DISCUSSION

TUATARA AND A NEW MORPHOMETRIC DATASET FOR RHYNCHOCEPHALIA: COMMENTS ON HERRERA-FLORES ET AL.

by FELIX VAUX1,2, MARY MORGAN-RICHARDS1, ELIZABETH E. DALY1 and STEVEN A. TREWICK1

1Wildlife & Ecology Group, Massey University, Private Bag 11222, Palmerston North, 4442, New Zealand; m.morgan-richards@massey.ac.nz, evolves@massey.ac.nz, elizabeth.daly@horizons.govt.nz, s.trewick@massey.ac.nz
2Current address: Coastal Oregon Marine Experiment Station, Department of Fisheries & Wildlife, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Drive, Newport, OR USA; vaux.evo@gmail.com

© The Palaeontological Association doi: 10.1111/pala.12402

It has recently been suggested that a ‘living fossil’ can be identified because it is both morphologically conservative and exhibits a significantly slower rate of morphological evolution compared to related lineages (Herrera-Flores et al. 2017a). As an exemplar, variation among known rhynchocephalians was investigated, and it was concluded that the New Zealand tuatara Sphenodon punctatus Gray is a living fossil species (Herrera-Flores et al. 2017a). We respond here to Herrera-Flores et al., questioning the biological meaning and basis for describing a ‘living fossil’, and we draw attention to limitations in their previous investigation of morphological conservatism among rhynchocephalians. To assist future studies of dentary-shape-variation data among rhynchocephalians and morphological conservatism, we provide a new geometric morphometric dataset for tuatara.

THEORETICAL PROBLEMS WITH A ‘LIVING FOSSIL’

Charles Darwin (1859, p. 486) recognized that living species possessed seemingly primitive morphological features and might, ‘fancifully be called living fossils’. As a concept a ‘living fossil’ is an oxymoron that has been widely criticized in scientific and education literature (Schopf 1984; Freeman & Herron 2007; Mobley et al. 2009; Casane & Laurenti 2013; Futuyma 2013; Mathers et al. 2013; Grandcolas et al. 2014; Werth & Shear 2014; Bennett et al. 2017). Putative living fossils are typically extant taxa that are considered to have undergone little morphological and inferred ecological change from a temporally distant ancestor known from palaeontological evidence (Grandcolas et al. 2014; Grandcolas & Trewick 2016). The amount of change considered to be ‘little’ is subjective as it varies wildly among organisms due to different evolutionary rates, durations in the fossil record, and taxonomic paradigms employed (Vaux et al. 2016; Bennett et al. 2017). Also, limited morphological change over time might not indicate evolutionary conservatism, but trait fluctuation around a mean state due to constraining selection (Charlesworth & Lande 1982; Eldredge et al. 2005; Hunt 2007; Davis et al. 2014; Grandcolas et al. 2014; Pennell et al. 2014a). The morphology of fossils represents only a fraction of the genetic, behavioural, ecological, internal morphological and physiological diversity that has evolved in a lineage (Grandcolas et al. 2014; Vaux et al. 2016, 2017a).

Most species gain recognition as ‘living fossils’ because they belong to groups that contained greater species diversity in the past (see references in Grandcolas et al. 2014) but this reasoning is barren because taxon rarity on a lineage reflects rates of extinction. Indeed, almost all biodiversity is extinct. Phylogenetic relicts are not especially uncommon, and the perception of rarity and the timing of species origination are sensitive to subjective taxonomy (Forey et al. 2004; Haile-Selassie & Simpson 2013; Vaux et al. 2016). Phylogenetic relicts at all systematic levels from genera to phyla attract attention because they represent the sole living representatives of taxonomic clades known better in the fossil record (e.g. ginkgo Werth & Shear 2014; vampire squid Yokobori et al. 2007; hoatzin Mobley et al. 2009; red panda Mallick 2015). For example, humans could justifiably be considered to be living fossils under current taxonomy due to the extinction of ~16 other species within Homo (Schopf 1984).
Some putative living fossils are inferred to have persisted over many millions of years as a single evolutionary lineage segment (e.g. *Ginkgo biloba*, see *Werth & Shear* 2014). However, a taxon that appears to straddle a long period of evolutionary time almost certainly represents inadequate sampling of the fossil record, where evolutionary change and multiple evolutionary lineage segments have simply not been documented thoroughly (*Werth & Shear* 2014; *Vaux et al.* 2016). Palaeontologists are acutely aware of this issue, and it is one reason why rare specimens are not always classified below the genus level and certain taxa are treated as merely placeholders (e.g. *Archaeopteryx* *Rauhut et al.* 2018; see discussion by *Forey et al.* 2004).

