



Insufficient Evidence for Multiple Species of *Tyrannosaurus* in the Latest Cretaceous of North America: A Comment on “The Tyrant Lizard King, Queen and Emperor: Multiple Lines of Morphological and Stratigraphic Evidence Support Subtle Evolution and Probable Speciation Within the North American Genus *Tyrannosaurus*”

Thomas D. Carr¹ · James G. Napoli² · Stephen L. Brusatte³ · Thomas R. Holtz Jr.^{4,5} · David W. E. Hone⁶ · Thomas E. Williamson⁷ · Lindsay E. Zanno^{8,9}

Received: 4 April 2022 / Accepted: 24 June 2022 / Published online: 25 July 2022
© The Author(s) 2022

Abstract

The Late Cretaceous dinosaur *Tyrannosaurus rex* was recently split into three species based on the premise that variation in the *T. rex* hypodigm is exceptional, indicating cryptic species and “robust” and “gracile” morphs. The morphs are based on proportional ratios throughout the skeleton. The species are claimed to be stratigraphically separate, with an early robust species followed by robust and gracile descendants. There are problems with the hypothesis: the taxon diagnoses are based on two features that overlap between the species; several skulls cannot be identified based on the diagnoses; proportional comparisons between *Tyrannosaurus* and other theropods are based on incomparable samples; the tooth data are problematic; the stratigraphic framework divides the Hell Creek Formation into thirds, without the stratigraphic position of each specimen, or independent age control showing the subdivisions are coeval over the entire geographic area; previous work found variation in *T. rex*, but it cannot be parsed into discrete categories. We tested for “gracile” and “robust” morphs by analyzing the femoral and tooth ratios that were published in the multiple species study using agglomerative hierarchical clustering. The results found that each set of ratios are explained by one cluster, showing that dimorphism is not supported. We tested for exceptional variation of the femoral ratio of *Tyrannosaurus*; we calculated the mean intraspecific robusticity for 112 species of living birds and 4 nonavian theropods. The results showed that the absolute variation in *Tyrannosaurus* is unexceptional and it does not indicate cryptic diversity. We conclude that “*T. regina*” and “*T. imperator*” are subjective junior synonyms of *T. rex*.

Keywords *Tyrannosaurus* · Clustering · Cretaceous · Taxonomy · Variation · Robusticity

Abbreviations

AMNH FARB	American Museum of Natural History, Fossil Amphibians, Reptiles, and Birds, New York	CMNH	Cleveland Museum of Natural History, Cleveland
BHI	Black Hills Institute of Geological Research, Inc., Hill City	LACM	Los Angeles County Museum of Natural History, Los Angeles
BMRP	Burpee Museum of Natural History, Rockford	MOR	Museum of the Rockies, Bozeman
CM	Carnegie Museum, Pittsburgh	PIN	Paleontological Institute, Moscow
		UWBM	Burke Museum, Seattle
		TCM	The Children’s Museum, Indianapolis

Thomas D. Carr and James G. Napoli are co-lead authors.

✉ Thomas D. Carr
tcarr@carthage.edu

Extended author information available on the last page of the article

Introduction

Species are hypotheses based on the best available evidence. Testing and re-testing species identity is a worthwhile venture given that an oversplit taxon, or several

species erroneously lumped into one, will have cascading effects upon macroevolutionary, macroecological, and paleobiological hypotheses (e.g., Benson et al., 2022). However, the destabilization of taxonomy has serious implications for science and society and should be avoided if not backed up by rigorous empirical and replicable evidence (e.g., Kaiser et al., 2013). Since it was named (Osborn, 1905), the validity of *Tyrannosaurus rex* was accepted, given its great size relative to its closest relatives and limited temporal span. As the fossils of *T. rex* accumulated in museum and university collections, studies of variation within the species focused on sexual dimorphism (Carpenter, 1990; Carpenter & Smith, 2001; Larson, 1994, 1997, 2008a) with subsequent rigorous statistical and quantitative analyses failing to recover evidence of dimorphism in *T. rex* (Carr, 2020; Mallon, 2017; Saitta et al., 2020). Since then, less attention has been paid to whether *T. rex* consists of more than one species (Larson, 2008a). This issue was recently brought to the fore by the study of Paul et al. (2022), which argued that what is currently called '*Tyrannosaurus rex*' is actually three separate species, which can be diagnosed by differences in the femur and anterior dentary teeth (Fig. 1).

In principle, we agree with Paul et al. (2022) that there is good reason to question whether the hypodigm presently considered to comprise *T. rex* might in fact contain multiple species. The contemporary ceratopsid genus *Triceratops* was recognized on morphological grounds to contain two species, *T. horridus* and *T. prorsus* (Forster, 1996) and these species were later found, based on high-quality stratigraphic assessments, to be chronologically separated (Scannella et al., 2014). Preliminary work suggests that another contemporary dinosaur, *Pachycephalosaurus*, may also contain two successive species in this same interval (Fowler, 2017.) Furthermore, the proposal of coexisting species of *Tyrannosaurus* is defensible, given the well-documented overlap of ecologically similar congeneric species in modern and extinct faunas that extends even to predators at high trophic levels (Napoli et al., 2019, 2021). However, some tyrannosaur species can persist for over 1 Myr, as indicated by, for example, the vertical distribution of *Gorgosaurus* throughout the 1.7 Myr Dinosaur Park Formation, in contrast to the rapid turnover (~300 ka) seen among ornithischians in that unit (Béland & Russell, 1978; Currie, 2005; Currie & Russell, 2005; Mallon et al., 2012; Sternberg, 1950). Although the proposal that there could be multiple species of *Tyrannosaurus* in the late Maastrichtian of the North American Western Interior is not inherently implausible, the question here is whether the case has been sufficiently demonstrated, the null hypothesis (i.e., that all specimens belong to the single species *T. rex*) appropriately rejected, and alternative hypotheses that could explain the variation observed, also rigorously tested and rejected.

In terms of the question of multiple species, the hypothesis of no difference (the simplest explanation) is the one to test. However, rejecting a hypothesis of no difference does not rule out alternative explanations for variation observed. A multitaxon hypothesis would only be defensible if discrete groups were recovered that could not be explained by hypotheses of ontogenetic, sexual, geographic, or individual variation (Brusatte et al., 2016). In other words, landing on a multispecies hypothesis is not so straightforward—it may be that differences observed between sister species might be masked by other sources of variation, as is seen in extant archosaurs (Bonnar et al., 2008; Mallon, 2017). In short, the multiple species hypothesis faces the same challenges as claims of sexual dimorphism in taxa where we lack the appropriate a priori knowledge.

The spectrum of “robust” and “gracile” sexual morphs in *T. rex* was first proposed by Carpenter (1990) based on ischium shape and form of the cervical column. He proposed that specimens with steeply posteroventrally extending ischia were females, whereas those with shallowly angled ischia represented males. He identified gracile and robust morphs based on a qualitative comparison of two cervical series. The association of a shallow ischium with a gracile cervical series led Carpenter (1990) to identify the robust specimens as female and the gracile specimens as male. Larson (1994) identified more qualitative features he thought corresponded to these robust and gracile morphs, and added the form and position of the first hemal arch as informative for differentiating sexes. Carpenter and Smith (2001) followed Larson (1994) in noting gross differences between humeri that might indicate sexual dimorphism. An effort to quantify sexual dimorphism in *T. rex* was done by Larson (2008a) where he compared the length and height of the lacrimal pneumatic foramen, and length and circumference or height of the humerus, ilium, femur, and metatarsal II. However, no quantitative analyses of the data were done to test the hypothesis of dimorphism (Larson, 2008a). Brochu (2003), Molnar (2005), Mallon (2017), and Carr (2020) offered skeptical assessments of sexual dimorphism in *T. rex*, concluding that variation was best explained by individual variation or growth. Using extant archosaurs as a model, Erickson et al. (2005) showed that hemal arch position and form are not useful in distinguishing male and female extinct archosaurs.

