

Research



Cite this article: White HE, Clavel J, Tucker AS, Goswami A. 2020 A comparison of metrics for quantifying cranial suture complexity.

J. R. Soc. Interface **17**: 20200476.

<http://dx.doi.org/10.1098/rsif.2020.0476>

Received: 16 June 2020

Accepted: 8 September 2020

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

evolution, computational biology

Keywords:

suture, complexity, mammal, morphology, morphometrics, skull

Author for correspondence:

Heather E. White

e-mail: h.white@nhm.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5127032>.

A comparison of metrics for quantifying cranial suture complexity

Heather E. White^{1,2,3}, Julien Clavel^{1,4}, Abigail S. Tucker² and Anjali Goswami^{1,3}

¹Department of Life Sciences, Natural History Museum, London SW7 5BD, UK

²Centre for Craniofacial and Regenerative Biology, King's College London SE1 9RT, UK

³Division of Biosciences, University College London WC1E 6DE, UK

⁴Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, F-69622, Villeurbanne, France

id HEW, 0000-0002-0874-6259; AST, 0000-0001-8871-6094

Cranial sutures play critical roles in facilitating postnatal skull development and function. The diversity of function is reflected in the highly variable suture morphology and complexity. Suture complexity has seldom been studied, resulting in little consensus on the most appropriate approach for comparative, quantitative analyses. Here, we provide the first comprehensive comparison of current approaches for quantifying suture morphology, using a wide range of two-dimensional suture outlines across extinct and extant mammals ($n = 79$). Five complexity metrics (sinuosity index (SI), suture complexity index (SCI), fractal dimension (FD) box counting, FD madogram and a windowed short-time Fourier transform with power spectrum density (PSD) calculation) were compared with each other and with the shape variation in the dataset. Analyses of suture shape demonstrate that the primary axis of variation captured attributes other than complexity, supporting the use of a complexity metric over raw shape data for sutural complexity analyses. Each approach captured different aspects of complexity. PSD successfully discriminates different sutural features, such as looping patterns and interdigitation amplitude and number, while SCI best-captured variation in interdigitation number alone. Therefore, future studies should consider the relevant attributes for their question when selecting a metric for comparative analysis of suture variation, function and evolution.

1. Background

The quantitative analysis of shape is a well-established approach to robustly address questions across a breadth of disciplines and subdisciplines in biology, with broad application in the study of evolution, ecomorphology, development, biogeography, taxonomy and phylogenetics [1–10]. The ability to capture complex shape data has been greatly expanded by advances in geometric morphometric techniques which improve upon traditional morphometrics by providing the ability to capture information about where the parts of the shape are located with respect to each other in a Cartesian plane. These techniques have grown a particular interest and have been successful in the study of vertebrate morphology during the past three decades. Within the system of the vertebrate skull, the overwhelming focus of morphometric studies has been the bony structures themselves, with little attention given to the sutural joints between these cranial bones [11–22].

As sites of intramembranous bone growth, sutures are actively responsible for determining this cranial morphology via the facilitation of postnatal craniofacial development [23]. An intrinsic relationship clearly exists between sutures and cranial morphology, primarily due to sutures providing the principal growth site, although the relationship is not restricted to this and extends to other critical functional roles of sutures; facilitating brain development by allowing adequate space for brain expansion, fusing for protection of the developed brain, providing

shock absorption during locomotion and facilitating feeding by providing joint mobility [24–26]. These suture functions have a clear impact on the cranial morphology. Despite their functional importance, suture morphology has seldom been studied in a comparative, quantitative framework and is considerably underrepresented in comparison to the well-studied morphology of the cranium. This stems from the lack of well-defined homologous anatomical landmarks for non-osseous and essentially featureless structures [27] while the open outline nature of sutures further adds to the inherent difficulty in quantifying their shape [28,29].

Suture morphology is known to be highly variable from suture to suture, across development, between the sexes and across species. When comparing suture to suture, those in the facial region have been reported to have straight morphologies compared to the interdigitated sutures of the braincase [30]. Looking across development, sutures have been suggested to progress from straight to highly interdigitated morphologies [31]. Sexual dimorphism of sutures has also been identified, with males exhibiting greater levels of complexity for facial sutures, thought to reflect the pressures of head-to-head fighting [32]. Interspecific species variation has been described at different levels, both within genus (*Cebus*) and within infra-order (Caviomorpha), where suture morphology ranges from low to high suture interdigitations [33,34]. Beyond this level, very little work has been conducted that addresses interspecific variation across larger clades, although it has been postulated that mammals have more intricate and interdigitated sutures than reptiles [30]. Evidently, a huge degree of suture variation has been observed at multiple levels of comparison.

The few papers that have quantified this variation in suture morphology to date, have suggested a wide range of drivers shaping suture morphological variation including innovation over evolutionary time, novel biomechanical relationships to ecological variation, mechanical strain from compressive forces, patterns of growth and development, and responses to environmental stress [29,30,35–38]. More specifically, the aspects of complexity shaping this morphological variation have been associated with a number of biological pressures. These biological pressures create heightened biomechanical stress across the skull, with the resultant effect being a deformation of the sutures to an interdigitated morphology [39]. Interdigitations act as shock absorbers by increasing the surface area for collagen and compression-resisting fibre attachments to dissipate the strain experienced by surrounding bones [39,40]. Biomechanical stress shaping this interdigitated suture morphology is thought to be emergent from a number of biological inputs: head-to-head fighting, tougher foods, chisel tooth digging, jaw opening and closing during mastication, and masticatory muscle mass [30,33,34,40–42]. However, very little is known about the forces generated from these sources of biomechanical stress and the subsequent impact on suture morphology [43]. Unmistakeably interdigitations play a huge part in the biological relevance of suture morphology. However, as both amplitude and frequency form this interdigitated morphology [29], it is unclear from the current studies whether one or both of these aspects of interdigitation are functionally relevant to the suture and thus biologically speaking important. Nevertheless, the suture morphological variation appears to reflect adaptations to mechanical demands from the environment, habitat and ecological pressures. The wider significance of how these pressures and biomechanical stresses impact suture morphology is

currently limited by the absence of comparative studies in a phylogenetic framework which are necessary to place the biological interpretations into a broader evolutionary context [37].

Variation in suture morphology has been captured in a range of biological systems, from the vertebrate cranium to the ammonoid shell using methods that attempt to quantify complexity [28–30,33,34,37,44]. The implementation of a complexity metric, in contrast with using shape data, such as semi-landmarks, offers the advantage of generating a single quantitative value that provides the ability to easily make direct comparisons across species, suture locales and developmental ages. Moreover, shape incorporates many other attributes beyond complexity, complicating functional interpretations across diverse systems. For the purpose of our study, we consider complexity to measure how a suture deviates from its simplest form, a straight line, across its overall path. By contrast, shape captures additional aspects such as orientation and a reversal of the same morphology. While multiple methods have been developed to describe suture morphology as a complexity metric with this in mind, the transformation of suture morphology into meaningful quantitative data still does not come without its challenges. The intricate nature of sutures produces patterns that are hugely diverse even within an individual suture. Sutures reflect a tremendous degree of natural complexity, making their morphology inherently difficult to quantify. The challenges, however, are not limited to the suture morphology, but are further complicated by the methods themselves. The methods previously applied, operate on a variety of mathematical approaches, each of which adopts different assumptions, therefore resulting in variation for the definition and aspects of complexity captured. Moreover, many previous methods attempt to quantify morphology by means of a univariate statistical complexity value, whereby disparate morphologies have been reported to be present with near-identical complexity values, meaning it is impossible to reconstruct the original suture morphology [45]. An outline of the approaches and equations used here can be found in the electronic supplementary material (electronic supplementary material, text S1).

Despite the many methods proposed, there is no consensus regarding which approach is the most appropriate for capturing suture morphology as a quantified value of complexity, and in particular if the different approaches capture various aspects of the suture shape complexity which may be appropriate for addressing different biological questions. While we know that complexity in the form of an interdigitated shape has functional roles at the sutures, other aspects of complexity may also be functionally significant but have not been recognized in the limited previous work on this topic. Direct comparison of the strengths and weaknesses of the alternative approaches and identification of the different aspects of complexity each approach captures, will inform further research into suture morphology, development and evolution. In addition, the direct comparison of the approaches will offer the potential to integrate the results of studies implementing different complexity metrics. Here, we compare five methods for quantifying suture complexity: sinuosity index (SI), suture complexity index (SCI), fractal dimension (FD) box counting approach, FD madogram approach and windowed short-time Fourier transform (STFT) with a power spectrum density (PSD) calculation, using a dataset sampled across a diverse range of mammal suture morphologies. Mammal crania exhibit an incredible variety of shape and

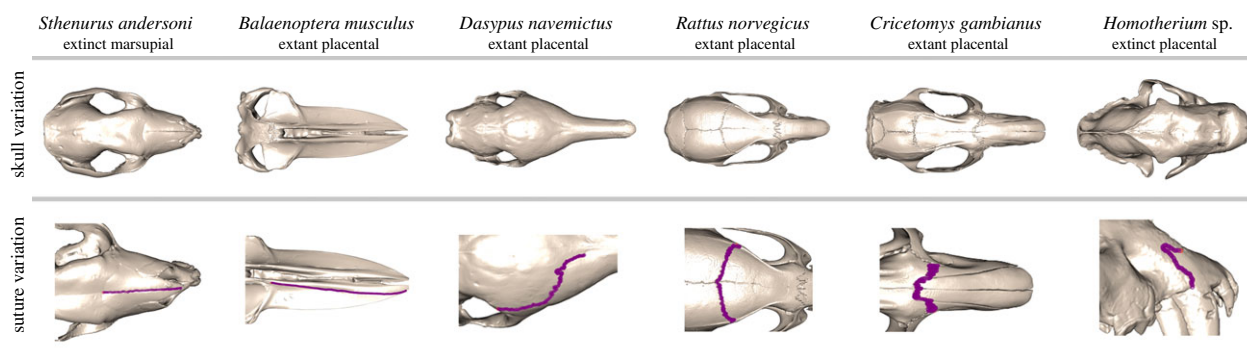


Figure 1. An example of skull and suture variation captured by the dataset from a broad range of mammalian taxa.

