Evolutionary biomechanics: hard tissues and soft evidence?

Sarah Broyde[1], Matthew Dempsey[1], Linjie Wang[4], Philip G. Cox[2,3], Michael Fagan[4] & Karl T. Bates[1*].

[1]Department of Musculoskeletal Biology, Institute of Aging and Chronic Disease, University of Liverpool, The William Henry Duncan Building, 6 West Derby Street, Liverpool L7 8TX, UK;

[2]Department of Archaeology, University of York, PalaeoHub, Wentworth Way, Heslington, York YO10 5DD, UK

[3]Hull York Medical School, University of York, PalaeoHub, Wentworth Way, Heslington, York YO10 5DD, UK

[4]Department of Engineering, University of Hull, Hull, HU6 7RX, UK.

*Correspondence to: k.t.bates@liverpool.ac.uk.

Key words: macroevolution, biomechanics, multi-body dynamics, finite element analysis, rodent mastication

SUMMARY

Biomechanical modelling is considered a powerful tool for quantifying the evolution of functional performance in extinct animals to understand key anatomical innovations and selective pressures driving major evolutionary radiations. However, the fossil record is composed predominantly of hard parts, forcing palaeontologists to estimate or subjectively reconstruct soft tissue properties in such models. Rarely are these reconstruction approaches validated on extant animals, despite soft tissue properties being highly determinant of functional performance. The extent to which soft tissue reconstructions and biomechanical models accurately predict quantitative or even qualitative patterns in macroevolutionary studies is therefore unknown. Here, we modelled the masticatory system in extant rodents to objectively test the ability of current soft tissue reconstruction methods correctly identify quantitative and qualitative differences between macroevolutionary morphotypes. Baseline models generated using measured soft tissue properties yielded differences in muscle proportions, bite force and bone stress expected between extant sciuromorph, myomorph and hystricomorph rodents. However, predictions from models generated using reconstruction methods typically used in fossil studies varied widely from high levels of quantitative accuracy to a failure to correctly capture even relative differences between macroevolutionary morphotypes. Our novel experiment emphasises that correctly reconstructing even qualitative differences between taxa in a macroevolutionary radiation is challenging using current methods. Future studies of fossil taxa should incorporate systematic assessments of reconstruction error into their hypothesis testing and, moreover, seek to expand primary data sets on muscle properties in extant taxa to better inform soft tissue reconstructions in macroevolutionary studies.

1. Introduction

Changes in functional morphology or biomechanics have fundamentally underpinned some of the most significant evolutionary transitions in the history of life. Colonisation of the land by the earliest tetrapods [1-2], mammalian origins and diversification [3-6], the evolution of locomotion in dinosaurs and birds [7-24] and functional and ecological shifts in human ancestors [25-32] represent some extensively studied examples. These evolutionary events, and the anatomical adaptations associated with them, are central to understanding major adaptive radiations in earth history and the interplay of biological evolution with other aspects of the earth system (e.g. climate, tectonics). The last two decades has seen widespread adoption of sophisticated mathematical-computational approaches, such as finite element and multi-body dynamics analysis, to study functional morphology in extinct animals and the biomechanics of evolutionary transitions recorded in the fossil record. These approaches realise a number of benefits relative to more traditional comparative (qualitative) approaches [33-34], but perhaps most importantly they are able to deliver absolute measures of function and performance in fossil animals (e.g. energy costs, maximal performance) to quantitatively test hypotheses about how anatomical innovations enabled major behavioural or niche adaptions over geological time.

Mathematical-biomechanical approaches yield quantitative predictions of animal performance by combining general models of Newtonian physics and solid mechanics with mathematical descriptions of tissue behaviour and physiology. In doing so they incorporate all the major causative anatomical and physiological factors that underpin mechanical function, and in living animals these approaches have been shown to deliver accurate predictions of metabolic energy costs in walking [e.g. 26], maximal locomotor [e.g. 9, 14, 22] and bite performance [e.g. 35-36] among other parameters. However, in living animals most, if not all, anatomical and physiological input parameters required to build biomechanical models can be measured from the species under study. One challenging aspect in their use on extinct animals is that they require precise specification of

numerical values for soft tissue parameters that are rarely, or never, preserved in fossils. Biomechanical modelling studies of extinct animals have subsequently employed a diverse range of approaches to estimate absolute values for soft tissue parameters in fossil organisms, ranging from standardised properties based on estimated mean values for all living taxa, scaling values from supposed analogous extant animals, and computer-aided design approaches to reconstruct the size and geometry of soft tissues directly in the fossil themselves. Sensitivity analyses have been carried out in small number of these studies and have consistently shown that large errors in soft tissue parameters will lead to significant inaccuracy in function or performance predictions [12-14, 20, 22, 36-37]. However, it remains qualitatively and quantitatively uncertain what the likely error magnitudes are for such soft tissue reconstructions: in other words, it is unclear whether or not the uncertainty surrounding soft tissue parameters is yielding such significant errors that biomechanical studies lack the resolution required to accurately reconstruct functional consequence of anatomical change and test hypotheses about macroevolutionary radiations observed in the fossil record.

In this study we take the most direct and comprehensive approach to-date to assess how inaccuracy and imprecision in soft tissue reconstruction currently impact upon our ability to identify quantitative and even qualitative differences between extinct taxa, and therefore our ability to recognize adaptive trends and evolutionary changes in the fossil record. To do this we first carry out multiple types of biomechanical modelling on extant taxa that are known to exhibit quantitative and qualitative functional differences using real (measured) soft tissue data. Subsequently we repeat this multi-modal biomechanical analysis by substituting real (measured) soft tissues properties with values derived from reconstructive methods typically used on fossil animals. Comparing the functional predictions generated using ‘real’ versus reconstructed soft tissue data not only allows us to examine inaccuracy and imprecision quantitatively, but perhaps more fundamentally allows us to examine if known qualitative differences between extant taxa are preserved by current soft tissue

reconstruction methods. Quantitative error is expected and perhaps will always be unavoidable in fossil animals, but the ability to reliably identify qualitative differences between extinct taxa is fundamental to evolutionary studies that seek to identify adaptations or trends across fossil lineages and major evolutionary transitions in the history of life [1-69]. Prior to this study this fundamental premise, underpinning an entire field of research [1-69], has not been extensively tested.

2. Material and Methods

(a) Case study: evolutionary biomechanics of the rodent masticatory system

Masticatory biomechanics in rodents is an area of study that has received a considerable amount of attention (reviewed in [70]) and one that provides a useful opportunity for addressing the issues raised above. The Rodentia is the largest order of extant mammals, comprising over 2,500 living species [71]. Despite this diversity, almost all rodents can be assigned to one of three groups based on the morphology of their masticatory musculature, specifically the masseteric complex. These three morphotypes are all thought to be derivations of the ancestral morphology (present in a single living species, the mountain beaver), and are referred to as the 'sciuromorph' (squirrel-like), 'myomorph' (mouse-like) and 'hystricomorph' (porcupine-like) conditions [72]. Each of these derived morphotypes represents an extension of the masseter on to the rostrum: in sciuromorph species, the lateral masseter originates from an expanded zygomatic plate; in hystricomorphs, the zygomatico-mandibularis (ZM) extends through the orbit and an enlarged infraorbital foramen; and myomorphs show a combination of both the sciuromorphous and hystricomorphous conditions [73-74]. Furthermore, each of these configurations of the musculature is associated with a characteristic cranial morphology (e.g. presence of the zygomatic plate, size of the infraorbital foramen), allowing recognition of the morphotypes in fossil rodents as well. It has long been recognised that the rodent muscle morphotypes do not represent monophyletic clades [75]. Each muscle arrangement has evolved at least twice independently within the rodents, and previous analyses have indicated that

each conveys different functional capabilities i.e., sciuromorphy enables efficient gnawing at the incisors, hystricomorphy leads to efficient molar chewing, and myomorphy provides greatest efficiency at both feeding modes [76-77] Thus, the rodents are an ideal case study for testing the accuracy with which muscle anatomy can be estimated from skeletal morphology, and the impact of such estimations on inferences of function.

Detailed biomechanical analyses of the rodent masticatory system were previously undertaken by [77-78] who conducted finite element analysis (FEA) on the skulls of the eastern grey squirrel (*Sciurus carolinensis*), the brown rat (*Rattus norvegicus*), and the domesticated guinea pig (*Cavia porcellus*) representing the sciuromorph, myomorph and hystricomorph conditions, respectively. The benchmark ('measured' or 'extant') input data for the current study was provided by these earlier studies, including the 3D reconstructions of the skull and mandible of each species from microCT scans, the material properties of the bone and teeth (determined by nano-indentation), and data on the masticatory muscles. Volume reconstructions of each muscle were generated from diceCT scans of the squirrel, rat and guinea pig [74] Muscle physiological cross sectional areas (PCSAs) were calculated by dividing each muscle volume by the average fibre length (Tables S1-3).