Paul et al. (2022) split *T. rex* into three species based on the evidence of stratigraphic position, robusticity, and the basal length ratio of the second and third dentary teeth. Relative to previous work, the authors have moved the “robust”—“gracile” spectrum from sexual dimorphism to species identity. In other words, the authors could have made the claim for discovering sexual size dimorphism in *T. rex*, but differences in mesial dentary tooth morphology and stratigraphic position persuaded them of a taxonomic

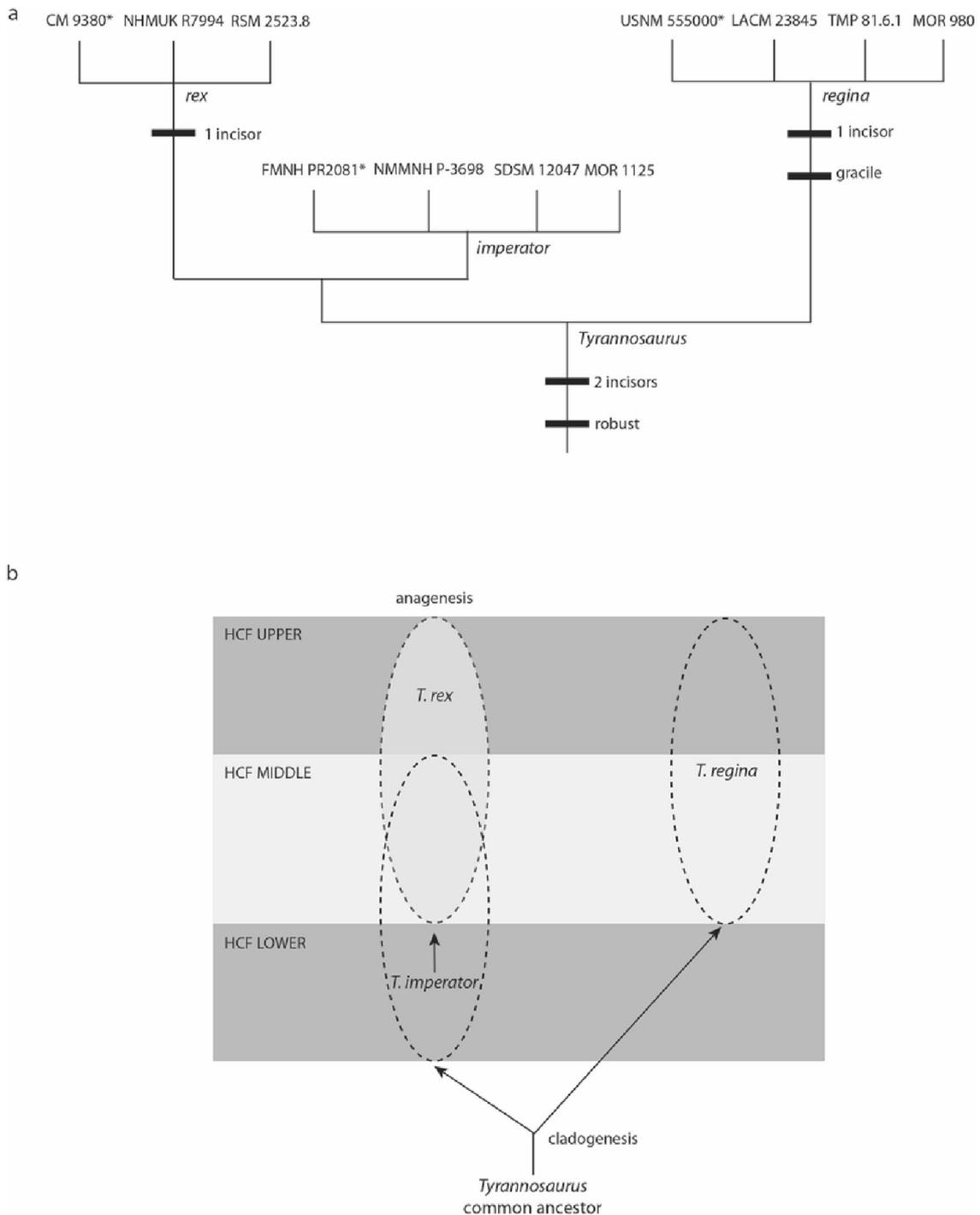


Fig. 1 The multiple species hypothesis of Paul et al. (2022) presented as **a** a dendrogram showing specimens, taxa, synapomorphies, and autapomorphies; and **b** a phylogeny showing the stratigraphic ranges of the taxa as ellipses, and the speciation modes of cladogenesis and anagenesis overlaid on the stratigraphic framework of the Hell Creek

Formation. Paul et al. (2022) did not independently test the trichotomy hypothesis by a numerical cladistic analysis; although the dendrogram drafted here is an arrangement it does accurately reflect the subjective schema of those authors. Specimens shown are limited to those in public trusts. Asterisks indicate type specimens

hypothesis. Taxonomic revisions in and of themselves are standard practice in paleontology, but for a species with as many good specimens as *T. rex*, any significant taxonomic

revision must be fully consistent with any and all available evidence. For example, an analysis of 1,850 variable characters from throughout the skull and skeleton for over 40

specimens of *T. rex* found that most variation in the species represents growth and individual variation (Carr, 2020).

Since reproductive isolation cannot be directly observed in fossil taxa, paleontology is limited to the Morphological Species Concept (MSC), where species are identified by the presence of one or more autapomorphies. Also in paleontology, chronospecies (unbranching chronologically successive species) are also morphologically defined. The schema of Paul et al. (2022) includes two chronospecies (*T. imperator* → *T. rex*) and one morphological species (*T. regina*). However, they did not independently test their schema with a cladistic analysis. In the chronospecies lineage, both taxa have “robust” femora, whereas the number of incisiform dentary teeth decreases from two to one. The morphological species is gracile and has one incisor, presumably derived independently from *T. rex*. In the absence of a cladistic analysis, it is unclear which states represent the plesiomorphic condition.

Here, we first discuss issues with the taxon diagnoses, sample size, and stratigraphic framework of Paul et al. (2022) and then we examine their three main results and conclusions: (1) “Skeletal robusticity is consistent within specimens but is not correlated with absolute size and presumed maturation,” (2) “*Tyrannosaurus* femoral proportions do show unusual variation,” and (3) “incisiform dentary tooth arrangement correlates with femoral robusticity and also appears to change with time.” Also at issue is that a significant number of specimens (16 of 37) are in commercial or private hands, which raises the issue of replicability of the results published in Paul et al. (2022).

Materials and Methods

Issues with Fuzzy C-Means and K-Means Clustering Used by Paul et al. (2022)

Paul et al. (2022) distinguish between “gracile” and “robust” morphotypes based on comparisons of various ratios of linear measurements of several cranial and postcranial bones; the most important of these are the ratio of femoral length to circumference, and the basal length of the third dentary tooth to the second tooth. Their justification of femoral robusticity as a taxonomically significant character in *T. rex* is deficient in both its statistical validity and broader contextualization. The authors used two statistical clustering algorithms—fuzzy C-means and K-means clustering—to justify a morphological “break” at a femoral robusticity of 2.4—specimens with a robusticity of less than this value were categorized as “robust”, and those with a greater value were categorized as “gracile”.

Paul et al. (2022) explicitly state that the justification for dividing specimens into two clusters was supported by these clustering algorithms. However, both fuzzy C-means and

K-means clustering define a desired number of clusters a priori, meaning that these procedures will find any number of clusters that the investigator seeks to justify. There are valid uses for these methods (for instance, attempting to classify specimens in a dataset with a known number of species), but they are unsuitable here. At best, their use makes this analysis circular; at worst, they can be used to fabricate a signal, because the desired number of groups will *always* be recovered. We do not accuse the authors of enforcing a desired result, but raise this point to state clearly that use of these clustering methods has the potential to be abused and should raise alarms for reviewers of future taxonomic work.

Rationale for Agglomerative Hierarchical Clustering Used in this Study

Agglomerative hierarchical clustering is much better suited to the question at hand, as it both recovers a dendrogram showing the relative closeness of all samples and can be used to determine the optimal number of clusters. We used this approach to test the author’s published data and results.

Issues with Taxonomic Comparisons of Paul et al. (2022)

Paul et al. (2022) further assert that the femur of *Tyrannosaurus* shows unusual proportional variation, but this statement is based on an improper comparative sample and is demonstrably false. Part of the comparative sample is based upon *Allosaurus* sp. from the Cleveland-Lloyd Dinosaur Quarry (CLDQ), which represents a highly temporally and geographically constrained death assemblage (on the order of decades; Gates, 2005) interpreted to reflect population-level differences of a single species (Carpenter, 2010). Also, the CLDQ assemblage is dominated by subadults (Gates, 2005), and so the *Allosaurus* sample might not represent the full range of variation for adults. In contrast, the included specimens of *Tyrannosaurus* come from a wide geographic distribution (e.g., Alberta to Wyoming) and a longer temporal interval—therefore, they are not one biological population, and should be expected to show more variation even if the only sources of variation are geographic and temporal clines. The comparative sample is further provided by other tyrannosaurids, but each is represented by fewer specimens than *Tyrannosaurus*, and therefore they should also be expected to show lower variation.

There are two further reservations to note here. The first is that all taxa used for comparison are themselves extinct, and therefore themselves of unknowable taxonomic status. While there is no data to suggest that, e.g., *Albertosaurus sarcophagus* is chimeric, it is theoretically possible for any and all of these taxa to contain multiple biological species.

Furthermore, the comparisons drawn concern the percent difference between femoral robusticity scores, which themselves are normalized data calculated by dividing femoral circumference by femoral length. This procedure normalizes normalized data and may artificially inflate perceived differences. Because the data are already normalized to a dimensionless unit, differences between taxa can be directly compared. A better comparison is modern avian theropods of known taxonomic status.