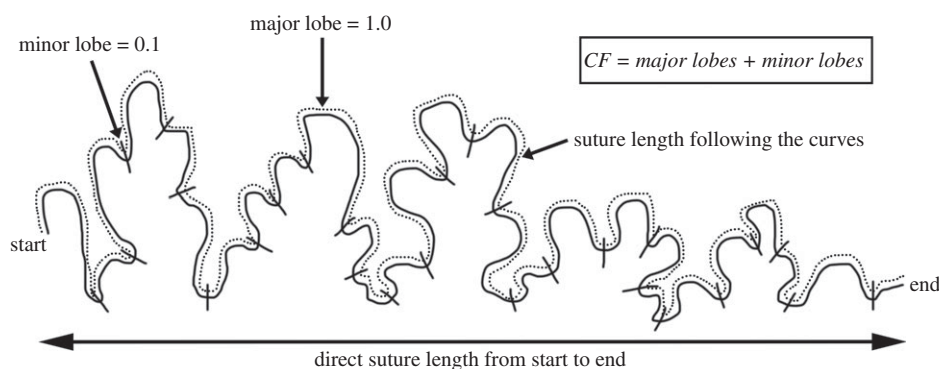


Figure 2. Depiction of SI and SCI calculation [51].

suture morphology, likely related to their complex mastication, and dietary and ecological diversity [46]. To accompany this cranial diversity, a wide range of suture morphologies are observable across mammals. Moreover, they are abundant and easily available in museum collections, thus providing the ideal sample dataset for comparing the metrics for quantifying suture morphology and complexity. Determining the aspects of complexity captured by the various metrics and which is the most appropriate approach for quantifying cranial suture complexity on this diverse empirical dataset will allow for future detailed study of the function, development and evolution of sutures across mammals.

2. Methods

2.1. Data collection

A sample of 79 extinct ($n = 34$) and extant ($n = 45$) mammals was generated to assess and test approaches available for quantifying suture morphological complexity (electronic supplementary material, table S1). Multiple sutures were sampled across this sample dataset rather than a single homologous suture, as the goal was to determine the most appropriate metric for a sample with a broad range of morphologies and not to assess the morphological variation within the sample dataset (figure 1). Spirit preserved specimens were scanned using the X-Tek HMX ST 225 μ CT scanner (Nikon, Tokyo, Japan) and Go!Scan 20 (Creaform), Go!Scan 50 (Creaform) or EDGE ScanArm HD (FARO) scanners were used for osteological specimens. Scans were reconstructed in Avizo v.9.3 (FEI, Hillsboro, OR, USA), Geomagic Wrap (3D Systems), and VXEelements v6.0 (Creaform) and cleaned in Geomagic Wrap leaving only the skull elements in the three-dimensional isosurface. Two-dimensional images ($n = 79$) of the multiple different sutures across the sample dataset were captured using uniform positioning to minimize the

impact of parallax in the *rgl* R package [47]. Two-dimensional semi-landmarks were manually positioned on the suture images using the *StereoMorph* R package [48] and resampled at 500 semi-landmarks per suture to capture the open suture outline. Studies should first determine the number of manual landmarks required to capture the entire suture curve profile and resample to a number of landmarks that does not risk losing morphological complexity. This is particularly important as fractal-like sutures could be more sensitive to the number of landmarks positioned. Generalized Procrustes analysis was performed on the resampled two-dimensional semi-landmarks to remove all non-shape elements using the 'gpagen' function in the *geomorph* R package [49,50]. Using the two-dimensional superimposed semi-landmarks, a principal component analysis (PCA) was performed to determine the major axes of shape variation captured by the sample dataset and whether these reflected variations in complexity.

2.2. Complexity analysis

The five complexity metrics were applied to the Procrustes superimposed two-dimensional semi-landmarks. SI calculations were implemented in base R (v.3.6.0; R Core Team, 2019) using a standard distance between two points equation (electronic supplementary material, textS2: equation 7) in order complete the SI equation (electronic supplementary material textS2: equation 1, figure 2) [51]. SI has also previously been referred to as total sutural interdigitation [52] and relative suture length [33]. SCI used the SI value calculated in base R (v.3.6.0; R Core Team, 2019) and a complexity factor multiplier which was calculated from the interdigitation lobes (major and minor) [53] (figure 2). FD was calculated in the *fractaldim* R package [54], using both the box counting and madogram methods by implementing the 'fd.estim.boxcount' and 'fd.estim.madogram' functions, respectively. STFT was computed using the 'stft' function in the *e1071* R package [55]. The PSD of each suture was calculated using the STFT results, by averaging the squared STFT coefficients

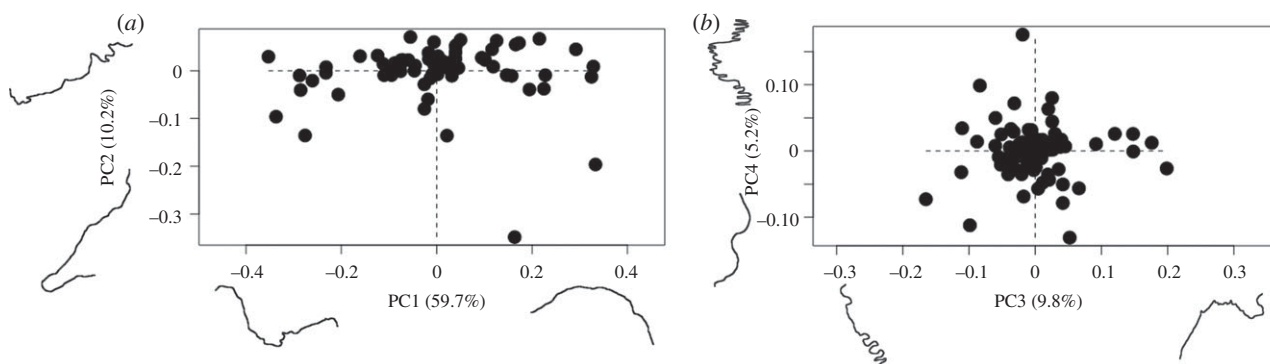


Figure 3. Principal components analysis of two-dimensional semi-landmark shape data with extreme suture morphologies for: (a) PC1 and PC2; (b) PC3 and PC4.

over each frequency across the local transforms and summing the averages at each harmonic, as described by Allen (2006) [29].

2.3. Comparative analysis of the complexity metrics

The appropriateness and effectivity of each metric was assessed using multiple comparative approaches. Comparisons between suture shape data and complexity scores were made to determine whether the metrics captured functionally relevant suture morphologies. Functionally relevant aspects of suture morphology include, but are not limited to, the multiple features of interdigitations (number and amplitude) thought to be shaped by biomechanical stresses as a result of behaviour and diet. Correspondence between complexity score and suture shape PC score were quantified with Pearson's correlation coefficient (r) and visualized with a heatmap using the *'corrplot'* function from the *corrplot* R package [56]. The significance ($p < 0.05$) of the correlations were tested using the *'cor.mtest'* function [54]. A heatmap reflecting complexity scores for each specimen was mapped onto the PCA of Procrustes superimposed two-dimensional semi-landmarks to further determine whether the complexity metrics captured the functionally relevant aspects of morphological variation. Similarly, correlations between the five complexity metrics were quantified and visualized using the *'corrplot'* function in the *corrplot* R package, significance was tested using the *'cor.mtest'* function [56]. Finally, we conducted a further PCA on the specimen complexity scores, rather than the shape data, using the *factoextra* R package [57]. The contribution of each variable ($n = 5$; complexity metrics) on the PC axes (i.e. the scaled PC loadings) were plotted on a correlation circle. The PC loadings were superimposed onto the PCA, of the first two axes, thus representing both variables ($n = 5$) and individuals ($n = 79$) using the *'fviz_pca_biplot'* function, in order to assess how the approaches, relate to each other and the complexity across the sample dataset. Further detail on the methodology is provided in the electronic supplementary material (electronic supplementary material, text S2). The datasets generated and/or analysed throughout our study are available in the Github repository: https://github.com/HeatherEWhite/suture_metrics_comparison [58].

3. Results

3.1. Shape analysis

Shape analysis refers to variation in the shape variables themselves, the Procrustes superimposed semi-landmarks. PCA of the Procrustes superimposed two-dimensional semi-landmarks (figure 3) identified 19 principal components (PCs) that explained 99% of the shape variation (electronic supplementary material, table S2). Of these PCs, the first four each accounted for greater than 5% of the overall variation, with PC1, PC2, PC3 and PC4 accounting for 59.7%, 10.2%,

9.8% and 5.2%, respectively. The main axis of variation (PC1) was associated with an inversion of the suture morphology, with the positive aspect of PC1 reflecting an 'n' shaped morphology and the negative aspect reflecting a 'u' shaped morphology (figure 3 and electronic supplementary material, figure S2). Variation along PC2, PC3 and PC4 more clearly reflected differences in the complexity of shape, such as looping patterns, irregularity and straight to interdigitated sutures (figure 3 and electronic supplementary material, figure S2).