Morphological evolution is a crucial aspect of palaeontology and evolutionary biology, independent of any consideration of ‘living fossils’. Many recent studies have taken model-based approaches to analyse morphological change through time, and such methods are appropriate to address morphological conservatism and the origin of relict lineages (e.g. paleoTS and Geiger; *Hunt* 2007; *Harmon et al.* 2008; *Pennell et al.* 2014b; *Hunt et al.* 2015).

**TUATARA AS A ‘LIVING FOSSIL’**

The tuatara (*Sphenodon punctatus*) is among the most frequently cited animal phylegenetic relicts (*Fraser* & *Sues* 1994; *Jones* 2008; *Jones & Sues* 2009a; *Meloro & Jones* 2012; *Cree* 2014; *Grandcolas et al.* 2014). It is often described as a living fossil (*Hay et al.* 2008; *Mitchell et al.* 2008; *Benton* 2014; *Klein & Scheyer* 2017), although fossil remains of tuatara are all fewer than 10,000 years old (*Hay et al.* 2008). Fossil jaw fragments from New Zealand Miocene deposits have been assigned to *Sphenodontinae* (a putative subgroup within Rhynchocephalia), but the material is insufficient to identify to genus level (*Jones et al.* 2009a). Given the brief fossil record of tuatara, it is not possible to infer morphological stasis for the evolutionary lineage it represents.

*Herrera-Flores et al.* (2017a) investigated morphological variation in the dentary bones of rhynchocephalian reptiles and concluded that tuatara have a conservative morphology among their extinct relatives. The authors focused especially on the comparison of morphological variation with inferred diet. A phenogram based on morphology (*Herrera-Flores et al.* 2017a, fig. 1), indicated that tuatara are equally related to fossil genera that might have eaten plants and insects (*Clevosaurus* *Bonaparte & Sues*; *Fraser* 1988), bivalves or crabs (*Oenosaurus* *Rauhut et al.*, 2012) and seaweed (*Ankylosphenodon* *Reynoso*, 2000). A principal component analysis (PCA) scatterplot was presented too (reproduced in Fig. 1; *Herrera-Flores et al.* 2017a, fig. 3) indicating that the tuatara occupies a region of morphospace shared with putatively distant relatives that are argued to have been omnivorous, insectivorous and piscivorous (*Herrera-Flores et al.* 2017a). Since tuatara were the ‘seventh most similar to the average shape (out of 29 taxa)’, it was concluded that the species is ‘conservative’ for dentary shape as well as diet (*Herrera-Flores et al.* 2017a).

Treating diet as a categorical variable is problematic, even for living species for which behavioural and dietary data are available from observation, as habitats and feeding opportunity can vary dramatically among individuals and time (*Arthur et al.* 2008; *López-Alfaro et al.* 2014; *Williams et al.* 2015; *Sato et al.* 2016). Tuatara themselves are primarily insectivorous (*Walls* 1981; *Cree et al.* 1999; *Ussher* 1999; cf. *Herrera-Flores et al.* 2017a) but also feed on crustaceans, reptiles and fish, as well as the eggs and juveniles of small seabirds (*Cartland-Shaw et al.* 1998; *Cree et al.* 1999; *Ussher* 1999). Many animal species also change dietary preference with ontogeny. For example, juvenile *Clevosaurus brasiliensis* *Bonaparte & Sues* may have been restricted to insectivory, but it is likely that adults also ate plants (Mat & *Sato* 2015). Classifying both *Sphenodon* and *Clevosaurus* as omnivores (*Herrera-Flores et al.* 2017a) risks oversimplifying real intraspecific diversity, and introducing inaccuracy where no controls or sampling restrictions are enforced for ontogeny. A compounding issue is the accuracy of dietary estimates for fossil taxa, and inferences about extinct rhynchocephalians have been criticized (de *Vivar Martínez & Soares* 2015). Studies of other reptiles also demonstrate that dentary morphology without additional evidence, such as preserved stomach contents or coprolites, can mislead interpretation of diet (e.g. some herbivory-adapted dinosaurs ate crustaceans and rotten wood; *Chin et al.* 2017).