Rationale for Taxonomic Comparisons in this Study

We rectified these reservations through analysis of a comprehensive dataset of osteological measurements of extant birds published by Field et al., (2013; Supp. Info. 1). Following Paul et al. (2022), femoral robusticity was calculated by dividing femoral circumference by femoral length. After pruning the dataset of species represented by fewer than three individuals, it consists of 112 species of living birds (represented by 613 specimens) and 4 species of theropods (*T. rex*, *Tarbosaurus bataar*, *A. sarcophagus*, and *Gorgosaurus libratus*; represented by 52 specimens). We also applied agglomerative hierarchical testing to their dentary tooth data to test for the clusters that they recovered.

Results and Discussion

Synonymy of “*T. imperator*” and “*T. regina*” with *T. rex*

The multiple species hypothesis of Paul et al. (2022) is based on unequal character weighting, where a small set of presumed diagnostic characters was not rigorously tested. The splitting of one taxon into several must be supported by an independent test of all of the available evidence; Paul et al. (2022) have based their hypothesis on a vanishingly small subset of the available data and on the assumption of discrete robust and gracile morphs. These and other issues with the multiple species hypothesis are discussed and tested below. Although Paul et al. (2022) discussed the alternative hypotheses of sexual dimorphism, ontogeny, and individual variation, we focus our attention on the evidence they give for the multiple species hypothesis.

Diagnoses

The diagnoses of the three species of *Tyrannosaurus* are limited to the form of the femur and the first two dentary teeth (Paul et al., 2022). The diagnosis of each species is vague and they overlap with each other. For example, *T. rex* is “generally robust...[femoral length to circumference] ratio of about 2.4 or less...usually one...incisiform dentary tooth”

(Paul et al., 2022). The qualifiers “generally” and “usually” imply overlap between the species, which undermines the utility of the diagnoses and defies taxonomic best-practices. Indeed, the basal crown ratios do overlap between taxa in the case of RSM 2523.8 (see below). Also, the authors do not have a clear definition or diagnosis of the clade *Tyrannosaurus*. This is an issue because the diagnoses of the various species of *Tyrannosaurus* could apply to any tyrannosaurid if their femur ratios and number of small mesial dentary teeth happen to match. In light of these issues, we provide a diagnosis of *T. rex* here.

Systematic Paleontology

Tyrannosauridae.

Tyrannosaurus Osborn (1905)

Tyrannosaurus rex Osborn (1905)

Dynamosaurus imperiosus Osborn (1905)

Gorgosaurus lancensis Gilmore (1946)

Gorgosaurus megagracilis Paul (1988)

Tyrannosaurus imperator Paul et al. (2022).

Tyrannosaurus regina Paul et al. (2022).

Tyrannosaurus incertae sedis Paul et al. (2022).

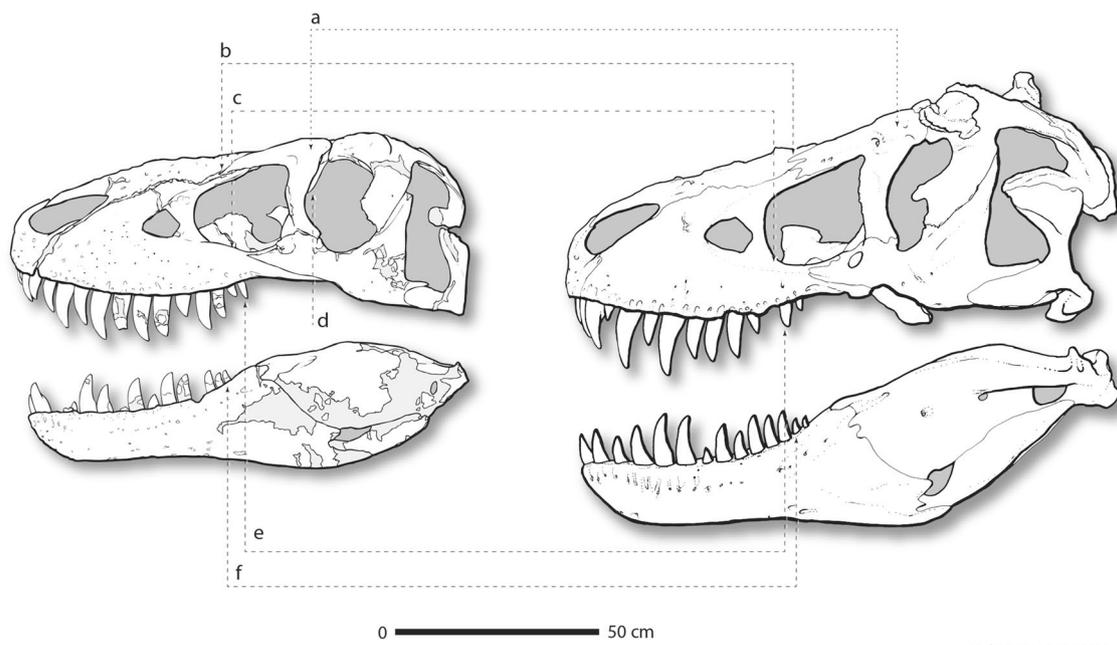
Tyrannosauridae incertae sedis Paul et al. (2022).

Type species: *Tyrannosaurus rex* Osborn (1905).

Diagnosis

Tyrannosaurus rex can be distinguished from its sister species *Ta. bataar* on the basis of: giant body size in skeletally mature individuals where the femur length is 1.2 m or greater; orbitotemporal region of the skull is wide (snout width to temporal width measured across quadratojugals jugals is 53% or less); presence of a rostradorsal process on the rostral ramus of the lacrimal and a corresponding caudolateral process of the nasal, producing an interlocking suture between the bones in lateral view; absence of a lateral swelling on the supraorbital process of the lacrimal; absence of a subcutaneous flange along the ventral margin of the antorbital fossa (= external antorbital fenestra); short ventral process of the prefrontal; absence of a deep ventral flange on the vomer; a minimum of 11 maxillary and 12 dentary teeth in skeletally mature individuals; long arm (humerus to femur ratio is 29%) and metatarsus relative to femur length (metatarsal III to femur ratio is 51%) (Currie, 2003; Fig. 2; Table 1; Brochu, 2003).

Another issue is the authors could not identify to the species level several well-preserved and nearly complete skulls (AMNH FARB 5027, LACM 23844, MOR 008, UWBM 99000) that they list as incertae sedis. We concede that one of the *incertae sedis* specimens (AMNH FARB 5027) includes both dentaries, but it is contained in a glass case,



© Thomas D. Carr 2022

Fig. 2 Comparison of the skulls of *T. rex* (right) and *Ta. bataar* (left) showing several species differences in the supraorbital process (**a**) and rostral ramus (**b**) of the lacrimal, horizontal ramus of the maxilla

(**c**), ventral ramus of the prefrontal (**d**), and maxillary (**e**) and dentary (**f**) tooth counts. See text and Table 1 for details

Table 1 Summary of differences that distinguish *T. rex* from its sister species, *T. bataar*

	<i>Tyrannosaurus rex</i>	<i>T. bataar</i>
Body size, proxy: femur length	Exceeds 1.30 m	Maximum 1.12 m (Currie, 2003)
Skull, orbitotemporal region, width	Wide	Narrow
Nasal, caudolateral process, presence	Present	Absent
Maxilla, subcutaneous flange, presence	Absent	Present
Lacrimal, rostradorsal process, presence	Present	Absent
Lacrimal, supraorbital process, lateral bulge, presence	Absent	Present
Prefrontal, ventral process, length	Short	Long
Vomer, ventral flange, presence	Absent	Present
Maxillary teeth, minimum number among skeletally mature individuals	11	12
Dentary teeth, minimum number among skeletally mature individuals	12	15
Humerus, antebrachium, metacarpus, metatarsus—length relative to femur	Long	Short

The size-independent features are also seen in juveniles and so, if they were sympatric, the two species would be easily identified

making it inaccessible to measurement. In contrast, several incomplete specimens, including one without a skull (TCM 2001.90.1), were identified to species. The authors state these unidentifiable specimens are of uncertain proportions or stratigraphic position.

Although any taxonomic revision can cause issues, the scientific priority, following the spirit of the ICZN, is to make the taxonomy stable and practicable. Following that principle, the multiple species hypothesis causes more

uncertainty than clarity: Of the 53 specimens listed in the taxonomic section, only five are confidently assigned to *T. rex*, 11 to “*T. imperator*” and seven to “*T. regina*.” A further three are *T. rex*?, one “*T. imperator*”?, four *T. rex* or “*T. imperator*”, seven *T. rex* or “*T. regina*”?, five cannot be determined and seven are listed as incertae sedis. So 30 specimens are uncertain and only 23 are assigned to one of the three species. If the commercially-held BHI specimens are excluded, this worsens to only

13 confidently assigned and 25 not—so two thirds of the specimens are of uncertain identity. In effect, their taxonomic revision results in instability: the diagnoses take many specimens that, until now, were easily classified as *T. rex* and render them uncertain.