3.2. Complexity analysis

Complexity analysis captures variation in the complexity scores across the dataset, which are estimated from the suture shapes, but not the shapes themselves. The five methods produced different ranges of complexity across the sample dataset, with the largest value indicative of the most complex suture: 1.024 to 4.861 for SI; 0.205 to 87.963 for SCI; 1.055 to 1.120 for FD (box counting method); 1.452 to 1.657 for FD (madogram method); 1.449 to 1.928 for PSD, extreme morphologies for each metric are presented in electronic supplementary material, figure S3. Full results for each suture are presented in electronic supplementary material, table S3, with a subset of sutures representative of the range of suture complexity detailed further in figure 4.

Sutures recognized as 'simple' (figure 4*a,b*) following SI, SCI, FD (box counting), FD (madogram) and PSD analysis presented with a largely straight morphology and very few interdigitations. For each of the simple sutures, all metrics indicated a low complexity value. The FD and PSD complexity values were very similar to one another, with greater differences observed for the SI, SCI and FD (box counting) metrics.

Sutures recognized as 'moderately complex' (figure 4*c,d*) by all the applied metrics presented with a morphology that deviated away from a straight line accompanied by a greater number of interdigitations that were small in magnitude. All methods quantified these morphologies with higher values than for the simple sutures.

The 'most complex' sutures, recognized by the majority of applied metrics, exhibited large magnitude interdigitations, irregular repetitions and looping patterns (figure 4*e,f*). These more complex sutures had higher SI values than the moderately complex sutures. SCI indicated that suture 31 (specimen ID 31), which had interdigitations of a heightened magnitude, was highly complex compared to the relatively simple suture 46 (specimen ID 46), which presented with an irregular looping pattern. PSD, however, showed the opposite, with suture 46 (specimen ID 46) being the most complex of the two, although suture 31 (specimen ID 31) was still

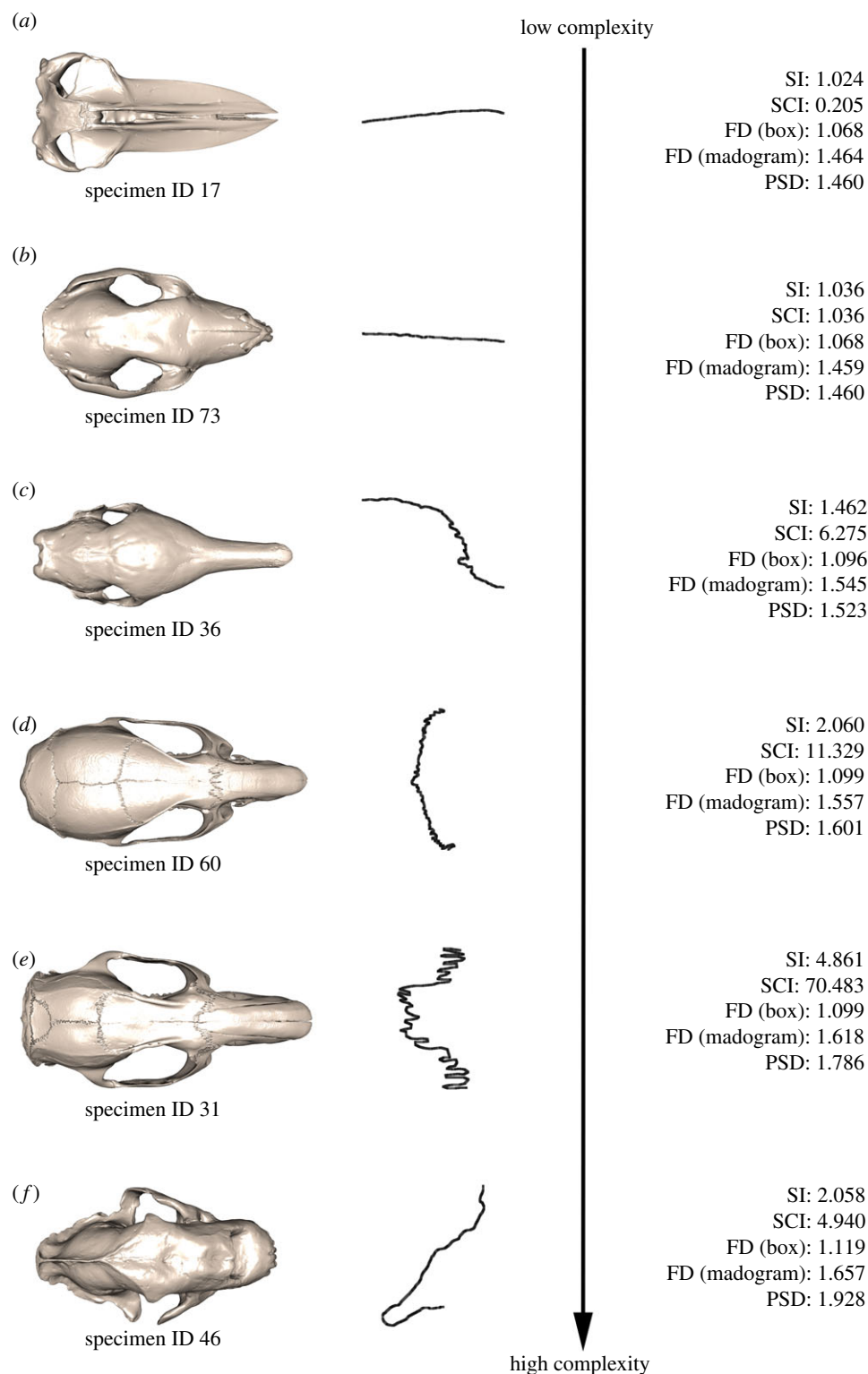


Figure 4. Suture complexity variation from low (a) to high (f) complexity within the mammal skull sample ($n = 79$). Results presented for all five methods tested: SI, SCI, FD box counting, FD madogram, PSD. Specimen ID and their species associations are detailed in electronic supplementary material, table S1.

represented as complex. FD box counting and FD madogram values returned very similar values for the two sutures considering them both to be similarly highly complex sutures. Evidently, while both sutures were identified as complex by the majority of metrics their morphologies were largely disparate. These differences in morphology appeared to be better reflected by the SCI and PSD values, as these both ascribed quite different complexity scores to the two sutures compared to the values determined by SI, FD box counting and FD madogram. Scenarios, where the five complexity metrics disagree, undoubtedly become more evident with the more highly

complex sutures as a result of greater shape variety in the more complex sutures which is not always limited to interdigitations.

3.3. Comparative analysis of the metrics

Here, we consider how variation in shape compares to variation in complexity, while addressing the relationships between each complexity metric. PC scores for all specimens were extracted from the PCA of suture shape variation (figure 3), for the PCs that explained greater than 5% of the

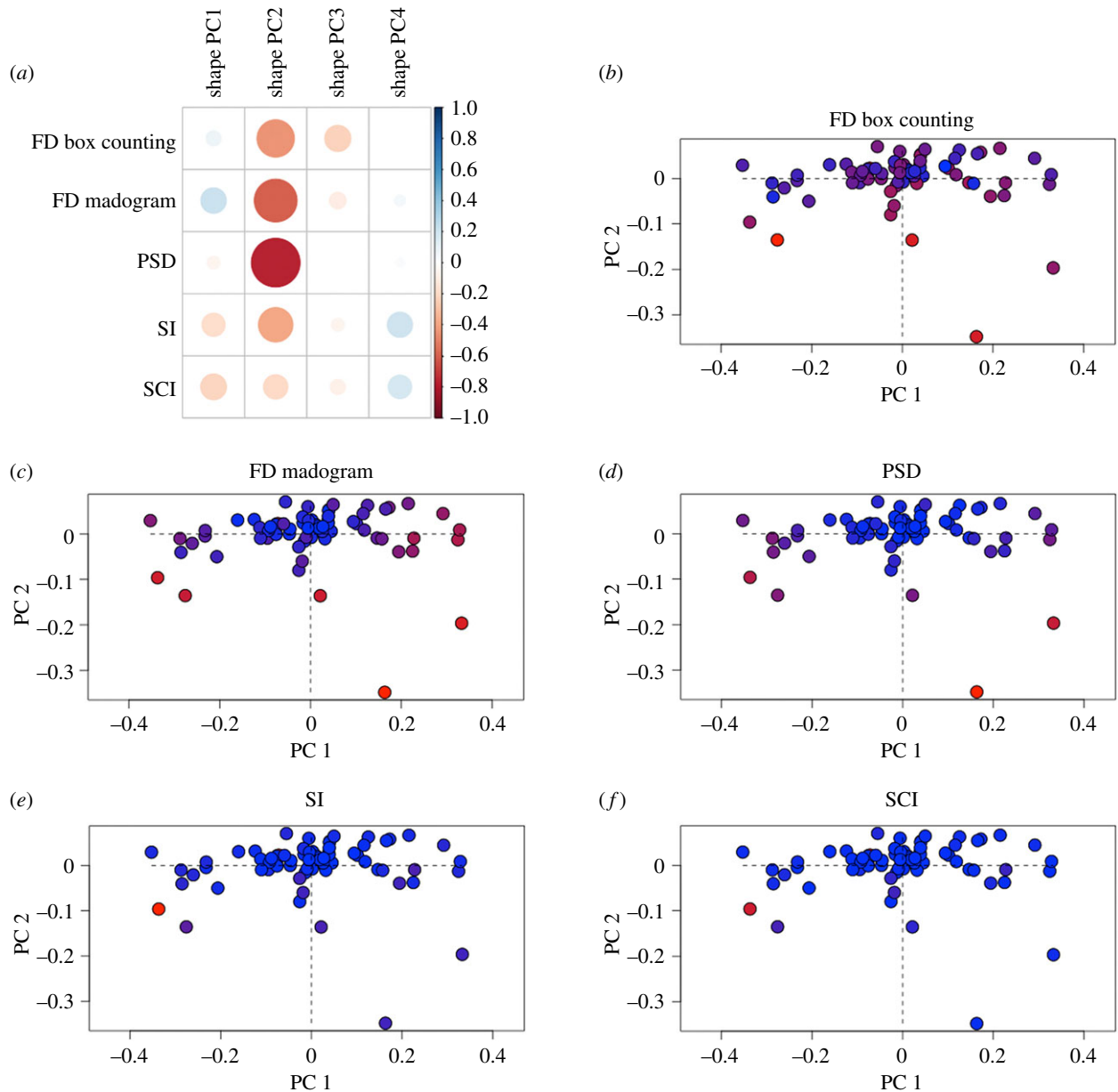


Figure 5. Principal components analysis of two-dimensional semi-landmark data and its relationship to the complexity methods: (a) correlation plot between the principal component scores of the shape data and the complexity scores from the five complexity metrics; (b–f) PCA of shape data with heatmaps indicating the complexity scores for FD box counting, FD madogram, PSD, SI and SCI complexity scores, respectively. Complexity scores for each of the five methods were subdivided into 10 equal bins to define the heatmap colour associations, with red indicating the most complex sutures and blue indicating the least complex sutures.

