[Papers in Palaeontology, 2023, e1487]

Machine learning confirms new records of maniraptoran theropods in Middle Jurassic UK microvertebrate faunas

by SIMON WILLS[1,2,*], CHARLIE J. UNDERWOOD[2] *and* PAUL M. BARRETT[1]

[1]Fossil Reptiles, Amphibians & Birds Section, Natural History Museum, Cromwell Road, South Kensington, London SW7 5BD, UK; s.wills@nhm.ac.uk; p.barrett@nhm.ac.uk

[2]Department of Earth & Planetary Sciences, Birkbeck College, Malet Street, London WC1E 7HX, UK; c.underwood@bbk.ac.uk

*Corresponding author

Typescript received 19 May 2022; accepted in revised form 23 December 2022

Abstract: Current research suggests that the initial radiation of maniraptoran theropods occurred in the Middle Jurassic, although their fossil record is known almost exclusively from the Cretaceous. However, fossils of Jurassic maniraptorans are scarce, usually consisting solely of isolated teeth, and their identifications are often disputed. Here, we apply different machine learning models, in conjunction with morphological comparisons, to a suite of isolated theropod teeth from Bathonian microvertebrate sites in the UK to determine whether any of these can be confidently assigned to Maniraptora. We generated three independent models developed on a training dataset with a wide range of theropod taxa and broad geographical and temporal coverage. Classification of the Middle Jurassic teeth in our sample against these models and comparison of the morphology indicates the presence of at least three distinct dromaeosaur morphotypes, plus a therizinosaur and troodontid in these assemblages. These new referrals significantly extend the ranges of Therizinosauroidea and Troodontidae by some 27 myr. These results indicate that not only were maniraptorans present in the Middle Jurassic, as predicted by previous phylogenetic analyses, but they had already radiated into a diverse fauna that pre-dated the break-up of Pangaea. This study also demonstrates the power of machine learning to provide quantitative assessments of isolated teeth in providing a robust, testable framework for taxonomic identifications, and highlights the importance of assessing and including evidence from microvertebrate sites in faunal and evolutionary analyses.

Key words: Maniraptora, Theropoda, Middle Jurassic, teeth, machine learning.

MANIRAPTORA is a diverse and speciose clade of theropod dinosaurs that includes some of the most familiar small-bodied predators of the Cretaceous Period, such as *Velociraptor* and *Deinonychus*. In addition to these iconic dromaeosaurids, the clade also includes troodontids, scansoriopterygids, oviraptorosaurs, therizinosaurs, alvarezsaurs and the only living dinosaurs, birds. During the Cretaceous they occupied a varied range of niches ranging from obligate herbivores to arboreal insectivores, as well as cursorial predators. Maniraptoran remains are best known from the northern hemisphere, but they achieved a wide geographic distribution that also encompassed South America, Africa and Madagascar (Ding *et al.* 2020).

Although maniraptoran remains are known almost exclusively from the Cretaceous, ghost lineages derived from phylogenetic analyses indicate that it is likely that their initial radiation occurred in the Middle Jurassic (Holtz 2000; Rauhut 2003; Xu *et al.* 2010; Carrano *et al.* 2012; Rauhut & Foth 2020). This date is bracketed by discoveries of earlier-branching coelurosaurs, such as tyrannosauroids, in Middle Jurassic deposits (Rauhut *et al.* 2010). However, the Jurassic maniraptoran record is frustratingly incomplete: a handful of named taxa are known from the Late Jurassic (*Archaeopteryx*, scansoriopterygids and possibly *Ornitholestes*) and there is one possible Middle Jurassic representative, *Eshanosaurus*, the identification and dating of which remains contentious (Kirkland & Wolfe 2001; Xu *et al.* 2001; Barrett 2009). Nevertheless, fragmentary, generically indeterminate remains of some maniraptoran subclades, such as possible dromaeosaurs, have been reported from Middle Jurassic microvertebrate sites in Europe and Asia (Evans & Milner 1994; Metcalf & Walker 1994; Averianov *et al.* 2005; Prasad & Parmar 2020). However, due to the disarticulated nature of the material, it has not been possible to identify these specimens beyond clade level and these identifications have been questioned, even at this coarse level of taxonomic resolution (Benson 2010*a*; Foth & Rauhut 2017; Ding *et al.* 2020; Sellés *et al.* 2021). This, and issues relating to the dating of some sites, has meant that these discoveries have usually been excluded from, or overlooked by, broader evolutionary analyses. As a result,

doi: 10.1002/spp2.1487

they have had little impact on determining the divergence times or palaeobiogeographic relationships of the major maniraptoran lineages. Consequently, the discovery of temporally well-constrained maniraptoran material from the Jurassic is of critical importance to more accurately constrain the timing of this major diversification event and shed light on early maniraptoran evolution.