Also, several juvenile and subadult specimens are regarded as incertae sedis; based on what is seen in *Ta. bataar*, at least one species-level autapomorphy is seen throughout growth in this sister taxon (e.g., subcutaneous flange of the maxilla; Tsuihiji et al., 2011). If species characters also appear early in growth in the *T. rex* hypodigm (and in fact, the autapomorphic skull shape of *T. rex* is emergent in juveniles—Carr, 2020), then the best preserved fossils of juveniles (BMRP 2002.4.1, CMNH 7541, LACM 238471) should be referable to any one of the three taxa. To test this hypothesis, we obtained the ratio of dentary teeth 3:2 in BMRP 2002.4.1 and compared that with the ratios in the three taxa. The ratio is 1.6, which supports referral to *T. imperator* if the presence of a slender second dentary tooth is a reliable diagnostic character that emerges early in growth (but see below). Alternatively, we take the point of Wiens et al. (2005) that diagnostic characters tend to emerge late in growth; i.e., if all young *T. rex* have slender second dentary teeth, it is improbable that this is a diagnostic character of *T. imperator*; rather, it is the juvenile state of the character that later changes (in this case, the basal tooth length increases) as the animal matures.

If a species diagnosis is to have any value, then every specimen that includes a complete skull, but not the femur, or vice versa, should be identifiable even if the other relevant information is missing. In this case, the diagnostic features are limited to one cranial character (dentary tooth 3:2 ratio) and one postcranial character (femur 1:c ratio), which taken separately are not sufficient to distinguish the taxa. For example, isolated dentaries of *T. rex* and *T. regina*, and isolated femora of *T. rex* and *T. imperator*, would be confused for each other since their ‘diagnostic’ features are identical. It remains to be established if the diagnostic characters, as set out by the authors, are invariant throughout growth; if they are invariant then their case for multiple species might be strengthened. If stratigraphic data are necessary to identify taxa, as is expressed by the authors, then the identification is circular if morphology is not taken into account (Álvaro & Esteve, 2020). For example, the presumed stratigraphic separation between the troodontids *Laetia navenatrix mcmasterae* and *Stenonychosaurus inequalis* was the basis of referring specimens to one or the other taxon (van der Reest & Currie, 2017), until it was found that they are not stratigraphically (or morphologically) separate and represent the same taxon, *S. inequalis* (Cullen et al., 2021).

Sample Size

In terms of statistically detecting dimorphism, sample size is problematically low in Paul et al. (2022). Only 25 specimens were used in the regression analysis of femoral ratios, and only 6 or 8 specimens were used in the analyses of each “species.” Modeling has shown that sample sizes of 100 specimens might not detect sexual differences as great as 28% (Godfrey et al., 1993 in Mallon, 2017). Modeling also shows that small sample size ($n = < 50$) affects the precision of quantitatively estimating sexual dimorphism (Kościński & Pietraszewski, 2004). Hone and Mallon (2017) suggest that a sample size of at least 35 specimens of each sex is required for sexual size differences to be statistically detectable, assuming a high alligator level of dimorphism.

The diagnoses of the three species are based on two morphological characters, femoral robusticity and the ratio of two dentary teeth (Paul et al., 2022). This is not to say that naming fossil taxa on a limited number of traits is inherently incorrect or flawed; when data are limited because remains are fragmentary, this is understandable in context. However, this low number of diagnostic characters stands in contrast to the numerous well-documented specimens of *Tyrannosaurus* available and 1850 anatomical traits that have been scored for more than 40 specimens (Carr, 2020). Furthermore, the tooth count of individual cranial bones is known to be variable in multiple theropod groups, including tyrannosaurines (Carr et al., 2017), and so selecting the presence or absence of a single dentary tooth is a trait that varies intra- and interspecifically is not a strong character for the basis of taxonomic identity (Carr, 2020).

Stratigraphic Framework

The hypothesis of Paul et al. (2022) depends on the precise and accurate stratigraphic position of the specimens in their study. However, these data are based on Larson (2008b), personal communication, and a meeting abstract (Kaskes et al., 2016). The data in Larson (2008b) are imprecise (several are personal communications, or vague, e.g., from the “lower half” of the formation) and the personal communication is taken at face value; this renders independent verification of these stratigraphic assessments difficult to test by other researchers. Precise stratigraphic data are required to put fossils into a robust framework (e.g., Scannella et al., 2014). The stratigraphic framework itself is vague, where the Hell Creek Formation is blocked out into lower, middle, and upper thirds by Paul et al. (2022)—an approach of insufficient resolution to provide a result comparable to the *Triceratops* speciation study (Scannella et al., 2014). This is not to say that taxa cannot be separated into broad stratigraphic bins that are taxonomically informative and laterally

continuous (Mallon et al., 2012), rather, the stratigraphic data of specimens must be reliably accurate.

Additionally, Paul et al. (2022) did not demonstrate that their “lower, middle, and upper” subdivisions of the Hell Creek and its regional stratigraphy equivalents were comparable to previous well-established stratigraphy of these units (Eberth & Kamo, 2019; Hartman et al., 2002; Wilson et al., 2014). The Hell Creek represents a formation in which considerable research effort has already established a geochronology and zonation based on lithology, macrofloral, and palynological data (Fastovsky & Bercovici, 2016), paleomagnetism, and radioisotopic dating (Hicks et al., 2002): Ideally, any research program focusing on evolutionary changes within this unit should be addressed within this context if it is to be integrated into the overall picture of environmental transformations through the late Maastrichtian Western Interior.

As noted above, the authors were unable to refer several excellent specimens to any one of the three taxa. In some cases they were unable to narrow down identifications owing to “uncertain stratigraphic position” or, conversely, they ambiguously narrowed down identifications to two taxa based on “high stratigraphic placement” (Paul et al., 2022). However, the taxonomic identification of a specimen must be independent of stratigraphic position or the identification becomes circular (Álvarez & Esteve, 2020). Instead, unique and derived anatomical features should be used to identify specimens.

“Skeletal Robusticity is Consistent Within Specimens But is Not Correlated with Absolute Size and Presumed Maturation”

We applied agglomerative hierarchical clustering to the data presented by Paul et al. (2022) using unweighted pair-group arithmetic average (UPGMA) agglomerative hierarchical clustering implemented via the R package ‘cluster’ (Maechler et al., 2022). We recovered a dendrogram (Fig. 3a) that mostly matches their three putative species, but importantly also indicates that the optimal number of clusters in the dataset is one (Fig. 3c). In other words, there is indeed structure in the dataset (as in any biological dataset), but there is no statistical justification for assigning the specimens to different species corresponding to this structure.

“Tyrannosaurus Femoral Proportions Do Show Unusual Variation”

Calculating the robustness of each extant avian theropod yields a mean intraspecific femur robusticity range of 0.66 ± 0.35 —greater than the absolute variation observed in *T. rex* (0.61) and every other tyrannosaurid included (Fig. 4). The median robustness range is slightly smaller at

0.59. While it is true that *T. rex* has higher variation in femur robusticity than the other three tyrannosaurids, a taxonomically controlled and statistically large sample shows it to be unexceptional—in fact, some modern birds are over three times as variable in this character. The undesirable statistical properties of ratios have been known for decades (Atchley et al., 1976) – namely, ratios change the structure of the underlying dataset and can create spurious correlations. To ensure that this did not bias our results, we regressed femoral circumference against femur length for all 112 species of theropods in our dataset, and computed the R-squared value for each regression as a metric of dispersion. As in the ratio data, *T. rex* resolved as having a degree of variation that is entirely typical of modern avian theropods (Suppl. Info. 2) under both ordinary least-squares and standardized major axis regression. Therefore, the argument that the range of robusticity scores for *T. rex* is abnormal (and thus suggestive of cryptic diversity) can be rejected.

“Incisiform Dentary Tooth Arrangement Correlates with Femoral Robusticity and Also Appears to Change with Time”

The number of mesial incisiform teeth was used by Paul et al. (2022) in their diagnoses of the three taxa. Specifically, they considered the ratio of the “diameter” of the third tooth to the second tooth as taxonomically meaningful. Although not defined in the article, “diameter” corresponds to the mesiodistal basal crown length (WSP, personal communication to TDC, March 15, 2022). The authors considered a ratio of tooth 3:tooth 2 of 1.2 or greater as indicating the second tooth is small enough to be considered a second incisor; i.e., a ratio closer to (or less than) 1.0 indicates a large second tooth that is not an incisor. However, comparisons of the mesial dentition between specimens is not straightforward: (1) owing to ontogenetic tooth loss the homology of mesial tooth identity between specimens is unclear, (2) it is not clear if the measurements used by Paul et al. (2022) are from successive teeth from the same side, opposite sides, or are taken from tooth sockets, or some combination of the three, and (3) it is not clear if the 1.2 ratio is a statistically meaningful threshold.