overall shape variation (PC1, PC2, PC3 and PC4) (electronic supplementary material, table S4). Correlation analysis between these PC scores and the method complexity scores shows that PC2 aligns most strongly with all of the complexity metrics, except SCI (figure 5, electronic supplementary material, table S5). Of the five implemented complexity metrics, PSD has the strongest correlation to PC2 ($r = -0.766$, $p < 0.001$). When the specimen complexity scores are mapped onto the shape morphospace using heatmaps, marked differences across the different metrics relative to suture shape are evident. Both SI and SCI returned relatively few highly complex sutures (figure 5b,c), but these always fell in the negative region of PC1 and PC2. For FD box counting, FD madogram and PSD metrics the most complex sutures clustered in the negative region of PC2, while the least complex sutures appeared at the positive end of PC2. By contrast, there was little association between complexity scores and the shape variation described by PC1 for these three metrics.

Complex sutures were present on either end of PC1 for both FD madogram and PSD, whereas for FD box counting, it was the sutures with lower complexity that appeared at both extremes of the PC1 axis. PCA of the two-dimensional semi-landmarks using heatmaps to indicate complexity (figure 5b–f), support the correlation analyses (figure 5a) that indicated the complexity metrics were most strongly correlated with the shape variation of PC2.

Direct comparisons across the five implemented complexity metrics indicated that there were very few strong correlations ($r > 0.7$) (figure 6a, electronic supplementary material, table S6). The strongest correlation was unsurprisingly between SI and SCI ($r = 0.954$, $p < 0.001$), as SCI is a modification of SI. Interestingly, there was a further strong correlation between the FD madogram method and PSD method ($r = 0.823$, $p < 0.001$). This correlation was much stronger than the correlation between the two FD metrics even though a similar fractal method underlies both of the FD approaches, suggesting

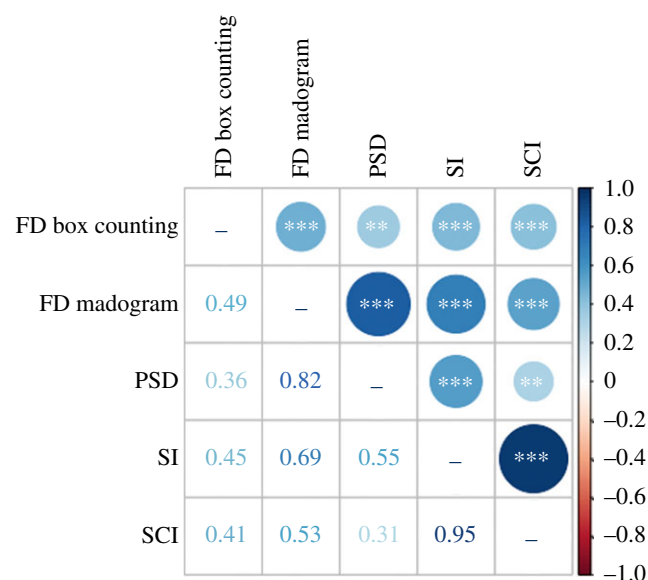


Figure 6. Correlation plot comparing complexity metrics indicating in the lower half the correlation coefficients (r) and in the upper half the strength of the correlation by colour and circle size, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

the two fractal metrics are largely different thus capturing variations in the suture shape complexity. Furthermore, the 95% confidence intervals for these two comparisons (SI versus SCI and FD madogram versus PSD) were small and thus emphasized a strong correlation. Only moderate correlations ($0.3 < r < 0.7$) were identified for all other complexity metric comparisons, with correlation coefficients ranging between 0.314 and 0.689. These moderate correlation coefficients presented with a wider 95% confidence interval than the stronger correlations, emphasizing uncertainty in estimating the correlation coefficient between the majority of metrics. All results were statistically significant ($p < 0.01$) (figure 6b).

PCA performed using the method complexity scores and plotted with each metric as a variable (figure 5) indicated that the five metrics captured a different aspect of complexity, evidenced by the diverging variables which reflect opposite PC loadings on the first components (figure 7a,b). The first three PCs explained 95% of the overall complexity variation (electronic supplementary material, table S7), with the first PC accounting for 65% of the variation (figure 7a). The loadings (i.e. the correlations between the PC scores and the complexity data) of the five metrics on PC1 are all positively related and show that this major axis of variation mostly explains the range of values from the five metrics, with the least complex sutures at the positive aspect and more complex sutures at the negative aspect (figure 7c; electronic supplementary material, figure S4a). However, the loadings on PC2 (17.8%) show contrasting correlation patterns between the metrics on this axis, seen by the diverging variables (figure 7a). Specifically, SI and SCI are loaded together towards the negative aspect of PC2, whereas FD madogram and PSD are associated in the opposite direction of PC2. A strong correlation between the variation of PC2 and the SI, SCI, FD madogram and PSD methods was identified from the four variables approaching the outer circle limit. Of these loadings, PSD and SCI contribute the most to the overall variation in PC2 (electronic supplementary material, figure S4b). By contrast, FD box counting accounted for very little of the variation explained by PC2 but is instead highly

loaded and aligns best with PC3, which explains 13.7% of the overall variation (figure 7b, electronic supplementary material, figure S4c). No other PC axis explained $\geq 5\%$ of the variance in the complexity metrics.

The opposite association between the complexity metrics along PC2, from the PCA of suture complexity (figure 7c), indicate that suture morphological complexity is multi-faceted, whereby each of the various metrics are effective at capturing different characteristics of suture morphological complexity. Individual specimen complexity scores plotted onto this PCA (figure 7c) show the extreme morphological complexities captured by the various metrics. Specimen ID 7 (figure 7h) reflects the extreme morphological complexity of SCI characterized by numerous interdigitations along the suture length with a heightened magnitude and a slight irregular pattern. Specimen ID 31 (figure 7g) highlights the extreme of SI morphological complexity characterised by a regular repeating pattern, with numerous interdigitations of high magnitude. Specimen ID 8 (figure 7f) reflects the extreme morphological complexity of the FD box counting method and is characterized by a regular patterning, with many small interdigitations. Specimen ID 64 (figure 7e) reflects the extreme morphological complexity of the FD madogram method, characterized by an irregular pattern with a heightened number of high magnitude interdigitations captured by FD madogram complexity. Specimen IDs 46 and 77 (figure 7d) represent the extreme morphological complexity of the PSD method, whereby PSD perceives morphological complexity as a largely irregular pattern, with folds of high magnitude and a looping morphology. Furthermore, along PC2 where the methods diverge, the spread of the specimens is positively associated to the FD madogram and PSD method and negatively associated to the SI and SCI methods.

4. Discussion

Sutures play a critical role in the development and functioning of the skull throughout an organism's life [23–26,59–61], and thus suture morphology is important for understanding skull morphology and function. Complexity, in particular, has been discussed as a critical variable for skull function [30], but approaches that quantify complexity and their relationship to suture shape have not been directly compared. We tested and compared five complexity metrics (SI, SCI, FD box counting, FD madogram and PSD) based on different mathematical approaches for quantifying suture complexity across a range of mammalian cranial suture morphologies. While there was some correspondence across methods, they each captured different aspects of suture morphological complexity, with PSD showing a particularly strong correlation with PC2, which captured diverse aspects of complexity represented in this sample dataset. Therefore, studies should consider which metric, or a combination of metrics is most suitable depending on the aspect of suture complexity that is of interest.