(b) Quantitative soft tissue reconstructions

Our soft tissue reconstructions focus on two critical parameters that govern muscle force generation capacity and subsequently play a highly determinate role in bite force magnitudes and the magnitude and distribution of stress/strain in the skull: muscle mass (or volume) and fibre length (FL). Under static maximal biting conditions typically analysed in fossil taxa, muscle force is calculated according to:

Eq. 1. Muscle force = physiological cross-sectional area (PCSA) * maximum isometric stress

With muscle mass (or volume) and FL determining PCSA in parallel-fibred muscles according to:

Eq. 2. Muscle PCSA = muscle volume/muscle FL

And in pennate muscles according to:

Eq. 3. PCSA = (muscle volume/muscle fibre length) * COS(pennation angle)

A number of independent studies of masticatory performance and evolution in extinct animals have used computer-based approaches to reconstruct the volume of masticatory muscles around and within 3D digital models of fossil skulls for the purposes of calculating PCSA and ultimately bite force [e.g. 35-36, 57, 68]. Similar approaches have been used with limb muscles in studies seeking to constrain locomotor performance in exemplar extinct species [e.g. 17, 23, 68] or reconstruct postural evolution through fossilized evolutionary lineages [24]. Only a small number of these studies have attempted to assess the accuracy of these approaches on living animals and found varied degrees of precision (e.g. 4-22% error relative to measured values in the same extant species [17, 23, 36, 68]). Furthermore, independent studies carried out by different teams using identical methods of reconstruction have produced highly disparate estimates of muscle volumes for the same fossil specimens (e.g. total masticatory muscle volume differing by 41% in [35-36]). However, the source of inaccuracy and discrepancies between studies, and their impact on our ability to the evolution of performance metrics like bite force, have not yet been assessed.

Here we developed a protocol for muscle volume sculpture (Fig. 1) based on methods used in previous fossil studies [e.g. 35-36, 57, 68]. This protocol was formalised in an instruction sheet (see ESM1), which outlined the specific modelling approach to be used and anatomical diagrams on

which to base the 3D muscle sculptures around 3D bone models, which are similar to those used in qualitative muscle reconstructions of fossils. As noted above, previous application of similar methods to the same fossil specimens by independent research teams have produced highly disparate muscle volumes (see discussion [37]). We therefore conducted the first analysis of inter-investigator variability in muscle volume sculpture, with three of the authors independently generating muscle volumes in all three rodent models following only the instruction sheet (ESM1). This analysis provides the first quantitative insight into the potential for investigator subjectivity in soft tissue reconstruction to lead to disparate interpretations of functional evolution across evolutionary transitions (Tables S4-6). A brief discussion of investigator expertise and experience is provided in the ESM.

Different approaches to muscle FL estimation has also led to highly disparate functional predictions in extinct animals [37]. A recent review highlighted the relative paucity of masticatory muscle architecture data relative to other body regions, and suggested that combining such data with information on maximal range of motion and muscle length change in-vivo might provide statistical basis for muscle FL estimation in the masticatory muscles of fossil forms [37], as has been attempted based on small data sets in locomotion studies [19, 21]. However, in the absence of such data, we utilised several approaches used in a recent study [37], which cover different scenarios or assumptions about the nature of muscle architecture in the extinct group under analysis. First, we generated FLs for each muscle under the assumption that all muscles were non-pennate (i.e. parallel fibred), and that FLs were equal to muscle length (measured as the distance between the centroids of the origins and insertions in the 3D models derived from diceCT scans [70] In this scenario, the PCSAs of all muscles are calculated according to Eq. 2 (see above). For each investigator, these models are referred to as iteration A. Second, we generated an iteration of models which differed only in their specification of the medial pterygoid muscle. This muscle consistently shows a pennate architecture in rodents [70] and in the three taxa studied here average measured pennation angles

range from 20-25 degrees (Tables S1-3). Our second iteration of the models therefore represented the medial pterygoid muscle with a pennation angle of 25 degrees in all three taxa with calculated PCSA for this muscle according to Eq. 3. The average ratio of measured FL to muscle length across the three taxa was used to calculate the FLs for the medial pterygoids in this iteration (hereafter referred to as iteration B). Finally, we generated a third iteration of possible FLs and PCSAs, which are considered to be maximal reasonable deviations from the first iteration (iteration A). In this third iteration, all muscles were modelled as pennate, with a pennation angle of 25 degrees, the maximum value measured in these three rodents. The average ratio of measured FL to muscle length in each muscle across the three taxa was used to calculate the FLs for all muscles and subsequently PCSA (using Eq. 3) for this iteration (hereafter referred to as iteration C). While this might be considered an extreme deviation for the known muscle architecture of the three rodents under study, we argue this approach is important for three reasons. First, it must be acknowledged that in fossil taxa the precise values for architectural parameters are completely unknown and therefore assuming a high degree of uncertainty is the most objective approach. Second, in at least some cases, the extinct taxa under study have no direct functional analogue among extant taxa and thus their quantitative soft tissue properties may be expected to differ also. Third, at present there is relatively little quantitative data of cranial muscle architecture in extant taxa [37] and so the full range of values for extant groups are unlikely to be well sampled. These three FL and PCSA iterations were applied to the three muscle volume sculptures generated independently by the three investigators, yielding nine fossil models per taxon (27 fossil model iterations in total) to be evaluated relative to the model using real (measured) muscle values in multi-body dynamics (MDA) and finite element (FE) models.

(c) Multi-body dynamics (MDA) analysis

We used the open source forwards dynamic package GaitSym (version 2013) to construct MDA models and simulate maximal muscle contraction and symmetrical incisor bite forces in all three

rodent models (Fig. 1) following the approach of [36-37] (see also additional description in ESM). Muscle geometries (origins, insertions and approximate lines of action) were based on physical dissection and contrast-enhanced micro-CT reconstructions of the specimens being modelled [70] and were standardised across all model iterations. The physiological characteristics of muscles were standardised across all taxa and model iterations, as is typical in fossil studies. From these base models, we subsequently generated 10 MDA models for each taxon. For each taxon we generated an 'extant' model, where muscle FLs and masses, and subsequently PCSAs, were measured directly from specimens being modelled [70]. The remaining 9 models consisted of three per investigator, in which each investigators' muscle volumes were used to generate three models according to the three fibre architecture iterations (A, B and C) explained above. All soft tissue input values for the 27 fossil iterations are tabulated in supplementary material (Table S7-9).

(d) Finite element analysis

We re-analysed the existing FE models [77-78] (Fig, 1) of our three rodent taxa in ANSYS Mechanical APDL 2019 R1 using the newly generated muscle force values from our MDA models. See the ESM for slight modifications made to the models in ANSYS. We also standardized the tissue material properties of the models (Table S10) across these taxa (applying the guinea pig properties to all models), as is standard in analyses of fossils (refs). To compare the stresses predicted by the different model iterations we uniformly divided each cranium into 10 sections anteroposteriorly (Fig. S3). The mean Von Mises stress of all elements in each section were extracted and calculated for every loading scenario's simulation. FE models, and the extant iterations of our MDA models, are available to download from http://datacat.liverpool.ac.uk/id/eprint/1184.

3. Results

(a) Muscle volume reconstruction

The total (summed) masticatory muscle mass reconstructed by investigator 1 yielded errors of 14.5%, 9.7% and 3.1% for the guinea pig, rat and squirrel (Fig. 2; Tables S4-6). Investigator 2 produced lower errors of 1.8%, 3% and -2.8% for the guinea pig, rat and squirrel, while investigator 3 produced greater errors of 57.8%, 15.3% and 93.8% (Fig. 2; Tables S4-6). Error magnitudes for individual muscles varied more widely, from less than 1% up to 552% (Fig. 2; Tables S4-6). Visual inspection suggests no common pattern among muscles in terms of error magnitudes, although on the whole there was a greater tendency to overestimate rather than underestimate muscle volume (Fig. 2; Tables S4-6). Regression analysis provides no support for size effects (e.g. systematically larger errors in bigger or smaller muscles) in error magnitudes (Fig. S4).

The three investigators also vary considerably in relative accuracy of the reconstructed total muscle volume and the relative volumes of individual homologous muscles across the three species. Measurements indicate that guinea pigs have the highest summed masticatory muscle volume (3654 mm^3), followed by the squirrel (3431 mm^3) and then the rat (2461 mm^3). Investigators 1 & 2 recovered this relative pattern correctly, but the reconstructions by investigator 3 produced qualitative error with the squirrel being reconstructed with greater overall masticatory muscle volume than the guinea pig (Tables S4-6). In terms of the relative sizes of individual muscles, investigator 1 produced 36% correct relative placements, versus 84% and 52% in the reconstructions of investigators 2 & 3.