Among adult *T. rex*, the number of dentary teeth is variable and ranges from 12 to 14 teeth (Carr, 2020). In adults, it appears that the first tooth is lost completely and is represented by a small divot at the rostradorsal corner of the dentary ahead of the first open alveolus (Carr, 2020; Fig. 5). If the divot is correctly identified, then the first three open alveoli represent sockets two to four, not one to three. If the loss of mesial alveoli accounts for the difference in number, then it is difficult to establish the homology between the mesial teeth of specimens with different tooth counts. Therefore, what appears to be the “second” tooth position in two

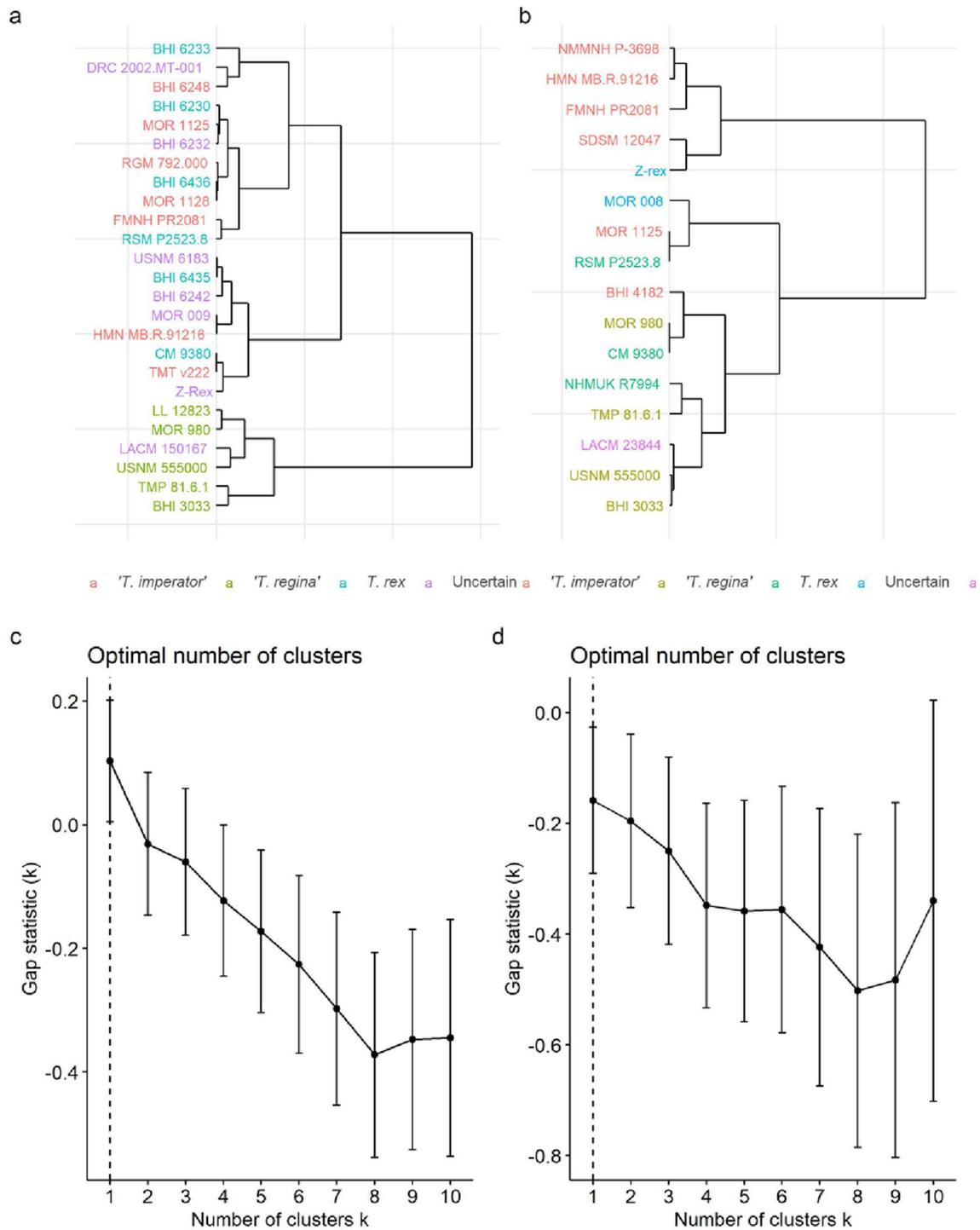


Fig. 3 Results of hierarchical clustering analysis, including **a** a dendrogram recovered through UPGMA agglomerative hierarchical clustering of femoral robusticity values from Paul et al. (2022); **b** an analogous dendrogram recovered through clustering analysis of dentary tooth ratio data; **c** a gap statistic plot showing the optimal number of clusters to describe femoral robusticity; **d** a gap statistic plot similarly

showing the optimal number of clusters for describing tooth ratios. The taxonomy proposed by Paul et al. (2022) is generally similar to these results, although *T. rex* and '*T. imperator*' do not form distinct morphological clusters, and clusters are not congruent between the proposed diagnostic characters. The gap statistic plots both indicate that a one-cluster model is the best supported

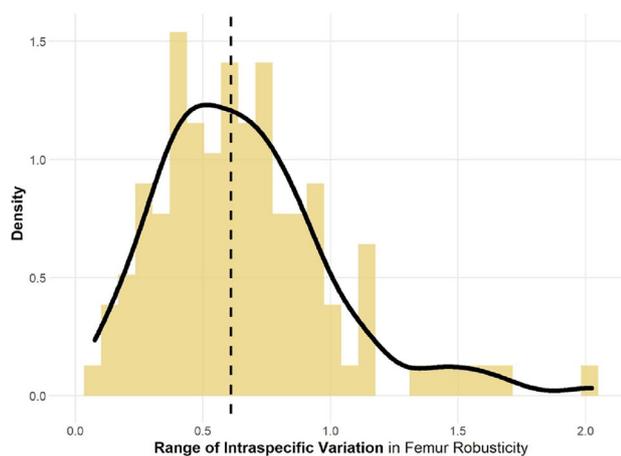


Fig. 4 Histogram and density function describing the range of intraspecific variation seen in modern theropods and four tyrannosaurids. The vertical dotted line corresponds to the variation in robusticity seen in *T. rex*, which is well within the range of variation seen in living birds; the dark curve is the kernel density

specimens might really be the third tooth in one specimen (first mesial tooth lost developmentally) but the fourth in another (first two mesial teeth lost developmentally). Ergo, the ratios might be incomparable because the teeth are developmentally nonhomologous.

It is worth pointing out that in their Table 2, the corresponding columns of tooth diameter measurements are almost certainly transposed where, in most cases, the diameters of the second tooth exceed those of the third tooth. In *Tyrannosaurus*, as in all other tyrannosaurids, dentary tooth size increases sequentially from the first tooth to the fourth or fifth position; therefore, the large measurements under the heading “2nd Dentary tooth base diameter” are really of the third tooth and vice versa. Also, the ratio of the second to third teeth are listed under the heading “Hum ratio,” which implies “Humerus ratio” (Paul et al., 2022).

After rectifying these issues, we tested their results by comparing the measurements we have for the same specimens. We found that in several cases teeth were incomplete (not indicated by the authors), were completely missing, leaving an empty socket (not indicated by the authors), and we found that all of our corresponding measurements were lower than what the authors published (Table 2). Of note is that they did figure a specimen that is missing the first three teeth (Paul et al., 2022: Fig. 3a), but they give tooth measurements for it in their Table 2. Therefore, we conclude that the tooth data presented in Table 2 of Paul et al. (2022)—in the case of the figured specimen and those for which we have data—actually represent measurements of teeth from opposite sides, empty alveoli, or some combination of each (Table 2).

In the one case where we were able to obtain the dentary tooth 3:2 basal crown length ratio for sequential teeth of the same side, the result was significantly different. Paul et al. (2022) obtained a ratio of 0.98 (indicating one incisor) for MOR 980, which they identified as “*T. regina*” based in part on the ratio; in contrast, we obtained a ratio of 1.26 (indicating two incisors since the third tooth is significantly larger than the second tooth; Table 2), which in their schema would identify it as a “gracile” “*T. imperator*.”

When their tooth ratio data are arranged in sequential order, the ratios are continuous; it appears that the 1.2 ratio is an arbitrary threshold that was not based on a biological discontinuity. Consistent with this observation are two specimens, MOR 1125 and RSM 2523.8, which are referred to “*T. imperator*” and *T. rex*, respectively. Although their incisor count is not given, based on their high tooth ratios (Paul et al., 2022, Table 2) they have two incisors, which is diagnostic, in part, for “*T. imperator*,” and so the identification of MOR 1125 is consistent with the diagnosis. However, the presence of two incisors in RSM 2523.8 conflicts with its diagnosis as *T. rex*, which should only have one incisor. Accordingly, the authors moderate their identification of this specimen with a question mark. The discrepancy of RSM 2523.8 might account for the qualifying nature of the taxon diagnoses, where specific incisor counts are described as “usually” occurring. This issue is emphasized by the agglomerative clustering results where the dendrograms recovered for the femoral and dentary tooth data do not match, indicating character discrepancy (Fig. 3).

