From success to persistence: Identifying an evolutionary regime shift in the diverse Paleozoic aquatic arthropod group Eurypterida, driven by the Devonian biotic crisis

James C. Lamsdell1,2,3 and Paul A. Selden4,5

1Division of Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024
2Department of Geology and Geography, West Virginia University, 98 Beechurst Avenue, Brooks Hall, Morgantown, West Virginia 26506
3E-mail: james.lamsdell@mail.wvu.edu
4Paleontological Institute and Department of Geology, University of Kansas, 1475 Jayhawk Boulevard, Lawrence, Kansas 66045
5Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom

Received October 27, 2015
Accepted October 21, 2016

Mass extinctions have altered the trajectory of evolution a number of times over the Phanerozoic. During these periods of biotic upheaval a different selective regime appears to operate, although it is still unclear whether consistent survivorship rules apply across different extinction events. We compare variations in diversity and disparity across the evolutionary history of a major Paleozoic arthropod group, the Eurypterida. Using these data, we explore the group’s transition from a successful, dynamic clade to a stagnant persistent lineage, pinpointing the Devonian as the period during which this evolutionary regime shift occurred. The late Devonian biotic crisis is potentially unique among the “Big Five” mass extinctions in exhibiting a drop in speciation rates rather than an increase in extinction. Our study reveals eurypterids show depressed speciation rates throughout the Devonian but no abnormal peaks in extinction. Loss of morphospace occupation is random across all Paleozoic extinction events; however, differential origination during the Devonian results in a migration and subsequent stagnation of occupied morphospace. This shift appears linked to an ecological transition from euryhaline taxa to freshwater species with low morphological diversity alongside a decrease in endemism. These results demonstrate the importance of the Devonian biotic crisis in reshaping Paleozoic ecosystems.

KEY WORDS: Biotic crises, macroecology, macroevolution, mass extinction, morphological disparity, phylogeny.

Mass extinctions are known to be extraordinary events during which the normal rules of natural selection do not apply, and it has been suggested that a different selective regime acting upon clade-level properties operates during these periods (Jablonski 2005). This suggests that taxonomic survivorship fits the paradigm of “nonconstructive selectivity” (Raup 1984) caused by nonrandom extinction but with survivorship determined by traits that are not otherwise strongly selected for during periods of background extinction. Of the “Big Five” Phanerozoic mass extinctions (Raup and Sepkoski 1982; Sepkoski 1986), the late Devonian is unusual in being characterized by a drop in speciation rates rather than any increase in extinction (Bambach et al. 2004), although two
discrete extinctions, the Kellwasser and Hangenberg Events at the beginning and end of the Famennian (House 1985, 2002), have been posited. Causal factors for these events remain elusive (Chen et al. 2005), although marine transgression and widespread anoxia are frequently evoked (Bond and Wignall 2008; Carmichael et al. 2016). The cause of the anoxia is thought to be increased productivity due to the influx of nutrients from escalated continental weathering, resulting in eutrophication of the extensive shallow seas (Kaiser et al. 2016; Ma et al. 2016; Rakocinski et al. 2016). However, anoxia was prevalent before (Boyer et al. 2014) and after (Zaton et al. 2014) both the Kellwasser and Hangenberg Events, while some localities exhibit no evidence of anoxia during the events themselves (George et al. 2014), suggesting that anoxia alone could not be responsible for any mass extinction events. The rapid onset of glaciation during the shift out of greenhouse conditions, which prevailed through most of the Devonian (Slavik et al. 2016), has recently been suggested to be the driving factor for the late Devonian extinctions (McGhee et al. 2013) with cooling driven by the radiation of terrestrial plants (Myrow et al. 2014). Meanwhile, selenium depletion has been found associated with a number of extinction events throughout the Phanerozoic, including horizons at the end-Emsian and during the Famennian (Long et al. 2016). The pervasive nature of anoxia and repeated periods of selenium depletion throughout the Middle and Upper Devonian suggests a protracted period of environmental stress from the Emsian onwards with further perturbations such as rapid glaciation adding stress to an ecosystem already in crisis. Furthermore, biogeographic and phylogenetic studies of Laurentian species have suggested that the reduction in speciation is due to a decrease in provincialism caused by increases in interbasinal species invasions (Abe and Lieberman 2009; Stigall 2012). It has been proposed that these invasions were facilitated by transgressive marine pulses that enabled range expansion of ecological generalists, thereby eliminating vicariant speciation (May 1995; Abe and Lieberman 2009; Stigall 2012). Understanding the processes occurring during the late Devonian could, therefore, be important for modern conservation efforts in view of current fears of a human-driven sixth mass extinction (Barnosky et al. 2014) and faunal homogenization through biotic invasions and habitat loss (McKinney and Lockwood 1999).

Eurypterids are aquatic arthropods that form important components of Paleozoic ecosystems and are known to fill a variety of ecological roles (Anderson et al. 2014; McCoy et al. 2015). The group has a well-resolved phylogeny (Lamsdell 2013; Lamsdell et al. 2015), an evolutionary history ranging from the Middle Ordovician (467 mya) to the late Permian (254 mya), and comprises two constituent clades (Eurypterina and Stylonurina). Eurypterids, therefore, present an ideal case study for evaluating the impact of Paleozoic mass extinctions, particularly the late Devonian, which is known to have severely impacted the group (Lamsdell and Braddy 2010). Prior to the Devonian, eurypterids had successfully weathered one mass extinction event and appeared to be a thriving, successful group. Studying their decline, therefore, affords a rare opportunity to observe the ecological and evolutionary collapse of an otherwise globally dominant faunal component.

In this article, we analyze trends in taxonomic diversity and morphological disparity of the two eurypterid clades, Stylonurina and Eurypterina, along with rates of speciation and extinction across their recorded history. We explore the timing and impact of the Devonian biotic crisis, and compare it to the end-Ordovician and Kasimovian extinction events to examine how different forcing mechanisms influenced the trajectory of eurypterid evolution. In this way, we use past mass extinctions as natural experiments to better understand the processes occurring during our current biodiversity crisis.