(b) Muscle FL and PCSA

Muscle architecture iteration A overestimated muscle fibre length in all muscles in this analysis (Fig 3; Tables S11-13). That is, muscle length always exceeded measured fibre lengths in the masticatory muscles of all three taxa. Overestimation ranged from +55% to +205% in the squirrel,

+29% to +292% in the guinea pig, and +20% to +203% in the rat (Fig 3; Tables S11-13). By utilizing the average muscle length to FL ratio to derive fibre length, muscle architecture iteration C yielded much lower errors in predicted fibre lengths, with errors ranging from -27.3% to +40%, -6.6% to +86.4% and -42.84% to +17.5% in the squirrel, guinea pig and rat (Fig 3; Tables S11-13). Therefore, given accurate muscle volumes, muscle architecture iteration A will always tend to underestimate PCSA, while muscle architecture iteration C will yield lower errors but overestimate PCSA in some muscles while underestimating it in others.

Because PCSA is a function of muscle volume and fibre length, and muscle volume varied considerably and non-systematically across the investigators (Fig 3; Tables S11-13), this parameter shows a complex pattern across the nine fossil model iterations. However, on the whole muscle architecture iteration A tended to underestimate PCSA in all models (all species, all investigators) even where investigators had overestimated muscle volume (Figs. 2-3; Table S4-6) due to the relatively large errors resulting from the assumption that fibre length was equal to muscle length (see above: Fig. 3). Interestingly, maximum underestimations of PCSA were quite similar across species (-81.7% to -96%) and all occurred in models of investigator 3. Where overestimation of PCSA did occur, investigator 3 again yielded the highest errors in all three species, with magnitudes of +283.6%, +94.1% and +39.13% in the squirrel, guinea pig and rat (Fig 3; Tables S11-13).

The range of PCSA error magnitudes in models using muscle architecture iteration C were greater (Fig 3; Tables S11-13), despite the fact that this iteration matched real (measured) fibre lengths more closely than iteration A (Fig 3; Tables S11-13). The range in error magnitudes varied considerably across the three species, ranging from -80.5% to +714%, -92.3% to +240.5% and -65.1 to +80.3% in the squirrel, guinea pig and rat (Fig 3; Tables S11-13). Muscle architecture iteration C yields highly varied levels of inaccuracy in PCSA within and between investigators,

although generally errors noticeably lower in investigator 2 for all species (Fig 3; Tables S11-13).

Because we have modelled static biting, and thus modelling muscle contraction under isometric conditions, PCSA is directly proportional to muscle force in this analysis. It is therefore worth evaluating frequency with which the model iterations correctly predict the relative PCSA of homologous muscles across the three species. Investigator 1 correctly ordered individual taxa in terms of relative PCSA seven out of 24 (29%) times in their muscle architecture iteration A, and eight out of 24 (33.3%) times in iteration C. Despite relatively high quantitative errors, investigator 3 correctly ordered individual taxa in terms of relative PCSA 18 out of 24 (63%) times in both muscle architecture iterations A and C. In line with their relatively lower absolute errors in PCSA, investigator 2 correctly ordered individual taxa in terms of relative PCSA 18 out of 24 (75%) times in both muscle architecture iterations A and C.

(c) Bite forces in MDA models

Our initial MDA models, using measured (real) muscle properties yielded maximal static incisor bite forces of 47.9 N, 56.8 N and 70.2 N for the guinea pig, rat and squirrel models (Fig. 4; Table S14). Individual muscle forces and associated errors for all model iterations are tabulated in Table S14. The three model iterations of investigator 1 yielded quantitative errors in incisors bite force ranging between -65.9% to +16.9% of bite forces from the extant models. All model iterations from investigator 2 underestimated bite force, by between -63% to -6.7%, while the models reconstructed by investigator 3 ranged from -52.2% to +30.6% of the values from the extant model (Fig. 4). Within each investigator, the lowest bite forces and largest absolute errors were recovered in iteration A, where the overestimation of fibre lengths yielded underestimates of PCSA and subsequently maximum isometric muscle force (Fig. 4; Tables S11-13). Reconstructing the medial

pterygoid with more representative pennate architecture and shorter fibre lengths, and subsequently use of Eq.3 to calculate PCSA, increased its maximum isometric force and thus incisor bite force, leading to very small improvements (1-5%) in absolute accuracy (Fig. 4; Table S14). This reduced underestimation in bite force in investigator 2 to -6.7 to -17.6%, and overall error in investigator 1 to -18.6% to +9.8% across the three taxa (Fig. 4; Table S14). However, in investigator 3, iteration C reversed the -35 to -62% underestimated error seen in iterations A and B to slightly lower magnitudes of overestimated error (+13 to +30.6%; Fig. 4; Table S14).

The three investigators also vary considerably in the accuracy with which their models correctly predicted the relative bite forces of the three species. None of the model iterations generated by investigator 1 placed all three taxa in the correct order in terms of relative bite force. Investigator 1's models did consistently predict higher bite forces in the rat compared to the guinea pig, but only iteration C correctly predicted higher forces in the squirrel compared to the guinea pig. Iterations A and C by investigator 3 correctly identified the squirrel as generating the highest bite force of the three taxa, but incorrectly predicted relatively higher bite forces in the guinea pig (Fig. 4; Table S14). Iteration C by investigator 3 and all three iterations (A-C) by investigator 2 correctly predicted relative bite forces across the three species (Fig. 4; Tables S13).

(d) Stress and strain in FE models

FE models loaded using outputs from the 'extant' MDA models indicate that the rat experiences the highest stresses, followed by the squirrel and then the guinea pig along the entire skull length (Fig. 5a-d). The most striking pattern among fossil model iterations is the variation in stress magnitudes. With the exception of small regions of the rat and guinea pig models in iteration C of investigator 2 (Fig. 5b, d & e), all fossil models produced by investigators 1 and 2 underestimate stress relative to

343 the extant models (Fig. 5a-b). Error is higher in the models of investigator 1, where stress
344 magnitudes are less than one-third of that seen in extant models in some regions of the skull (Fig.
345 5a, d & f). The models of investigator 3 showed a more complex pattern of error, with all model C
346 iterations overestimating stress magnitudes throughout the skull, while iterations A and B vary in
347 the nature and magnitude of error across the three rodent taxa (Fig. 5c). For example, iterations A
348 and B of the guinea pig model slightly underestimate stress in most regions, but overestimate stress
349 in between 30-45% skull length (Fig. 5c).

350

351 Despite extremely high variation in stress magnitudes, the qualitative pattern or distribution of
352 stress across the skull seen in the extant models is mostly preserved in the fossil model iterations
353 (Fig. 5) with relatively subtle deviations. A notable exceptions to this is the absence of the sharp
354 increase in stress, or stress peak, between 20-50% skull length in all three fossil iterations of the
355 squirrel model of investigator 1, which changes the stress distribution in the zygomatic arch relative
356 to the extant model and the models of the guinea pig and rat (Fig. 5). However, while the qualitative
357 pattern of stress distribution across the three rodents are mostly preserved across the fossil
358 iterations, the pattern of absolute stress magnitude (i.e. rat > squirrel > guinea pig) is not always
359 recovered (Fig. 5). The aforementioned error in the squirrel models of investigator 3, along with
360 general underestimation of stress therein, means that the relative stress patterns recovered in the
361 squirrel and guinea pig are qualitatively reversed (Fig. 5a, d & f). The models of investigator 3 most
362 preserve qualitative differences between the morphotypes, but iteration C exaggerates the
363 quantitative differences, while iterations B and C underestimate them (Fig. 5c).

364

365 **4. Discussion and Conclusions**

366 Soft tissue reconstructions and biomechanical models provide quantitative measures of functional
367 performance in extinct taxa and thereby offer a unique insight into the evolution of life on Earth [1-

69]. These quantitative measures of function and performance (e.g. energy costs, running speeds, bone strain and safety factors) represent the most direct basis for understanding how anatomical innovations enabled major behavioural or niche adaptions over geological time, and for testing hypotheses about the selective ecological pressures driving major evolutionary radiations [1-69]. Constructing accurate biomechanical models of extant taxa, where (theoretically) all anatomical and physiology parameters can be measured directly, is challenging and some level of abstraction and hence inaccuracy is expected, even in highly detailed models [79]. Greater quantitative error should be expected in extinct animals and arises from the need to progressively reconstruct (i.e. estimate) absolute values for soft tissue parameters like muscle size and architecture that underpin their force generating capabilities (Fig. 1). Some studies, of both living and fossils animals, have used sensitivity analyses to formally acknowledge quantitative error arising from uncertain and often subjectively reconstructed soft tissues parameters [12-14, 20, 22, 36-37]. While this approach undoubtedly represents good practice and demonstrates the sensitivity of simulated predictions to particular input parameters, sensitivity analyses on finalised biomechanical models do not inherently constrain the actual likely magnitude of error within a specific set of fossil soft tissue reconstructions, and subsequently the biomechanical models generated thereafter. Thus, sensitivity analysis, by itself, may not provide a direct test of our ability to reconstruct soft tissue properties and subsequently to progressively estimate quantitative and even qualitative differences between extinct taxa.