4.1. A complexity metric could be more useful than raw shape data for studying sutures

The complexity metrics were most strongly correlated with the shape variation of PC2, rather than PC1 which accounted for approximately 60% of the overall shape variation. The shape

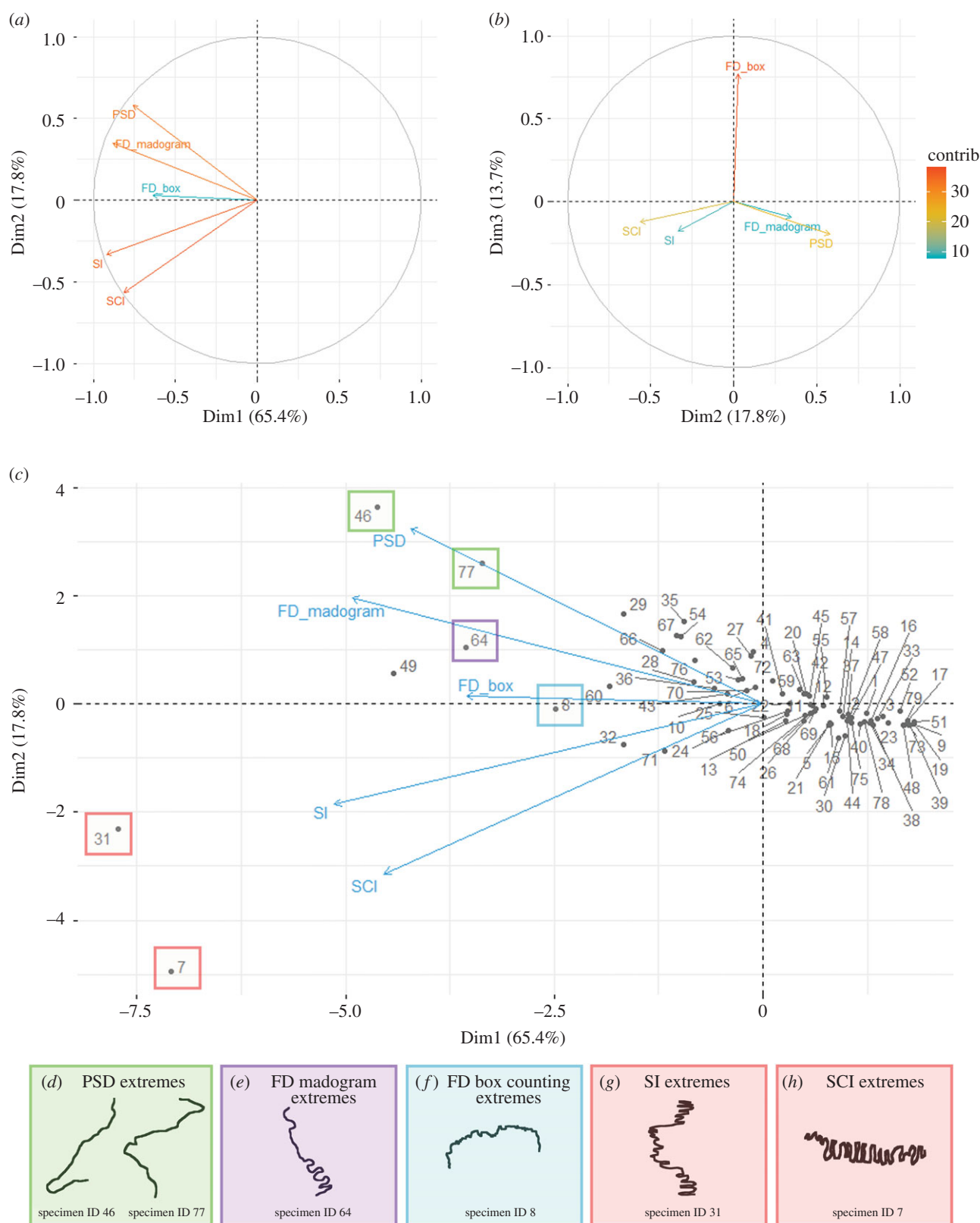


Figure 7. Principal components analysis of suture complexity, with variables reflecting the five complexity metrics (SI, SCI, FD box counting, FD madogram, PSD): (a) variable correlation plot for PC1 and PC2; (b) variable correlation plot for PC2 and PC3; (c) PCA for PC1 and PC2, plotted with the individual specimens (PC scores); (d–h) PSD, FD madogram, FD box counting, SI and SCI morphological complexity extremes, respectively. Specimen ID and their species associations are detailed in electronic supplementary material, table S1.

variation of PC1, from the PCA of two-dimensional semi-landmarks, captured an inversion of shape or mirror image ('n' versus 'u' shaped morphology at the opposite extremes of PC1 on figure 3a) and did not appear to capture functionally relevant variation in complexity, such as looping patterns, straight morphologies and interdigitations. The largest variation in shape complexity, rather than a reversal of the same

shape, was captured by PC2. As nearly all of the complexity metrics correlated most strongly with this axis of variation, it is clear that the complexity metrics are capable of capturing the predominant variation in morphological complexity.

Indeed, the complexity metrics are only extracting part of the suture shape information, but these metrics are most highly correlated with the features of suture shape that have

the highest functional relevance to the suture. This suggests that using the raw shape data alone to understand sutural complexity and morphological variation could be misleading and instead supports the use of a complexity score. The use of a complexity score which provides a single value is a useful tool to make direct comparisons across diverse datasets. Attempts to use complexity scores to make comparisons of suture complexity across different species from various published studies has previously been attempted [30,37]. However, these studies have been performed on various taxa using various methods, thus with little overlap, meaning comparative data are sparse despite the consensus to evaluate complexity based on interdigitations. Therefore, determining a single complexity metric, as our study aimed to do, will prove incredibly useful to make such comparisons. Moreover, the comparisons made by our study of the previously implemented complexity metrics, identify which aspects suture shape each approach captures. This will further prove useful for future studies, to aid in the determination of the most appropriate metric dependent on the biological questions and empirical dataset proposed by the study.

4.2. Different complexity metrics capture different sutural features

Comparing the distributions of complexity scores for the five metrics and their correlations with each other and with suture shape allows us to identify which aspects of suture complexity each approach captures. PSD greatly contributed to the major axes of complexity score variation and therefore describes a very high level of variability in suture complexity across a range of mammal suture morphologies. Moreover, from the PCA analysis of raw shape data, PSD correlated most strongly with PC2, which largely reflected variation in suture complexity rather than a mirroring of the suture shape, as captured by PC1 (figure 3*a*). While the other metrics also correlated strongly with PC2, PSD appeared to be more sensitive to differences in suture shapes, such as looping, at higher end of complexity scores. As a result, we observed that PSD discriminates among disparate morphologies that were scored as equally complex by non-Fourier methods (FD box counting and FD madogram). This difference between PSD and FD approaches, shown here with mammal cranial sutures, supports previous analysis of ammonoid sutures where a difference between SCI and PSD was also identified [29,45,62]. Therefore, PSD may be the most appropriate metric when diverse aspects of complexity are present, as in mammal cranial sutures. It may be unsurprising that our analyses identified PSD as the most appropriate method for quantifying suture complexity, given that the method is based on a windowed STFT approach [29,45]. The Fourier foundation of PSD suggests that this approach is mathematically robust, as the statistical transformations, such as windowing techniques, account for profiles in nature, such as suture morphology, that are discrete and non-stationary [63]. Therefore, PSD provides a mathematical description of the suture morphology, by implementing a Fourier transform approach, that is unique to each suture morphology [45].

Of the two FD methods implemented (box counting and madogram), it was perhaps unsurprising that the madogram method most closely followed the loadings of the PSD variable, given its reported efficiency and robustness over other FD methods, including the box counting and ruler dimension methods [64]. Interestingly, the FD ruler dimension approach

is more commonly applied to suture complexity data [30,37,42] than the FD box counting approach [65], despite it having been argued that such an approach should not be implemented on observational data [63]. Despite its robustness, the FD madogram method has not been previously applied to the suture complexity problem. As FD madogram has been previously identified to be the most robust of the FD methods, this supports our findings of a stronger correlation between FD madogram and PSD, than the FD box counting approach and PSD. However, it was surprising that the stronger correlation was observed between the FD madogram and PSD metrics than between the two FD metrics which are based upon very similar fractal methods.

The FD box counting metric mapped well to both the features reflected by PC2 and PC3 of shape variation. Of the approaches, this metric appears least sensitive to the extreme forms sampled within the dataset thus producing well-behaved results. However, it has been described as less robust and efficient than the FD madogram metric [64]. There is, however, a critical issue in the use of fractal analysis for patterns in nature, such as suture morphology, because natural forms are well documented to be intricate and often do not follow the self-similarity assumptions of fractal analysis [29]. More specifically, serrated sutures generally do not follow self-similarity but can be very intricate. Of the FD approaches, FD madogram has been suggested to align better with the self-similarity assumptions [63] suggesting a potential advantage over the other FD methods.

Regardless of the issues with the FD methods, previous studies still identified FD as a better objective descriptor of suture complexity than SCI [65]. While SCI calculates complexity based on a length ratio measurement (SI), it also includes a subjective and arbitrary measure of interdigitation number and indentation. This likely results in the insensitivity we observed to other aspects of geometric complexity, which is also an issue for the SI approach that provides the foundation for SCI. The resultant effect of a linear length measurement for complexity is that shallower lobes would give similar complexity values to fewer deeper lobes [53] in contrast to PSD. Nevertheless, SCI is effective at determining complexity based on interdigitation number, due to its inclusion of a complexity factor multiplier [53]. This disparity between SCI and PSD was also identified in a previous study [29] comparing SCI and PSD for ammonoid suture morphology.

The lack of very strong correlations among the tested methods indicates that the five methods capture different aspects of suture morphological complexity. Suture morphology is hugely diverse from straight to incredibly intricate morphologies, each with varying complexity scores. Of the features forming suture morphological complexity, previous work has implied that interdigitations are the most biologically relevant feature of a suture form. This is because interdigitations are thought to reflect the biomechanical stress associated with diet [30,33,34,42], extrinsic mechanical factors [40,43,59,66,67] and also suture development [68]. From the PCA analysis of complexity, SI and SCI were most strongly associated with the interdigitated morphologies. For these methods, the more complex sutures were those with a greater number of interdigitations.