**Methods**

**PHYLOGENETIC DATA**

For the study, recent phylogenetic matrices for the Stylonurina (Lamsdell 2013) and Eurypterina (Lamsdell et al. 2015) were used. Both analyses were expanded to make species sampling more comprehensive and then combined into a single matrix, resulting in an analysis comprising 226 characters coded for 150 taxa. Of the 250 currently valid species of eurypterids (Dunlop et al. 2015), 145 were included in the phylogenetic analysis, of which 106 were Eurypterina and 39 were Stylonurina (see the data matrix; the data matrix is also available in the online MorphoBank database [O’Leary and Kaufman 2012] under the project code p2384, accessible at http://morphobank.org/permalink/?P2384, and through DRYAD at http://dx.doi.org/10.5061/dryad.24767). The 105 eurypterids not included in the phylogenetic analysis comprise taxa that are extremely poorly known and fragmentary, commonly consisting only of isolated carapaces or appendages. The majority of eurypterid higher taxa were represented in the analysis, with a few exceptions: Doropterus, which is known only from an isolated telson and may not represent a eurypterid (Tetlie 2007a; Lamsdell and Selden 2013); Eocarcinosoma, known from a single small carapace (Caster and Kjellesvig-Waering 1964); Marsupipiterus, also known from a single telson but which has been considered to either show similarities to basal Eurypterina (Braddy et al. 1995; Lamsdell 2011) or be a phyllocarid crustacean (Rolfe 1963); Necrogaummarus, previously considered to be a crustacean or myriapod but since shown to be a pterygotid eurypterid likely synonymous to either Pterygotus or Erettopterus (Selden 1986); and Unionopterus, which is some form of adelophthalmoid, although the original description is somewhat lacking and the holotype and only known specimen is currently lost (Tetlie and Van Roy 2006).
Individual species were coded to remove the potential for the inclusion of paraphyletic or polyphyletic groups. The application of paraphyletic or polyphyletic groups in biodiversity studies can have drastic effects on our understanding of mass extinction events. Erroneous patterns of pseudoextinction can be retrieved when monophyly is not recognized, whereas incorrect assumptions of monophyly can dampen the observed loss of biodiversity when studies are conducted at higher taxonomic levels (Patterson and Smith 1989). The data matrix was subjected to cladistic analysis using TNT (Tree analysis using New Technology: Goloboff et al. 2008), employing random addition sequences followed by tree bisection–reconnection (TBR) branch swapping with 100,000 repetitions with all characters unordered and of equal weight (for a recent review of the issues surrounding differential character weighting see Congreve and Lamsdell 2016). The strict consensus tree was constructed (Figs. 1 and S1) and used as the framework for all subsequent diversity and disparity analyses.

**DIVERSITY**
The phylogenetic tree was plotted against geological time, based on an existing eurypterid database (Lamsdell and Braddy 2010), updated for recent discoveries. Taxon occurrence was extended according to implied ghost ranges (Norell 1992) and parsimony reconstructed ancestors included as inferred from the tree topology (Brusatte et al. 2011), thereby correcting for potential stratigraphic biases (see discussion of phylogenetic diversity estimates [PDEs] next). Where reconstructed character states conflicted between different optimization models (accelerated or decelerated transformation), they were considered equivocal and coded as unknown. These inferred ancestors were assigned ranges based on the basic time-scaling methods of Norell (1992) and Smith (1994). Although this method can result in zero-duration branch lengths (Bell and Lloyd 2015), this is a nonissue for the current study as analyses requiring some estimate of branch length, such as rates of morphological evolution, were not performed. Basic time-scaling methods are acceptable protocols for identifying times of divergence and assessing relative variation in diversity (Bapst 2014), in which they outperform minimum branch length scaling methods (e.g., Brusatte et al. 2008). This resulted in the earliest occurrences being capped at the Darriwilian, in accordance with the occurrence of the earliest known eurypterid fossils (Lamsdell et al. 2015). Although pre-Darriwilian eurypterids most likely exist, given that the earliest recorded eurypterids are not the most phylogenetically basal, this measure affects a most conserva-
tive estimate of eurypterid diversity throughout the Paleozoic. Two measures of diversity were calculated, a taxic diversity estimate (TDE; see Levinton 1988) derived from the raw species counts and a PDE (Norell 1993) incorporating ghost range extensions and inferred ancestors based on the phylogenetic tree, for stage-level time intervals (Table S1). A number of processes—such as uneven geographic sampling, unequal outcrop availability, and differential preservation potentials in different depositional environments (Newell 1959; Simpson 1960; Raup 1972)—are known to cause inherent bias in the fossil record, and so corrections must be applied to the raw diversity counts to attain a numerical estimate of the relative amount of diversity in each time interval (Alroy 2010a). When only bulk diversity counts are available, methods such as shareholder quorum sampling (SQS) can be used to dampen the impact of these biases (Alroy 2010b); however, when a phylogenetic framework is available, this can be used to generate a PDE to provide an estimate of minimum relative diversity (Norell 1992, 1993; Smith 1994). Although a fundamentally different approach than SQS, PDEs also reduce the impact of sampling biases and have been shown to outperform TDEs using both simulated (Lane et al. 2005) and actual (Butler et al. 2009) datasets. PDEs are known to exhibit some artifacts due to the Signor–Lipps effect, whereby the terminal portion of a taxon’s range is truncated through increasing rarity of individuals decreasing the probability of sampling (Signor and Lipps 1982; Wagner 2000), thereby dampening the impact of abrupt mass extinction events (Lane et al. 2005). However, SQS is equally impacted by the Signor–Lipps effect (Alroy 2014) and alternative methods such as use of the three-tier (3T) log ratio (Alroy 2010c) are noisy when turnover rates are high (Alroy 2014), as would be expected with species-level data and stage-level time bins. Furthermore, it has been noted that the record of mass extinctions is biased toward widespread, dominant taxa, and that apparent extinctions in the fossil record often mark critical drops in abundance (Hull et al. 2015), a phenomenon that is also observable when looking at the record of modern threatened and endangered species (Plotnick et al. 2016). Failure to sample the termination of a lineage may then be widespread and therefore not impart as much of a directional bias into diversity studies as feared, while it can also be argued that terminally dwindling populations are already functionally extinct from an ecosystem perspective. It has also been shown that gradual extinctions can appear abrupt (Wang et al. 2012), sometimes mediated by clustering of last observed occurrences at sequence stratigraphic horizons mediated by sedimentary hiatuses and changes in water depth (Holland and Patzkowsky 2015). The Signor–Lipps effect, therefore, appears to be less ubiquitous, and thus less directionally biasing, than initially conceived. Similar biases caused by Lazarus taxa—taxa that dwindle in abundance resulting in apparent extinction only to recover and reappear in the fossil record (Jablonski 1986)—that the 3T sampling method explicitly seeks to remove are also accounted for in PDEs as any gaps in the ranges of the operational taxonomic units (OTUs—in this case, species) are filled as part of ghost range extension.