Beyond diet estimated from dentary morphology, little consideration was given to other known differences among the rhynchocephalians sampled by *Herrera-Flores et al.* (2017a). Predominant habitat differs among species of Rhynchocephalia from terrestrial to aquatic (de *Vivar Martínez & Soares* 2015; *Klein & Scheyer* 2017; *Bever & Norell* 2017) and there is a considerable size range (*Apesteguía et al.* 2014; de *Vivar Martínez & Soares* 2015). Notably, although tuatara exhibit similar dentary shape to the earliest known species *Brachyrhinodon* (Triassic, 250–200 million years ago), body shape differs with tuatara having a larger trunk and shorter limbs (Fraser & *Benton* 1989). Overall, there are good reasons why the well-preserved relatives of tuatara, alive millions of years ago, have been interpreted as representing diversity spanning genera and higher ranks. Therefore, focusing on a single trait is unlikely to provide a reliable assessment of evolutionary change in general (*McCormack et al.* 2010; *Funk et al.* 2012), or estimation of phenotypic stasis (*Hunt* 2007).
LIMITATIONS OF THE MORPHOMETRIC APPROACH

Herrera-Flores et al. used a two-dimensional, landmark-based geometric morphometric method to analyse morphological variation in dentary bones sampled from the extant tuatara and 30 extinct Rhynchocephalian taxa. Principal component analysis (PCA) and a disparity analysis based on Procrustes distances were interpreted as evidence of morphological conservatism (Herrera-Flores et al. 2017a). They interpreted the tuatara dentary as being ‘close to the centroid of morphospace’, indicating ‘that the morphology of the modern tuatara is conservative’ and it is also an ‘average rhynchocephalian’. Given this result, the species was interpreted as being morphologically ‘average’ and conservative among fossil relatives (Herrera-Flores et al. 2017a). This result appeared to fulfil one criterion for the proposed definition of a living fossil (Herrera-Flores et al. 2017a). In fact, closer examination of all five statistically significant principal components (PCs) for 31 taxon specimens reveals a range of the proximity to the centroid. In terms of ranked position, we found that the tuatara was 5th, 24th, 18th, 10th and 17th closest specimen to the centroid for PCs 1–5 respectively.

The geometric morphometric method used by Herrera-Flores et al. is statistically powerful, and similar landmark approaches have been used to investigate morphological variation in many different organisms (e.g. Dactylorhiza orchids, Shipunov & Bateman 2005; Achinanthishdium diatoms, Potapova & Hamilton 2007; Placostylus snails, Dowle et al. 2015; Tanganyikan cichlid fish, Clabaut et al. 2006). However, the accuracy of statistical results is always dependent on the precision of the data used. The fossil record for rhynchocephalia is sparse, and so just one dentary bone per species was used to represent 240 million years of evolution (Herrera-Flores et al. 2017a). A single dentary bone was used in the final analysis for tuatara as well, even though this living species is represented by curated specimens in museums worldwide. Thus, the conclusion of morphological conservatism for tuatara was based on the shapes of one specimen per species.

Images used for traditional and geometric morphometric analysis need to be produced in a consistent manner with similar quality photographs, rather than a mixture of material. Researchers are aware that differences in lens distortion and camera resolution, and minor changes in the positioning or lighting of an object can all introduce non-biological variation (Arnquist & Mårtensson 1998; Collins & Gazley 2017). Wherever possible experimental error is quantified (Dowle et al. 2015; Vaux et al. 2017b) because combining datasets generated by different authors can introduce further confounding variation (Arnquist & Mårtensson 1998; Fruciano 2016; Fruciano et al. 2017). The mixture of images used by Herrera-Flores et al. (2017a), was a likely source of non-biological variation. Fourteen photographs of dentary bones belonging to tuatara were initially used, but all fossil rhynchocephalian material was digitized from images produced in previous literature (Herrera-Flores et al. 2017a). Of the 30 images used, 23 were illustrated reconstructions, rather than direct digital images of intact specimens (Gilmour 1909; Fraser 1988; Ortega et al. 2009; Apesteguia & Carballido 2014). Additional shape artefacts are likely to have been introduced through digitizing PDF copies of previously published photographs and reconstructions, rather than using original images (Herrera-Flores et al. 2017b). Thus, most of the data in the analysis represented species singletons obtained from re-scanned images of fossil reconstructions, with no controls for errors.