In this study we have taken a novel approach to evaluating the accuracy and precision of soft tissue and biomechanical reconstructions of extinct animals, and the ability of current methods to accurately capture a functional macroevolutionary radiation (Figs 2-5). The rodent masticatory system has evolved three distinct morphotypes (sciuromorph, hystricomorph and myomorph) with osteological, myological and functional characteristics that lead to disparate specializations in food processing in each morphotype (see section 2(a) above). The rat, representative of the myomorph

condition, has a temporalis muscle 1.6x larger than the squirrel (sciuromorph) and 1.7x guinea pig (hystricomorph) [70] (Tables S4-6). Despite this significant real (measured) difference in size, only one of the three investigators sculpted the rat with the largest temporalis muscle and ordered the three morphotypes successfully in relative temporalis size (Fig. 2; Tables S4-6). The medial and lateral pterygoids were also reconstructed disproportionately in relative terms by all three investigators: two of the three investigators correctly reconstructed the guinea pig with the largest medial pterygoid, but incorrectly reconstructed the squirrel as having the smallest volume for this muscle (Fig. 2; Tables S4-6). The other investigator incorrectly reconstructed the squirrel with the largest medial pterygoid, and rat with the smallest (Fig. 2; Tables S4-6). None of the investigators correctly reconstructed the squirrel with the largest lateral pterygoid volume (Fig. 2; Tables S4-6). However, despite often large magnitudes of quantitative error (Fig. 2; Tables S4-6), the qualitative proportions of a number of muscles (e.g. posterior deep masseter, posterior and infraorbital zygomaticomandibularis) were correctly reconstructed by two and sometimes all three investigators. Overall the investigators averaged 70.3%, 12.3% and 94.57% error at the individual muscle level (Fig. 2), providing clear evidence that studies utilising volume sculpture approaches to assess the evolution of muscle proportions and performance should incorporate an assessment of error in their hypothesis testing.

Bite force, and the mechanical efficiency of biting, are crucial adaptive functional distinctions between the three rodent morphotypes [77-78]. Our extant MDA models with real (measured) muscle properties predict the highest incisor bite forces in the squirrel, followed by the rat and then guinea pig (Fig. 4; Table S14), which is consistent with previous studies [77-78]. Here we show, for the first time, that accuracy with which such a qualitative macroevolutionary pattern is recovered by palaeontological methods varies across investigators and across different model iterations according to the reconstruction of muscle architecture (Fig. 4). The impact of subjectivity, largely related to sculpture of muscle volumes (Fig. 2; Tables S4-6), is manifested in the highly disparate relative

accuracy in bite forces across the investigators: investigator 1 did not capture the true macroevolutionary pattern in any iteration, while investigator 2 correctly recovered the expected pattern across morphotypes in all cases (Fig. 2; Tables S4-6). This difference reflects the considerably lower levels of qualitative and quantitative error in muscle volumes sculpted by investigator 2 (Fig. 2; Tables S4-6). However, the pattern of relative error in bite force seen in investigator 3 demonstrates that even recovering qualitative differences between taxa is not simply a matter of accurately reconstructing muscle size (or its linear equivalents like maximum isometric stress). Muscle force is proportional to PCSA (Eq. 1), which is a function of muscle volume and fibre architecture (Eqs 2 and 3). The first and second model iterations of investigator 3, in which muscles are reconstructed with parallel fibred architecture and fibre lengths equivalent to muscle length, led to incorrect relative bite forces, and failure to capture the true functional macroevolution pattern that has evolved across rodent morphotypes (Fig. 2; Tables S4-6). However, use of average ratios of muscle fibre length to overall length to calculate fibre length, and subsequently use of Eq.3 to calculate PCSA, led to investigator 3's muscle volumes correctly recovering the true macroevolutionary pattern across rodent morphotypes (Fig. 2; Tables S4-6). This emphasises the complex interaction between estimation of muscle size, architecture and force generating capabilities, and highlights that simple sensitivity tests in which muscle size or force is scaled uniformly up or down may be insufficient in macroevolutionary studies (see further discussion below).

These issues regarding both quantitative and qualitative error in masticatory muscle anatomy and bite force translate directly into analyses of absolute and relative stress in FE models (Fig. 5). To our knowledge this is the first study to explicitly examine the likely magnitudes of error in FE models capturing a macroevolution radiation resulting from disparate reconstructions of muscle force generating properties (see further discussion below). As with muscle volumes (Fig. 2) and bite forces (Fig. 4) our data provides clear evidence that current approaches to soft tissue reconstruction

446 can not only recover the correct qualitative or relative differences between taxa, but also generate
447 stress magnitudes and distributions that are quantitatively consistent with models loaded using real
448 (measured) muscle data (Fig. 5b, d-e). While this is encouraging, the errors noted in muscle
449 volume, architecture and bite force predictions (Figs 2-4) inherently mean that many of the fossil
450 model iterations yield highly inaccurate stress magnitudes, and in some instances produce
451 magnitudes and distributions that are qualitatively dissimilar to the extant models and thus do not
452 correctly capture the true qualitative macroevolutionary pattern. Cox et al. [77] noted that stress
453 patterns along the zygomatic arch are different between the three rodent morphotypes, which our
454 extant models capture here (Fig. 5a-d). The magnitude of the stress differences in this region of the
455 skulls varies across model iterations, particularly those of investigator 3 where relative differences
456 between rodents are exaggerated and underestimated by different iterations (Fig. 5c).
457 Underestimation of stress in the zygomatic arch in the models of investigator 2 means that the
458 relative stress magnitudes between the squirrel and guinea pig models are incorrectly represented in
459 this key region (Fig. 5a, d, f). Cox et al [77] also note that the rat shows a pattern of elevated stress
460 around the origin of the temporalis muscle compared to the guinea pig and squirrel models, which is
461 causatively associated with this taxon's larger temporalis muscle (Fig. 2; Tables S4-6). The extent
462 to which this pattern is recovered in the fossil models presented here varies according to the
463 accuracy of temporalis muscle reconstruction. As noted above, only one of the investigators
464 correctly reconstructed the relative size of the temporalis muscle across the three rodent
465 morphotypes (Fig. 2; Tables S4-6).

466

467 To put our study and its conclusions into context, we surveyed 67 published studies that utilised
468 quantitative soft tissue reconstruction alone or in combination with biomechanical models to
469 examine evolutionary changes in functional morphology in fossil taxa [2-32, 35-69]. Our goal was
470 not to provide exhaustive coverage of all relevant papers, but to sample enough studies to provide
471 coverage of most major taxonomic groups, body regions (limbs, skulls, necks etc.) and

472 methodological approaches. We assessed two aspects of quantitative soft tissue reconstruction in
473 these studies; first, whether the study used a method of quantitative soft tissue reconstruction
474 associated with muscle force properties that had been validated in equivalent models of extant
475 animals. Specifically, we assessed whether extant taxa had been used to either demonstrate that an
476 approach yields quantitative results that are highly comparable to measured soft tissue data, and/or
477 to provide an expected level of error in the final predictions that are used to quantitively constrain
478 predictions (and hypothesis testing) in extinct animals. Second, we assessed whether sensitivity
479 analysis was used to explicitly test for uncertainties in final predictions associated with the
480 reconstruction of soft tissue force generating capabilities in fossil taxa. Our subjective judgement of
481 these criteria lead us to suggest that only around 35% of studies have utilised methods of numerical
482 soft tissue reconstruction that have been validated for precision and accuracy in extant animals, and
483 only around 32% of studies have used any kind of sensitivity analysis in their assessments of the
484 force generating capacity of muscles in extinct animals. In the latter aspect (sensitivity analysis) this
485 figure of 32% can be considered optimistic as we choose to be maximally inclusive and include
486 studies that our present results (Figs 2-5) would suggest are insufficient in terms of sensitivity
487 testing. For example, a number of assessments of bite mechanics in extinct animals provided
488 minimum and maximum estimates of bite force by either selecting extreme low and high values for
489 maximum isometric stress [44-45] or by adding a model iteration in which a correction factor was
490 applied to increase muscle force [46] across all muscles. These sensitivity analyses were limited to
491 bite force predictions and not carried forward to FE analyses of the fossil taxa, presumably because
492 all muscle forces were varied uniformly. As our results demonstrate, uniform error in the
493 reconstruction of individual muscles, even within one taxon, should not be expected (Figs 2-3), and
494 the magnitude of non-unform error across muscles results in unpredictable and differential
495 consequences in functional predictions (Fig. 4-5). Breaking these studies down in body regions and
496 biomechanical approaches reveals a clear signal in the tendency to quantitatively validate and
497 recognise soft tissue error in biomechanical predictions. Studies of limbs more frequently applied at

498 least some of their reconstructions approaches to extant animals (90%) and carried out sensitivity
499 analyses on their reconstructions of fossil taxa (55%), while studies of skulls have done so much
500 less frequently (7% and 21% respectively). This same disparity is reflected in MDA (70% and 45%)
501 versus FEA (2.9% and 17%) approaches because the majority of locomotor studies have used
502 MDA, while FEA is most common in analyses of skulls.