Interdigitation complexity is not, however, defined simply as the number of interdigitations. The SI and SCI methods focus on this one aspect and disregard the other features that form these interdigitations, such as amplitude. While SI

and SCI captured well the number of interdigitations, PSD offered the ability to capture both heightened interdigitation number and greater interdigitation amplitude as complex [29]. FD box counting and FD madogram also captured interdigitation magnitude to a greater extent than SI and SCI, as well as capturing interdigitation number. Therefore, while all approaches are able to account for the interdigitations formed as a result of biomechanical stress from external inputs including diet and behaviour [33,40], PSD and both FD metrics likely better capture the details interdigitations are comprised of. As interdigitation is thought to reflect biomechanical stress produced from a number of biological pressures [30,33,34,39–42], it is likely that complexity score reflects skull function. More specifically, sutures exhibiting a higher complexity score are likely to possess a greater ability to absorb stress from external dietary and behavioural inputs. Moreover, we expect that a greater degree of complexity may accumulate with developmental age, as functional pressures on skulls increase. Beyond interdigitations, further aspects of complexity, such as irregularity, are incorporated into the PSD method definition of complexity [29] but have been seldom discussed in previous studies. It is possible that these previously disregarded additional aspects of complexity are biologically important or entirely irrelevant to the suture functioning, but this is currently unknown.

Each metric has strengths and weaknesses relative to different aspects of suture complexity. Therefore, in future studies, the appropriate metric should be determined based on the questions posed. For the sample dataset of mammalian cranial sutures, spanning a range of morphologies, used in our study, we prefer PSD, due to its ability to discriminate between disparate morphologies based on multiple aspects of complexity. However, a combination of metrics could also prove useful, such as PSD and FD box counting, to ensure the range of suture morphologies across the major axes of complexity variation (PC2 and PC3) are captured. Understanding the features of suture complexity captured by each of the methods will assist in making direct comparisons across the literature where various approaches have previously been employed.

While we found PSD to be a promising metric for comparative quantitative analysis of suture morphological complexity, we only compared mammalian taxa, and thus other methods may prove more appropriate for other systems. Nevertheless, we have no reason to believe that the PSD method would be unsuccessful for other taxa, in particular since it has been previously identified as a useful method for ammonoid sutures [29]. Furthermore, this study compared sutures covering a broad range of morphologies and thus may not be as powerful for detecting minor differences in sutures, such as those that may be encountered in an intraspecific dataset for a single suture, depending on the attributes of interest. From our analysis, it appears as though the sampled percentage of the sutural segment plays a role in determining the suture complexity score. Different regions of sutures may display different complexities. For example, sampling a smaller segment of suture 46 (specimen ID 46) would likely generate a lower complexity score. Therefore, it is imperative that future studies consider and are consistent with the scale of sutural complexity analysis by separating out macro-/meso-/micro-scale suture sampling. Moreover, our analysis is currently limited to two-dimensional data, due to the currently available implementations of each method. In many systems, analysis of two-dimensional data closely

approximates that of three-dimensional data [69], but it is not possible to assess this for suture morphology at present. It is hoped that the recent advances in three-dimensional geometric morphometrics [21,70–72] will quickly broaden the quantification of sutures in three dimensions. When this becomes possible, it will be important to compare the results of two-dimensional and three-dimensional studies of suture morphology and assess whether the same methods work well with both approaches, as has occurred in recent years with studies of skull and body shape [73–75]. Given that sutures provide the main joints of the skull, consideration of their three-dimensional complexity could be crucial in understanding their overall functioning. Despite our use of X-ray microtomography and laser surface scanning to produce the dataset, all five complexity metrics could equally be implemented on high-quality (two-dimensional) photographs. When implementing the metrics on two-dimensional datasets, it would be critical to consider the possible effects of parallax. As a result of this two-dimensional implementation, each of the compared complexity metrics can be easily applied to a wide range of datasets and systems covering an array of scientific fields.

Finally, while many of the methods used here have been previously applied to quantifying suture complexity in a limited number of taxa, including caimans, primates, caviomorph rodents and ammonoids [30,65,76], this is the first study to apply and compare these methods across a wide range of extant and extinct mammal species and a breadth of suture morphologies. It is surprising that there have been so few prior quantitative studies of suture morphology in mammals, given the huge diversity of the group in terms of cranial morphology, diet, habitat and life history, all of which are known to generate variety in suture morphology [34,77,78]. Moreover, mammals are extremely well studied in terms of quantification of cranial morphology [79–84], cranial development [85–87] and cranial mechanics [88,89]. Therefore, the work presented here will be instrumental in bringing the quantitative analysis of suture morphology into the already rich field of mammalian cranial evolution. Our future work will apply the results of this study to a large-scale analysis of suture development and evolution across mammals.

5. Conclusion

Suture morphology has received relatively little attention in comparative biological studies, in comparison to the abundance of work focused on cranial morphology. While it has been well recognized that the quantification of cranial morphology with geometric morphometrics can offer invaluable insights into evolutionary patterns, ecomorphology, development, taxonomy and phylogenetics. It is also known that suture development and morphology play active roles in facilitating postnatal cranial development, skull function, feeding and shock absorption during locomotion. Indeed, it is for all these functional reasons that ecomorphological reconstructions, taxonomy and phylogenetic reconstructions have been performed for the skull. Consequently, it is highly likely that shape analysis of sutures can offer important insights for reconstructing the evolution and development of the skull.

In this study, we have provided the first comprehensive comparison of the available methods for quantifying suture morphology. Our analyses found that windowed STFT with

a PSD calculation best captured the multiple aspects of suture complexity observed across diverse mammalian taxa and suture forms. PSD captured all aspects that form interdigitations, which are biologically important as they are strongly associated with high biomechanical stresses. Sutures with high PSD scores likely possess a greater ability to absorb stress from external biological pressures, than those with lower PSD scores. Nevertheless, all the metrics studied here capture some key aspects of sutural complexity, while all providing a univariate score which can be easily implemented for comparisons across sutures and taxa. Therefore, future studies attempting to quantify suture complexity, should first outline which features of suture shape are most relevant to address the questions of the specific study, in order to determine the most appropriate complexity metric. Quantifying suture morphology with a complexity metric in a comparative phylogenetic framework has great potential for bringing the understudied field of suture morphology into the vibrant fields of evolutionary and developmental biology.

Data accessibility. The semi-landmark data analysed during the current study are available in the electronic supplementary .csv file. Computer code and final raw data used during this study are available within a public GitHub repository at [https://github.com/](https://github.com/HeatherEWhite/suture_metrics_comparison)

HeatherEWhite/suture_metrics_comparison [58]. All code and raw data are available to download, complete with MIT licence. Please cite this paper and the Zenodo doi: <https://doi.org/10.5281/zenodo.4037220> when using the data or raw code.

Authors' contributions. H.E.W. and A.G. designed the methods comparison study. H.E.W. undertook the data collection, performed the analyses and wrote the first draft of the manuscript. J.C. assisted with the PCA of complexity metrics. All authors provided insightful discussion and ideas, interpreted analyses and results, contributed to the writing of the manuscript, and read and approved the final version.

Competing interests. We declare we have no competing interests.

Funding. H.E.W. was supported by the Biotechnology and Biosciences Research Council through the London Interdisciplinary Doctoral Training Programme (BBSRC LIDo DTP) (grant no. BB/ M009513/1). J.C. was supported by Marie-Sklodowska-Curie Individual Fellowship (IF 797373-EVOTOOLS). A.S.T. was supported by a Wellcome Trust Investigator Awards (102889/Z/13/Z). A.G. was supported by a European Research Council grant STG-2014-637171.

Acknowledgements. We would like to thank Anne-Claire Fabre, Ryan Felice, Marcela Randau, Carla Bardua, Andrew Knapp, Ellen Coombs, João Leite and Eve Noirault (members of the Goswami Lab) who offered invaluable discussions throughout the process of this study. We would particularly like to thank Eve Noirault, who worked tirelessly to scan and reconstruct the scans used by this study. Finally, we would like to thank the anonymous reviewers for their thoughtful feedback and comments.