Landmark-based geometric morphometric analyses use fixed landmarks placed at biologically homologous positions that are anatomical structures of shared origin present among all individuals or taxa (Bookstein 1991; Öyvind & Harper 2008, pp 115–116; Webster & Sheets 2010; Gunz & Mitteroecker 2013). Semi-landmarks are used to describe the curves or surfaces of an object, where homology between samples is difficult to identify. To be comparable, a curve must be anchored by at least one landmark, with semi-landmarks slid between two landmarks (or semi-landmarks) along a sequence of points (Perez et al. 2006; Sheets et al. 2006; Gunz & Mitteroecker 2013). The placement of a semi-landmark upon a curve itself is arbitrary (Webster & Sheets 2010), and so it is crucial that placement procedures are consistent between samples if mathematical descriptions of curves are to be compared (Gunz & Mitteroecker 2013). This is usually achieved using a ‘fan’ or ‘comb’ that enforces equal angular or linear spacing along a curve (Sheets et al. 2006) or by placing points evenly around a curve (Collins et al. 2013). In the ‘living fossil’ analysis, semi-landmarks were anchored between fixed landmarks and slid using Procrustes distances. However, some of these semi-landmarks seem to have been placed on biologically homologous positions, akin to the treatment of fixed landmarks. It is therefore unclear how semi-landmarks were placed consistently and why some of the biologically homologous positions used for semi-landmarks were not treated instead as fixed landmarks (see Herrera-Flores et al. 2017a, b).

Each PC is a mathematically independent axis of variation among specimens, and therefore morphological differences represented by each PC are equally valid to investigate, but it is useful to reduce interpretations to a smaller number of ‘meaningful’ PCs. It is common practice for the broken-stick test to be applied to the eigenvalues of each PC to establish statistical significance (Jackson 1993; Bro & Smilde 2014) although there are
many other component retention methods (Cangelosi & Goriely 2007). If more than the first two PCs are significant, the distribution of specimens on the further axes can potentially reveal additional relationships (Bro & Smilde 2014). Herrera-Flores et al. did not mention the number of statistically significant PCs for their dataset, and only presented scatterplots using PC1 (38.5% (38.8% in original paper)) and PC2 (15.1%). Using the broken-stick test implemented in the R package vegan (Oksanen et al. 2015; R Core Team 2016), we found that PCs 1–5 were all statistically significant in the dataset produced by Herrera-Flores et al. The additional PCs show different patterns of dentary-shape-variation with, for example, the combination of PC2 and PC5 (5.9%) revealing a more distant position of tuatara from the centroid of sampling (Fig. 1). Determining a meaningful distance from a centroid is largely subjective, and we agree with the authors that there is no methodological consensus for demonstrating morphological conservatism (Herrera-Flores et al. 2017a). However, all meaningful axes of variation need to be explored before an average can be considered.

Lastly, Herrera-Flores et al. also plotted a ‘phylogeny’ onto their PCA analysis as a phylomorphospace (Herrera-Flores et al. 2017a, fig. 3c). This approach risks circularity however, because the phylogeny (more accurately a phenogram) was derived from the same dentary-shape-variation data used to estimate the principal components. Separating derived, convergent and ancestral morphological traits in a phylogeny built from these same traits is not possible (de Queiroz 1996).

A NEW MORPHOMETRIC DATASET FOR TUATARA DENTARY BONES

We provide a new dataset for dentary shape and size variation in tuatara, adapting a previous protocol used by de Vivar Martínez & Soares (2015) to analyse dentary morphology in Clevosaurus brasiliensis. Sampling of further fossil species is required before morphological change can be investigated among other rhynchocephalians, and the new data from tuatara cannot be used alone to address the hypotheses considered by Herrera-Flores et al. In addition, the method presented here was developed to digitize tuatara, and it therefore may not be suitable for all rhynchocephalian dentary bones, and there is certainly room for modification and improvement. The intention of this method and new dataset is to assist with future investigations of morphological evolution in Rhynchocephalia. The new dataset also considers error associated with photography and landmark placement, and the variation associated with comparing left and right dentary specimens.