Taxic diversity, including that derived through PDEs, can also be biased by the Lagerstätten effect, whereby sites of exceptional preservation inflate relative abundance through increased sampling (Butler et al. 2009). These artifacts are, in part, thought to be caused by taxonomic oversplitting of exceptionally preserved material; however, much work has been done in recent years to reduce this phenomenon among eurypterids (e.g., Bradly 2000; Tetlie 2006; Lamsdell 2011). A number of taxa that are likely synonymous but have yet to be officially merged—for example, Parahughmilleria hefteri and Parahughmilleria major (Stormer 1973), considered to be ontogenetic variants of one another; Moselopterus ancylotelson and Moselopterus elongatus (Stormer 1974), M. elongatus probably representing an undeformed specimen of M. ancylotelson; and Eurypterus remipes and Eurypterus lacustris, considered chronospecies by Tetlie et al. (2007) and therefore synonyms according to any species definition criterion—were treated as synonyms for all analyses. Taxonomic oversplitting is therefore not likely to bias the current study, and so the impact of the Lagerstätten effect will be dampened. Comparing the eurypterid diversity estimates with a recent compendium of fossil Lagerstätten (Retallack 2011) shows that, aside from large peaks in the Cambrian (a period from which no eurypterids are known), overall Lagerstätten frequency has remained relatively constant throughout the Phanerozoic. When narrowing the comparison to only localities preserving cuticle and nonbiomineralized soft tissue, the post-Cambrian Paleozoic record is again shown to be relatively constant, with the exception of notable peaks in the late Silurian and late Devonian. The Devonian peak does not manifest in the eurypterid diversity estimates; however, the late Silurian peak is apparent in the TDE, but not the PDE as it is subsumed into a generally high Silurian diversity (Fig. 2). Diversity estimates of late Silurian eurypterids may therefore be magnified in the TDE; however, the lack of a corresponding peak in diversity during the late Devonian uptick in Lagerstätten occurrence suggests the decrease in eurypterid abundance during the Devonian is real. It is worth noting also that eurypterids are explicitly used as indicators of Lagerstätten-type preservation (Retallack 2011) and so the late Silurian peak in exceptional preservation occurrence is largely driven by an increase in the number of eurypterid localities; as such, it is impossible to separate whether this increase is due to an increase in suitable preservational environments or simply because eurypterids are more abundant at this time. Furthermore, eurypterid occurrences are not limited solely to sites of exceptional preservation, and eurypterids are frequently preserved throughout their geologic history in a variety of lithologies, including sandstones, where cuticle is ephemeral or entirely
absent (e.g., Cope 1886; Kjellesvig-Waering 1950, 1961; Leutze 1961; Lamsdell et al. 2009; Lamsdell et al. 2013a,b; Budil et al. 2014). As such, the increase in eurypterid abundance during the Silurian is considered to be a real signal, although the magnitude of this increase may be somewhat exaggerated. To test this, average ghost lineage duration (Cavin and Forey 2007) was calculated for each time bin to explore which peaks in diversity result in a reduction in average ghost range duration (representing a genuine biological radiation) and which result in no reduction in average ghost range duration and therefore are a result of the Lagerstätten effect. Cavin and Forey, however, assumed a “budding” mode of speciation, arbitrarily assigning the range of the inferred ancestral taxon to the basal-most member of a clade—something that may prove impossible if the clade does not have a pectinate topology. To circumvent these issues, we simply accounted for the inferred ancestral taxon as a separate ghost range at the earliest occurrence of the clade itself.

A final concern, that PDEs assume a “true” cladogram for the study group exists (Lane et al. 2005), is refutable based on the fact that no cladogram is ever considered to represent the “true” tree of organism relationships but is instead a testable hypothesis based on the available data and can be further refined in the future. In this manner, it is no different than studies that use SQS and 3T metrics, as these are also hypotheses about taxon distributions given the available data. If multiple hypotheses (cladograms) existed for the relationships of the study group then the analyses could be performed on all the available cladograms and compared, or cladograms could be evaluated for inclusion prior to the analysis. In this case, however, only a single dataset exits, although analysis of this dataset does result in multiple cladograms. As noted above, these cladograms are summarized through strict consensus and this subsequent tree was used as the framework for this study so as not to bias the results toward any of the equally supported most probable trees. Because conflicts are summarized through polytomies, this results in the loss of an inferred ancestor and so represents the most conservative reconstruction of unsampled diversity and disparity. Both TDE and PDE measures show congruent patterns of diversity shifts across the Paleozoic (Fig. 2), and so the PDE was used for the diversity/disparity comparisons. Using the PDE, estimates of diversity are directly comparable to the disparity data, which also incorporates ghost range extensions and inferred ancestral taxa. The results and interpretations are not altered by using the TDE, nor when excluding inferred ancestral taxa from the disparity analyses.