References

- Swiderski D, Zelditch M, Fink W. 2000 Phylogenetic analysis of skull shape evolution in marmotine squirrels using landmarks and thin-plate splines. *Hystrix: It. J. Mamm* **11**. (doi:10.4404/hystrix-11.1-4136)
- Cucchi T, Fujita M, Dobney K. 2008 New insights into pig taxonomy, domestication and human dispersal in Island South East Asia: molar shape analysis of *Sus* remains from Niah Caves. *Sarawak. Int. J. Osteoarchaeol.* **19**, 508–530. (doi:10.1002/oa.974)
- Figueirido B, Palmqvist P, Pérez-Claros JA. 2008 Ecomorphological correlates of craniodental variation in bears and paleobiological implications for extinct taxa: an approach based on geometric morphometrics. *J. Zool.* **277**, 70–80. (doi:10.1111/j.1469-7998.2008.00511.x)
- Herrel A, Fabre A-C, Hugot J-P, Keovichit K, Adriaens D, Brabant L, Van Hoorebeke L, Cornette R. 2012 Ontogeny of the cranial system in *Laonastes aenigmamus*. *J. Anat.* **221**, 128–137. (doi:10.1111/j.1469-7580.2012.01519.x)
- Klingenberg CP, Marugán-Lobón J. 2013 Evolutionary covariation in geometric morphometric data: analyzing integration, modularity, and allometry in a phylogenetic context. *Syst. Biol.* **62**, 591–610. (doi:10.1093/sysbio/syt025)
- Goswami A, Smaers JB, Soligo C, Polly PD. 2014 The macroevolutionary consequences of phenotypic integration: from development to deep time. *Phil. Trans. R. Soc. B* **369**, 20130254. (doi:10.1098/rstb.2013.0254)
- Mayer C, Metscher BD, Müller GB, Mitteroecker P. 2014 Studying developmental variation with geometric morphometric image analysis (GMIA). *PLoS ONE* **9**, e115076. (doi:10.1371/journal.pone.0115076)
- Dumont M, Wall CE, Botton-Divet L, Goswami A, Peigne S, Fabre AC. 2015 Do functional demands associated with locomotor habitat, diet, and activity pattern drive skull shape evolution in musteloid carnivores? *Biol. J. Linn. Soc.* **117**, 858–878. (doi:10.1111/bij.12719)
- Wroczyn C, Neubauer TA, Meyer J, Piller WE. 2016 Shape variation in neotropical *Cytheridella* (Ostracoda) using semilandmarks-based geometric morphometrics: a methodological approach and possible biogeographical implications. *PLoS ONE*, e0168438. (doi:10.1371/journal.pone.0168438)
- Larouche O, Zelditch ML, Cloutier R. 2018 Modularity promotes morphological divergence in ray-finned fishes. *Sci. Rep.* **8**, 7278. (doi:10.1038/s41598-018-25715-y)
- Stayton CT. 2005 Morphological evolution of the lizard skull: a geometric morphometrics survey. *J. Morphol.* **263**, 47–59. (doi:10.1002/jmor.10288)
- Pierce SE, Angielczyk KD, Rayfield EJ. 2009 Shape and mechanics in thalattosuchian (Crocodylomorpha) skulls: implications for feeding behaviour and niche partitioning. *J. Anat.* **215**, 555–576. (doi:10.1111/j.1469-7580.2009.01137.x)
- Brusatte SL, Sakamoto M, Montanari S, Harcourt Smith WEH. 2011 The evolution of cranial form and function in theropod dinosaurs: insights from geometric morphometrics. *J. Evol. Biol.* **25**, 365–377. (doi:10.1111/j.1420-9101.2011.02427.x)
- Foth C, Brusatte SL, Butler RJ. 2012 Do different disparity proxies converge on a common signal? Insights from the cranial morphometrics and evolutionary history of Pterosauria (Diapsida: Archosauria). *J. Evol. Biol.* **25**, 904–915. (doi:10.1111/j.1420-9101.2012.02479.x)
- Fabre AC, Cornette R, Huyghe K, Andrade DV, Herrel A. 2014 Linear versus geometric morphometric approaches for the analysis of head shape dimorphism in lizards. *J. Morphol.* **275**, 1016–1026. (doi:10.1002/jmor.20278)
- Lu X, Ge D, Xia L, Huang C, Yang Q. 2014 Geometric morphometric study of the skull shape diversification in Sciuridae (Mammalia, Rodentia). *Integr. Zool.* **9**, 231–245. (doi:10.1111/1749-4877.12035)
- Parr WCH, Wilson LAB, Wroe S, Colman NJ, Crowther MS, Letnic M. 2016 Cranial shape and the modularity of hybridization in dingoes and dogs; hybridization does not spell the end for native morphology. *Evol. Biol.* **43**, 171–187. (doi:10.1007/s11692-016-9371-x)
- McCurry MR, Evans AR, Fitzgerald EMG, Adams JW, Clausen PD, McHenry CR. 2017 The remarkable convergence of skull shape in crocodylians and toothed whales. *Proc. R. Soc. B* **284**, 20162348. (doi:10.1098/rspb.2016.2348)
- Heck L, Wilson LAB, Evin A, Stange M, Sánchez-Villagra MR. 2018 Shape variation and modularity of skull and teeth in domesticated horses and wild equids. *Front. Zool.* **15**, 14. (doi:10.1186/s12983-018-0258-9)
- Felice RN, Goswami A. 2018 Developmental origins of mosaic evolution in the avian cranium. *PNAS* **115**, 555–560. (doi:10.1073/pnas.1716437115)
- Bardua C, Wilkinson M, Gower DJ, Sherratt E, Goswami A. 2019 Morphological evolution and modularity of the caecilian skull. *BMC Evol. Biol.* **19**, 30. (doi:10.1186/s12862-018-1342-7)

22. Evans K, Vidal-Garcia M, Tagliacollo V, Taylor S, Fenolio D. 2019 Bony patchwork: mosaic patterns of evolution in the skull of electric fishes (Apteronotidae: Gymnotiformes). *Integr. Comp. Biol.* **59**, 420–431. (doi:10.1093/icb/icz026)
23. Opperman LA. 2000 Cranial sutures as intramembranous bone growth sites. *Dev. Dyn.* **219**, 472–485. (doi:10.1002/1097-0177(2000)9999:9999<:AID-DVDY1073>3.0.CO;2-F)
24. Herrel A, Aerts P, De Vree F. 2000 Cranial kinesis in geckoes: functional implications. *J. Exp. Biol.* **203**, 1415–1423.
25. Curtis N, Jones MEH, Evans SE, O'Higgins P, Fagan MJ. 2013 Cranial sutures work collectively to distribute strain throughout the reptile skull. *J. R. Soc. Interface* **10**, 20 130 442. (doi:10.1098/rsif.2013.0442)
26. Jin S-W, Sim K-B, Kim S-D. 2016 Development and growth of the normal cranial vault: an embryologic review. *J. Korean Neurosurg. Soc.* **59**, 192–196. (doi:0.3340/jkns.2016.59.3.192)
27. Toussaint N, Redhead Y, Liu W, Fisher EMC, Hallgrímsson B, Tsubulewicz VJ, Schnabel JA, Green JBA. 2019 A landmark-free morphometrics pipeline: application to characterise a Down syndrome model skull phenotype. *BioRxiv*. (doi:10.1101/711259)
28. Canfield DJ, Anstey RL. 1981 Harmonic analysis of cephalopod suture patterns. *Math. Geosci.* **13**, 23–35. (doi:10.1007/BF01032007)
29. Allen EG. 2006 New approaches to Fourier analysis of ammonoid sutures and other complex, open curves. *Paleobiology* **32**, 299–315. (doi:10.1666/04042.1)
30. Monteiro LR, Lessa LG. 2000 Comparative analysis of cranial suture complexity in the genus *Caiman* (Crocodylia. Alligatoridae). *Rev. Brasil Biol.* **60**, 689–694. (doi:10.1590/S0034-7108200000400021)
31. Curtis N, Witzel U, Fagan MJ. 2014 Development and three-dimensional morphology of the zygomaticotemporal suture in primate skulls. *Folia Primatol.* **85**, 77–87. (doi:10.1159/000357526)
32. Jaslow CR. 1989 Sexual dimorphism of cranial suture complexity in wild sheep (*Ovis orientalis*). *Zool. J. Linn. Soc.* **95**, 273–284. (doi:10.1111/j.1096-3642.1989.tb02312.x)
33. Byron CD. 2009 Cranial suture morphology and its relationship to diet in *Cebus*. *J. Hum. Evol.* **57**, 649–655. (doi:10.1016/j.jhevol.2008.11.006)
34. Buezas G, Becerra F, Vassallo A. 2017 Cranial suture complexity in caviomorph rodents (Rodentia; Ctenohystica). *J. Morphol.* **278**, 1125–1136. (doi:10.1002/jmor.20699)
35. Schopf TJM. 1976 Environmental versus genetic causes of morphologic variability in bryozoan colonies from the deep sea. *Paleobiology* **2**, 156–165. (doi:10.1017/S0094837300003432)
36. Henderson JH, Chang LY, Song HM, Longaker MT, Carter DR. 2005 Age-dependent properties and quasi-static strain in the rat sagittal suture. *J. Biomech.* **38**, 2294–2301. (doi:10.1016/j.jbiomech.2004.07.037)
37. Nicolay CW, Vaders MJ. 2006 Cranial suture complexity in white-tailed deer (*Odocoileus virginianus*). *J. Morphol.* **267**, 841–849. (doi:10.1002/jmor.10445)
38. Wood C. 2015 The age-related emergence of cranial morphological variation. *Forensic Sci. Int.* **251**, e1–220.e20. (doi:10.1016/j.forsciint.2015.03.030)
39. Herring SW. 1972 Sutures — a tool in functional cranial analysis. *Acta. Anat.* **83**, 222–247. (doi:10.1159/000143860)
40. Jaslow CR, Biewener AA. 1995 Strain patterns in the horncores, cranial bones and sutures of goats (*Capra hircus*) during impact loading. *J. Zool.* **235**, 193–210. (doi:10.1111/j.1469-7998.1995.tb05137.x)
41. Herring SW, Teng S. 2000 Strain in the braincase and its sutures during function. *Am. J. Phys. Anthropol.* **112**, 575–593. (doi:10.1002/1096-8644(200008)112:4<575::AID-AJPA10>3.0.CO;2-0)
42. Byron CD, Borke J, Yu J, Pashley D, Wingard CJ, Hamrick M. 2004 Effects of increased muscle mass on sagittal suture morphology and mechanics. *Anat. Rec.* **279A**, 676–684. (doi:10.1002/ar.a.20055)
43. Rafferty KL, Herring SW. 1999 Craniofacial sutures: morphology, growth, and *in vivo* masticatory strains. *J. Morphol.* **242**, 167–179. (doi:10.1002/(SICI)1097-4687(199911)242:2<167::AID-JMOR8>3.0.CO;2-1)
44. Boyajian G, Lutz T. 1992 Evolution of biological complexity and its relation to taxonomic longevity in the Ammonoidea. *Geology* **20**, 983–986.
45. Gildner RF. 2003 A Fourier method to describe and compare suture patterns. *Palaeontol. Electron.* **6**, 1–12.
46. Jones KE, Safi K. 2011 Ecology and evolution of mammalian biodiversity. *Philos. Trans. R. Soc. Lond. B* **366**, 2451–2461. (doi:10.1098/rstb.2011.0090)
47. Adler D, Murdoch M. 2019 rgl: 3D Visualization Using OpenGL. R package version 0.100.26 <https://cran.r-project.org/web/packages/rgl/index.html>.
48. Olsen A, Haber A. 2018 StereoMorph: Stereo Camera Calibration and Reconstruction. R package version 1.6.2. <https://cran.r-project.org/web/packages/StereoMorph/index.html>.
49. Adams DC, Collyer ML, Kaliontzopoulou A. 2019 geomorph: Geometric Morphometric Analyses of 2D/3D Landmark Data. R package version 3.1.3. <https://cran.r-project.org/web/packages/geomorph/index.html>.
50. Adams DC, Otárola-Castillo E. 2013 Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393–399. (doi:10.1111/2041-210X.12035)
51. Westerman GEG. 1971 Form, structure and function of shell and siphuncle in coiled Mesozoic ammonoids. *Life Sci. Contrib. R. Ont. Mus.* **78**, 1–39.
52. Anton SC, Jaslow CR, Swartz SM. 1992 Sutural complexity in artificially deformed human (*Homo sapiens*) crania. *J. Morphol.* **214**, 321–332. (doi:10.1002/jmor.1052140307)
53. Saunders WB. 1995 The ammonoid suture problem: relationships between shell and septum thickness and suture complexity in Paleozoic ammonoids. *Paleobiology* **21**, 343–355. (doi:10.1017/S0094837300013348)
54. Sevcikova H, Percival D, Gneiting T. 2014 fractaldim: Estimation of fractal dimensions. R package version 0.8-4 <https://cran.r-project.org/web/packages/fractaldim/index.html>.
55. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, Chang C-C, Lin C-C. 2019 e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7.2 See <https://cran.r-project.org/web/packages/e1071/index.html>.
56. Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla J. 2017 corplot: Visualization of a Correlation Matrix. R package version 0.84 <https://cran.r-project.org/web/packages/corplot/index.html>.
57. Kassambara A, Mundt F. 2017 factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package 1.0.5. See <https://cran.r-project.org/web/packages/factoextra/index.html>.
58. White HE, Clavel J, Tucker AS, Goswami A. 2020 A comparison of metrics for quantifying cranial suture complexity. *Github*: https://github.com/HeatherWhite/suture_metrics_comparison. Accessed 18 September 2020. Release for publication. *Zenodo*: <https://doi.org/10.5281/zenodo.4037220>.
59. Jaslow CR. 1990 Mechanical properties of cranial sutures. *J. Biomech.* **23**, 313–321. (doi:10.1016/0021-9290(90)90059-c)
60. Markey MJ, Main RP, Marshall CR. 2006 *In vivo* cranial suture function and suture morphology in the extant fish *Polypterus*: implications for inferring skull function in living and fossil fish. *J. Exp. Biol.* **209**, 2085–2102. (doi:10.1242/jeb.02266)
61. Roston RA, Roth VL. 2019 Cetacean skull telescoping brings evolution of cranial sutures into focus. *Anat. Rec.* **302**, 1055–1073. (doi:10.1002/ar.24079)
62. Lutz TM, Boyajian GE. 1995 Fractal geometry of ammonoid sutures. *Paleobiology* **21**, 329–342. (doi:10.1017/S0094837300013336)
63. Fox CG. 1989 Empirically derived relationships between fractal dimension and power law form frequency spectra. *Pure App. Geophys.* **131**, 211–239. (doi:10.1007/978-3-0348-6389-6_12)
64. Gneiting T, Sevcikova H, Percival DB. 2012 Estimators of fractal dimension: assessing the smoothness of time series and spatial data. *Stat. Sci.* **27**, 247–277. (doi:10.1214/11-STS370)
65. Skrzat J, Walocha J. 2003 Application of fractal dimension in evaluation of cranial suture complexity. *Harmonic Fractal Image Anal.* **2003**, 39–41.
66. Daniel TL, Helmuth BS, Saunders WB, Ward PD. 1997 Septal complexity in ammonoid cephalopods increased mechanical risk and limited depth. *Paleobiology* **23**, 470–481. (doi:10.1017/S0094837300019849)
67. Olóriz F, Palmqvist P, Pérez-Claros JA. 2002 Morphostructural constraints and phylogenetic overprint on sutural frilling in Late Jurassic ammonites. *Lethaia* **35**, 158–168. (doi:10.1111/j.1502-3931.2002.tb00076.x)

