PALEOBIOLOGY | Paleontological SOCIETY
A PUBLICATION OF THE

Article

A biased fossil record can preserve reliable phylogenetic signal

C. Henrik Woolley* ⓘ *, Jeffrey R. Thompson, Yun-Hsin Wu, David J. Bottjer, and
Nathan D. Smith*

Abstract.—The fossil record is notoriously imperfect and biased in representation, hindering our ability to place fossil specimens into an evolutionary context. For groups with fossil records mostly consisting of disarticulated parts (e.g., vertebrates, echinoderms, plants), the limited morphological information preserved sparks concerns about whether fossils retain reliable evidence of phylogenetic relationships and lends uncertainty to analyses of diversification, paleobiogeography, and biostratigraphy in Earth's history. To address whether a fragmentary past can be trusted, we need to assess whether incompleteness affects the quality of phylogenetic information contained in fossil data. Herein, we characterize skeletal incompleteness bias in a large dataset (6585 specimens; 14,417 skeletal elements) of fossil squamates (lizards, snakes, amphisbaenians, and mosasaurs). We show that jaws + palatal bones, vertebrae, and ribs appear more frequently in the fossil record than other parts of the skeleton. This incomplete anatomical representation in the fossil record is biased against regions of the skeleton that contain the majority of morphological phylogenetic characters used to assess squamate evolutionary relationships. Despite this bias, parsimony- and model-based comparative analyses indicate that the most frequently occurring parts of the skeleton in the fossil record retain similar levels of phylogenetic signal as parts of the skeleton that are rarer. These results demonstrate that the biased squamate fossil record contains reliable phylogenetic information and support our ability to place incomplete fossils in the tree of life.

*C. Henrik Woolley and Yun-Hsin Wu. Dinosaur Institute, Natural History Museum of Los Angeles County,
Los Angeles, California 90007, U.S.A.; and Department of Earth Sciences, University of Southern California,
Los Angeles, California 90089, U.S.A. E-mail: chwoolle@usc.edu, yhwu@usc.edu*

*Jeffrey R. Thompson. Department of Earth Sciences, Natural History Museum, London SW7 5BD, U.K.; and UCL
Centre for Life's Origins and Evolution, University College London, London WC1E 6BT, U.K. E-mail: jeffery.
thompson@nhm.ac.uk*

*David J. Bottjer. Department of Earth Sciences, University of Southern California, Los Angeles, California 90089,
U.S.A. E-mail: dbottjer@usc.edu*

*Nathan D. Smith. Dinosaur Institute, Natural History Museum of Los Angeles County, Los Angeles, California
90007, U.S.A. E-mail: nsmith@nhm.org*

Introduction

The fossil record is a fundamental natural historical archive for understanding evolution and Earth's history. Without information from fossils, we would have considerably less knowledge concerning the extinction and emergence of lineages (Sun et al. 1998; Valentine et al. 1999; Field et al. 2020), the nature of biotic crises (Sun et al. 2012; Lyson et al. 2019; Petsios et al. 2019), and major phenotypic innovations in Earth's flora and fauna (Daeschler et al. 2006; Murdock and Donoghue 2011; Stein et al. 2012). But behind the illuminating biological information contained in the fossil record is the reality that fossil data are inherently incomplete (Darwin 1859; Foote and Sepkoski 1999; Smith 2001; Kidwell and Holland 2002), due to geological factors (Raup 1976; Smith and McGowan 2005), factors related to fossil preservation (e.g., taphonomic bias; Sansom et al. 2010), and asymmetrical research interest and sampling intensity among workers (e.g., sampling bias; Smith 1994, 2001). These biases have been characterized in the fossil record of a variety of organismal groups (Crane et al. 2004; Brocklehurst et al. 2012;

0094-8373/22

Dean et al. 2016; Driscoll et al. 2019), and a growing number of sampling proxies can be used to account for the multitude of factors that distort paleobiological data. As a result, our ability to explore major questions related to biodiversity in the fossil record is enhanced, but methodological problems remain (Dean et al. 2016; Sakamoto et al. 2017). Furthermore, the patterns of biases associated with a majority of fossil groups are uncharacterized, and the association between biased records and phylogenetic data content remains underexplored (Sansom et al. 2010, 2017; Sansom 2015; Brocklehurst and Benevento 2020).

Taphonomic and sampling biases still represent consistent barriers to reconstructing the phylogenetic relationships of fossil organisms (Patterson 1981; Sansom et al. 2010, 2017; Sansom 2015; Brocklehurst and Benevento 2020), which are essential to the evolutionary frameworks that synthetic studies of biodiversity (Raup 1979), biostratigraphy (Norell and Novacek 1992), and paleobiogeography (Benson et al. 2013) rely upon (Sakamoto et al. 2017). Incomplete fossils with missing morphological data have been shown to decrease the accuracy of inferred phylogenies (Guillerme and Cooper 2016; Vernygora et al. 2020), but this effect is lessened when fossil taxa are scored alongside extant taxa and more morphological characters are coded (Guillerme and Cooper 2016) or when the number of taxa included in the analysis is increased (Vernygora et al. 2020).

The most straightforward way to minimize the negative effects of missing fossil morphological data on inferring evolutionary relationships is by utilizing the most complete and best-preserved fossil specimens. Most animal and plant fossil records, however, are made up of fragmentary remains and disarticulated parts (Crane et al. 2004; Brocklehurst et al. 2012; Dean et al. 2016), and by focusing primarily on complete specimens for phylogenetic analyses, the majority of extinct biodiversity and available fossil data in natural history collections is often excluded. Consequently, important fossil taxa known from fragmentary material remain underutilized in phylogenetic analyses, and we are left with considerable uncertainty in the evolutionary relationships of most known fossil organisms, which obscures our understanding of major evolutionary patterns in Earth's history (Sansom et al. 2010; Dornburg et al. 2015; Sansom 2015).

Maximizing the vast amount of evolutionary information available in fossil collections necessitates a quantitative characterization of the biases present in this record. Furthermore, if we wish to accurately assess the evolutionary relationships of incompletely preserved fossil specimens, we need to know whether the phylogenetic data preserved in these incomplete samples are reliable and consistent. We compiled a large dataset from natural history collections (6585 specimens; 14,417 skeletal elements) to characterize skeletal representation bias in the extensive and diverse fossil record of squamates (lizards, snakes, amphisbaenians, and mosasaurs). We then applied parsimony- and model-based estimates of phylogenetic signal to test a major question fundamental to paleobiology: Does the observed fossil record contain reliable phylogenetic information?

Materials and Methods

Sampling the Fossil Squamate Skeleton in Natural History Collections

Despite their biases, natural history collections are the most data-rich record of the evolutionary history of squamates and are the basis for all morphological and most occurrence-based analyses of their fossil record. Although natural history collections have been used to assess the nature of bias and asymmetrical preservation in the squamate fossil record, these assessments are either limited to general observations (Nydam 2013; Rage 2013) or restricted taxonomically to mosasaurs (Driscoll et al. 2019). In this study, we quantify which regions of the squamate skeleton are more prevalent in natural history collections and, by extension, their observed fossil record.

To understand which regions of the fossil squamate skeleton appear more frequently in museum collections, we divided the skeleton into seven discrete regions: (1) jaws and palatal bones in the skull (all bones in the mandible, premaxilla, maxilla, vomer, palatine, pterygoid, and teeth); (2) posterior cranial bones (nasal, frontal, parietal, jugal, quadrate,

braincase, etc.); (3) axial elements (i.e., vertebrae and ribs); (4) pectoral girdle; (5) pelvic girdle; (6) appendicular elements; and (7) dermal elements (Fig. 1, Supplementary Fig. S1). Using these criteria, we sampled for the number of occurrences of fossil squamate skeletal elements belonging to each region. We surveyed 6585 fossil squamate specimens and tallied 14,417 occurrences of individual skeletal elements across three major fossil squamate body plans: giant, fully aquatic/marine squamates (mosasaurs); small to medium-sized, four-limbed, terrestrial lizards; and small to medium-sized, limb-reduced/limbless squamates (e.g., snakes, amphisbaenians). Our surveyed specimens range in age from the Late Jurassic to the late Pleistocene (Fig. 2A).

This study used in-person observations of fossil squamate collections and surveyed readily available digital collections databases. In-person sampling (institutions: American Museum of Natural History [AMNH], New York, N.Y., U.S.A.; Denver Museum of Nature & Science [DMNH], Denver, Colo., U.S.A.; and Yale Peabody Museum of Natural History [YPM], New Haven, Conn., U.S.A.) was carried out by identifying squamate fossils to the lowest taxonomic level and with as specific anatomical terminology as possible. Taphonomic bias and preservation mode played key roles in the precision of taxon identification and anatomical assignment. Incomplete specimens for which not enough diagnostic information was preserved were often restricted to identification at order/suborder/family taxonomic levels, using generalized anatomical terminology (e.g., "partial jawbone," "trunk vertebra," "metapodial"). More complete specimens could be identified at genus/species taxonomic levels, with specific anatomical terminology (e.g., "anterior right maxilla," "proximal left femur").