Dinosaur teeth, including those of theropods, were continually shed and replaced throughout the animal's life and are highly resistant to chemical alteration and abrasion (Argast *et al.* 1987; Currie *et al.* 1990; Farlow *et al.* 1991). As a result, they are abundant in many Mesozoic deposits and sometimes represent the only evidence recording the dinosaur species-richness at such sites (e.g. Evans & Milner 1994; Fiorillo & Currie 1994; Larson & Currie 2013; Gates *et al.* 2015). The comparatively simple structure of theropod teeth has made identifications difficult historically, given that traditional taxonomic characters lack the resolution for distinguishing the teeth of closely related clades. However, apomorphy-based identifications, and statistical and morphometric analyses, have now been developed that offer solutions to this problem (Currie *et al.* 1990; Farlow *et al.* 1991; Smith *et al.* 2005; Larson 2008; Larson & Currie 2013; Williamson & Brusatte 2014; Hendrickx *et al.* 2015*a*; Gerke & Wings 2016; Young *et al.* 2019; Chiarenza *et al.* 2020). Most recently, the use of machine learning procedures has been shown to produce accurate group-discrimination when applied to morphological data (Hoyal Cuthill *et al.* 2019; MacLeod & Kolska Horwitz 2020). Wills *et al.* (2021) applied this technique to a diverse sample of theropod teeth and demonstrated that these methods lead to higher classification accuracies than more traditional statistical analyses.

Here, we apply these new methods to a large sample of isolated theropod teeth from a series of UK Middle Jurassic microvertebrate sites. Using machine learning and morphological-based approaches we demonstrate that many of these teeth can be referred with confidence to three distinct maniraptoran lineages (Dromaeosauridae, Troodontidae, Therizinosauroidea). These represent some of the earliest, or the earliest, records of these clades known from anywhere in the world, and their presence confirms the predictions of numerous phylogenetic analyses. They indicate that multi-taxic maniraptoran faunas were established by the Bathonian, millions of years earlier than the well-sampled biotas from the Late Jurassic (e.g. Yanliao biota) or late Early Cretaceous (e.g. Jehol biota) that previously represented the best windows on the initial diversification of the clade.

Institutional abbreviations. GLRCM, Museum of Gloucester, UK; NHMUK, Natural History Museum, London, UK.

GEOLOGICAL SETTING

Rapid changes in sedimentary facies took place during the Middle Jurassic in the region that is now the UK, with the shallow marine conditions that prevailed during the Early Jurassic giving way to more varied environments, ranging from open shallow-water marine in the south of England to increasingly non-marine strata in the East Midlands, Yorkshire and Scotland (Fig. 1). Deposition took place in a series of rifted basins with intervening structural highs and carbonate shelves developed on the margins of these landmasses. In southern and central England, there were emergent landmasses in the areas that are now South-West England, Wales and the London area. The generally north–south seaway between these consisted of open marine conditions in the south, a lagoon and mudflat complex in the north and a series of oolitic shoals separating these. Sealevel fluctuations throughout the Bathonian often caused pauses in marine sedimentation with occasional localized emergence accompanied by the development of hardgrounds, palaeosols and terrigenous sediment influxes (Palmer & Jenkyns 1975; Palmer 1979; Horton *et al.* 1995; Wyatt 1996; Underwood 2004; Hesselbo 2008; Barron *et al.* 2012; Wills *et al.* 2019). These changing conditions created a mosaic of different environments that were populated by a series of diverse Bathonian vertebrate faunas. Although the remains of large-bodied terrestrial taxa are relatively rare, several important microvertebrate localities have yielded large numbers of small vertebrate remains, including sharks, bony fish, mammals, turtles, crocodilians, choristoderes, pterosaurs, squamates and amphibians (Freeman 1976*a*, 1976*b*, 1979; Metcalf *et al.* 1992; Evans & Milner 1994; Wills *et al.* 2014, 2019). Dinosaur teeth are common and some of these were referred tentatively to various coelurosaurian theropod clades (Freeman 1976*a*, 1976*b*, 1979; Metcalf *et al.* 1992; Evans & Milner 1994; Wills *et al.* 2014, 2019). A summary of each main locality is provided below.