MATERIAL AND METHOD

Sampling

We sampled 86 dentary bones (44 left, 42 right) of Sphenodon punctatus held at the National Museum of New Zealand Te Papa Tongarewa (Vaux et al. 2018, table S1). Of these, 9 bones were of recent origin, collected from populations of tuatara living on small offshore islands, whereas 71 bones were Holocene fossils (<10 000 years old) collected at sites on the North Island and South Island (Hay et al. 2010; Fig. 2). Many of the sampled fossil bones came from geological sites or collections previously sampled for radiocarbon dating (14C), DNA sequencing or morphological trait comparison (Hay et al. 2003; Hay et al. 2008; Jones et al. 2009a). As a result, 29 fossils from six sample collections ranged between 1130 and 5190 years old (Vaux et al. 2018, table S1), based on the radiocarbon dating of bones from the same provenance (Hay et al. 2008). We did not attempt to account for differing age of specimens or ontogeny.

We first investigated the potential for experimental error during photography and digitization by repositioning and re-photographing the left and right copy of the dentary from one individual multiple times (five times left, four times right), and then re-digitized one left image four more times. Combined with 79 observations from left and right dentaries taken from separate
individuals, this yielded an Error dataset of 92 observations. A Paired dataset comprising our total sample of 86 dentaries, which included six pairs of left and right bones each taken from the same individual, was used to investigate differences between left and right bones.

Our Main dataset consisted of 80 extant and Holocene fossil tuatara dentary bones, randomly excluding either the left or right dentaries from individuals where both sides were available. For an initial inspection of variation in the Main dataset, specimens were classified into three broad geographic regions encompassing mainland and nearshore islands (Fig. 2A): (1) North: Hawke’s Bay to Northland; (2) Central: the Wairarapa, the Cook Strait, Marlborough and Tasman; (3) South: Canterbury and West Coast.

**Geometric morphometric method**

Dentary bones were positioned on fine-grade silica sand of contrasting colour. Specimens were photographed in the external lateral view using a Canon EF-S 600D camera with an 18–55 mm IS II lens. All photographs were taken from a height of 54.5 cm using a Compact Action Manfrotto tripod (MKCOMPACTACN), with a 50 mm scale bar included in each digital image. All positioning, photography and subsequent digitization was conducted by

![Sampling and morphometric landmarks used for the analysis of tuatara dentary bones. A, sample regions for our morphometric sampling of tuatara dentary bones, grouped into three broad geographic regions: North, Central and South. B, a live tuatara, using its dentary to ingest a New Zealand common gecko *Woodworthia maculata* (Gray) (Photo: Tony Jewell). C, landmarks used to digitize morphological variation in the dentary bones of the tuatara; six fixed landmarks (stars and diamonds) are individually numbered; landmarks 1, 2, 5 and 6 (stars) were also used by de Vivar Martínez & Soares (2015) to digitize dentaries of *Clevosaurus brasilienensis*, whereas fixed landmarks 3 and 4 (diamonds) and semi-landmarks 7–19 (circles) are unique to this tuatara dataset. The 13 semilandmarks used to describe the ventral edge of the dentary were identified using the comb overlaid onto the dentary. The dentary can be subdivided into ‘primary’ and ‘secondary’ bone, and it is fused with the angular bone. Although the majority of digitized specimens were dislocated, we label and illustrate the articulation of the dentary with the coronoid process and surangular bones. Colour online.
TABLE 1. Definitions used for the geometric morphometric landmarks digitized for tuatara dentary bones.

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The most superior point on the anterior end of the dentary</td>
</tr>
<tr>
<td>2</td>
<td>The inflexion point of the dentary as it descends in front of the coronoid process, before the most posterior tooth</td>
</tr>
<tr>
<td>3</td>
<td>Most superior point of the dentary immediately after the curve of the dentary parallel to the coronoid process, which grows parallel with the coronoid process</td>
</tr>
<tr>
<td>4</td>
<td>Inflexion point of the dentary in the curve after the coronoid process and before the suture of the dentary, angular and surangular</td>
</tr>
<tr>
<td>5</td>
<td>Most posterior point of the dentary, and the suture of the dentary, angular and surangular</td>
</tr>
<tr>
<td>6</td>
<td>The most anterior-ventral point of the ‘secondary bone’ of the dentary. An explanation of this feature observed in tuatara was provided by Kieser et al. (2009), and de Vivar Martínez &amp; Soares (2015)</td>
</tr>
<tr>
<td>7-19</td>
<td>Semi-landmarks describing the ventral curve of the dentary, anchored between landmarks 5 and 6</td>
</tr>
</tbody>
</table>