68. Byron CD. 2006 Role of the osteoclast in cranial suture waveform patterning. *Anat. Rec. A Discov. Mol. Cell Evol. Biol.* **288**, 552–563. (doi:10.1002/ar.a.20322)
69. Cardini A. 2014 Missing the third dimension in geometric morphometrics: how to assess if 2D images really are a good proxy for 3D structures? *Hystrix: It. J. Mamm* **25**, 73–81. (doi:10.4404/hystrix-25.2-10993)
70. Bookstein FL. 1997 Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Image Anal.* **1**, 225–243. (doi:10.1016/s1361-8415(97)85012-8)
71. Gunz P, Mitteroecker P, Bookstein FL. 2005 Semilandmarks in three dimensions. In *Modern morphometrics in physical anthropology* (ed. DE Slice), New York, NY: Kluwer Academic/Plenum.
72. Goswami A, Watanabe A, Felice RN, Bardua C, Fabre A-C, Polly PD. 2019 High-density morphometric analysis of shape and integration: the good, the bad, and the not-really-a-problem. *Integrat. Comp. Biol.* **59**, 669–683. (doi:10.1093/icb/icz120)
73. Breno M, Leirs H, Van Dongen S. 2011 Traditional and geometric morphometrics for studying skull morphology during growth in *Mastomys natalensis* (Rodentia: Muridae). *J. Mammal.* **92**, 1395–1406. (doi:10.1644/10-MAMM-A-331.1)
74. Buser TJ, Sidlauskas BL, Summers AP. 2018 2D or Not 2D? Testing the utility of 2D vs. 3D landmark data in geometric morphometrics of the sculpin subfamily Oligocottinae (Pisces; Cottoidea). *Anat. Rec. (Hoboken)* **301**, 806–818. (doi:10.1002/ar.23752)
75. Gabelaia M, Tarkhishvili D, Adriaens D. 2018 Use of three-dimensional geometric morphometrics for the identification of closely related species of Caucasian rock lizards (Lacertidae: Darevskia). *Biol. J. Linn. Soc.* **125**, 709–717. (doi:10.1093/biolinnean/bly143)
76. Gildner RF, Ackerly S. 1985 A Fourier technique for studying ammonoid sutures. *Geol. Soc. Am. Abstr. Prog.* **17**, 592.
77. Byron CD, Hamrick M, Yu J. 2006 Cranial suture morphology: understanding how dietary strategy and brain size influence primate craniofacial bone growth (abstract). *Am. J. Phys. Anthropol. Suppl* **42**, 71.
78. Rager L, Hautier L, Forasiepi A, Goswami A, Sánchez-Villagra MR. 2014 Timing of cranial suture closure in placental mammals: phylogenetic patterns, intraspecific variation, and comparison with marsupials. *J. Morphol.* **275**, 125–140. (doi:10.1002/jmor.20203)
79. Goswami A. 2006 Cranial modularity shifts during mammalian evolution. *Am. Nat.* **168**, 270–280. (doi:10.1086/505758)
80. Goswami A. 2006 Morphological integration in the carnivoran skull. *Evol.* **60**, 169–183. (doi:10.1111/j.0014-3820.2006.tb01091.x)
81. Goswami A. 2007 Phylogeny, diet and cranial integration in australodelphian marsupials. *PLoS ONE* **2**, e995. (doi:10.1371/journal.pone.0000995)
82. Cooper N, Purvis A. 2009 What factors shape rates of phenotypic evolution? A comparative study of cranial morphology of four mammalian clades. *J. Evol. Biol.* **22**, 1024–1035. (doi:10.1111/j.1420-9101.2009.01714.x)
83. Hautier L, Weisbecker V, Goswami A, Knight F, Kardjilov N, Asher RJ. 2011 Skeletal ossification and sequence heterochrony in xenarthran evolution. *Evol. Dev.* **13**, 460–476. (doi:10.1111/j.1525-142X.2011.00503.x)
84. Cardini A, Polly D. 2013 Larger mammals have longer faces because of size-related constraints on skull form. *Nat. Commun.* **4**, 2458. (doi:10.1038/s41467-018-07156-3)
85. Chai Y, Jiang X, Ito Y, Bringas P, Han J, Rowitch DH, Soriano P, McMahon AP, Sucov HM. 2000 Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127**, 1671–1679.
86. Sánchez-Villagra MR, Goswami A, Weisbecker V, Mock KS. 2008 Conserved relative timing of cranial ossification patterns in early mammalian evolution. *Evol. Dev.* **10**, 519–530. (doi:10.1111/j.1525-142X.2008.00267.x)
87. Koyabu D *et al.* 2014 Mammalian skull heterochrony reveals modular evolution and a link between cranial development and brain size. *Nat. Commun.* **5**, 3625. (doi:10.1038/ncomms4625)
88. Dumont ER, Piccirillo J, Grosse IR. 2005 Finite-element analysis of biting behavior and bone stress in the facial skeletons of bats. *Anat. Rec. A Discov. Mol. Cell Evol. Biol.* **283**, 319–330. (doi:10.1002/ar.a.20165)
89. Sharp AC. 2015 Comparative finite element analysis of the cranial performance of four herbivorous marsupials. *J. Morphol.* **276**, 1230–1243. (doi:10.1002/jmor.20414)