**RATES OF SPECIATION AND EXTINCTION**

Rates of speciation and extinction were calculated for each clade at the stage level from the PDE range data. Numerous speciation and extinction metrics have been proposed for rate calculation of modern and fossil taxa, most commonly based on the deterministic exponential model of taxon growth (Sanderson and Donoghue 1996; Nee 2004, 2006). Rates calculated from paleontological
data are commonly derived from datasets without any phylogenetic structure that generate per-capita rates normalized for standing diversity (Alroy 2014). The most widespread of these methods for rate calculation is the generalized boundary-crosser (BC) method (Foote 2000a,b) that incorporates counts of taxa either crossing both the bottom and top boundaries of an interval or crossing one boundary but not the other. However, as with raw counts of taxic diversity, the BC method can be influenced by the occurrence of Lazarus taxa and other sampling biases, and alternative methods such as the 3T (Alroy 2010c) and gap filler (GF; Alroy 2014) equations have been proposed to ameliorate these effects. Despite this, simulations have shown that much of the bias in BC analyses stems from the use of higher taxa, with taxonomic practices of overlapping of long-ranging, rare genera biasing the results (Alroy 2014)—an issue largely circumvented by the use of species-level data. The issue of Lazarus taxa is again removed through the utilization of phylogenetically corrected ranges. Deterministic rates have also been used in paleobiological analyses where phylogenetic frameworks are available (Hulbert 1993; Rode and Lieberman 2005; Abe and Lieberman 2009; Stigall 2010). These rates are derived from the standard equation for exponential growth as first used by Maurer (1989)—and therefore are not modifications to Foote’s (2000a, b) BC equation as claimed by Rode and Lieberman (2005) and Abe and Lieberman (2009). These methods use the phylogenetic tree to correct for sampling biases and remove the effect of Lazarus taxa (Smith 1994), and are more suitable for species-level data, which are more likely to result in short-ranged taxa with a greater preponderance of singletons.

Comparison of calculated rates allows a broader basis for interpreting the generality of observed biodiversity patterns (Stigall 2010). Per-capita rates for speciation ($\hat{p}$) and extinction ($\hat{q}$) were calculated for each time bin following the BC method of Foote (2000a,b): 

$$\hat{p} = -\ln \left( \frac{N_{bt}}{N_t} \right) / \Delta t,$$

$$\hat{q} = -\ln \left( \frac{N_{bt}}{N_{bt}} \right) / \Delta t,$$

where $N_{bt}$ is the number of taxa that cross both the lower and upper boundary of the time bin, $N_t$ is the number of taxa that cross the lower time bin boundary from the previous time bin, $N_{bt}$ is the number of taxa that cross the upper time bin boundary and persist into the subsequent time bin, and $\Delta t$ is the length of duration (in million years) of the time bin. Deterministic rates of speciation rate ($S$) and extinction rate ($E$) were calculated using the equations of Maurer (1989) as set out by Hulbert (1993):

$$S = \frac{\ln (N_0 + o_0) - \ln N_0}{\Delta t},$$

$$E = \frac{\ln (N_0 + o_0) - \ln N_t}{\Delta t},$$

where $N_0$ is the number of taxa that cross the lower time bin boundary from the previous time bin, $N_t$ is the number of taxa that cross the upper time bin boundary and persist into the subsequent time bin, $o_0$ is the number of taxa that originate within the time bin, and $\Delta t$ is again the length of duration (in million years) of the time bin. Therefore, $N_0$ is equivalent to $N_b$, $N_t$ is equivalent to $N_e$, and $N_0$ combined with $o_0$ is equivalent to Foote’s (2000a,b) $N_{tot}$, representing the total diversity of a time bin. Both methods suffer from edge effects caused by the sudden origination or extinction of a large amount of taxa in the first and last time bins, respectively; while an edge effect is unlikely in the last time bin (the Wuchiapingian) as eurypterid diversity has already drastically winnedow down to only a handful of species, an edge effect is likely in the first time bin (the Darriwilian), and this should be considered when reviewing the rates retrieved for the Darriwilian.

**DISPARITY**

Analyses of disparity (morphological diversity) have become increasingly common in the paleobiological literature over the last couple of decades, in particular those utilizing discrete character matrices (Foote 1994, 1999; Wills et al. 1994; Wills 1998; Lofgren et al. 2003; Wesley-Hunt 2005; Brusatte et al. 2008; Young et al. 2010; Thorne et al. 2011; Bapst et al. 2012; Hughes et al. 2013; Ruta et al. 2013; Hopkins and Smith 2015; Lamsdell 2016; Oyston et al. 2016). These studies have employed a number of different distance metrics to generate Euclidean pairwise distances between all taxa (see Lloyd 2016); however, the two most commonly used are generalized Euclidean distance (GED: Wills et al. 1994) and Gower’s coefficient (GC: Gower 1971). The two metrics differ primarily in their handling of missing data, with GED inserting a weighted mean fractional univariate distance based on those distances that are calculable, while GC simply rescales calculable distances based on the amount of information available. Data are handled during this process in a phenetic manner, with no assumptions of hierarchy or homoplasy, and as such it is not circular to compare results from disparity analyses to the phylogenetic hierarchy retrieved from the same character matrix. Analyses of disparity formulated from discrete character matrices have been shown to converge on a common signal with other forms of morphometric data (Foth et al. 2012; Hetherington et al. 2015) and so results are comparable with those attained from, for example, landmark data.

Eurypterid disparity measures were calculated from the cladistic character data through GED. Along with the observed species, reconstructed ancestors (Brusatte et al. 2011) were included as necessitated by tree topology. Ancestors were coded into the phylogenetic matrix with their characters reconstructed using parsimony-based ancestral state reconstruction in Mesquite (Maddison and Maddison 2015) and outgroup non-eurypterid taxa
(Weinbergina and the chasmataspidids Chasmataspis, Loganamaraspis, Octoberaspis, and Diploaspis) were removed from the matrix prior to analysis. Autapomorphies (character states unique to single taxa) were included as, unlike in parsimony analyses where they contribute no information to the inference of relationships, unique character states can be important for determining the relative distance between taxa in morphospace. However, with the inclusion of inferred ancestors and the fact that autapomorphies were frequently coded as a single state within a multistate character, the only continuous characters in the analysis were those rendered invariable after the removal of outgroup taxa, and so these characters were also excluded from the analysis.