Finally, inspection of their diagram of incisors (Paul et al., 2022, Fig. 3) shows one specimen with empty tooth sockets and another with the first two teeth erupted. The exemplar of two small incisors (Paul et al., 2022, Fig. 3a) shows tooth sockets that are nearly the size of the third, whereas the exemplar of one small incisor (CM 9380) indeed shows a first tooth (and alveolus) that is truly a fraction of the size of the second and third alveoli that are similar in size to each other (Paul et al., 2022, Fig. 3B). Taken at face value, the first example simply has large mesial teeth, whereas the second does indeed have a small first tooth. Therefore, as presented by Paul et al. (2022), the conflicting variation of the mesial dentary teeth data are unresolved and therefore should be considered uninformative for diagnosing taxa.

Agglomerative Hierarchical Clustering Test of Basal Crown Length Ratios

Despite the issues discussed above, we tested their tooth ratio data (Paul et al., 2022, Table 2) using agglomerative hierarchical clustering to determine the optimal number

Fig. 5 The mesial end of the tooth row of the right dentary of the holotype of *T. rex* (CM 9380) showing the divot that represents the bone-filled first dentary alveolus

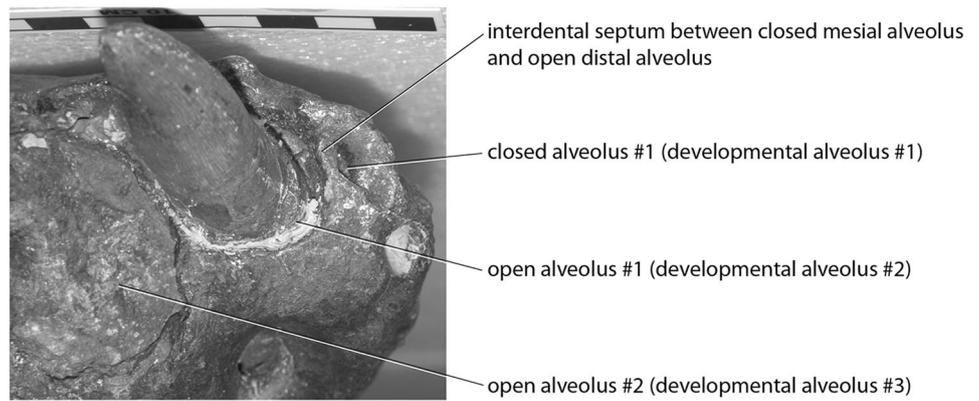


Table 2 Comparison of measurements of the “tooth base diameter” in Paul et al. (2022) with our basal crown length data for several specimens

Specimen	2nd tooth (Paul et al., 2022)	2nd tooth (this paper)	3rd tooth (Paul et al., 2022)	3rd tooth (this paper)	3:2 ratio (Paul et al., 2022)	3:2 ratio (this paper)
RSM 2523.8 teeth	38	34.6 [L] 37.8 [R]	48	? [L] 34.5 + [R]	1.26	? (?)
RSM 2523.8 alveoli	n/a	36.1 [L] 38.9 [R]	n/a	45.9 [L] 52.2 [R]	n/a	1.27 [L] 1.34 [R]
CM 9380 teeth	55	? [L] 38.1 [R]	54	44.6 [L] ? [R]	0.98	?
CM 9380 alveoli	n/a	? [L] 32.8 [R]	n/a	? [L] 50.8 [R]	n/a	1.55 [R]
NHMUK R7994 teeth	42	38.8 [L] ? [R]	50	? [L, R]	1.19	?
NHMUK R7994 alveoli	n/a	49.8 [L] ? [R]	n/a	49 [L] ? [R]	n/a	0.98 [L]
MOR 980 teeth	52	37.8 [L] 33.7 [R]	51	? [L] 42.3 [R]	0.98	1.26 [R]
LACM 23,844	45	36.5 [L]	49	? [L]	1.09	?

All measurements in millimeters. As published, Paul et al., (2022, Table 2) reversed the measurement columns, which we have corrected here. All measurements reported by Paul et al. (2022) were stated as measurements of teeth; however, we have found in many instances that the second and third tooth are often missing and so the measurements of Paul et al. (2022) are either measurements of alveoli or of teeth from opposite sides. We find that tooth lengths of opposite sides are different and that alveoli lengths exceed tooth lengths, making composite ratios unreliable. This table shows (1) the extreme difference in measurements between Paul et al. (2022) and our measurements of the same specimens, (2) tooth measurements are given by Paul et al. (2022) for empty sockets, indicated in our data with question marks (“?”), (3) the side a tooth measurement was taken is not given by Paul et al. (2022), (4) the difference between basal tooth length and alveolus length when both measurements can be compared, and (5) the difference in tooth ratio between Paul et al. (2022) and ours for MOR 980

of clusters. We found that the optimal number of clusters in the dataset is one (Fig. 3b, d). Based on their data, there is no statistical justification for assigning the specimens to different clusters.

Replicability

Nearly half of the specimens (16 out of 37) used in the study of Paul et al. (2022) are held by commercial fossil companies or are privately owned. The practice of obtaining data from fossils that are not in public trusts is strongly discouraged to ensure replication of observations, as set out in the ethics guidelines of the Society of Vertebrate Paleontology and the instructions to authors of the Journal of Vertebrate Paleontology, the professional standard-bearers for vertebrate paleontologists.

The recent auction of “Stan,” a fossil held by a commercial company until its recent auction to a private collector in October, 2020 (Greshko, 2022; Vogel, 2020), emphasizes the fact that commercial stockrooms are not on the same footing as recognized public trusts. All of the commercially-held fossils are also vulnerable to sale and the privately owned fossils are simply off limits as a matter of scientific practice. The fact that so many of the fossils in the study of Paul et al. (2022) are not in public trusts is highly problematic because their results are not replicable unless the ethics of vertebrate paleontology are transgressed, and in some cases, like when specimens disappear into a private collection, not at all.

Recommendations and Caveats

For a case such as *Tyrannosaurus* with a good number of well-preserved specimens we propose the following procedure to test the hypothesis that one fossil species might really be several:

1. Clearly state the null and alternative hypotheses, and give the rationale for the alternative. In the case of a multiple species hypothesis, the null is one species that lacks sexual dimorphism. Alternatively, if sexual dimorphism is assumed to be plesiomorphic for dinosaurs, then the null would be to expect dimorphism. Untested “key” taxonomic characters are identified in the rationale for the alternative hypothesis; ultimately, all available evidence must be brought to bear on testing the hypothesis.
2. Complete a quantitative cladistic analysis of variation of individual specimens to test for the presence of two or more synapomorphy- (or synontomorphy-) based groups. A linear cladogram (or ontogram) indicates the test of multiple groups has failed; alternatively, recovery of several distinct groups indicates support. All available morphological evidence must be brought to bear on the analysis; untested key “taxonomic” characters must be avoided.
3. Complete a rigorous statistical analysis of a large sample measurement data to test the hypothesis of quantitative (group) differences in the data; the agglomerative hierarchical test is appropriate here.
4. Obtain precise (meter-level) stratigraphic data to set the specimens in a rigorous geochronological framework incorporating previously established geochronologic, lithostratigraphic and biostratigraphic data to independently test that hypothesized sympatric taxa are spatiotemporally discrete (note that this is supportive evidence for taxonomic separation and should not be used to justify taxonomic constructs, which must be rigorously justified morphologically or molecularly). This approach would distinguish chronospecies (= anagenetic lineages) from morphological species produced by cladogenesis.
5. When sympatric species and sexual dimorphism are not known a priori, they can be confused for each other; clear criteria must be established to distinguish between them. In the case of size dimorphism in *T. rex* considered as evidence for sexual dimorphism on the one hand, or for species differences on the other, we offer two heuristics to assess such cases of ambiguity: Size dimorphism is the best evidence for sexual dimorphism, in the absence of discrete morphological character differences. Size dimorphism is the best evidence for species dimorphism, in the presence of discrete morphological

traits throughout growth that differentiate two or more species.

