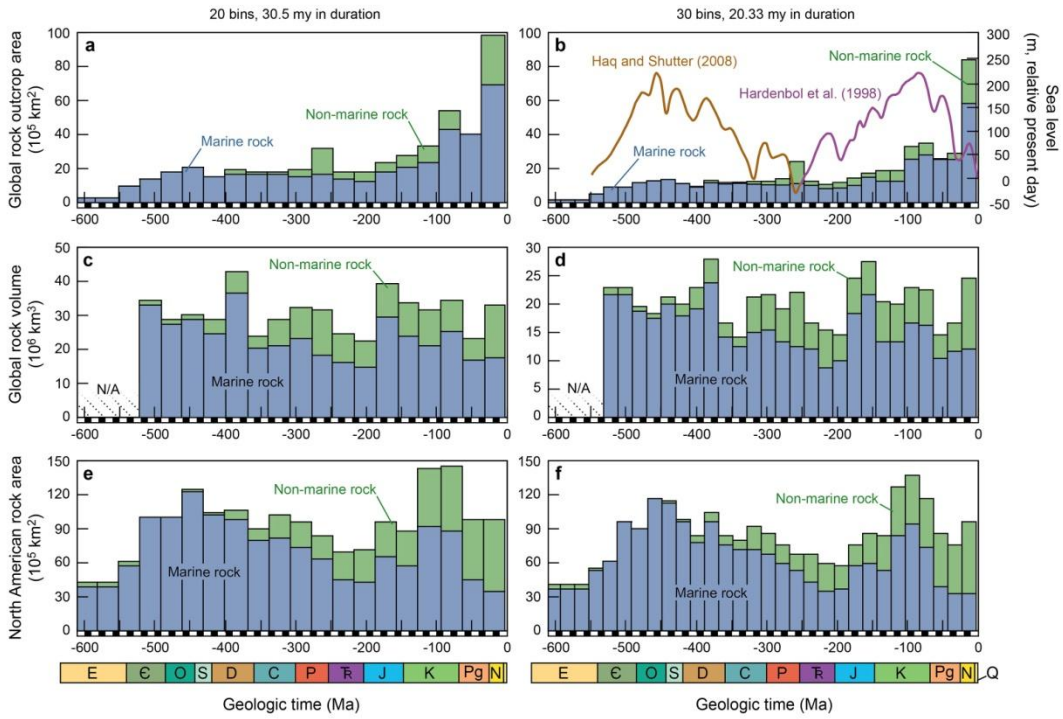


Figure 10

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

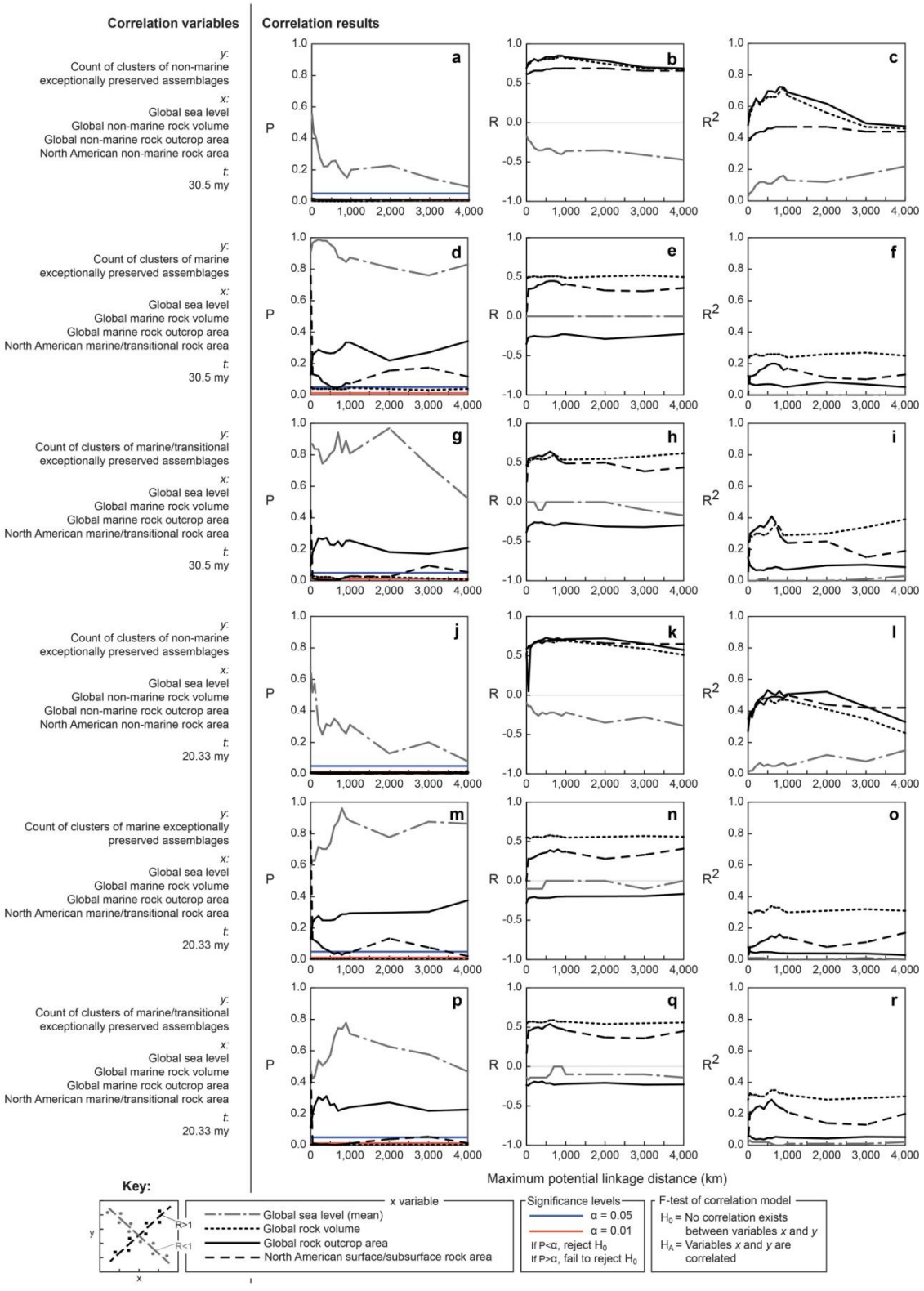


Figure 11

ACCEPTED MANUSCRIPT