In the resulting matrix, 41% of the character codings were missing or inapplicable, although these were not evenly distributed throughout the matrix, resulting in some taxa having significantly fewer characters coded than others. This resulted in GED being used over GC, as GC will result in incalculable distances if there are few comparable characters for a specific distance, whereas GED ensures that all taxa are included in the analysis as long as there is at least one coded character in common between two taxa. Pairwise (taxon to taxon) distances were calculated with negative eigenvalue results for some axes removed through the addition of a constant, calculated computationally following Cailliez (1983), and subjected to ordination through principal coordinates analysis (PCO) using scripts in R (Hughes et al. 2013). One characteristic of analyses of discrete character data is that the variance tends to get spread relatively evenly over each of the axes (Lloyd 2016); as such, it can be more informative to look for discontinuities among the distribution of encompassed variance per axis. Three discontinuities are evident on the distribution of variance generated from the eurypterid data, one after the second axis, another after the fourth, and the last after the tenth (Fig. 3). However, these axes only represent 2.7, 4.4, and 8.3% of the total variance, respectively. As such, phylomorphospace plots were produced using the first two axes in Scatterplot3d (Liggs and Maechler 2003), as these two axes are representative of the overall changes in spatial distribution that the morphospace exhibits in each time bin, however, all statistical tests were performed using data from all axes.

Disparity was quantified through the sum of range (SOR), which describes the total range of morphospace occupation represented by the volume of a hypercube that encompasses all of the morphotypes (Foote 1991) thereby showing the size of dissimilarity within a morphospace, and the sum of variance (SOV), which is the sum of variances along the morphospace axes (Van Valen 1974) and acts as a measure of the distances between the points (taxa) in morphospace. SOR and SOV were calculated for each time bin using all axes and presented as the mean value of 1000 bootstrap replicates; SOR and SOV values were considered distinct from one another if there was no overlap of their standard error (SE). Changes in morphospace occupation between time bins were quantified using the percentage difference in SOR, SOV, and position of the centroid (POC)—the median value along each morphospace axis (Dommergues et al. 1996)—resulting in percentage difference in SOR ($P_{\text{SOR}}$), percentage difference in SOV ($P_{\text{SOV}}$), and percentage difference in POC ($P_{\text{DPOC}}$) values for each time bin pair. These statistics can be used to quantify whether changes in morphospace are random (nonselective) or marginal or lateral (selective; Korn et al. 2013; see Table S2). As originations were being examined as well as extinctions, the presentation of the resulting figures was modified from that in Korn et al., with 100% (no change) taken as a baseline and variations taken as a positive or negative percentage from that (therefore a value of 116% in Korn et al. would be $+16\%$ herein, while a value of 78% would be $-22\%$), thereby permitting easier comparisons between extinctions and radiations. Although values for $P_{\text{SOR}}$ and $P_{\text{SOV}}$ are presented as either positive (representing an increase in SOR or SOV) or negative (representing a decrease), $P_{\text{DPOC}}$ simply indicates any deviation across the sum of all axes, and therefore is directionless, and as such $P_{\text{DPOC}}$ is shown as a positive value.

$P_{\text{SOR}}, P_{\text{SOV}},$ and $P_{\text{DPOC}}$ were calculated across the Katian/Hirnantian and Hirnantian/Llandovery stage boundaries, for the Kasimovian/Gzhelian stage boundary, and for each stage boundary from the Lochkovian to the Tournaisian, encompassing both the onset and breakdown of late Ordovician glaciation (Sheehan 2001; Congreve 2013), the Kasimovian rainforest collapse (Sahney et al. 2010a), and the entire duration of the Devonian. Changes in morphospace occupation caused through extinction were calculated by comparing the occupied morphospace of a time bin with that of only surviving taxa in the succeeding time bin, whereas morphospace changes through origination were calculated by comparing the occupied morphospace of a time bin with that of taxa originating in the succeeding time bin. In this manner, we can quantify whether both origination and extinction
are random or directional in regard to morphospace occupation, and whether any changes in morphospace occupation are driven by extinction, origination, or both. Multivariate statistical tests (PERMANOVA - permutational multivariate analysis of variance using the Euclidean distance measure) were performed to ascertain the statistical significance of overlap and separation of morphospace across all axes between time bins at the epoch and period level. Significance was estimated by permutation across groups with 10,000 replicates. Statistical analysis was performed both with and without Bonferroni correction. Bonferroni correction provides little statistical power and is probably too conservative (Sokal and Rohlf 1995; Moran 2003; Nakagawa 2004; Garamszegi 2006); instead, effect size is reported as a more accurate measure of robustness (Cohen 1988). Pairwise comparisons with Bonferroni correction generally exhibit less significance than those without, but still display the same general trend (Tables S3–S6).