Conclusions

Paul et al. (2022) have proposed a complex hypothesis based on three lines of weak evidence: Three species of *Tyrannosaurus* in the Hell Creek Formation and its lateral equivalents. In contrast, an assessment based on previous work (Carr, 2020; Mallon, 2017) and the contributions here, there is support for a simpler hypothesis: One species of *Tyrannosaurus* where the variation among adults is best explained by growth and individual variation; cryptic sexual dimorphism is not necessarily ruled out. We conclude:

1. Paul et al. (2022) have not provided sufficient evidence to support the hypothesis that *T. rex* can be split into three taxa.
2. There is no current evidence to suggest that more than one tyrannosaurid was present in Laramidia during the last ~ 1 million years of the Cretaceous. The short time duration of herbivorous dinosaurs (~ 300 ka; Mallon et al., 2012) is not seen in tyrannosaurids of other units, which are invariant (aside from ontogeny) for up to 1.7 Myr (Currie, 2005; Mallon et al., 2012). Therefore, it is not unreasonable to expect *T. rex* to be invariant for ~ 1 Myr.
3. New sympatric taxa must be diagnosed by unique and exclusive morphological characteristics appropriate to the specimens in hand.
4. A robust and high-resolution geochronologic and stratigraphic framework is required for a hypothesis of speciation for short time spans such as that represented by the Hell Creek Formation and lateral equivalents.
5. Scientists need to do their due diligence and vet the collections where specimens are stored; specimen availability (i.e., curation in a recognized institution) is essential for reproducibility and continued testing of taxonomic hypotheses.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11692-022-09573-1>.

Acknowledgements The authors thank W. Scott Persons (College of Charleston) for discussion. We thank Jordan Mallon and an anonymous referee for their comments which helped improve the manuscript and Campbell Rolian for their work as editor in handling the paper.

Author Contributions First draft written by SLB, TDC, TEW. Draft review, revisions, additions, and editing by SLB, TDC, TRH, DWEH, JGN, TEW, LEZ. Data collection by TDC and JGN. Data analysis by JGN. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Funding

acquisition not applicable. Figures 1, 2, and 5 by TDC; Figs. 3 and 4 by James Napoli.

Funding Not applicable: This work was not grant-supported.

Data Availability All data analyzed and generated during this study are included in this published article and its supplementary information file.

Declarations

Conflict of interest The authors declare that they have no financial or non-financial interests that are directly or indirectly related to this work. The authors declare that they have no conflict of interest.

Ethical Standards The authors declare that this work complies with the current laws of the country in which they were done. This work contains analyses of the data in Paul et al. (2022) that includes data from privately owned and commercially-held specimens.

Research Involving Animals, their Data or Biological Material Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Álvarez, J. J., & Esteve, J. (2020). Reply to Comment on: Álvarez J.J., Esteve, J. & Zamora, S. 2019. Morphological assessment of the earliest paradoxid trilobites (Cambrian Series 3) from Morocco and Spain [*Geological Magazine*] by Geyer G, Nowicki J, Żylińska A & Landing E. *Morphological Assessment of the Geological Magazine*, 157, 1971–1982. <https://doi.org/10.1017/S0016756820000217>
- Atchley, W. R., Gaskins, C. T., & Anderson, D. (1976). Statistical properties of ratios. I. Empirical Results. *Systematic Zoology*, 25, 137–148.
- Béland, P., & Russell, D. A. (1978). Paleocology of Dinosaur Provincial Park (Cretaceous), Alberta, interpreted from the distribution of articulated vertebrate remains. *Canadian Journal of Earth Sciences*, 15, 1012–1024.
- Benson, R. B. J., Brown, C. M., Campione, N. E., Cullen, T. M., Evans, D. C., & Zanno, L. E. (2022). Comment on “The influence of juvenile dinosaurs on community structure and diversity.” *Science*. <https://doi.org/10.1126/science.abj5976>
- Bonnan, M. F., Farlow, J. O., & Masters, S. L. (2008). Using linear and geometric morphometrics to detect intraspecific variability and sexual dimorphism in femoral shape in *Alligator mississippiensis* and its implications for sexing fossil archosaurus. *Journal of Vertebrate Paleontology*, 28, 422–431.
- Brochu, C. A. (2003). Osteology of *Tyrannosaurus rex*: Insights from a nearly complete skeleton and high-resolution computed tomographic analysis of the skull. *Society of Vertebrate Paleontology Memoir*, 7, 1–138.
- Brusatte, S. L., Carr, T. D., Williamson, T. E., Holtz, T. R., Jr., Hone, D. W. E., & Williams, S. A. (2016). Dentary groove morphology does not distinguish ‘*Nanotyrannus*’ as a valid taxon of tyrannosaurid dinosaur. Comment on: “Distribution of the dentary groove of theropod dinosaurs: Implications for theropod phylogeny and the validity of the genus *Nanotyrannus* Bakker et al. 1988”. *Cretaceous Research*. <https://doi.org/10.1016/j.cretres.2016.02.007>
- Carpenter, K. (1990). Variation in *Tyrannosaurus rex*. In K. Carpenter & P. J. Currie (Eds.), *Dinosaur systematics: Approaches and perspectives* (pp. 141–145). University of Cambridge Press.
- Carpenter, K. (2010). Variation in a population of Theropoda (Dinosauria): *Allosaurus* from the Cleveland-Lloyd Quarry (Upper Jurassic), Utah, USA. *Paleontological Research*, 14(4), 250–259.
- Carpenter, K., & Smith, M. (2001). Forelimb osteology and biomechanics of *Tyrannosaurus rex*. In D. H. Tanke & K. Carpenter (Eds.), *Mesozoic vertebrate life* (pp. 64–83). Indiana University Press.
- Carr, T. D. (2020). A high-resolution growth series of *Tyrannosaurus rex* obtained from multiple lines of evidence. *PeerJ*, 8, e9192.
- Carr, T. D., Varricchio, D. J., Sedlmayr, J. C., Roberts, E. M., & Moore, J. R. (2017). A new tyrannosaur with evidence for anagenesis and crocodile-like facial sensory system. *Scientific Reports*, 7, 44942.
- Cullen, T. M., Zanno, L., Larson, D. W., Todd, E., Currie, P. J., & Evans, D. C. (2021). Anatomical, morphometric, and stratigraphic analyses of theropod biodiversity in the Upper Cretaceous (Campanian) Dinosaur Park Formation. *Canadian Journal of Earth Sciences*, 58, 870–884.
- Currie, P. J. (2003). Allometric growth in tyrannosaurids (Dinosauria: Theropoda) from the upper Cretaceous of North America and Asia. *Canadian Journal of Earth Sciences*, 40, 651–665.
- Currie, P. J. (2005). Theropods, including birds. In P. J. Currie & E. B. Koppelhus (Eds.), *Dinosaur Provincial Park: A spectacular ancient ecosystem revealed* (pp. 367–397). Indiana University Press.
- Currie, P. J., & Russell, D. A. (2005). The geographic and stratigraphic distribution of articulated and associated dinosaur remains. In P. J. Currie & E. B. Koppelhus (Eds.), *Dinosaur Provincial Park: A spectacular ancient ecosystem revealed* (pp. 537–569). Indiana University Press.
- Eberth, D. A., & Kamo, S. L. (2019). First high-precision U-Pb CA-ID-TIMS age for the Battle Formation (Upper Cretaceous), Red Deer River Valley, Alberta, Canada: Implications for ages, correlations, and dinosaur biostratigraphy of the Scollard, Frenchman, and Hell Creek Formations. *Canadian Journal of Earth Sciences*, 56, 1041–1051.
- Erickson, G. M., Lappin, A. K., & Larson, P. L. (2005). Androgynous rex—the utility of using chevrons for determining the sex of crocodylians and extinct archosaurs. *Zoology*, 4, 277–286.
- Fastovsky, D. E., & Bercovici, A. (2016). The Hell Creek Formation and its contribution to the Cretaceous-Paleogene extinction: A short primer. *Cretaceous Research*, 57, 368–390.
- Field, D. J., Lynner, C., Brown, C., & Darroch, S. A. F. (2013). Skeletal correlates for body mass estimation in modern and fossil flying birds. *PLoS ONE*, 8(11), e82000. <https://doi.org/10.1371/journal.pone.0082000>
- Forster, C. A. (1996). Species resolution in *Triceratops*: Cladistic and morphometric approaches. *Journal of Vertebrate Paleontology*, 16, 259–270.
- Fowler, D. W. (2017). Revised geochronology, correlation, and dinosaur stratigraphic ranges of the Santonian-Maastrichtian (Late Cretaceous) formations of the Western Interior of North America. *PLoS ONE*, 12(11), e0188426.