To broadly characterize fossil squamate morphological data available in museum collections globally, we sampled online electronic databases from six institutions on three continents (Institute for Vertebrate Paleontology and Paleoanthropology [IVPP], Beijing, China; Natural History Museum London [NHMUK], London, U.K.; Natural History Museum of Los Angeles County [LACM], Los Angeles,

Calif., U.S.A.; the Smithsonian National Museum of Natural History [USNM], Washington, D.C., U.S.A.; University of Florida Museum of Natural History [UFMNH], Gainesville, Fla., U.S.A.; and YPM). Our criteria for selecting these institutions was dependent on readily available digital vertebrate paleontology collections databases on the institutions' websites, with collections data that could be downloaded as .csv files or efficiently converted to a dataframe. The compilation of each dataframe was achieved on each institution's website by a simple keyword search for "Squamata" in the taxonomy category. An example of this workflow is given in Supplementary Figure S2, using YPM's website.

Once the dataframe was downloaded, we utilized the catalogued specimen labels to discern which elements of the fossil squamate skeleton were included under a given catalogue number. This was accomplished using keyword searches (Command/Control+F) for each bone in the squamate skeleton (Supplementary Data S1–S14) in Microsoft Excel. Because multiple skeletal elements are frequently included under a single specimen catalogue number, our search criterion was "number of occurrences" of a skeletal element on the specimen labels, rather than the number of individual elements with a catalogue number. If specific counts for a given skeletal element were indicated on a specimen label (e.g., vertebra [×35]), that counted number was included in the total specimens from the collection. If no count was included on a specimen label, then we counted the occurrence of that skeletal element as a single occurrence. We designed the keyword searches to be as inclusive as possible regarding the anatomical language on each specimen label. Some specimens had thorough anatomical descriptions (e.g., "posterior portion of the left maxilla"), while others were more general (e.g., "jaw fragment"). Regardless of level of detail on the label, if an element could be assigned to one of the seven anatomical regions we designated, we included it in the survey. Because our anatomical binning of phylogenetic characters for the assessments of phylogenetic signal uses the same seven anatomical regions (see "Parsimony-based Phylogenetic
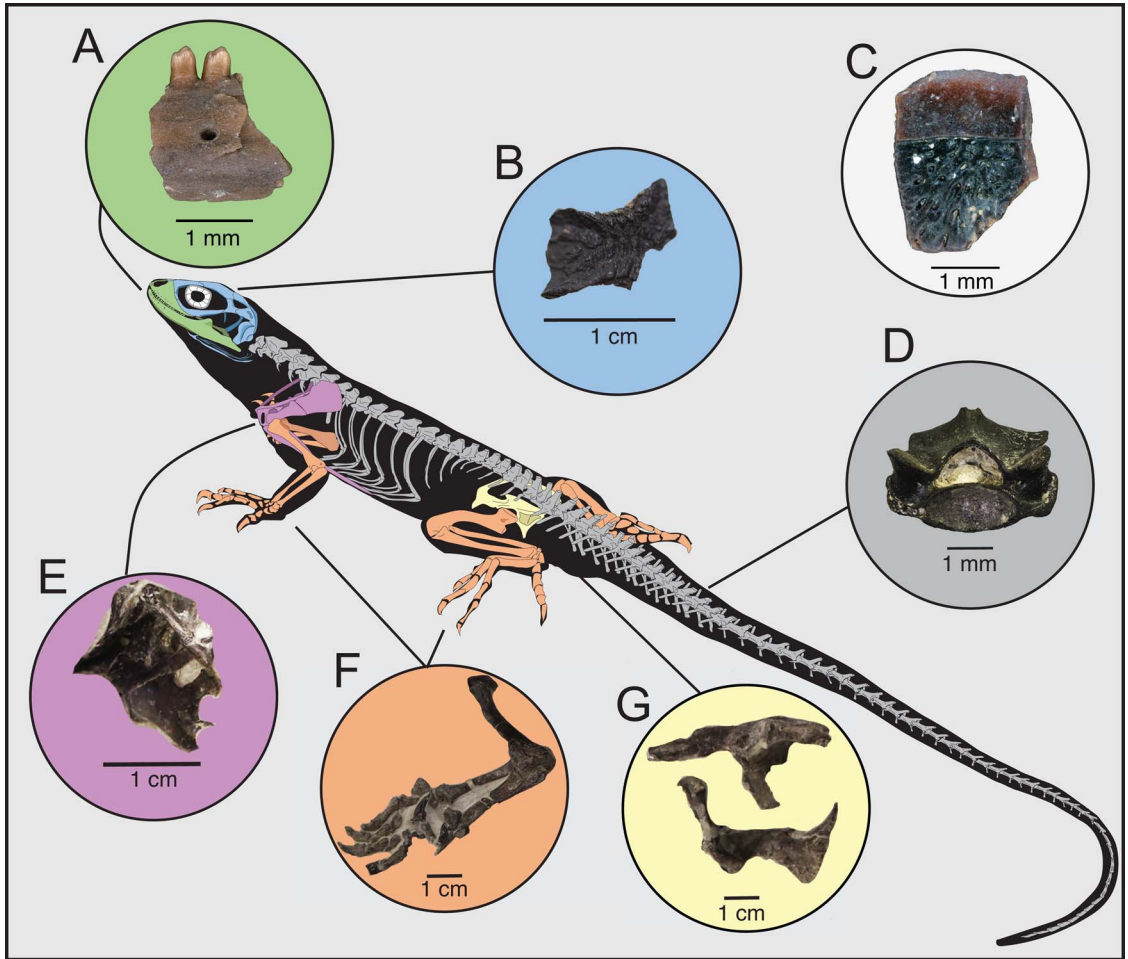
FIGURE 1. Schematic diagram of an example squamate skeleton (modeled after *Uta stansburiana*), with colorized anatomical regions used in this study for sampling of fossil squamate collections, and example fossil squamate elements. A, Jaws + palatal elements (example: DMNH EPV.119554 Scincomorpha partial left dentary; from Woolley et al. 2020). B, Posterior cranial elements (example: YPM [unnumbered], partial frontal). C, Dermal elements (osteoderms) (example: DMNH EPV.119455, Anguidae osteoderm; from Woolley et al. 2020). D, Axial elements (example: UU MAA 7173, *Ophisaurus* partial trunk vertebra; from Georgalis et al. 2019). E, Pectoral girdle (example: YPM 3230, *Polyglyphanodon sternbergi* scapulocoracoid). F, Appendicular elements (example: YPM 3230, *P. sternbergi* forelimb). G, Pelvic girdle (example: YPM 3230, *P. sternbergi* incomplete pelvis).

Signal"), this allowed us to include the specimen data with less specific labels in our characterization of bias in museum collections.