Results and Discussion
DIVERSITY, DISPARITY, AND SHIFTS IN MORPHOSPACE OCCUPATION THROUGH TIME

Diversity and disparity are decoupled throughout the evolutionary history of Eurypterida, with disparity remaining relatively constant (Figs. 4 and S2) irrespective of increases and decreases in diversity (Fig. 2). Similar patterns have been shown in graptolites (Bapst et al. 2012) and vertebrates (Ruta et al. 2013), thus emphasizing the importance of considering both diversity and disparity in studies of clade history. Eurypterids exhibited low levels of diversity in the Ordovician (485.4–443.8 mya) before undergoing a massive increase, reaching their peak diversity in the Silurian (443.8–419.2 mya), with decreases in average ghost range length during the Upper Ordovician and early Silurian showing that these radiations are at least in part real. Eurypterids do not appear to have been overtly impacted by the end-Ordovician mass extinction although rates of extinction underwent an increase in the Hirnantian (445.2–443.8 mya) that may correlate to the rapid onset and subsequent breakdown of extreme glaciation (Sheehan 2001; Congreve 2013). The clade exhibited a marked drop of diversity into the Devonian, driven in part by peaks of extinction in the Pridoli (423–419.2 mya) and Pragian (410.8–407.6 mya), with a subsequent steady decrease from the Emsian (407.6–393.3 mya) through to the Frasnian (382.7–372.2 mya). By the onset of the Carboniferous, only the Adelopthalmidae of the Eurypterina and the Hibbertopteridae and Mycteroptidae of the Stylonurina persist, with previously dominant clades having declined to extinction by the Middle Devonian (Carcinosomatidae are last known from the Lochkovian, Stylonuridae from the Pragian, Rhenopteridae and Eurypteridae from the Emsian, and the Pterygotidae from the Eifelian). The SOV also dropped during the Middle Devonian (393.3–382.7 mya), although no corresponding drop is seen in the SOR; diversity and SOV subsequently remained at relatively low levels for the remainder of the Paleozoic before dwindling further during the Permian (298.9–252.2 mya), until the eventual extinction of the group in the Wuchiapingian (259.8–254.2 mya). Such a pattern is consistent with that of survival without recovery (Jablonski 2001), making post-Devonian Eurypterida a “dead clade walking” (Jablonski 2002), albeit one that dwindles to a reduced ecological niche rather than suffering imminent extinction.
Morphospace occupation is not significantly distinct from that in the Permian. Morphospace occupation is seen to shift during the Devonian, with the position of species in morphospace shifting to one of the lateral extremes during the Middle Devonian from a more central position during the Ordovician–Silurian. This correlates with the decrease in morphospace mutability, with the morphospace occupation of each epoch from the Middle Devonian onwards being statistically indistinguishable from one another (Tables S7 and S9). This shift is not reflected in the SOR (Figs. 4 and S2) as the overall range of morphospace occupied remains relatively unchanged, with two groups located in disparate regions of morphospace (Fig. 5); however, the amount of variance has decreased as the average distance between species decreases, with the two groups being tightly clustered; this is reflected in the decrease in SOV, which occurs during the Middle Devonian (Figs. 4 and S2).

Quantifying the change in morphospace occupation between stages due to extinction (Korn et al. 2013), however, reveals that extinction was random throughout the Devonian (Tables 2 and S10), with \( P_{SOR} \) remaining well below the >10 values that would indicate a nonrandom extinction, while \( P_{SOV} \) remains below the >13 value threshold and \( P_{DPOC} \) is fixed at around 1–2, well within random levels. Quantification of extinction during the twin pulses of the end-Ordovician mass extinction (Sheehan 2001; Congreve 2013), the Kellwasser and Hangenberg events (House 1985, 2002), and the Kasimovian rainforest collapse (Sahney et al. 2010a) reveals that random reduction of morphospace occupation is a consistent response among eurypterids during extinction events. The subsequent lateral shifts in morphospace occupation during the Devonian are therefore due to differential patterns of origination rather than extinction; origination became increasingly skewed throughout the Early Devonian, with \( P_{SOV} \) and \( P_{DPOC} \) values dropping at each subsequent stage boundary (Table 2 and S10). Negative values for \( P_{SOV} \) during origination indicate that the originating species are more clustered in morphospace than the species in the previous time bin, indicating that only a subset of morphospace was being repopulated by new species after the random extinction events. As the threshold values required for a meaningful shift in morphospace drop as the magnitude of change increases (Korn et al. 2013), it is significant that the \( P_{SOR} \), \( P_{SOV} \), and \( P_{DPOC} \) values for origination during the Devonian stages are consistently higher than those for extinction despite the magnitude of change due to origination being less than half of that due to extinction. Therefore, selective patterns of speciation, rather than extinction, were responsible for changes in morphospace occupation across the Devonian.

**SHIFT IN EVOLUTIONARY REGIME AND THE DEVONIAN BIOTIC CRISIS**

Eurypterids underwent a major shift in evolutionary regime during the Devonian, corresponding to the migration and subsequent...
reduction of morphospace occupation in Eurypterina and Stylonurina. During the Devonian, Stylonurina and Eurypterina independently underwent an ecological transition from consisting of predominantly marine taxa to comprising purely freshwater species (O’Connell 1916; Lamsdell and Braddy 2010), while retaining their cosmopolitan distribution. This resulted in a decrease in diversity and an apparent stagnation in their evolution. It is possible that, due to their widespread distribution but limited founding pool of closely related species, each clade suffered a reduced morphological variability akin to the loss of species genetic variability due to small founder populations (Mayr 1963). Rapid recovery has been shown to somewhat mitigate such a loss in variability at the level of both populations (Nei et al. 1975) and clades, whereby the rate at which a group invades available ecological space after mass extinctions can greatly influence its chances of long-term recovery (Sahney et al. 2010b). Much of the late Paleozoic, however, is characterized by sluggish macroevolution, in which origination rates are depressed following the formation of glaciers in the Mississippian, resulting in species having broad ecological niches, widespread geographic distributions, and large and relatively stable populations (Stanley and Powell 2003; Powell 2005). Although some lineages underwent rapid recovery in the early Carboniferous (Smithson et al. 2016), prior to the onset of the Mississippian glaciation, eurypterids showed only a muted recovery. This may have been, at least in part, due to the fact that while post-Devonian eurypterids comprised two disparate clades (having diverged over 110 million years prior), each with distinct feeding ecologies, both were limited to freshwater environments (Lamsdell and Braddy 2010). Interestingly, xiphosurids are also shown to invade freshwater environments at this time (Lamsdell 2016), while the majority of the earliest Carboniferous vertebrate faunas are also known from nonmarine localities (Sallan and Coates 2010; Sallan and Galimberti 2015), and nonmarine environments may have acted as refugia for lineages impacted by the biotic crisis that has been suggested to be primarily limited to the pelagic realm (Racki 1999). For eurypterids, the

### Table 1. PERMANOVA test results of Eurypterida (F_{5,321} = 1.298, \eta^2 = 0.016, P = 0.0001) from 100,000 permutations for statistical differences between taxa for each of the six period-level time bins based on PCO analyses.