RESULTS

Using the photography and digitization Error dataset (92 observations), the broken-stick test identified PCs 1–4 as statistically significant. These PCs represent 51.8%, 17.4%, 9.5% and 6.0% of shape variation respectively. Based on PCA ordination, three error treatments (left photo repeat photographs, right photo repeat photographs and left digitization repeats) overlapped in morphospace and could not be separated using 90% mean confidence ellipses (Fig. 3). Comparison of Procrustes variance contributed by the error treatment groups was low and statistically insignificant (<5% of variation; Table 2). Left repeated photographs and digitizations contributed 4.4% and 2.6% of variation among all left-sided specimens, and right repeated photographs contributed only 0.9% of variation among all right-sided specimens (Table 2). In contrast, all remaining observations (i.e. separate individuals, probably reflecting true biological variation) contributed 93.0% and 99.1% of variation in left and right-sided comparisons respectively (Table 2). Given that left digitization repeats contributed an intermediate percentage of
variation to left and right photography repeats, it is likely that photography and digitization error contribute a similar level of insignificant confounding variation to a dataset. If future analyses use the same morphometric method to estimate experimental error, it should also be possible to account for inter-dataset variation.

Using our Paired dataset, containing 86 observations, which included six pairs of left and right bones taken from same individuals, the broken-stick test identified PCs 1–4 as statistically significant. These PCs represented 44.6%, 19.4%, 11.3% and 7.1% of variation among observations respectively. PCA ordination indicated that the distances within shape space between paired bones were mostly short (Vaux et al. 2018, fig. S1). Our sampling of known, paired left and right bones was limited to only six individuals, and therefore further statistical analyses without additional sampling are problematic, but we observed no consistent directional trend in shape space for any PC between left and right bones (Vaux et al. 2018, fig. S1). Size (based on centroid size) was also almost identical between paired left and right dentary bones. Our current results therefore indicate that left and right bones from the same individual are similar in shape and size, and their shape does not differ in a consistent manner among individuals.

For the Main dataset, containing 80 extant and Holocene fossil tuatara bones, with paired left or right dentaries randomly excluded, PCs 1–4 were identified as being statistically significant by broken-stick test. These PCs represented 43.6%, 19.3%, 11.6% and 7.3% of variation among individuals respectively. PCA ordination using 90% mean confidence ellipses did not reveal shape difference between left and right dentary bones (Vaux et al. 2018, fig. S2). PCA ordination did however indicate some shape variation among the sampled locations across all significant PCs (Fig. 4; Vaux et al. 2018, fig. S3). Modern and Holocene fossil specimens appeared to differ for PC1, however, this was not evident from the remaining statistically significant PCs (collectively 38.2%; Fig. 4). Based on landmark PC loading across all axes, differences in length of the dentary, particularly at the posterior portion where the ‘primary’ dentary articulates with the coronoid process and surangular bones, contributed the most to variation among specimens (Vaux et al. 2018, fig. S4).

**DISCUSSION**

The variation observed in our morphometric dataset indicates that future investigations of variation in tuatara and Rhynchocephalia should be fruitful. Using our new tuatara dentary morphometric method, we estimated that experimental error during photography and landmark digitization was statistically insignificant (<5% of variation among observations; Fig. 3; Table 2). We found no evidence of fluctuating asymmetry between left and right dentary bones (Fig. 3; Vaux et al. 2018, figs. S1, S2). If extrapolated to other sphenodonts or rhynchocephalians, this suggests that morphometric studies combining left and right sampling dentary bones (e.g. de Vivar Martínez & Soares 2015; Herrera-Flores et al. 2017a) are unlikely to be confounded by asymmetric variation. This is a useful finding because choosing between left and right bones is unlikely to be an option in the sparse rhynchocephalian fossil record, and because asymmetry is an ongoing topic of morphometric research (Klingenberg 2015). The fact that we observed little variation among modern and recent specimens.
fossil (<10 000 years old) tuatara specimens, also indicates that modern and fossil material are similar enough to permit analyses of morphological evolution using programs such as paleoTS (Hunt 2007; Hunt et al. 2015), as soon as there are enough time point samples to model a series (e.g. if fossils older than 10 000 years are discovered). This finding means that topics such as morphological conservatism considered by Herrera-Flores et al. could be addressed in the future using geometric morphometric analyses.