Measurement of Phylogenetic Signal

In our survey of fossil squamate collections, we expect that taphonomic and anthropogenic sampling biases will lead to some regions of the skeleton being more frequently represented compared with others (Supplementary Fig. S3). In a phylogenetic context, this imbalanced representation of the skeleton may not align with the regions of the skeleton that are more heavily emphasized for morphological character selection in phylogenetic analyses. For example, the vertebrate skull is usually the most heavily emphasized and character-dense skeletal region in phylogenetic analyses, whereas other regions of the skeleton, such as the spinal column or appendicular regions, are comparatively less emphasized and less character dense. If a group of vertebrates has a fossil record in which few or no skull material is preserved/collected and instead is mostly
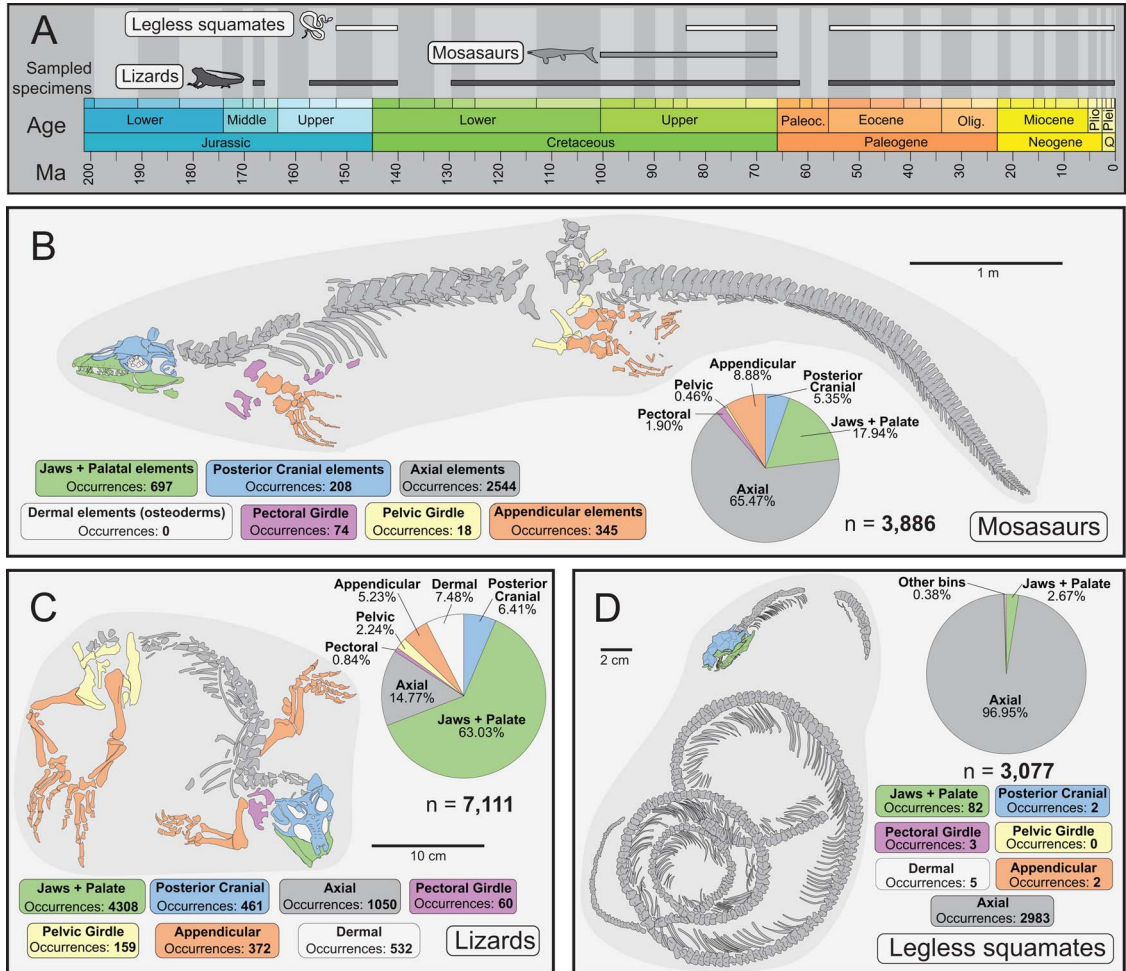
FIGURE 2.   Summary of sampled squamate collections (combined in-person and electronic databases) divided by schematic diagrams of predominant fossil squamate body plans. A, Sampled intervals in geological time for mosasaurs, lizards, and legless squamates. Silhouettes traced from publicly available renderings at PhyloPic (http://phylopic.org). B, Distribution of skeletal elements assigned to taxa within Mosasauria. Skeletal schematic drawing of LACM 128319, *Platecarpus tympanicus*, Smoky Hill Chalk Member (Niobrara Formation), late Santonian–earliest Campanian, Kansas; adapted from Konishi et al. (2012). C, Distribution of skeletal elements assigned to squamates excluding mosasaurs and legless squamates (i.e., lizards). Skeletal line drawing of YPM 3230, an articulated skeleton of *Polyglyphanodon sternbergi*, missing the tail, from the Maastrichtian North Horn Formation, Utah, USA. D, Distribution of skeletal elements assigned to legless squamates. Skeletal line drawing of SMFME 11332a *Palaeopython fischeri*, middle Eocene of Messel, Germany; adapted from Smith and Scanferla (2016).

represented by vertebrae, then the vertebrae are "overrepresented," given the smaller number of phylogenetic characters they can be scored for. Conversely, the skull material, or lack thereof, is "underrepresented" in the fossil record, given the larger quantity of phylogenetic characters it can be scored for. In this scenario, the accuracy of phylogenetic analyses for this fossil group is dependent on the fidelity of the overrepresented vertebral character evolution to the hypothesized topology of that group. To make sure these characters are reliable for reconstructing phylogeny, it is necessary to use measurements of phylogenetic signal to assess the quality of morphological character data both in overrepresented and underrepresented parts of the skeleton made available by the biased fossil record.

The presence of competing morphology-based hypotheses for the higher-level evolutionary

relationships of squamates (Gauthier et al. 2012; Simões et al. 2018) represents an opportunity to explore patterns of phylogenetic signal in the extensive fossil record (>242 Myr; Simões et al. 2018) of a major component of the modern vertebrate fauna (>11,182 extant species; Uetz et al. 2021). Broadly, phylogenetic signal describes the tendency of closely related species to resemble each other in a given trait, as the result of shared evolutionary history (Pagel 1999; Blomberg et al. 2003; Borges et al. 2019). Phylogenetic signal is measured by tracking how closely the evolution of a character trait aligns with a given evolutionary hypothesis. We used two topologically incongruent morphological phylogenetic datasets, one from Gauthier et al. (2012) (GEA) and the other from Simões et al. (2018) (SEA), to assess the reliability of phylogenetic data present in the squamate fossil record.

The GEA and SEA character matrices and .tre files of the strict consensus trees were downloaded from a publicly available phylogenetic database on G. T. Lloyd's website (Lloyd 2019). The time-calibrated Bayesian Majority-Rule Consensus tree from the SEA dataset used for calculation of the δ-statistic was obtained upon request from T. R. Simões. All phylogenetic datasets used in this study are available in Supplementary Data S15–S18 and S24–S29. To minimize the impact of highly incomplete fossils on estimates of phylogenetic signal, fossil taxa were removed from the GEA and SEA datasets using Mesquite (Maddison and Maddison 2018) and the *drop.tip* function in the ape package (Paradis and Schliep 2019) in R (R Core Team 2013), and only the extant taxa included in GEA and SEA datasets were considered herein. To effectively calculate measurements of character evolution and phylogenetic signal, morphological characters that remained in a constant state (i.e., zero changes across the entire evolutionary tree with dropped fossil taxa) were removed before carrying out analyses.

Parsimony-based Phylogenetic Signal.—To assess phylogenetic signal of skeletal elements in a parsimony-based framework, we calculated character consistency index (CI; Kluge and Farris 1969) and retention index (RI; Farris 1989) values corresponding to the seven

anatomical bins utilized in our sampling of fossil squamate collections. We used these values to measure differences in homoplasy (CI) and retained synapomorphy (RI) for characters corresponding to the seven regions of the fossil squamate skeleton. Individual CI and RI values for the 572 pruned GEA characters and 222 pruned SEA characters were calculated using the ape, TreeSearch (Smith 2018), and phangorn (Schliep 2011) packages in R (Supplementary Data S19–S22). The characters were binned according to anatomical region in our sampling of museum collections (Fig. 1, Supplementary Fig. S1), and the distributions (see Supplementary Figs. S4, S5) were then compared with one another using two nonparametric statistical tests: (1) the Mann-Whitney *U*-test; and (2) the Kolmogorov-Smirnov test (Supplementary Data S23). Because we performed multiple statistical comparisons of CI and RI values across all anatomical regions (GEA: $n = 37$; SEA: $n = 29$), statistical tests were run using a Bonferroni correction on the α value. Additionally, because the GEA and SEA datasets contain 30 overlapping operational taxonomic units (OTUs), we swapped the character matrices (Supplementary Data S76, S77) and their corresponding topologies (Supplementary Data S80, S81) to compare phylogenetic signal of the same character partitions using a fundamentally different topology (Supplementary Figs. S6–S8).

Model-based Phylogenetic Signal.—Recently, the δ-statistic (Borges et al. 2019) was developed as a means of utilizing model-based phylogenetic comparative methods to assess phylogenetic signal within categorical character data and has been recently applied in addressing paleobiological questions (e.g., Deline et al. 2020). Calculating the δ-statistic for a given morphological character relies on ancestral character state probabilities at each node in a phylogeny. The model operates under the expectation that the better a phylogeny is associated with a given trait, the better it is able to infer ancestral states with minimal uncertainty. The ancestral state probabilities simulate "entropy" at a given node: the higher the uncertainty in a given character state (i.e., less phylogenetic signal), the higher the entropy value at that node (for further

explanation, see Borges et al. 2019). δ is a measurement of the ratio of the distribution of higher entropy values to the distribution of lower entropy values for a given character across all nodes in a phylogeny. δ is higher when the distribution of entropies favors lower values (i.e., less ancestral state uncertainty) over higher values (i.e., more ancestral state uncertainty). Ancestral states for both the GEA and SEA characters were reconstructed using the R wrapper BTW (Griffin 2018) for the BayesTraits (Meade and Pagel 2016) program (Supplementary Fig. S9). Trees for the GEA dataset did not have branch lengths, so we time-calibrated the tree with a dataset of all fossil ages using the minimum branch length method (Laurin 2004). We dated 78 internal nodes dated using a previously published time-calibrated version of the GEA tree (Pyron 2017) and minimum branch lengths of 3, 5, and 7 Myr with the *BinTimePaleoPhy* function in the R package paleotree (Bapst 2012).