- Gates, T. A. (2005). The Late Jurassic Cleveland-Lloyd dinosaur quarry as a drought-induced assemblage. *Palaios*, 20(4), 363–375.
- Gilmore, C. W. (1946). A new carnivorous dinosaur from the Lance Formation of Montana. *Smithsonian Miscellaneous Collections*, 106, 1–19.
- Greshko, M. (2022). Stan the *T. rex* found! World's most expensive fossil finds home in new museum. *National Geographic Magazine*.
- Godfrey, L. R., Lyon, S. K., & Sutherland, M. R. (1993). Sexual dimorphism in large-bodied primates: The case of the subfossil lemurs. *American Journal of Physical Anthropology*, 56, 315–334.
- Hartman, J., Johnson, K. R., & Nichols, D. J. (Eds.). (2002). *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains* (Vol. 361, p. 520). Geological Society of America Special Paper.
- Hicks, J. F., Johnson, K. R., Obradovich, J. D., Tauxe, L., & Clark, D. (2002). Magnetostratigraphy and geochronology of the Hell Creek and basal Fort Union Formations of southwestern North Dakota and a recalibration of the age of the Cretaceous-Tertiary boundary. In J. H. Hartman, K. R. Johnson, & D. J. Nichols (Eds.), *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: An integrated continental record of the end of the Cretaceous* (Vol. 361, pp. 35–55). Geological Society of America Special Paper.
- Hone, D. W. E., & Mallon, J. C. (2017). Protracted growth impedes the detection of sexual dimorphism in non-avian dinosaurs. *Palaeontology*, 60, 535–545.
- Kaiser, H., Crother, B. I., Kelly, C. M., Luiselli, L., O'Shea, M., Ota, H., Passos, P., Schleip, W. D., & Wüster, W. (2013). Best practices: In the 21st century, taxonomic decisions in herpetology are acceptable only when supported by a body of evidence and published via peer-review. *Herpetological Review*, 44, 8–23.
- Kaskes, P., Schulp, A. S., Larson, P. L., Smith, J., Kuiper, K. F. & Abels, H. A. (2016). Placing Naturalis' *Tyrannosaurus rex* specimen in a taphonomic, paleoenvironmental, and integrated stratigraphic framework: Hell Creek Formation, Eastern Montana. In *Society of vertebrate paleontology 2016 annual meeting abstracts* (p. 313).
- Kościński, K., & Pietraszewski, S. (2004). Methods to estimate sexual dimorphism from unsexed samples: A test with computer-generated samples. *Anthropological Review*, 67, 33–55.
- Larson, P. (1994). *Tyrannosaurus* sex. In G. D. Rosenberg & D. L. Wolberg (Eds.), *Dino Fest. The paleontological society special publication no. 7* (pp. 139–155).
- Larson, P. (1997). The king's new clothes: A fresh look at *Tyrannosaurus rex*. In D. L. Wolberg, E. Stump, & G. D. Rosenberg (Eds.), *Dinofest international proceedings* (pp. 65–71). Academy of Natural Sciences.
- Larson, P. (2008a). Variation and sexual dimorphism in *Tyrannosaurus rex*. In P. L. Larson & K. Carpenter (Eds.), *Tyrannosaurus rex, the Tyrant King* (pp. 103–127). Indiana University Press.
- Larson, N. (2008b). One hundred years of *Tyrannosaurus rex*: The skeletons. In P. L. Larson & K. Carpenter (Eds.), *Tyrannosaurus rex, the Tyrant King* (pp. 1–55). Indiana University Press.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2022). Cluster: Cluster analysis basics and extensions. R package version 2.1.3.
- Mallon, J. C. (2017). Recognizing sexual dimorphism in the fossil record: Lessons from nonavian dinosaurs. *Paleobiology*, 43, 495–507.
- Mallon, J. C., Evans, D. C., Ryan, M. J., & Anderson, J. S. (2012). Megaherbivorous dinosaur turnover in the Dinosaur Park Formation (upper Campanian) of Alberta, Canada. *Palaeogeography, Paleoclimatology, Palaeoecology*, 350–352, 124–138.
- Molnar, R. E. (2005). Sexual selection and sexual dimorphism in theropods. In K. Carpenter (Ed.), *The carnivorous dinosaurs* (pp. 284–312). Indiana University Press.
- Napoli, J. G., Erickson, G. M., Hunt, T., & Norell, M. A. (2019). *Psittacosaurus amitabha*, a new species of ceratopsian dinosaur from the Ondai Sayr locality, Central Mongolia. *American Museum Novitates*, 3932, 1–36.
- Napoli, J. G., Ruebenstahl, A. A., Bhullar, B.-A.S., Turner, A. H., & Norell, M. A. (2021). A new dromaeosaurid (Dinosauria: Coelurosauria) from Khulsan, Central Mongolia. *American Museum Novitates*, 3982, 1–47.
- Osborn, H. F. (1905). *Tyrannosaurus* and other Cretaceous carnivorous dinosaurs. *Bulletin of the American Museum of Natural History*, 21, 259–265.
- Paul, G. S. (1988). *Predatory dinosaurs of the world: A complete illustrated guide*. Simon & Schuster.
- Paul, G. S., Persons, W. S., & Van Raalte, J. (2022). The tyrant lizard king, queen, and emperor: Multiple lines of morphological and stratigraphic evidence support subtle evolution and probable speciation within the North American genus *Tyrannosaurus*. *Evolutionary Biology*. <https://doi.org/10.1007/s11692-022-09561-S>
- Saitta, E. T., Stockdale, M. T., Longrich, N. R., Bonhomme, V., Benton, M. J., Cuthill, I. C., & Makovicky, P. J. (2020). An effect size statistical framework for investigating sexual dimorphism in non-avian dinosaurs and other extinct taxa. *Biological Journal of the Linnean Society*, 131, 231–273.
- Scannella, J. B., Fowler, D. W., Goodwin, M. B., & Horner, J. R. (2014). Evolutionary trends in *Triceratops* from the Hell Creek Formation, Montana. *Proceedings of the National Academy of Sciences*, 111, 10245–10250.
- Sternberg, C. M. (1950). Steveville—West of the fourth meridian, Alberta. *Geological survey of Canada topographic map 969A. 1/31,680 scale (1 inch to 1/2 mile)*.
- Tsuihiji, T., Watabe, M., Tsogtbaatar, K., Tsubamoto, T., Barsbold, R., Suzuki, S., Lee, A. H., Ridgely, R. C., Kawahara, Y., & Witmer, L. M. (2011). Cranial osteology of a juvenile specimen of *Tarbo-saurus bataar* (Theropoda, Tyrannosauridae) from the Nemegt Formation (Upper Cretaceous) of Bugin Tsav, Mongolia. *Journal of Vertebrate Paleontology*, 31, 497–517.
- van der Reest, A. J., & Currie, P. J. (2017). Troodontids (Theropoda) from the Dinosaur Park Formation, Alberta, with a description of a unique new taxon: Implications for deinonychosaur diversity in North America. *Canadian Journal of Earth Sciences*, 54, 919–954.
- Vogel, G. (2020, October 7). Stan the *T. rex* sells for a record \$32 million at auction: Buyer unknown, but paleontologists worry researchers will lose access. *Science Insider*. <https://doi.org/10.1126/science.abf1554>
- Wiens, J. J., Bonett, R. M., & Chippindale, P. T. (2005). Ontogeny discombobulates phylogeny: Paedomorphosis and higher-level salamander relationships. *Systematic Biology*, 54, 91–110.
- Wilson, G. P., Clemens, W. A., Jr., Horner, J. R., & Hartman, J. H. (Eds.). (2014). *Through the end of the Cretaceous in the type locality of the Hell Creek Formation in Montana and adjacent areas* (p. 503). Geological Society of America Special Paper.

Authors and Affiliations

Thomas D. Carr¹  · James G. Napoli²  · Stephen L. Brusatte³  · Thomas R. Holtz Jr.^{4,5}  · David W. E. Hone⁶ · Thomas E. Williamson⁷  · Lindsay E. Zanno^{8,9} 

James G. Napoli
jnapoli@amnh.org

Stephen L. Brusatte
Stephen.Brusatte@ed.ac.uk

Thomas R. Holtz Jr.
tholtz@umd.edu

David W. E. Hone
d.hone@qmul.ac.uk

Thomas E. Williamson
Thomas.Williamson@state.nm.us

Lindsay E. Zanno
lindsay.zanno@naturalsciences.org

¹ Department of Biology, Carthage College, 2001 Alford Park Drive, Kenosha, WI 53140, USA

² Richard Gilder Graduate School and Division of Paleontology, American Museum of Natural History, Central Park West at 79th St, New York, NY 10024, USA

³ School of Geosciences, Grant Institute, University of Edinburgh, James Hutton Road, Edinburgh EH9 3FE, UK

⁴ Department of Geology, University of Maryland, 8000 Regents Drive, College Park, MD 20742, USA

⁵ Department of Paleobiology, National Museum of Natural History, Washington, D.C. 20013, USA

⁶ School of Biological and Behavioural Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

⁷ New Mexico Museum of Natural History and Science, 1801 Mountain Road NW, Albuquerque, NM 87104, USA

⁸ Paleontology, North Carolina Museum of Natural Sciences, 11 W. Jones St., Raleigh, NC 27601, USA

⁹ Department of Biological Sciences, Campus Box 7617, North Carolina State University, Raleigh, NC 27695, USA