503

504 This crude appraisal of the frequency with which current studies explicitly incorporate error in soft
505 tissue properties, in some way, into functional assessments of extinct animals is concerning given
506 the new systematic assessment of muscle property reconstruction (Figs 2-3), muscle kinetics (Fig.
507 4) and bone stress (Fig. 5) we present here. Quantitative uncertainty and error will perhaps always
508 remain unavoidable in evolutionary biomechanics, but an ability to identify qualitative similarities
509 and differences across fossil lineages, and between extinct taxa and extant groups with known
510 behaviours is fundamental to our understanding of palaeoecology and ecosystem dynamics,
511 adaptive radiations and selective extinctions and functional constraints on biological evolution [1-
512 69]. Our novel analysis highlights that correctly reconstructing qualitative differences between taxa
513 in a macroevolutionary radiation is challenging and that both false positive and negative results are
514 possible using current approaches to quantitative soft tissue reconstruction. Our results provide
515 quantitative evidence that studies of fossil taxa should incorporate a systematic assessment of
516 reconstruction error into their experimental procedures and hypothesis testing and provide clear
517 incentive for an expansion of primary data sets on muscle properties in extant taxa to better inform
518 soft tissue reconstructions in macroevolutionary studies.

519

520 **Acknowledgements**

This work was funded by research projects grant from a NERC standard grant (NE/G001952/1) to P.G.C and M.J.F, a BBSRC responsive mode grant to M.F. and K.T.B (BB/R016380/1; BB/R016917/1; BB/R017190/1), and a NERC doctoral dissertation grant to M.D.

References

[1] Pierce SE, Clack JA, Hutchinson JR. 2012 Three-dimensional limb joint mobility in the early tetrapod Ichthyostega. *Nature* **486**, 523-526.

[2] Neenan, JM, Ruta M, Clack JA, Rayfield EJ. 2014 Feeding biomechanics in Acanthostega and across the fish-tetrapod transition. *Proc. Roy. Soc. B* **281**, 20132689.

[3] Lautenschlager S, Gill PG, Luo ZX, Fagan MJ, Rayfield EJ. 2018 The role of miniaturization in the evolution of the mammalian jaw and middle ear. *Nature* **561**, 533-537.

[4] Fahn-Lai P, Biewener AA, Pierce SE. 2020 Broad similarities in shoulder muscle architecture and organization across two amniotes: implications for reconstructing non-mammalian synapsids. *PeerJ* **8**, e8556.

[5] Adams NF, Rayfield EJ, Cox PG, Cobb SN, Corfe IJ. 2019 Functional tests of the competitive exclusion hypothesis for multituberculate extinction. *Roy. Soc. Open Sci.* **6**, 181536.

[6] Lautenschlager S, Gill P, Luo ZX, Fagan MJ, Rayfield EJ. 2016) Morphological evolution of the mammalian jaw adductor complex. *Biological Reviews* **92**, 1910-1940.

[7] Bishop PJ et al., 2018 Cancellous bone and theropod dinosaur locomotion. Part III – Inferring posture and locomotor biomechanics in extinct theropods, and its evolution on the line to birds. *PeerJ* **6**, e5778.

[8] Pontzer H, Allen V, Hutchinson JR. 2009 Biomechanics of running indicates endothermy in bipedal dinosaurs. *PLoS ONE* **4**, e7783.

[9] Gatesy SM, Baeker M, Hutchinson JR. 2009 Constraint-based exclusion of limb poses for reconstructing theropod dinosaur locomotion. *J. Verte. Pal.* **29**, 535-544.

[10] Hutchinson JR, Ng-Thow-Hing V, Anderson FC. 2007 A 3D interactive method for estimating body segmental parameters in animals: application to the turning and running performance of *Tyrannosaurus rex*. *J. Theo. Bio.* **246,** 660-680.

[11] Snively, E et al. 2019. Lower rotational inertia and larger leg muscles indicate more rapid turns in tyrannosaurids than in other large theropods. *PeerJ* **2**, e6432.

[12] Hutchinson JR, Anderson FC, Blemker S, Delp SL. 2005 Analysis of hindlimb muscle moment arms in *Tyrannosaurus rex* using a three-dimensional musculoskeletal computer model. *Paleobiology* **31**, 676-701.

554 [13] Hutchinson JR. 2004 Biomechanical modeling and sensitivity analysis of bipedal running
555 ability. II. Extinct taxa. *J. Morph.* **262**, 441-461.
556 [14] Hutchinson JR, Garcia M. 2002 *Tyrannosaurus* was not a fast runner. *Nature* **415**, 1018-1021.
557 [15] Persons WS, Currie PJ. 2014 Duckbills on the run: the cursorial abilities of giant Late
558 Cretaceous ornithopods and implications for tyrannosaurs avoidance strategies. In: Hadrosaurs (eds.
559 Eberth DA and Evans DC). Indiana University Press, Indianapolis, IN. 449-458.
560 [16] Persons WS, Currie PJ. 2011. Dinosaur speed demon: the caudal musculature of Carnotaurus
561 sastrei and implications for the evolution of South American abelisaurids. *PLoS ONE* **6**, e25763.
562 [17] Persons WS, Currie PJ. 2011. The tail of Tyrannosaurus: reassessing the size and locomotive
563 importance of the M. caudofemoralis in non-avian theropods. *The Anat. Rec.* **294**, 119–131.
564 [18] Sellers WI, Pond SB, Brassey CA, Manning PL, Bates KT. 2017 Investigating the running
565 abilities of Tyrannosaurus rex using stress-constrained multibody dynamic analysis. *PeerJ* **5**, e3420.
566 [19] Sellers WI, Coria RA, Manning PL, Margetts L. 2013 March of the Titans: the locomotor
567 capabilities of sauropod dinosaurs. *PLoS ONE* **8**, e78733
568 [20] Bates KT, Manning PL, Margetts L, Sellers WI. 2010 Sensitivity analysis in evolutionary
569 robotic simulations of bipedal dinosaur running. *J. Verte. Pal.* **30**, 458-466.
570 [21] Sellers WI, Lyson T, Margetts L, Stevens K, Manning PL. 2009. Virtual palaeontology: gait
571 reconstruction of extinct vertebrates using high performance computing. *Pal. Elect.* **12**, 11A:26
572 [22] Manning PL, Sellers WI. 2007. Estimating dinosaur maximum running speeds using
573 evolutionary robotics. *Proc. Roy. Soc. B* **274**, 2711-2716.
574 [23] Bates KT, Benson, RBJ, Falkingham PL. 2012 The evolution of body size, stance and gait in
575 Allosauroidea (Dinosauria: Theropoda). *Paleobiology* **38**, 486-507.
576 [24] Allen V, Bates, KT, Zhiheng L, Hutchinson JR. 2013. Linking the evolution of body shape and
577 locomotor biomechanics in bird-line archosaurs. *Nature* **497**, 104-107.
578 [25] Crompton RH, et al. 2012 Human-like external function of the foot, and fully upright gait,
579 confirmed in the 3.66 million year old Laetoli hominin footprints by topographic statistics,
580 experimental footprint-formation and computer simulation. *J. Roy. Soc. Interface* **9**, 707-719.
581 [26] Sellers WI, Wang W, Crompton RH. 2005. Stride lengths, speed and energy costs in walking
582 of Australopithecus afarensis: using evolutionary robotics to predict locomotion of early human
583 ancestors. J. Roy. Soc. Interface **2**, 431-441.
584 [27] Wang W, Crompton RH, Carey TS, Günther MM, Li Y, Savage R, Sellers WI.
585 2004. Comparison of inverse-dynamics musculo-skeletal models of AL 288-1 Australopithecus
586 afarensis and KNM-WT 15000 Homo ergaster to modern humans, with implications for the
587 evolution of bipedalism. *J. Human Evo.* **47**, 453-478.