We initially ran these analyses using all nonconstant characters in the GEA and SEA datasets, as we had done for our parsimony-based measurements of phylogenetic signal (Supplementary Figs. S10–S13). However, in contrast to our parsimony-based measurements, we found a strong negative relationship between the percentage of missing/nonapplicable scores and the δ-statistic value for a given character (higher percentage of missing/nonapplicable scores correlates with lower δ-statistic values) (Supplementary Figs. S14–S16, Supplementary Discussion, Supplementary Data S47). For the analyses presented herein, we only considered characters for which all taxa were scored. Additionally, the extreme morphological disparities between limbed and limbless squamate taxa contributed to a large amount of missing/nonapplicable scorings when considering the two body plans together. Because of this, we ran our analyses separately for limbed and limbless squamate taxa for the GEA dataset (Supplementary Figs. S17–S19, S21) and limbed squamates for the SEA dataset (Supplementary Fig. S20). Because there are only 8 extant legless taxa in the SEA dataset, we could not analyze δ-statistic values for this subset of data; δ-statistic model sensitivity drastically decreases using <20 taxa (Borges et al. 2019).

We used the *drop.tip* function in the ape package (Paradis and Schliep 2019) in R (R Core Team 2013) to remove the appropriate taxa for each analysis. We then calculated node probabilities for each character state for the nonconstant, completely scored GEA characters and SEA characters. The δ-statistic was then calculated for each character using the set of resulting node probabilities for the GEA and SEA datasets for each character using the R script from Borges et al. (2019) (GitHub branch: mrborges23/delta_statistic) (Supplementary Data S30–S45, S52–S71). Bootstrap tests for differences between median δ-statistic values among anatomical distributions of phylogenetic characters were carried out using 10,000 samples with replacement in R (Supplementary Data S46). Because we performed multiple comparisons of δ-statistic values across all anatomical regions (GEA limbed: $n = 37$; SEA limbed: $n = 29$; GEA limbless: $n = 11$; GEA snakes: $n = 16$), bootstrap tests were run using a Bonferroni correction on the α value. Similarly to our parsimony-based assessments of phylogenetic signal, we swapped character matrices (Supplementary Data S78, S79) and topologies (Supplementary Data S80, S81) using 23 overlapping legged OTUs among the GEA and SEA datasets. Code to repeat parsimony- and model-based analyses of phylogenetic signal is available at the GitHub branch chwoolle/PhylogeneticSignal.

Results

Characterizing Bias in the Squamate Fossil Record

Similar patterns of skeletal region representation are found in the fossil record of the three major squamate body plans (Fig. 2B–D). Among mosasaurs (Fig. 2B), 83.41% of the observed fossil elements came from the axial (65.47%, $n = 2544$) and jaws + palatal (17.94%, $n = 697$) regions of the skeleton. Among lizards (Fig. 2C), 77.8% of the observed fossil elements came from the axial (14.77%, $n = 1050$) and jaws + palatal (63.03%, $n = 4308$) regions of the skeleton. The overwhelming majority (96.95%, $n = 2983$) of fossil legless squamate skeletal elements (Fig. 2D) belong to the axial region of the skeleton, whereas jaws + palatal elements

(2.67%, $n = 82$) make up almost all of the remaining occurrences. These results reveal that the majority of fossil squamate morphological data (84.28%; Supplementary Fig. S3A) in museum collections is from the axial and jaws + palatal regions of the skeleton, whereas other regions of the skeleton (posterior cranial, dermal, pectoral girdle, pelvic girdle, and appendicular elements) do not occur as frequently. Additionally, notable discrepancies in morphological data are present between specimens sampled in person and specimens sampled using electronic collections databases, with in-person sampling yielding a higher proportion of dermal and appendicular skeletal elements than sampled in electronic databases (Supplementary Fig. S3A). These differences emphasize the importance of increased physical access to natural history collections and caution against overreliance on databases as a panacea for paleobiological analyses (see Supplementary Text).

We interpret the distribution of fossil squamate skeletal data using the following terminology: the jaws + palatal and axial regions of the fossil squamate skeleton are comparatively overrepresented in natural history collections, whereas the posterior cranial region of the skull, pectoral and pelvic girdles, appendicular elements, and dermal elements are comparatively underrepresented. For example, the low occurrence of squamate posterior cranial skeletal elements in collections (4.74% of skeletal element occurrences) contrasts with an outsized proportion of morphological characters sourced from posterior cranial skeletal elements in the GEA and SEA datasets (43.61% and 34.29%, respectively; Supplementary Fig. S3B). Therefore, in both raw numbers and within the context of phylogenetic data, the posterior cranial region of the skull is underrepresented in the squamate fossil record. Similarly, the axial region of the skeleton, which makes up the largest portion of fossil skeletal data (46.79%) but a small portion of phylogenetic characters in the GEA and SEA datasets (4.10% and 14.99%, respectively), is overrepresented in the squamate fossil record. This decoupling of skeletal partitions represented in fossil data from the partitions emphasized in phylogenetic datasets necessitates a test of

whether overrepresented regions in the fossil record contain as much, more, or less phylogenetically informative character data than underrepresented regions.

Analyses of Phylogenetic Signal

Parsimony-based Assessments of Phylogenetic Signal.—Overall, we found no statistically significant difference between the distributions of CI and RI values (Fig. 3B,C,E,F) among characters corresponding to either overrepresented or underrepresented skeletal regions (Mann-Whitney U, Kolmogorov-Smirnov tests, all p-values $> \alpha$; Supplementary Data S23). If we parse out the character CI and RI distribution per squamate anatomical bin (Supplementary Figs. S4, S5, Supplementary Data S15–S23), we observe that distribution shapes and median CI values (SEA dataset) and RI values (GEA dataset) of characters corresponding to the pectoral girdle are statistically significantly different from other regions of the skeleton. Among characters in the other anatomical bins (jaws + palate, axial, posterior cranial, pelvic, appendicular, dermal, other characters), the amount of homoplasy and retained synapomorphy of characters is not significantly different when compared with one another in both GEA and SEA datasets (Supplementary Data S23).