Without additional sampling to permit more comprehensive statistical analyses, our current sample of 80 bones (Main dataset) is insufficient to determine the biological significance of dentary shape and size variation in tuatara. Nonetheless, our dataset permits the first geometric morphometric comparison of extant and fossil tuatara. Minor shape differences were observed among samples from modern and fossil geographical locations (Fig. 4; Vaux et al. fig. S3). PCA results indicated that modern and fossil specimens mostly differ for PC1 (43.6% of variation among samples), but samples overall could not be distinguished due to homogeneity across the remaining significant PCs 2–4 (collectively 38.2%; Fig. 4). Specimens

**Fig. 4.** Shape variation of tuatara dentary bones sampled across time (A) and space (B). Principal component (PC) scatterplots illustrate shape variation among 80 sampled tuatara dentary bones (Main dataset). Variation is shown for the comparison of PC1 (43.6% of variation) and PC2 (19.3%), as well as PC3 (11.6%) and PC4 (7.3%). 90% mean confidence ellipses for groups are shown. A, temporal origin: sampling of modern and Holocene fossils <10 000 years-old. B, spatial origin (see Fig. 2A). Colour online.
CONCLUSION

The objective and approach of identifying a 'living fossil' is of limited evolutionary insight. We find that the previously available data (Herrera-Flores et al. 2017a) are insufficient to justify the statistical interpretation of tuatara as morphologically conservative among ancestors or a living fossil. However, with appropriate sampling and data, geometric morphometric methods provide powerful tools for testing explicit hypotheses about morphological evolution.

The new tuatara geometric morphometric dataset provided uses landmarks that should be compatible with the dentary bones of other rhynchocephalians (e.g. Cleosaurus brasilensis; de Vivar Martínez & Soares 2015) and we have demonstrated that experimental error and fluctuating asymmetry are unlikely to be sources of confounding variation (at least for this exact method in tuatara). To thoroughly investigate other hypotheses in tuatara, such as morphological stasis and population-level variation, further sampling is required.

The variation observed in our new tuatara dataset adequately demonstrates why the reduction of an entire species to a single specimen for evolutionary analysis, as conducted by Herrera-Flores et al. (2017a), is inappropriate (see similar discussion by de Vivar Martínez & Soares 2015). Given the criticisms we highlight about the analysis and data used by Herrera-Flores et al., we conclude that there is no advantage in using precision methods such a geometric morphometrics if the specimens used are deficient or suspect. We expect the dataset provided for tuatara to facilitate future morphometric analyses of Rhynchocephalia and efforts to investigate morphological conservatism.

Acknowledgements. We thank Thomas Schultz and the Museum of New Zealand Te Papa Tongarewa for providing us with access to the tuatara specimens, and Paulo R. Romo de Vivar Martínez (Federal University of Rio Grande do Sul) for sharing his morphometric dataset for Cleosaurus brasilensis. We are grateful for helpful comments by Katie S. Collins (University of Chicago) and James S. Crampton (GNS Science/Victoria University of Wellington) regarding the geometric morphometric methods. Thank you to Tony Jewell for sharing the photograph of the live tuatara specimen. Lastly, we thank the Editor, Andrew Smith and our two anonymous reviewers for their generous and detailed feedback regarding the manuscript and the geometric morphometric methods.

DATA ARCHIVING STATEMENT

Data, and supplementary table and figures for this study are available in the Dryad Digital Repository: https://doi.org/10.5061/dryad.ms7q2q4

Editor. Hannah O’Regan

REFERENCES


BEVER, G. S. and NORELL, M. A. 2017. A new rhynchocephalian (Reptilia: Lepidosauria) from the Late Jurassic of Solnhofen (Germany) and the origin of the marine Pleurosauridae. Royal Society Open Science, 4, 170570.