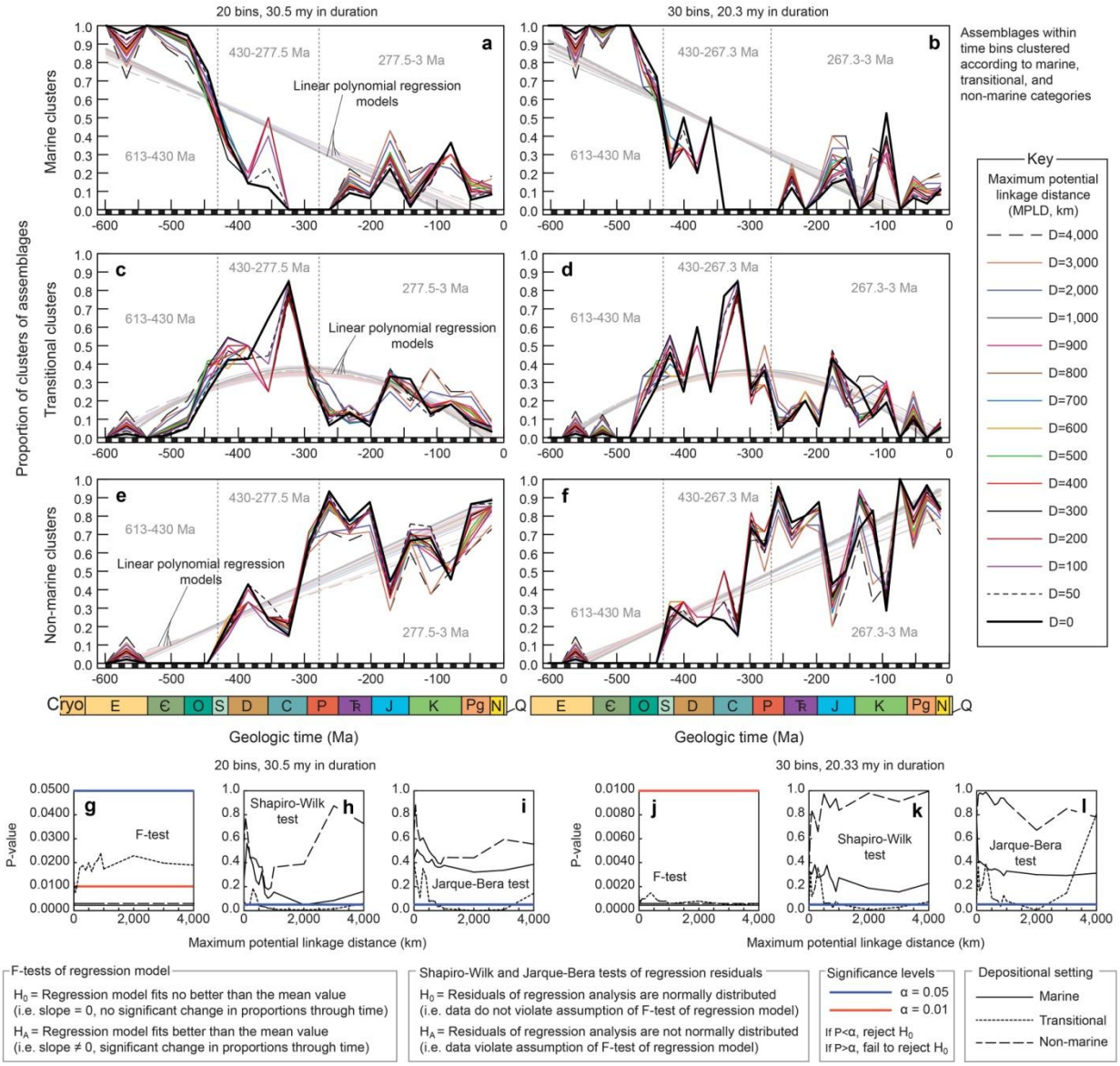


Figure 12

ACCEPTED MANUSCRIPT

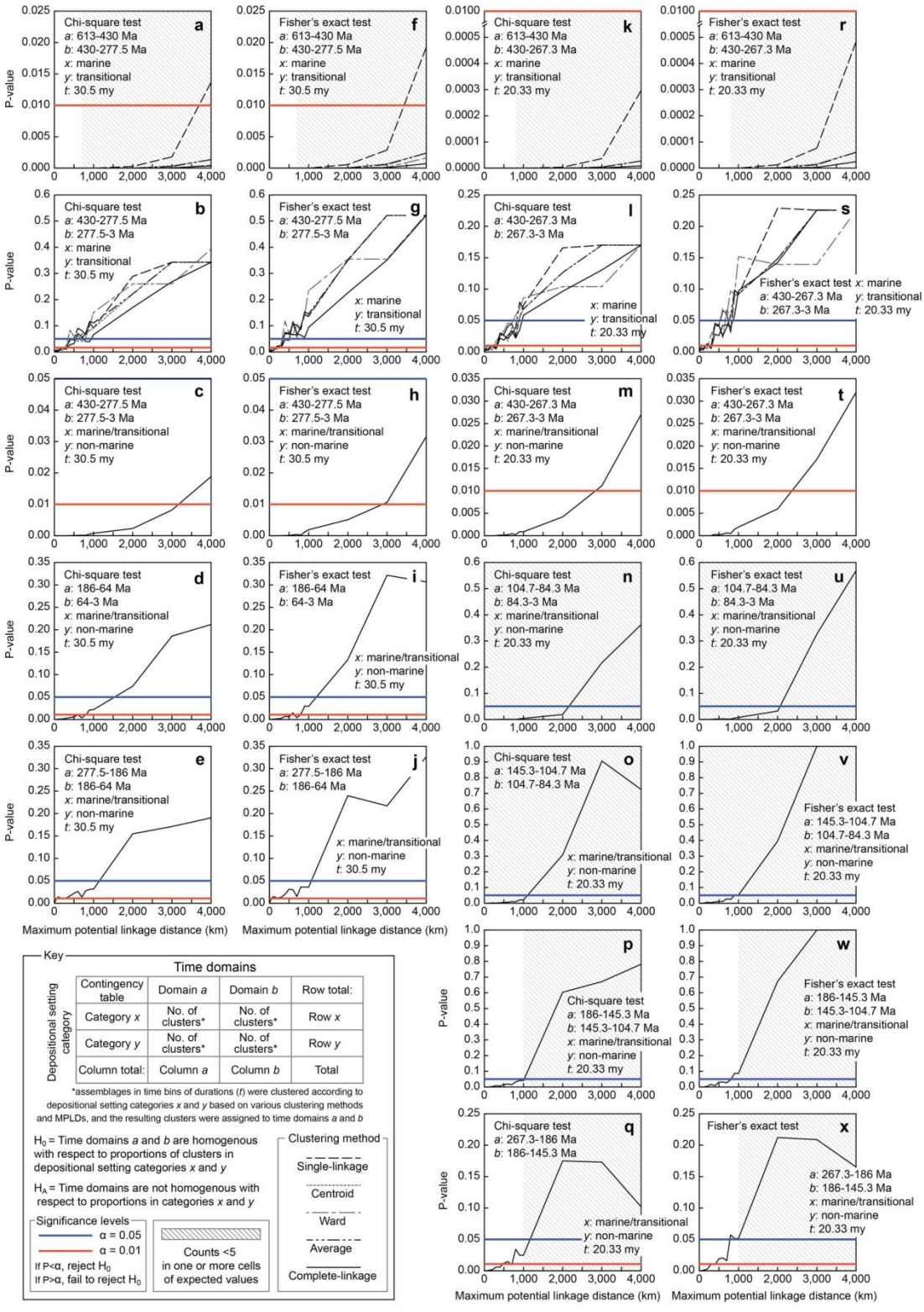


Figure 13

ACCEPTED MANUSCRIPT

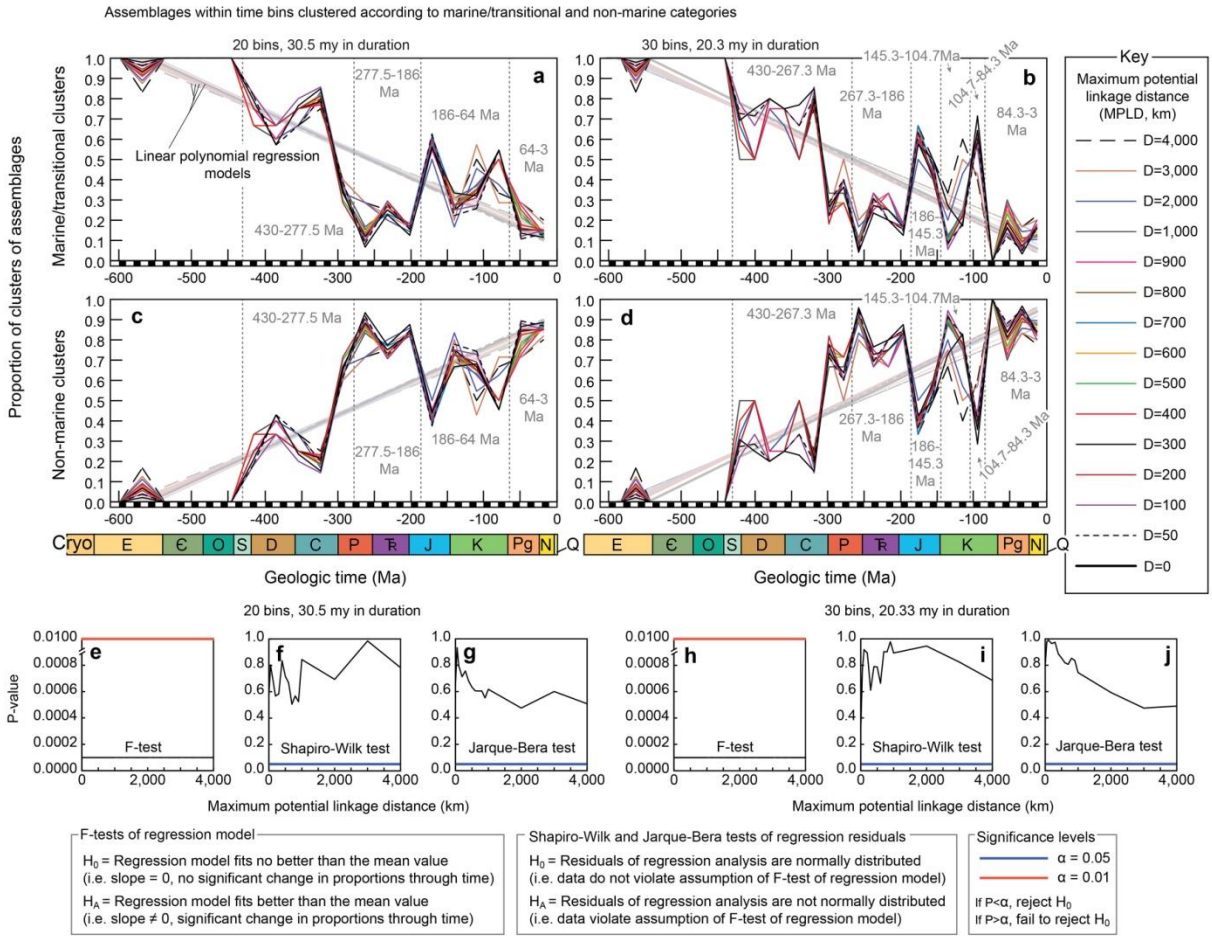


Figure 14

ACCEPTED

ACCEPTED MANUSCRIPT



Graphical abstract

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights

- *Lagerstätten* commonly occur in clusters within geographic and stratigraphic space
- These clusters significantly vary in number and facies through stratigraphy
- Such trends reflect global phenomena affecting fossil preservation and sampling

ACCEPTED MANUSCRIPT