With character matrices and topologies swapped (Supplementary Fig. S6), we found no statistically significant differences in median values and distribution shapes of overrepresented and underrepresented skeletal regions at-large (Mann-Whitney U, Kolmogorov-Smirnov tests, all p-values $> \alpha$; Supplementary Data S23). However, when we parse out the GEA matrix versus SEA topology results by anatomical region (Supplementary Fig. S7), we observe statistically significant differences in median value and distribution shape of the RI values of characters corresponding to the pectoral girdle and appendicular elements are statistically significantly different from other regions of the skeleton (Supplementary Data S23). The median and distribution shape of CI values of GEA matrix versus SEA topology characters corresponding to the pelvic girdle were statistically significantly different from values for other regions of the skeleton
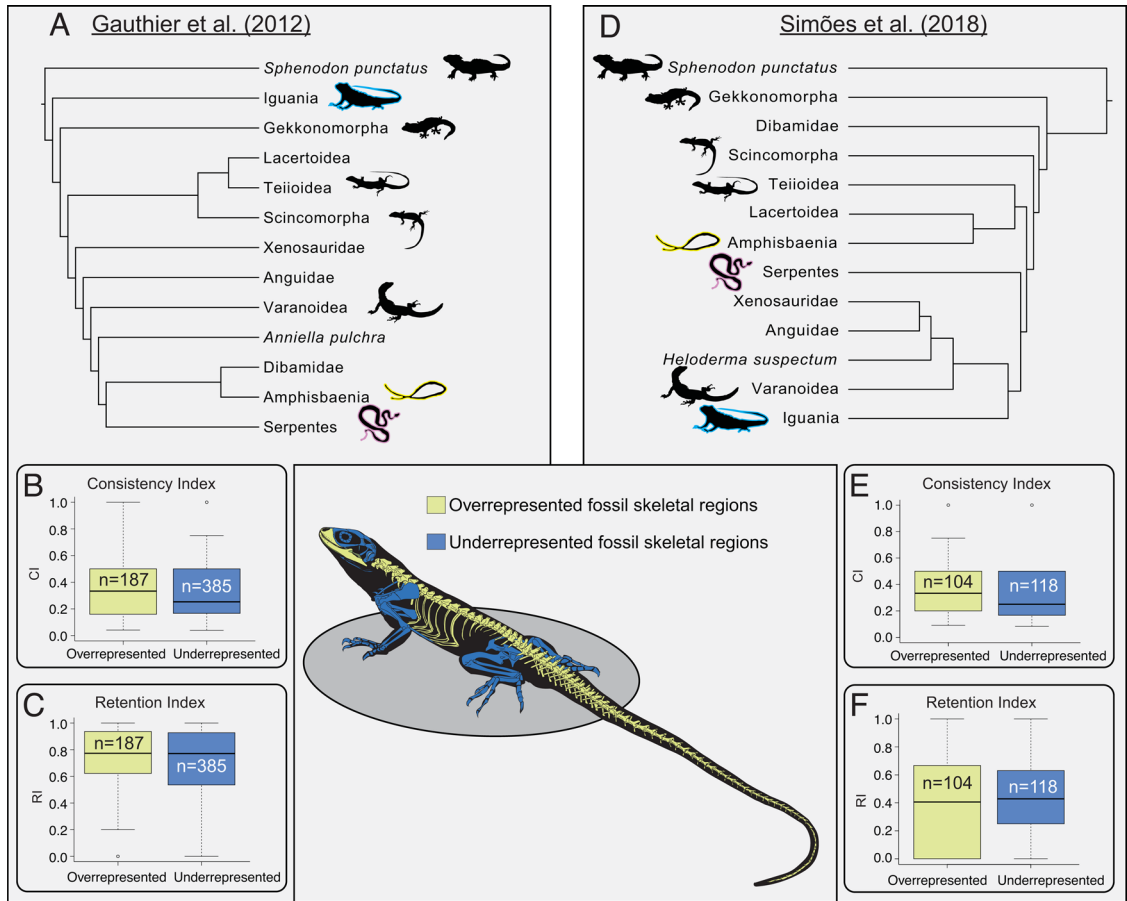
FIGURE 3. Summary of parsimony-based measurements of phylogenetic signal in the squamate fossil record used in this study. A, Hypothesis of higher-level squamate relationships according to Gauthier et al. (2012) (GEA). Silhouettes traced from publicly available renderings at PhyloPic (http://phylopic.org). Silhouettes with highlighted outlines indicate clades with notably different positions in the two hypotheses presented herein. B, Distribution of consistency index (CI) values for GEA characters corresponding to overrepresented skeletal elements in the squamate fossil record (see center panel illustrating overrepresented and underrepresented fossil skeletal regions mapped onto a schematic diagram of *Uta stansburiana*) and all underrepresented skeletal elements in the squamate fossil record. C, Distribution of retention index (RI) values for same GEA character bins as in B. D, Hypothesis of higher-level squamate relationships according to Simões et al. (2018) (SEA). E, Distribution of CI values of SEA characters. F, Distribution of RI values of SEA characters.

(Supplementary Data S23). For the SEA matrix versus GEA topology, there were no statistically significant differences among any of the anatomical regions (all *p*-values > α; Supplementary Data S23).

Model-based Assessments of Phylogenetic Signal.—For both the GEA and SEA datasets, characters with 0% missing data that correspond to overrepresented regions of the squamate skeleton in the fossil record exhibit: (1) similar overall variation in δ-statistic values and (2) similar median δ-statistic values compared with characters with 0% missing data

that correspond to underrepresented regions of the squamate skeleton in the fossil record (Figs. 4, 5). Two-sample bootstrap tests comparing 10,000 resamples with replacement of the median δ-statistic values of overrepresented and underrepresented characters, using a Bonferroni-corrected α (Supplementary Dataset S46), reveal that the differences are not statistically significant, regardless of evolutionary hypothesis (GEA limbed: $p = 0.4275$; SEA limbed: $p = 0.6978$; GEA limbless: $p = 0.1997$; GEA snakes: $p = 0.475$; Supplementary Dataset S46). Sensitivity analyses using each of the
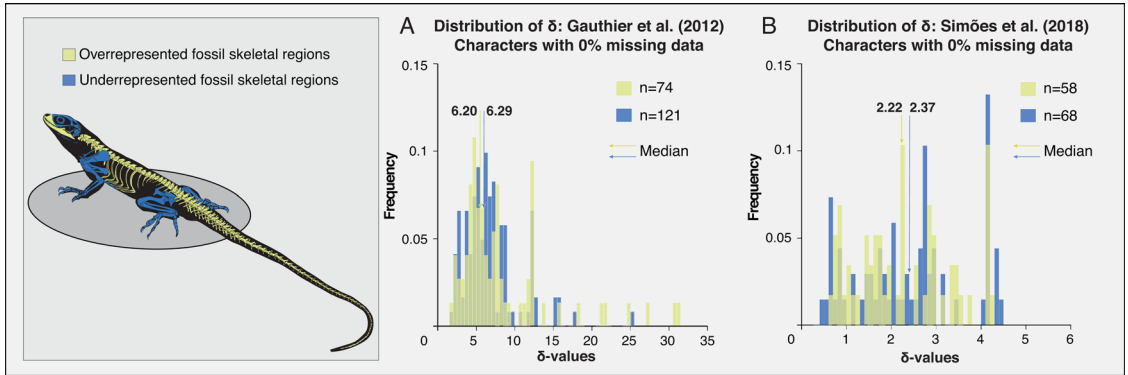
FIGURE 4. Distributions of δ-statistic for the pruned Gauthier et al. (2012) (GEA) and Simões et al. (2018) (SEA) legged squamate character dataset (prior on transition rates for ancestral node probability calculations = 0.01). Left panel: overrepresented and underrepresented fossil skeletal regions mapped onto a schematic diagram of *Uta stansburiana*. A, Comparative diagram of the distribution of δ-values between GEA characters (time-calibrated fossil tips + 78 internal nodes with minimum branch lengths of 5 Myr) corresponding to overrepresented (green) and underrepresented (blue) fossil squamate skeletal elements. B, Comparative diagram of the distribution of δ-values between SEA characters corresponding to overrepresented (green) and underrepresented (blue) fossil squamate skeletal elements.

minimum branch lengths were run to account for branch length uncertainty in the GEA dataset, though results were broadly the same (Supplementary Figs. S17–S19, Supplementary Data S30–S45). The SEA topology was already time-calibrated (Simões et al. 2018), therefore, sensitivity analyses concerning branch lengths were not run for that topology. Additionally, to assess the stability of our results under different model parameters, we ran a number of additional sensitivity analyses using different Uniform prior distributions (U) on character transition rates (U: 0–0.001; U: 0–0.1; U: 0–1.0). Results were largely insensitive to differences in model parameters (Supplementary Figs. S17–S20). When we swapped the GEA and SEA character matrices with their topologies (overlapping limbed squamate OTUs only), we found no statistically significant differences in the median δ-statistic values between any of the anatomical regions (all *p*-values > α; Supplementary Data S46, Supplementary Figs. S26–S28).