<table>
<thead>
<tr>
<th>Stage boundary</th>
<th>Ordovician</th>
<th>Silurian</th>
<th>Devonian</th>
<th>Carboniferous</th>
<th>Permian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordovician</td>
<td>—</td>
<td>1.1810</td>
<td>1.4900</td>
<td>1.7860</td>
<td>1.3390</td>
</tr>
<tr>
<td>Silurian</td>
<td>0.0009</td>
<td>—</td>
<td>1.2200</td>
<td>1.6610</td>
<td>1.2100</td>
</tr>
<tr>
<td>Devonian</td>
<td>0.0001</td>
<td>0.0004</td>
<td>—</td>
<td>0.9995</td>
<td>0.9998</td>
</tr>
<tr>
<td>Carboniferous</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.4798</td>
<td>—</td>
<td>0.7111</td>
</tr>
<tr>
<td>Permian</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.4244</td>
<td>0.9987</td>
<td>—</td>
</tr>
</tbody>
</table>

Nonsignificant P values (indicating no significant difference between time bins) shown in bold. Values in regular font are for the P values, those in italics are the raw F values.

### Table 2. Quantification of changes in morphospace across stage boundaries (Korn et al. 2013). Entries in bold represent stage boundaries (denoted by their first three letters) associated with mass extinctions.

<table>
<thead>
<tr>
<th>Stage boundary</th>
<th>Percentage of decrease</th>
<th>Extinction</th>
<th>Origination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kat/Hir</td>
<td>+0.2</td>
<td>48</td>
<td>16</td>
</tr>
<tr>
<td>Hir/Lla</td>
<td>-0.2</td>
<td>24</td>
<td>618</td>
</tr>
<tr>
<td>Loc/Pra</td>
<td>-0.5</td>
<td>41</td>
<td>31</td>
</tr>
<tr>
<td>Pra/Ems</td>
<td>-0.4</td>
<td>41</td>
<td>21</td>
</tr>
<tr>
<td>Ems/Eif</td>
<td>+0.9</td>
<td>70</td>
<td>22</td>
</tr>
<tr>
<td>Eif/Giv</td>
<td>+1.4</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>Giv/Fra</td>
<td>+0.2</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Fra/Fam</td>
<td>+0.7</td>
<td>14</td>
<td>200</td>
</tr>
<tr>
<td>Fam/Tou</td>
<td>+3.7</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>Kas/Gzh</td>
<td>+3.9</td>
<td>58</td>
<td>17</td>
</tr>
</tbody>
</table>

Ordovician: Kat = Katian, Hir = Hirnantian; Silurian: Lla = Llandovery; Devonian: Loc = Lochkovian, Pra = Praian, Ems = Emsian, Eif = Eifelian, Giv = Givetian, Fra = Frasnian, Fam = Famennian; Carboniferous: Tou = Tournasian, Kas = Kasimovian, Gzh = Gzhelian.

Values in regular font indicate changes due to extinction, whereas values in italics show changes due to origination. \( P_{SOV} \) quantifies the percentage difference in SOV between the first and second stage, \( P_{SOR} \) is the percentage difference in SOR, and \( P_{DPOC} \) is the percentage of change in the position of the centroid. See Table S11 for the raw values.
The drop in diversity during the late Devonian is one of the five largest declines in the Phanerozoic (Raup and Sepkoski 1982; Sepkoski 1986); however, extinction rate actually decreased throughout the late Devonian (Orzechowski et al. 2015), and both speciation and extinction rates decreased throughout the Devonian for Eurypterida (Fig. 6). Of the major late Devonian posited extinction events (House 1985, 2002), only the Hangenberg Event appears to be a real, abiotically driven mass extinction (Friedman and Sallan 2012), while the Kellwasser Event has been suggested to be a stratigraphic artifact of last occurrences caused by changes of water depth (Holland and Patzkowsky 2015), which characterize these horizons (McGhee 1996; Hallam and Wignall 1997). Instead, the long-term diversity decline appears to be due to decreased origination rather than elevated extinction (Bambach et al. 2004). It has been suggested that the drop in speciation rates was due to a decrease in endemism (May 1995; Stigall 2012), with a shutdown in vicariant speciation attributable to widespread interbasinal species invasions mediated by a global increase in sea level (Rode and Lieberman 2004; Abe and Lieberman 2009; Stigall 2010). Eurypterids show a high degree of endemism (Tetlie 2007b), possibly tied to their need to return to sheltered regions within very specific environmental tolerances to molt and possibly mate (Braddy 2001; Vrazo and Braddy 2011); similar population segregation is seen in modern horseshoe crabs, which show a strong correlation between genetic and geographic distance (King et al. 2005). Eurypterids show a marked drop in speciation from the Emsian, undergoing a recovery in origination rates during the Famennian. Other invertebrate groups in which a drop in speciation rates during the Devonian has been recorded—trilobites (Abe and Lieberman 2009), phyllocarids (Rode and Lieberman 2005), and brachiopods and bivalves (Stigall 2010)—are also highly endemic prior to the onset of the biotic crisis. However, the timing for the shutdown in speciation is different for each of these groups;
speciation stops for some trilobite groups from the Emsian onward (Abe and Lieberman 2009) and phyllocarids exhibit low levels of speciation from the Eifelian with a slight recovery in the Frasnian (Rode and Lieberman 2005), whereas the speciation rates of the studied bivalve and brachiopod clades only tail off during the Frasnian and Famennian (Stigall 2010). The differential onset of depressed rates of speciation is likely due to variations in ecology resulting in differences in population structure (Fordham et al. 2013) and ability to track environmental changes (Saupe et al. 2014) that, when combined with regional geographic differences, mean different groups have different ecological thresholds beyond which speciation shuts down.