Hornsleasow

This quarry exposes a complete section through the Bajocian–Bathonian-aged Chipping Norton Limestone Formation and the overlying Bathonian Sharp's Hill Formation (Richardson 1929; Channon 1950; Torrens 1969*a*; Sellwood & McKerrow 1974; Cox & Sumbler 2002). The microvertebrate horizon occurs in the Chipping Norton Limestone Formation, *Z. zigzag* Zone (Cope *et al.* 1980), as an 11 m × 1 m clay lens lying on a palaeokarst surface that can be traced throughout the quarry (Vaughan 1989; Metcalf *et al.* 1992; Metcalf & Walker 1994; Metcalf 1995).

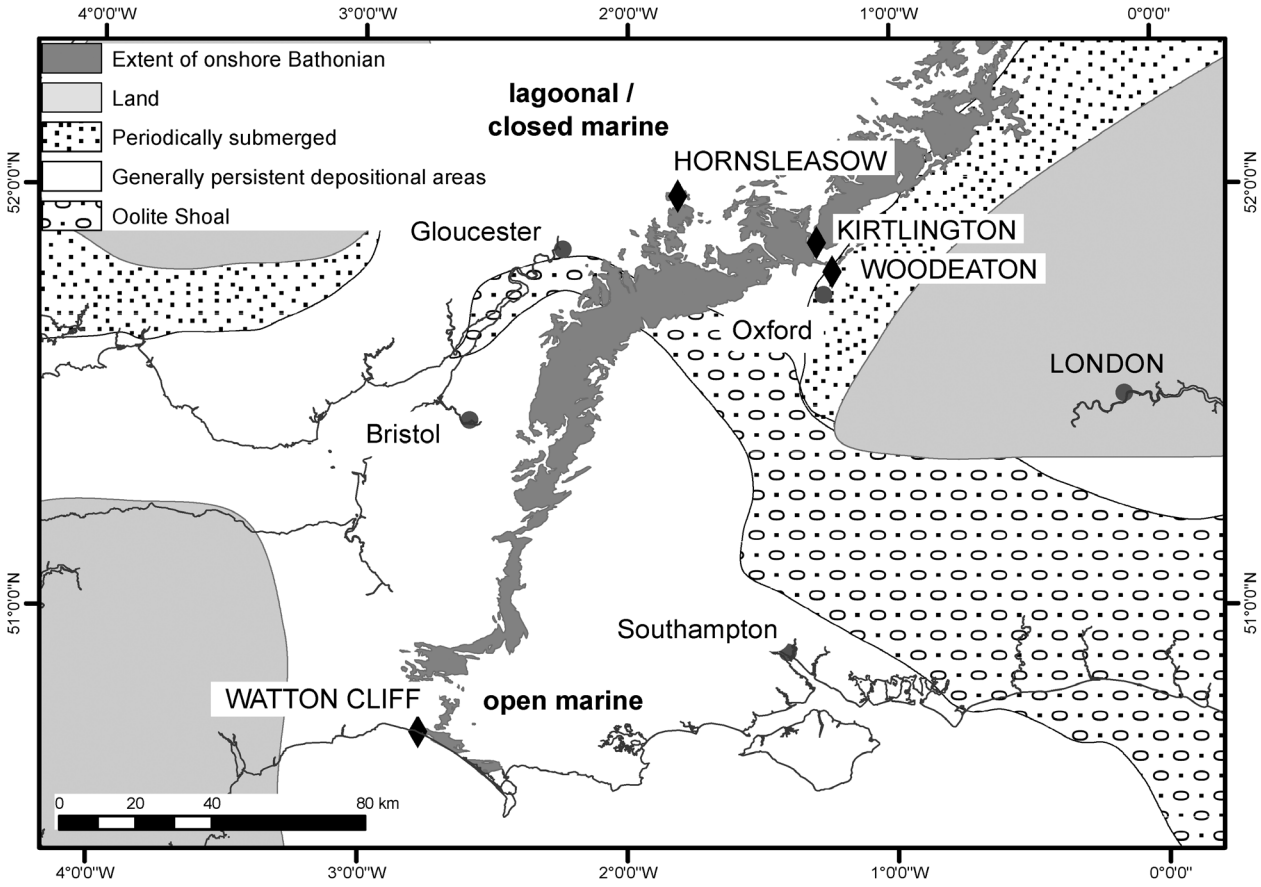

FIG. 1. Site localities and Middle Jurassic palaeogeography and depositional regimes of southern England. After Wills *et al.* (2019), Barron *et al.* (2012) and Underwood (2004).