Discussion

Biases in the Squamate Fossil Record

Considering that jaw + palatal bones in the skull and the axial region of the squamate skeleton contain dozens to hundreds of individual elements (e.g., vertebrae, ribs, isolated teeth), it is unsurprising that we see the highest raw numbers of skeletal elements from these regions among the fossil specimens surveyed (Fig. 2B–D, Supplementary Fig. S3A). In fact, the relative occurrences of all seven skeletal regions in the fossil records of mosasaurs (Fig. 2B) and legless squamates (Fig. 2D) appear to superficially coincide with the number of potentially "fossilizable" skeletal elements in these animals. The major anomaly, from a preservation/collecting perspective, comes from the fossil record of lizards, which disproportionately favors the jaw + palatal region of the skull (Fig. 2C). This may be due to the fact that tooth enamel is more resilient to taphonomic factors and that isolated lizard jaws can generally be easily distinguished from other small vertebrate teeth and tooth-bearing elements (combination of pleurodonty and heterodonty). As a result, based on raw numbers, the only clearly discernible taphonomic/collecting bias in the squamate fossil record comes from the portion pertaining to lizards. Future work utilizing established fossil completeness metrics (Brocklehurst et al. 2012; Dean et al. 2016; Driscoll et al. 2019; Woolley et al. 2021) would allow us to further characterize the taphonomic and/or collecting biases that lead to the distribution of the fossil squamate skeleton observed herein.
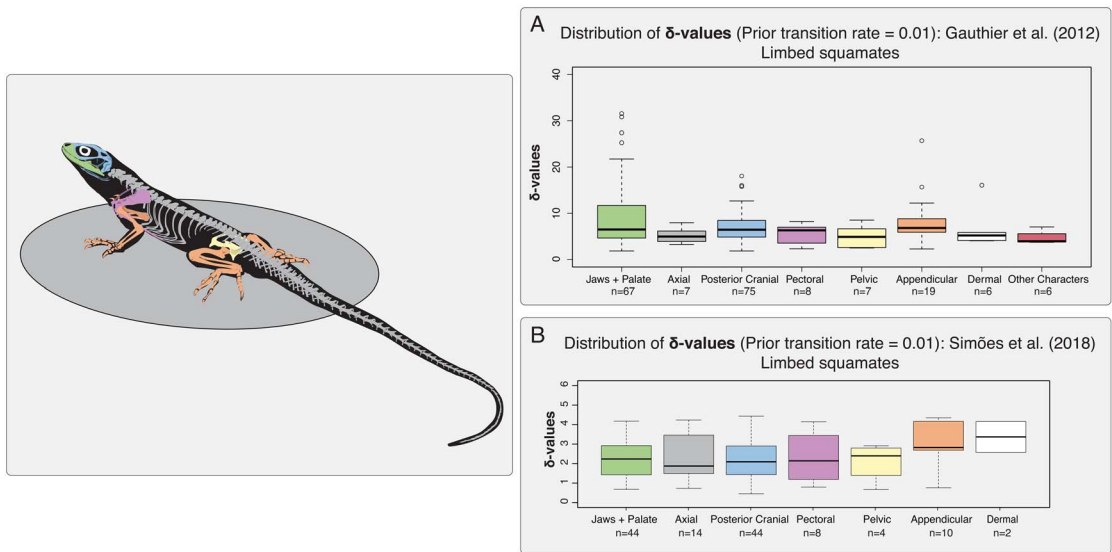
FIGURE 5. Distributions of δ-statistic among anatomical bins for the pruned Gauthier et al. (2012) (GEA) and Simões et al. (2018) (SEA) legged squamate character dataset (prior on transition rates for ancestral node probability calculations = 0.01). Left panel: Schematic diagram of an example squamate skeleton (modeled after *Uta stansburiana*), with colorized anatomical regions used in this study. A, Summary of distributions of δ-statistic values for characters with 0% missing data in the GEA dataset using only limbed squamate taxa. B, Summary of distributions of δ-statistic values for characters with 0% missing data in the SEA dataset, using only limbed squamate taxa.

If we are to consider squamate evolution as a whole, and our ability to place fossil squamates into broad, higher-level phylogenies alongside extant taxa, then the squamate fossil record is biased against the region of the skeleton containing the highest amount of phylogenetic data (the posterior cranial region of the skull, 43.61% of GEA characters and 34.29% of SEA characters; Supplementary Fig. S3B). Additionally, our survey shows that the majority of morphological phylogenetic characters available to score (GEA: 67.87%; SEA: 56.48%; Supplementary Fig. S3B) belong in squamate skeletal regions that are underrepresented in museum collections. This means that even though taphonomic and collecting biases largely align with the material in the skeleton available to fossilize (with the exception of fossil lizards), the observed record of fossil squamates is still biased against the regions of the skeleton that contain the majority of phylogenetic characters. Because the squamate fossil record is clearly missing a large quantity of useful phylogenetic data, it is imperative to assess the quality of the data that remain in characters associated with jaws + palatal elements and axial elements.

Homoplasy and Retained Synapomorphy in the Fossil Record

Results from our comparisons of CI and RI distributions among the GEA and SEA datasets demonstrate that overrepresented phylogenetic character data from the squamate fossil record are not any more likely to provide misleading evidence of phylogenetic relationships than character data from the rest of the skeleton. The lack of a significant difference in observed homoplasy between characters more prevalent in the squamate fossil record suggests that the phylogenetic placements of incomplete and fragmentary fossil taxa are neither more nor less reflective of evolutionary convergence than those of more complete, extant taxa. Similarly, the lack of significant difference in retained synapomorphy among characters in overrepresented versus underrepresented fossil squamate skeletal elements suggests that the phylogenetic placements of fragmentary fossil taxa are neither more nor less reflective of shared derived character states than more complete, extant taxa. Critically, this parsimony-based result is recovered regardless of hypothesis of squamate higher-level evolutionary relationships.

It is unclear what the exact cause is behind the lower CI and RI values for characters sourced from the pectoral girdle in the GEA phylogenetic dataset (Supplementary Fig. S4). Comparisons between CI/RI value and percentage of missing data per character (Supplementary Text, Supplementary Data S47) showed no meaningful relationship, unlike our model-based measurement of phylogenetic signal. This suggests that missing/non-applicable character data do not account for the patterns in CI/RI values among pectoral characters in the GEA dataset. It is possible that major ecomorphological transitions in squamate evolution, such as limb loss and/or specialized pectoral/limb morphologies may play a role in these differences. Alternatively, lower overall sample sizes of phylogenetic characters scored from this skeletal region (Supplementary Figs. S4, S5) could factor into these CI/RI distribution differences. Regardless of cause, the lack of significant differences in the distributions of overrepresented and underrepresented CI/RI values in the large samples sizes of phylogenetic characters in the skull (jaws + palatal and posterior cranial regions) appear to be the most influential on our results.

These results are generally consistent even when swapping character matrices and topologies (Supplementary Fig. S6), with the exception of the RI values of GEA appendicular characters mapped onto the SEA topology (Supplementary Fig. S7, Supplementary Data S23). The higher median RI value and higher variation in the range of RI values in the GEA appendicular region (Supplementary Fig. S4) are eradicated when mapped onto the SEA topology (Supplementary Fig. S7). This could be due to the possibility that GEA appendicular characters demonstrate high support for crownward nodes on the GEA topology, but because of the fundamental differences in evolutionary hypotheses, the characters do not support major nodes on the SEA tree. Further tests, including running separate phylogenetic analyses using each anatomical partition (e.g., Wencker et al. 2021), could be used in the future to understand the causes behind these differences in more detail.

Model-based Analyses of Phylogenetic Signal

Among GEA and SEA characters with no missing data, regions of the squamate skeleton that are underrepresented in the fossil record exhibit levels of median phylogenetic signal similar to overrepresented regions as measured by the δ-statistic. This is true when we consider only legged squamate taxa (Figs. 4, 5), only legless squamate taxa (Fig. 6), and snake taxa in the GEA dataset (Supplementary Fig. S21). By running separate analyses of legged taxa and legless taxa, we were able to consider a portion of characters corresponding to the pectoral girdle, pelvic girdle, and appendicular elements in legged squamates. However, even though we were able to survey representative characters for each of the seven anatomical regions, the strong negative correlation of missing data to δ-statistic values meant that we could only include roughly one-third of the total amount of characters (195 GEA characters, or 31.97% of total; 126 SEA characters, or 36.31% of total). In the absence of an appropriate benchmark study stating the minimum number of characters that can be used for the δ-statistic, it is difficult to tell how much our estimates of phylogenetic signal using the δ-statistic are impacted by low character sample size. However, our δ-statistic results are consistent with our parsimony-based results, which: (1) are derived from a much larger sample of characters (572 GEA characters, or 93.77% of total; 222 SEA characters, or 63.97% of total) and (2) have no discernible relationship with the amount of missing data (Supplementary Data S47). This suggests that smaller character sample sizes may not affect the sensitivity of the δ-statistic analyses as much as, for example, the number of included taxa (Borges et al. 2019). Further work to more rigorously assess the effect of small sample sizes on the δ-statistic is needed, but at present, we conclude that our analyses in this study are still an appropriate use of the method.