588 [28] Wroe S, Ferrara T, McHenry C, Curnoe D, Chamoli U. 2010 The craniomandibular mechanics
589 of being human. *Proc. Roy. Soc. B* **277**, 3579-3586.
590 [29] Strait DS et al. 2009 The feeding biomechanics and dietary ecology
591 of *Australopithecus africanus*. *PNAS* **106**, 2124-2129.
592 [30] Strait DS et al. 2010 The Structural Rigidity of the Cranium of *Australopithecus*
593 *africanus:* Implications for Diet, Dietary Adaptations, and the Allometry of Feeding Biomechanics.
594 *The Anat. Rec.* **293**, 583-593
595 [31] Ledogar JA et al. 2016 Mechanical evidence that *Australopithecus sediba* was limited in its
596 ability to eat hard foods. *Nature Comms* **7**, 10596.
597 [32] Wroe S et al. 2018 Computer simulations show that Neanderthal facial morphology represents
598 adaptation to cold and high energy demands, but not heavy biting. Proc. Roy. Soc. B 285,
599 20180085.
600 [33] Maidment SCR et al. 2014 Locomotion in Ornithischian Dinosaurs: An Assessment Using
601 Three-Dimensional Computational Modelling. *Biol Rev* **89**, 588-617.
602 [34] Hutchinson JR, Allen V. 2009 The evolutionary continuum of limb function from early
603 theropods to birds. *Naturwissenschaften* **96**, 423-448.
604 [35] Gignac PM, Erickson GM (2017) The biomechanics behind extreme osteophagy in
605 *Tyrannosaurus rex*. *Scientific Reports* DOI: 10.1038/s41598-017-02161-w.
606 [36] Bates KT, Falkingham PL. 2012 Estimating maximum bite performance in *Tyrannosaurus rex*
607 using multi-body dynamics. *Biology Letters* **8**, 660-664.
608 [37] Bates KT, Falkingham PL. 2018 The importance of muscle architecture in biomechanical
609 reconstructions of extinct animals: a case study using *Tyrannosaurus rex*. *J. Anat*. **233**, 625-635.
610 [38] Snively E, Cotton JR, Ridgely R, Witmer LM. 2013 Multibody dynamics model of head and
611 neck function in Allosaurus (Dinosauria, Theropoda). *Pal. Elect* **16**, 2.
612 [39] Snively E, Russell AP. 2007 Craniocervical feeding dynamics of *Tyrannosaurus*
613 *rex*. *Paleobiology* **33**, 610-638.
614 [40] Chambi-Trowell SAV, Whiteside DI, Benton M, Rayfield EJ. 2020 Biomechanical properties
615 of the jaws of two species of Clevosaurus and a reanalysis of rhynchocephalian dentary
616 morphospace. *Palaeontology* https://doi.org/10.1111/pala.12493.
617 [41] Coatham SJ, Vinther J, Rayfield EJ, Klug C. 2020 Was the Devonian placoderm Titanichthys a
618 suspension feeder? *Roy. Soc. Open Sci.* **7**, 200272.
619 [42] Ballell A, Moon BC, Porro LB, Benton MJ, Rayfield EJ. 2019 Convergence and functional
620 evolution of longirostry in crocodylomorphs. *Palaeontology* **62**, 867-887.

621 [43] Taylor AC, Lautenschlager S, Qi Z, Rayfield EJ. 2017 Biomechanical Evaluation of Different
622 Musculoskeletal Arrangements in *Psittacosaurus* and Implications for Cranial Function. The *Anat.*
623 *Rec.* **300**, 49-61.
624 [44] Button DJ, Barrett P, Rayfield EJ. 2016 Comparative cranial myology and biomechanics
625 of *Plateosaurus* and *Camarasaurus* and evolution of the sauropod feeding
626 apparatus. *Palaeontology* **59**, 887-913.
627 [45] Button DJ, Rayfield EJ, Barrett PM. 2014 Cranial biomechanics underpins high sauropod
628 diversity in resource-poor environments**.** *Proc Roy. Soc. B* **281**, 20142114**.**
629 [46] Foffa D et al. 2014 Functional anatomy and feeding biomechanics of a giant Upper Jurassic
630 pliosaur (Reptilia: Sauropterygia) from Weymouth Bay, Dorset, UK. *J. Anat*. **225**, 209-219.
631 [47] Lautenschlager S, Witmer LM, Altangerel P, Rayfield EJ. 2013 Edentulism, beaks and
632 biomechanical innovations in the early evolution of theropod dinosaurs. *PNAS* **110**, 20657–20662.
633 [48] Young MT, et al. 2012 Cranial biomechanics of Diplodocus (Dinosauria, Sauropoda): testing
634 hypotheses of feeding behaviour in an extinct megaherbivore. *Naturwissenschaften*, **99**, 637-643.
635 [49] Jasinoski SC, Rayfield EJ, Chinsamy A. 2010 Functional Implications of Dicynodont Cranial
636 Suture Morphology. *J. Morph.* **271**, 705-728.
637 [50] Jasinoski S C, Rayfield EJ, Chinsamy A. 2010 Mechanics of the scarf premaxilla-nasal suture
638 in the snout of Lystrosaurus. *J. Verte. Pal.* **30**, 1283-1288.
639 [51] Jasinoski SC, Rayfield EJ, Chinsamy A. 2009 Comparative feeding biomechanics of
640 Lystrosaurus and the generalized dicynodont Oudenodon. *The Anat. Rec.* **292***,* 862-874.
641 [52] Lautenschlager S, Figueirido B, Cashmore D, Bendel EM, Stubbs T. 2020 Morphological
642 convergence obscures functional diversity in sabre-toothed carnivores. *Proc. Roy. Soc. B.* **278**,
643 20201818.
644 [53] Montefeitro F, Lautenschlager S, Godoy P, Ferreira G, Butler R. 2020 A unique predator in a
645 unique ecosystem: modelling the apex predator within a Late Cretaceous crocodyliform-dominated
646 fauna from Brazil. *J. Anat.* **237**, 323-333.
647 [54] Ferreira G et al. 2020. Feeding biomechanics suggests progressive correlation of skull
648 architecture and neck evolution in turtles. *Scientific Reports* **10**, 5505.
649 [55] Figueirido B, Lautenschlager S, Pérez-Ramos A, Van Valkenburgh B 2018 Distinct predatory
650 behaviors in scimitar- and dirk-toothed sabertooth cats. *Current Biology* **28**, 3260-3266.
651 [56] Lautenschlager S, Witzmann F, Werneburg, I. 2016 Palate anatomy and morphofunctional
652 aspects of interpterygoid vacuities in temnospondyl cranial evolution, *Naturwissenschaften* **103**, 79.
653 [57] Lautenschlager S, Brassey C, Button D, Barrett P. 2016 Decoupled form and function in
654 disparate herbivorous dinosaur clades, *Scientific Reports* **6**, 26495.

655 [58] Lautenschlager S. 2013 Cranial myology and bite force performance of *Erlikosaurus andrewsi*:
656 a novel approach for digital muscle reconstructions. *J. Anat*. **222**, 260-272.
657 [59] Attard M, Johnston P, Parr WCH, Wroe S. 2016 Moa diet fits the bill: virtual reconstruction
658 incorporating mummified remains and prediction of biomechanical performance in avian giants.
659 *Proc. Roy. Soc. B* **283**, 2015-2043
660 [60] Oldfield CC et al. 2012 Finite Element Analysis of ursid cranial mechanics and the prediction
661 of feeding behaviour in the extinct giant *Agriotherium africanum*. *J. Zoology* **286**, 163-170.
662 [61] Wroe S. 2008 Cranial mechanics compared in extinct marsupial and extant African lions using
663 a finite-element approach. *J. Zoology* **274**, 332-339.
664 [62] Wroe S, Milne N. 2007 Convergence and remarkably consistent constraint in the evolution of
665 carnivore skull shape. *Evolution* **61**, 1251-1260.
666 [63] McHenry C, Wroe S, Clausen P, Moreno K, Cunningham E. 2007 Super-modeled sabercat,
667 predatory behaviour in Smilodon fatalis revealed by high-resolution 3-D computer simulation.
668 *PNAS* **104**, 16010-16015.
669 [64] Wroe S, McHenry C, Thomason J. 2005 Bite club: Comparative bite force in big biting
670 mammals and the prediction of predatory behaviour in fossil taxa. *Proc. Roy. Soc.* B **272**, 619-625.
671 [65] Snively E, Fahlke JM, Welsh RC. 2015 Bone-breaking bite force of *basilosaurus isis*
672 (Mammalia, Cetacea) from the late eocene of Egypt estimated by finite element analysis
673 e0118380. *PLoS ONE*, **10**, e0118380.
674 [66] Bell PR, Snively E, Shychoski L. 2009 A comparison of the jaw mechanics in hadrosaurid and
675 ceratopsid dinosaurs using finite element analysis. *The Anat. Rec.* **292**, 1338-1351.
676 [67] Serrano-Fochs S, De Esteban-Trivigno S, Marcé-Nogué J, Fortuny J, Fariña RA. 2015 Finite
677 Element Analysis of the Cingulata Jaw: An Ecomorphological Approach to Armadillo's Diets.
678 PLoS ONE 10, e0129953.
679 [68] Hutchinson JR, Bates KT, Molnar J, Allen V, Makovicky P. 2011 A computational and
680 comparative analysis of limb and body proportions in *Tyrannosaurus rex* with implications for
681 locomotion and growth. PLoS ONE 6, e26037.
682 [69] Sharp AC. 2014. Three dimensional digital reconstruction of the jaw adductor musculature of
683 the extinct marsupial giant *Diprotodon optatum*. *PeerJ* **2**, e514.
684 [70] Cox PG, Jeffery N. 2015 The muscles of mastication in rodents and the function of the medial
685 pterygoid. In: *Evolution of the Rodents: Advances in Phylogeny, Functional Morphology and*
686 *Development* (eds PG Cox, L Hautier). Cambridge: Cambridge University Press, pp 350-372.
687 [71] Burgin CJ, Colella JP, Kahn PL, Upham NS. 2018 How many species of mammal are there? *J.*
688 *Mamm*. **99**, 1-14.