The origins of the Devonian biotic crisis may, therefore, be traced to the beginning of the period. Sea-level varied consistently throughout the Silurian and began to stabilize during the onset of the Devonian, remaining generally high throughout the period (Haq and Schutter 2008). The Silurian is characterized by multiple pulses of ocean anoxia and heavy metal excursions (Vandenbrouke et al. 2015), resulting in an unstable environment with a high degree of biotic turnover. This boom and bust macroevolution can maintain high standing diversity while the system is maintained; however, a reduction in rates can impact groups differently depending on whether they were in a boom or bust phase. Eurypterids appear to have ended the Silurian on a bust, experiencing marked extinction during the Pridoli followed by another in the Pragian (Fig. 6). Surviving clades were subsequently winnowed by background extinction throughout the remainder of the Devonian, with only a handful maintaining the rates of speciation required to persist, resulting in the onset of the shift in occupied morphospace as of the Middle Devonian. A similar shift in morphospace over the same timeframe is seen in crinoids (Foote 1994), which also begin their recovery during the Famennian (Sallan et al. 2011).

The Devonian biotic crisis had a marked impact on multiple taxonomic groups, interrupting or altering evolutionary trajectories and resulting in noticeable shifts in evolutionary regime. Vertebrates, for example, experienced body-size reduction in response to the Hangenberg Event (Sallan and Galimberti 2015). Post-Devonian eurypterids had a more global distribution than their Silurian ancestors (Tetlie 2007b), and trilobites exhibit a similar biogeographic pattern (Fortey and Owens 1997). The trends observed among the studied arthropod groups are mirrored in other metazoan clades (McGhee 1996); the biotic response to the Devonian biodiversity crisis appears to be characterized by global homogenization, with many losers and few winners. If the primary causal factors for this trend are a breakdown in endemism mediated by a rise in sea level, this has implications for marine conservation efforts given projected sea-level rise over the next 100 years (Meehl et al. 2005; Rahmstorf 2007; Vermeer and Rahmstorf 2009).

ACKNOWLEDGMENTS
We thank M. Wills (University of Bath), M. Benton (University of Bristol), and S. Brusatte (University of Edinburgh) for discussion on disparity analysis, M. Hopkins (American Museum of Natural History) and V. McCoy (University of Leicester) for advice with R, and C. Congreve (Pennsylvania State University) and A. Falk (Centre College) for discussion. M. Webster and M. Friedman provided useful editorial comments, and M. Hopkins, M. Wills, and three anonymous referees are thanked for providing helpful comments during the review process. J. Lamsdell was supported by funding from a Lerner-Gray Postdoctoral Fellowship from the American Museum of Natural History to research macroevolutionary and macroecological trends among Paleozoic arthropods.

DATA ARCHIVING
The doi for the study data is 10.5061/dryad.24767 on DRYAD and 10.7934/P2384 on Morphobank.

LITERATURE CITED


Supporting Information
Additional Supporting Information may be found in the online version of this article at the publisher’s website:

Table S1. Stage-level counts for species diversity, showing both taxic (observed) and phylogenetic (modeled) diversity.

Table S2. Summary of changes in the sum of range (SOR), sum of variance (SOV), and position of the centroid (DPOC) measured as percentages for random, lateral, and marginal variations in morphospace under 50% (regular font), 75% (italicized font), and 87.5% (bold font) change as reported from the simulations of Korn et al. (2013).

Table S3. PERMANOVA test results of Eurypterida ($F(5,321) = 1.298, \eta^2 = 0.016, P = 0.0001$) from 100,000 permutations with Bonferroni correction for statistical differences between taxa for each of the six period-level time bins based on PCO analyses.

Table S4. PERMANOVA test results of Eurypterida ($F(14,494) = 1.092, \eta^2 = 0.028, P = 0.0003$) from 100,000 permutations with Bonferroni correction for statistical differences between taxa for each of the 14 epoch-level time bins based on PCO analyses.

Table S5. PERMANOVA test results of Eurypterida ($F(5,159) = 1.210, \eta^2 = 0.028, P = 0.0027$) from 100,000 permutations with Bonferroni correction for statistical differences between taxa for each of the six period-level time bins based on PCO analyses, excluding inferred ancestral taxa.

Table S6. PERMANOVA test results of Eurypterida ($F(14,257) = 0.976, \eta^2 = 0.042, P = 0.7527$) from 100,000 permutations with Bonferroni correction for statistical differences between taxa for each of the 14 epoch-level time bins based on PCO analyses, excluding inferred ancestors.

Table S7. PERMANOVA test results of Eurypterida ($F(14,494) = 1.092, \eta^2 = 0.028, P = 0.0003$) from 100,000 permutations for statistical differences between taxa for each of the 14 epoch-level time bins based on PCO analyses.

Table S8. PERMANOVA test results of Eurypterida ($F(5,159) = 1.210, \eta^2 = 0.028, P = 0.0027$) from 100,000 permutations for statistical differences between taxa for each of the six period-level time bins based on PCO analyses, excluding inferred ancestral taxa.

Table S9. PERMANOVA test results of Eurypterida ($F(14,257) = 0.976, \eta^2 = 0.042, P = 0.7527$) from 100,000 permutations for statistical differences between taxa for each of the 14 Epoch-level time bins based on PCO analyses, excluding inferred ancestors.

Table S10. Quantification of changes in morphospace excluding inferred ancestors across stage boundaries (Korn et al. 2013).

Table S11. Quantification of changes in morphospace across stage boundaries (Korn et al. 2013), raw values.

Figure S1. Phylogenetic tree of the Eurypterida.

Figure S2. Disparity metrics for Eurypterida from the Euclidean distance matrix derived from the character matrix, excluding inferred ancestors.