For both the GEA and SEA datasets, appendicular characters showcased the highest median δ-statistic value, in addition to an interquartile range with the highest δ-statistic values, among all anatomical bins that contained >2 characters (Fig. 5, Supplementary Figs. S22, S23, S27). Possible explanations for
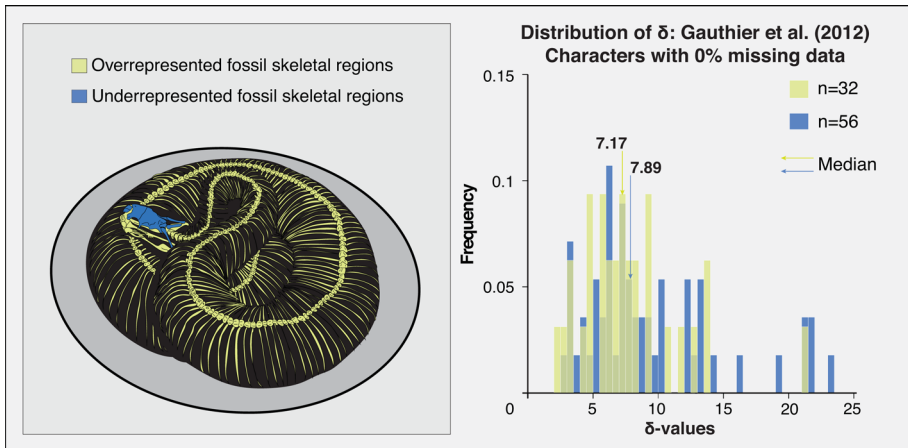
FIGURE 6.   Distributions of δ-statistic for the pruned Gauthier et al. (2012) (GEA) legless squamate character dataset (prior on transition rates for ancestral node probability calculations = 0.01). Left, Overrepresented and underrepresented fossil skeletal regions mapped onto a schematic diagram of *Crotalus atrox*. Right, Comparative diagram of the distribution of δ-statistic values between GEA characters (time-calibrated fossil tips with minimum branch lengths of 3 Myr) corresponding to overrepresented (green) and underrepresented (blue) fossil squamate skeletal elements.

this pattern among appendicular characters vary according to phylogenetic dataset. For the GEA dataset, a potential explanation for the high median δ-statistic values in appendicular characters has to do with the unique limb structure of the chamaeleonid taxa (*Brookesia brygooi* and *Chamaeleo laevigatus*) that were scored differently from all other limbed taxa. Because chamaeleonids are monophyletic in the GEA hypothesis, this results in higher δ-statistic values for the appendicular character bin, even though the characters do not illuminate relationships for any group of lizards beyond Chamaeleonidae. The character selection and taxa used in the SEA dataset are different, and the one sampled chamaeleonid taxon (*Trioceros jacksonii*) does not appear to be influencing the δ-statistic as strongly as the chamaeleonids in the GEA dataset. However, for both datasets, there are unique differences between the squamate limb and the limb of the outgroup taxon, *Sphenodon punctatus* (e.g., the presence of epiphyses on long bones), that polarize the characters and result in high δ-statistic values but are not necessarily informative on the interrelationships among limbed squamate taxa. In sum, the higher phylogenetic signal for appendicular characters is probably driven by the monophyly of a small subset of unique taxa (Chamaeleonidae), as well as the

in-group/out-group distinction among all squamates and *S. punctatus*.

Even with the character matrices and topologies swapped (Supplementary Figs. S27, S28), appendicular characters exhibit the highest median δ-statistic values and an interquartile range with the highest δ-statistic values among all anatomical regions. Why this differs from our appendicular RI values is unclear, especially given the median RI value of GEA appendicular characters was significantly lowered when mapped onto the SEA topology (Supplementary Fig. S7B). It is possible that the relatively smaller sample size of GEA characters in the appendicular region for the δ-statistic analyses excluded key characters that contributed to the lower median RI value and greater interquartile range. These conflicting results could also demonstrate greater plasticity in squamate limb morphology, where, apart from a few clades (e.g., Chamaeleonidae), limb morphological character evolution carries lower phylogenetic signal across the squamate tree of life.

δ-statistic values for characters associated with axial region of the squamate skeleton have variable distributions according to dataset and taxon subsets used. For both the GEA and SEA datasets in which only legged taxa were sampled, the axial characters have the lowest

(SEA) or third-lowest (GEA) median δ-statistic value out of all sampled anatomical regions (Fig. 5, Supplementary Figs. S22, S23, S27, S28). The same results hold for the swapped character matrices and topologies (Supplementary Figs. S27, S28). For the GEA dataset in which only snakes were sampled, the axial characters had higher median δ-statistic values than other regions (Supplementary Fig. S25). This suggests that, at least for the GEA dataset, axial characters are comparably more phylogenetically informative for snake taxa than they are for legged taxa or all legless taxa within the dataset. Within the context of the fossil record of snakes and other legless squamates (Fig. 2D), in which the vast majority of identified elements are from the axial region of the skeleton, this adds confidence to our ability to include incomplete fossil snakes into broader phylogenetic analyses.

As was the case with our parsimony-based measurements of phylogenetic signal, the anatomical bins with the largest sampling of characters (jaws + palatal and posterior cranial elements of the skull) appear to be influencing the results of our statistical comparisons the most. Bootstrap tests (Supplementary Data S46) indicate that there is no statistically significant difference between the median δ-statistic value of jaws + palatal and posterior cranial elements of the skull (all p-values > α), which correspond to overrepresented and underrepresented regions of the fossil squamate skeleton, respectively. This means that jaws + palatal and posterior cranial elements of the skull, which happen to be the most character-rich regions of the squamate skeleton, showcase levels of phylogenetic signal that are consistent with one another (Fig. 5, Supplementary Figs. S22–S25, S27, S28). It is important to consider this result alongside the observed fossil record of lizards (Nydam 2013; Rage 2013; Fig. 2C), which disproportionately preserves jaws + palatal elements of the skull (in particular, dentaries and maxillae). By demonstrating that these skeletal regions retain similar levels of phylogenetic signal as posterior cranial elements of the skull and other underrepresented regions of the fossil squamate skeleton, we can be more confident that including incomplete fossil lizard taxa in phylogenetic analyses will

not provide misleading evidence of evolutionary relationships.

A Biased Fossil Record Contains Reliable Phylogenetic Data

The arrival at similar assessments of phylogenetic signal using two non-independent but distinct phylogenetic datasets (GEA and SEA) gives us confidence that we are describing an actual natural phenomenon in the preservation of fossil data. Our parsimony- and model-based analyses show that the parts of the squamate skeleton most ubiquitously preserved and collected in the fossil record (jaws + palatal bones, vertebrae, and ribs) retain the same level of phylogenetic signal as other parts of the skeleton. Critically, these results are recovered regardless of hypothesis of squamate higher-level evolutionary relationships. This joint assessment of bias and phylogenetic signal in our sample of morphological data adds confidence to our ability to accurately infer the evolutionary relationships of fossil organisms that preserve disarticulated parts. Incorporating the world's abundance of incomplete, but potentially trustworthy fossils of vertebrates (Norell and Novacek 1992), echinoderms (Thompson and Denayer 2017), plants (Crane et al. 2004), and more into an evolutionary framework can bolster our ability to explore patterns of diversification, paleobiogeography, and biostratigraphy in deep time and can provide critical historical data for understanding today's worsening biotic crises (Ceballos et al. 2017). Using a quantitative phylogenetic framework to establish the reliability of fragmentary fossil material harnesses the full power of the fossil record in our effort to address major outstanding questions in Earth's biological history.

Acknowledgments

Data Availability Statement

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.m63xsj43b. C. H. Woolley, J. R. Thompson, Y. H. Wu, D. J. Bottjer, and N. D. Smith. 2021. Supplementary information for "A Biased Fossil Record Can Preserve Reliable Phylogenetic Signal."

Literature Cited

Bapst, D. W. 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. Methods in Ecology and Evolution 3:803–807.

Benson, R. B., P. D. Mannion, R. J. Butler, P. Upchurch, A. Goswami, and S. E. Evans. 2013. Cretaceous tetrapod fossil record sampling and faunal turnover: implications for biogeography and the rise of modern clades. Palaeogeography, Palaeoclimatology, Palaeoecology 372:88–107.

Blomberg, S. P., T. Garland Jr., and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Borges, R., J. P. Machado, C. Gomes, A. P. Rocha, and A. Antunes. 2019. Measuring phylogenetic signal between categorical traits and phylogenies. Bioinformatics 35:1862–1869.

Brocklehurst, N., and G. L. Benevento. 2020. Dental characters used in phylogenetic analyses of mammals show higher rates of evolution, but not reduced independence. PeerJ 8:e8744.

Brocklehurst, N., P. Upchurch, P. D. Mannion, and J. O'Connor. 2012. The completeness of the fossil record of Mesozoic birds: implications for early avian evolution. PLoS ONE 7(6):e39056.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences USA 114:E6089–E6096.

Crane, P. R., P. Herendeen, and E. M. Friis. 2004. Fossils and plant phylogeny. American Journal of Botany 91:1683–1699.

Daeschler, E. B., N. H. Shubin, and F. A. Jenkins. 2006. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. Nature 440:757–763.

Darwin, C. R. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London.

Dean, C. D., P. D. Mannion, and R. J. Butler. 2016. Preservational bias controls the fossil record of pterosaurs. Palaeontology 59:225–247.

Deline, B., J. R. Thompson, N. S. Smith, S. Zamora, I. A. Rahman, S. L. Sheffield, W. I. Ausich, T. W. Kammer, and C. D. Sumrall. 2020. Evolution and development at the origin of a phylum. Current Biology. 30:1672–1679.