689 [72] Brandt JF. 1855 Untersuchungen über die craniologischen Entwicklungsstufen und
690 Classification der Nage der Jetzwelt. *Mémoires de l'Academie Imperiale des Sciences de St*
691 *Pétersbourg Série* **6**, 1–365.
692 [73] Wood AE. 1965 Grades and clades among rodents. *Evolution* **19**, 115–130.
693 [74] Cox PG, Jeffery N. 2011. Reviewing the morphology of the jaw-closing musculature in
694 squirrels, rats, and guinea pigs with contrast-enhanced microCT. *The Anat. Rec.* **294**, 915–928.
695 [75] Wood AE. 1958 Are there rodent suborders? *Sys. Zoo.* **7**, 169–173.
696 [76] Druzinsky RE. 2010 Functional anatomy of incisal biting in *Aplodontia rufa* and sciuromorph
697 rodents—Part 2: sciuromorphy is efficacious for production of force at the incisors. *Cells Tissues*
698 *Organs* **192**, 50–63.
699 [77] Cox PG et al. 2012 Functional evolution of the feeding system in rodents. *PLoS ONE 7*,
700 e36299.
701 [78] Cox PG, Fagan MJ, Rayfield EJ, Jeffery N. 2011 Biomechanical performance of squirrel,
702 guinea pig and rat skulls: sensitivity analyses of finite element models. *J. Anat.* **219**, 696-709.
703 [79] Charles JP, Grant B, D'Août K, Bates KT. 2020 Subject-specific muscle properties from
704 diffusion tensor imaging significantly improve the accuracy of biomechanical models. *J. Anat.* **237**,
705 941-959.
706
707

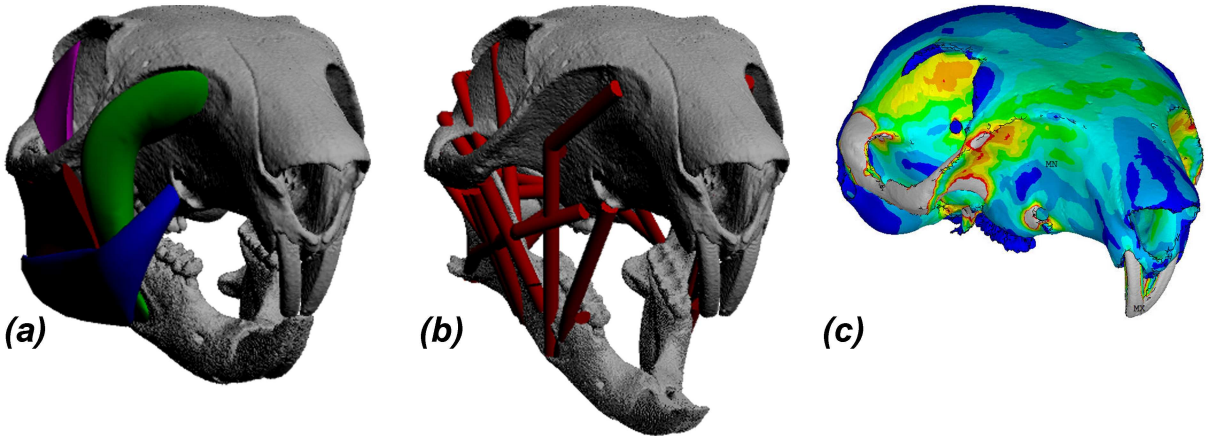
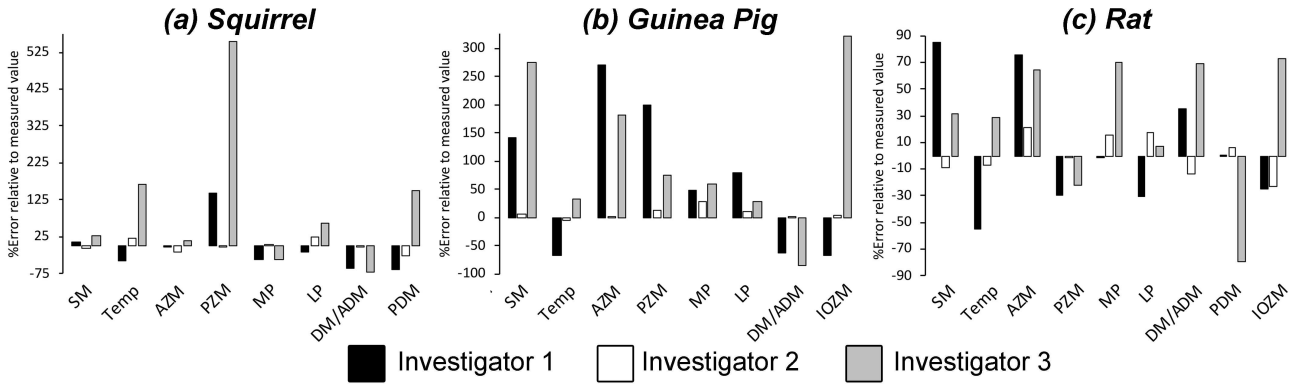


Figure 1. Quantitative soft tissue reconstruction and biomechanical modelling of rodent masticatory morphotypes. ***(a)*** Muscle volumes are reconstructed using 3D sculpture techniques, as commonly applied in fossils, with values combined with different estimates of fibre length to provide input values for biomechanical models. Incisor bite forces were predicted across 27 'fossil' model iterations of ***(b)*** MDA models for comparison to values predicted using real (measured) muscle data. ***(c)*** Predicted muscle forces from all model iterations were used to load FE models to compare stresses predicted in fossil models to those from models with real (measured) muscle properties.