Dornburg, A., M. Friedman, and T. J. Near. 2015. Phylogenetic analysis of molecular and morphological data highlights uncertainty in the relationships of fossil and living species of Elopomorpha (Actinopterygii: Teleostei). Molecular Phylogenetics and Evolution 89:205–218.

Driscoll, D. A., A. M. Dunhill, T. L. Stubbs, and M. J. Benton. 2019. The mosasaur fossil record through the lens of fossil completeness. Palaeontology 62:51–75.

Farris, J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5:417–419.

Field, D. J., J. Benito, A. Chen, J. W. Jagt, and D. T. Ksepka. 2020. Late Cretaceous neornithine from Europe illuminates the origins of crown birds. Nature 579:397–401.

Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record. Nature 398:415–417.

Gauthier, J. A., M. Kearney, J. A. Maisano, O. Rieppel, and A. D. B. Behlke. 2012. Assembling the squamate tree of life: perspectives from the phenotype and the fossil record. Bulletin of the Peabody Museum of Natural History 53:3–308.

Georgalis, G., A. Villa, I. Martin, D. Vasilyan, and M. Delfino. 2019. Fossil amphibians and reptiles from the Neogene locality of Maramena (Greece), the most diverse European herpetofauna at the Miocene/Pliocene transition boundary. Palaeontologia Electronica 22(3):1–99.

Griffin, R. H. 2018. BayesTraits Wrapper (btw). http://www.randi-griffin.com/projects/btw, accessed 7 July 2020.

Guillerme, T., and N. Cooper. 2016. Effects of missing data on topological inference using a total evidence approach. Molecular Phylogenetics and Evolution 94:146–158.

Kidwell, S. M., and S. M. Holland. 2002. The quality of the fossil record: implications for evolutionary analyses. Annual Review of Ecology and Systematics 33:561–588.

Kluge, A. G., and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. Systematic Biology 18:1–32.

Konishi, T., J. Lindgren, M. W. Caldwell, and L. Chiappe. 2012. *Platecarpus tympaniticus* (Squamata, Mosasauridae): osteology of an exceptionally preserved specimen and its insights into the acquisition of a streamlined body shape in mosasaurs. Journal of Vertebrate Paleontology 32:1313–1327.

Laurin, M. 2004. The evolution of body size, Cope's rule and the origin of amniotes. Systematic Biology 53:594–622.

Lloyd, G. T. 2019. Matrices > Squamates. http://www.graemetlloyd.com/matrsqua.html, accessed 4 April 2019.

Lyson, T. R., I. Miller, A. Bercovici, K. Weissenburger, A. Fuentes, W. Clyde, J. Hagadorn, M. Butrim, K. Johnson, and R. Fleming.

2019. Exceptional continental record of biotic recovery after the Cretaceous–Paleogene mass extinction. Science 366:977–983.

Maddison, W., and D. Maddison. 2018. Mesquite: a modular system for evolutionary analysis, Version 3.51. http://www.mesquite-project.org, accessed 4 April 2019.

Meade, A., and M. Pagel. 2016. BayesTraits, Version 3.01 [manual]. https://usermanual.wiki/Pdf/BayesTraitsV3Manual.1382538873/html, accessed 10 January 2022.

Murdock, D. J., and P. C. Donoghue. 2011. Evolutionary origins of animal skeletal biomineralization. Cells Tissues Organs 194:98–102.

Norell, M. A., and M. J. Novacek. 1992. The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history. Science 255:1690–1693.

Nydam, R. L. 2013. Squamates from the Jurassic and Cretaceous of North America. Palaeobiodiversity and Palaeoenvironments 93:535–565.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Patterson, C. 1981. Significance of fossils in determining evolutionary relationships. Annual Review of Ecology and Systematics 12:195–223.

Petsios, E., J. R. Thompson, C. Pietsch, and D. J. Bottjer. 2019. Biotic impacts of temperature before, during, and after the end-Permian extinction: a multi-metric and multi-scale approach to modeling extinction and recovery dynamics. Palaeogeography, Palaeoclimatology, Palaeoecology 513:86–99.

Pyron, R. A. 2017. Novel approaches for phylogenetic inference from morphological data and total-evidence dating in squamate reptiles (lizards, snakes, and amphisbaenians). Systematic Biology 66:38–56.

Rage, J.-C. 2013. Mesozoic and Cenozoic squamates of Europe. Palaeobiodiversity and Palaeoenvironments 93:517–534.

Raup, D. M. 1976. Species diversity in the Phanerozoic: an interpretation. Paleobiology 2:289–297.

Raup, D. M. 1979. Biases in the fossil record of species and genera. Bulletin of the Carnegie Museum of Natural History 13:85–91.

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Sakamoto, M., C. Venditti, and M. J. Benton. 2017. "Residual diversity estimates" do not correct for sampling bias in palaeodiversity data. Methods in Ecology and Evolution 8:453–459.

Sansom, R. S. 2015. Bias and sensitivity in the placement of fossil taxa resulting from interpretations of missing data. Systematic Biology 64:256–266.

Sansom, R. S., S. E. Gabbott, and M. A. Purnell. 2010. Non-random decay of chordate characters causes bias in fossil interpretation. Nature 463:797–800.

Sansom, R. S., M. A. Wills, and T. Williams. 2017. Dental data perform relatively poorly in reconstructing mammal phylogenies: morphological partitions evaluated with molecular benchmarks. Systematic Biology 66:813–822.

Schliep, K. P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27:592–593.

Simões, T. R., M. W. Caldwell, M. Tałanda, M. Bernardi, A. Palci, O. Vernygora, F. Bernardini, L. Mancini, and R. L. Nydam. 2018. The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps. Nature 557:706–709.

Smith, A. B. 1994. Systematics and the fossil record: documenting evolutionary patterns. Blackwell Scientific, London.

Smith, A. B. 2001. Large–scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London B 356:351–367.

Smith, A. B., and A. J. McGowan. 2005. Cyclicity in the fossil record mirrors rock outcrop area. Biology Letters 1:443–445.

Smith, K. T., and A. Scanferla. 2016. Fossil snake preserving three trophic levels and evidence for an ontogenetic dietary shift. Palaeobiodiversity and Palaeoenvironments 96:589–599.

Smith, M. R. 2018. TreeSearch: phylogenetic tree search using custom optimality criteria, R package version 1.0.1. doi: 10.5281/zenodo.1042590, accessed 23 August 2019.

Stein, W. E., C. M. Berry, L. V. Hernick, and F. Mannolini. 2012. Surprisingly complex community discovered in the mid-Devonian fossil forest at Gilboa. Nature 483:78–81.

Sun, G., D. L. Dilcher, S. Zheng, and Z. Zhou. 1998. In search of the first flower: a Jurassic angiosperm, Archaefructus, from northeast China. Science 282:1692–1695.

Sun, Y., M. M. Joachimski, P. B. Wignall, C. Yan, Y. Chen, H. Jiang, L. Wang, and X. Lai. 2012. Lethally hot temperatures during the Early Triassic greenhouse. Science 338:366–370.

Thompson, J. R., and J. Denayer. 2017. Revision of echinoids from the Tournaisian (Mississippian) of Belgium and the importance of disarticulated material in assessing palaeodiversity. Geological Journal 52:529–538.

Uetz, P., P. Freed, R. Aguilar, and J. Hošek. 2021. The reptile database. http://www.reptile-database.org, accessed 9 September 2021.

Valentine, J. W., D. Jablonski, and D. H. Erwin. 1999. Fossils, molecules and embryos: new perspectives on the Cambrian explosion. Development 126:851–859.

Vernygora, O. V., T. R. Simões, and E. O. Campbell. 2020. Evaluating the performance of probabilistic algorithms for phylogenetic analysis of big morphological datasets: a simulation study. Systematic Biology 69:1088–1105.

Wencker, L. C. M., E. Tschopp, A. Villa, M. L. Augé, and M. Delfino. 2021. Phylogenetic value of jaw elements of lacertid lizards (Squamata: Lacertoidea): a case study with Oligocene material from France. Cladistics 37:765–802.

Woolley, C. H., N. D. Smith, and J. J. Sertich. 2020. New fossil lizard specimens from a poorly-known squamate assemblage in the Upper Cretaceous (Campanian) San Juan Basin, New Mexico, USA. PeerJ 8:e8846.

Woolley, C. H., D. J. Bottjer, and N. D. Smith. 2021. The completeness of the fossil record of lizards and snakes: quantifying the effects of bias on global squamate diversity through time. Journal of Vertebrate Paleontology, Program and Abstracts, pp. 274–275.