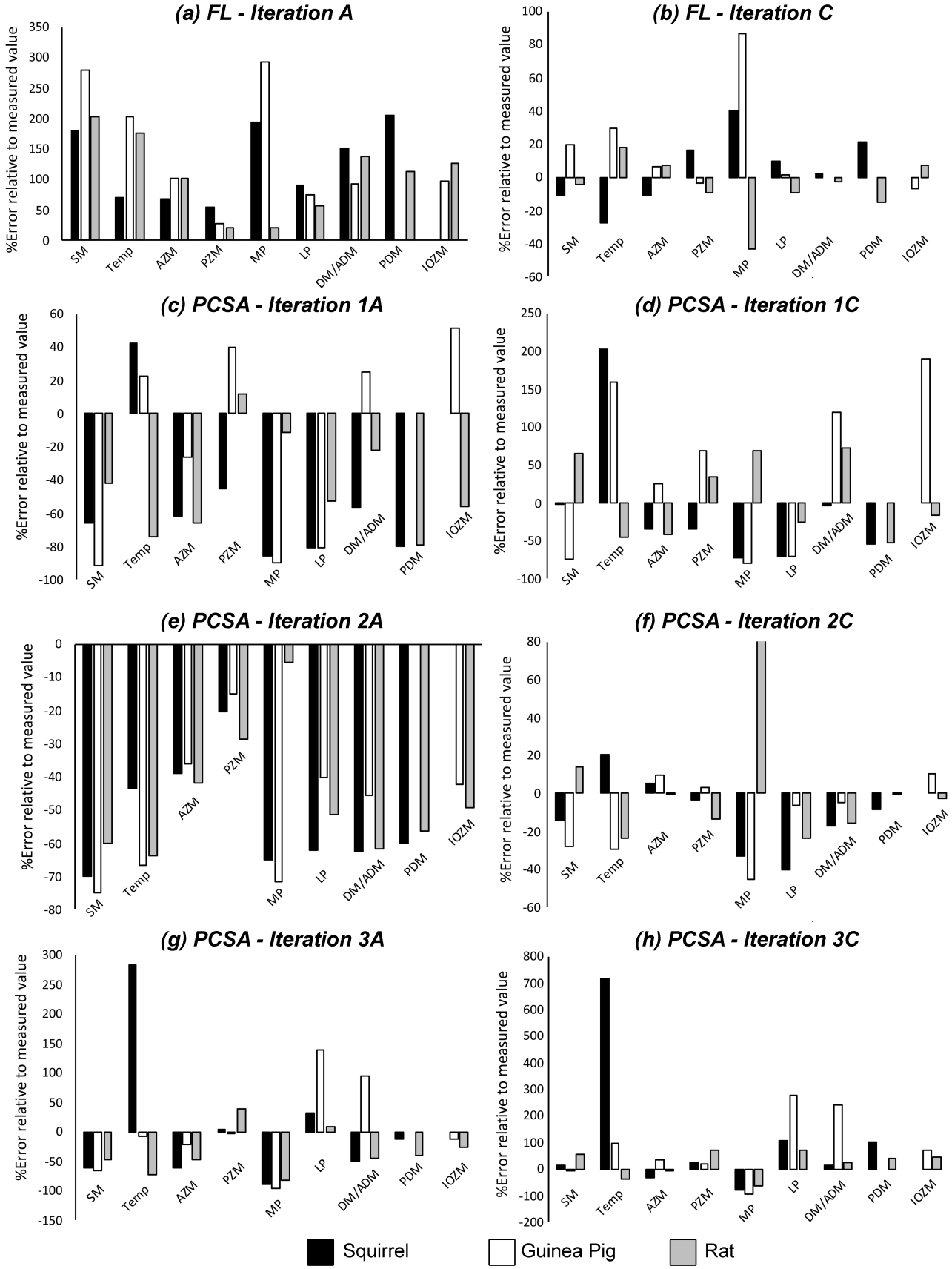
721

722 **Figure 2.** Error magnitudes in the sculptured muscle volume reconstructions by investigators 1, 2
723 and 3 for the ***(a)*** squirrel, ***(b)*** guinea pig and ***(c)*** rat. Abbreviations: SM, superficial masseter; Temp,
724 temporalis; AZM, anterior zygomatico-mandibularis; PZM, posterior zygomatico-mandibularis;
725 MP, medial pterygoid; LP, lateral pterygoid; DM/ADM, deep masseter/anterior deep masseter;
726 PDM, posterior deep masseter; Infraorbital zygomatico-mandibularis.

727

728

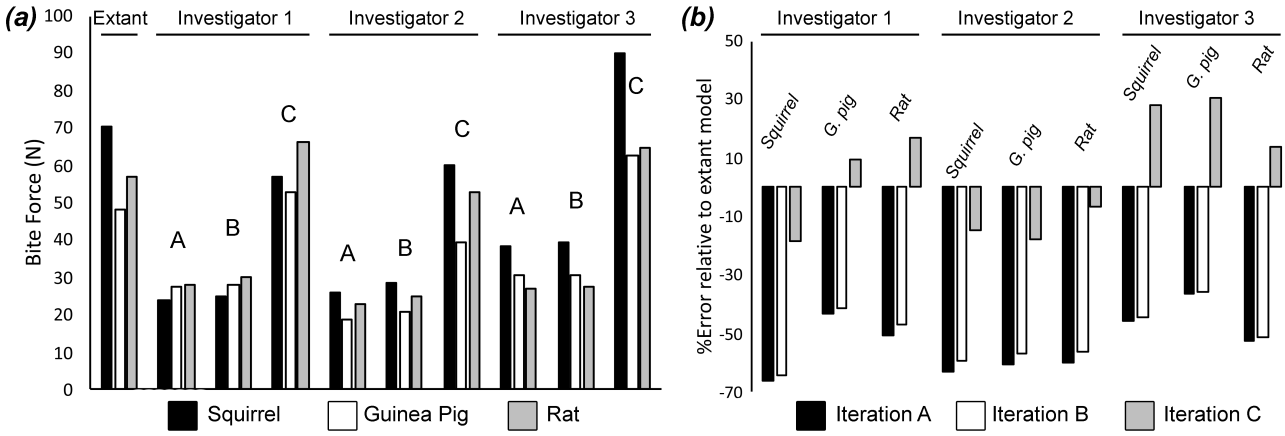
729



730

731 **Figure 3.** Error magnitudes in reconstructed ***(a-b)*** muscle fibre lengths and ***(c-h)*** PCSAs in the

732 three species.



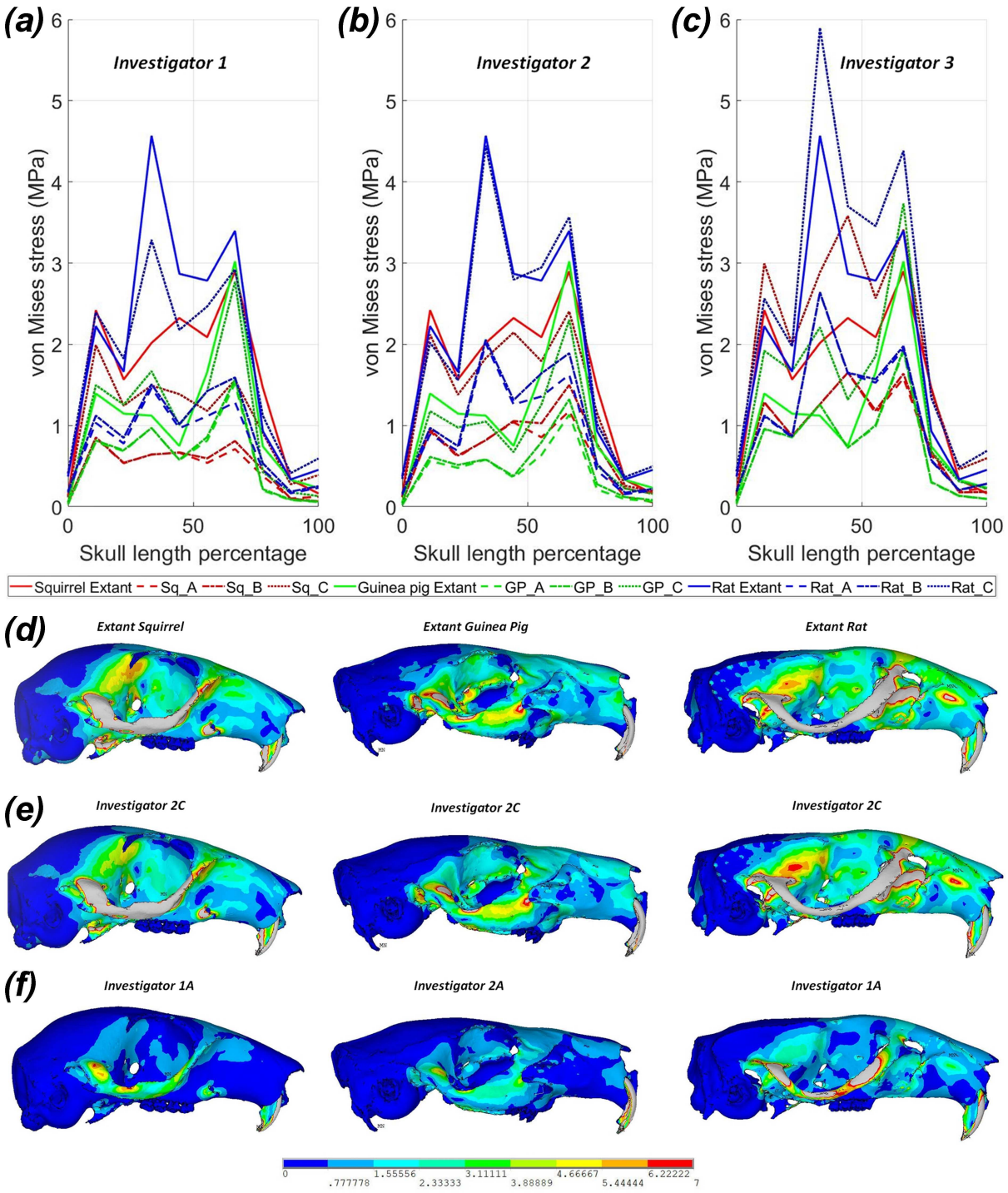
733

734 **Figure 4.** Comparison of ***(a)*** absolute bite forces and ***(b)*** percentage error magnitudes in bite forces
735 across the 'extant' and 'fossil' MDA models. ***(a)*** 'Extant' model iterations predict the highest
736 incisor bite forces in the squirrel, followed by the rat and then guinea pig. This qualitative pattern
737 across the morphotypes is recovered in all model iterations by investigator 2, by iteration C
738 investigator 3, but in none of the iterations by investigator 1. ***(b)*** Quantitative error varied
739 considerably, with most iterations tending to underestimate bite force.

740

741

742



743

744 **Figure 5.** Stress magnitudes and distributions (represented by von Mises stress) in the FE models
745 across the 30 model iterations. Stress magnitudes along the length of skull in the extant models are
746 compared to those of ***(a)*** investigator 1, ***(b)*** investigator 2 and ***(c)*** investigator 3 and demonstrate
747 significant quantitative and some qualitative error. Some reconstructions, such as ***(b, e)*** iteration C
748 those by investigator 2, show a close quantitative match to ***(d)*** the extant models, while some

749 reconstructions, such as ***(f)*** iteration A by investigator 1 contain both quantitative and qualitative

750 error in relative stress magnitudes and distribution across the morphotypes.

751