The clay unit developed following a flooding event that introduced the initial sediment into the karstic hollow, which subsequently became a coastal marsh pond supporting a wide variety of freshwater organisms (Metcalf 1995). The introduction of terrestrial vertebrate remains occurred both as a direct result of the initial flooding event and subsequent fluvial transport into the pond. This lens of non-marine sediments is over- and underlain by oolitic limestones that were deposited in fully (but lagoonal) marine environments.

Woodeaton

Woodeaton Quarry presents a continuous section through most of the Bathonian including the Rutland, White Limestone and the lower part of the Forest Marble formations, with lower horizons being also briefly exposed (Palmer 1973, 1974; Palmer & Jenkyns 1975; Horton *et al.* 1995; Wyatt 2002; Wills *et al.* 2019). Microvertebrates have been recovered from bed 23 of the Bladon Member, White Limestone Formation, *H. retrocostatum* Zone (Barron *et al.* 2012; Wills *et al.* 2019). This is a pale massive clay, marl or impure limestone that can be traced across the entire quarry face. Unlike other British Middle Jurassic microvertebrate sites, which represent shallow brackish to freshwater ponds, lakes or marginal marine settings of a restricted geographical extent, Woodeaton represents a larger scale brackish water lagoon of fluctuating salinity with periodic influxes of seawater that experienced seasonal aridity (Wills *et al.* 2019).

Kirtlington

This quarry exposes sections through the White Limestone Formation and overlying Forest Marble and Cornbrash formations with the microvertebrate horizon, the 'Mammal Bed', forming a thin and impersistent lens of unconsolidated brown marl at the boundary between the White Limestone and Forest Marble formations (*H. retrocostatum* Zone). While its exact correlation with Woodeaton is uncertain, it appears that the Kirtlington fauna is of a slightly younger age than the approximately

coeval section at Woodeaton (McKerrow *et al.* 1969; Wills *et al.* 2019). The Mammal Bed at Kirtlington formed in a shallow marginal marine environment during a period of marine regression along a shallow coastal plain region characterized by coastal lakes, swamps and lagoons (Freeman 1979; Palmer 1979; Evans & Milner 1994).

Watton Cliff

The Forest Marble Formation (*C. discus* Zone) section at Watton Cliff is composed of a 10 m thick lower sequence of clays and shales followed by 3–5 m of cross-bedded bioclastic limestones and 9 m of inter-bedded clays and siltstones (Woodward 1894; Strahan 1898; Torrens 1969*b*; Cope *et al.* 1980; Melville & Freshney 1982; Holloway 1983; Barron *et al.* 2012). The entire succession represents open marine conditions, probably with a moderate water depth, although with signs of weak storm influence (such as rippled sand lenses) throughout. There is a cross-stratified bioclastic unit, most of which is strongly cemented, with lenses and irregular patches that lack this cement forming a bioclastic gravel. These unconsolidated patches seem to represent either channels or burrows and commonly contain water-worn vertebrate material (Dineley & Metcalf 1999; Benton *et al.* 2005); similar material is also present (although harder to extract) in the cemented sediment. The Watton Cliff site represents deposition of a shell bank, possibly during storm-related events (Holloway 1983), in an open marine, clear water, shallow coastal sea on a gently sloping shelf, which was subject to continuous wave action in a tide-dominated system with runoff channels developing during emergent conditions. Terrestrial and freshwater organisms are present as allochthonous elements deposited alongside marine invertebrates and vertebrates (marine sharks and teleosaurid crocodilians) (Hunter & Underwood 2009).

MATERIAL AND METHOD

The material consists of isolated theropod teeth from four Middle Jurassic (Bathonian, Great Oolite Group) localities in the UK (Fig. 1): Woodeaton Quarry, Oxfordshire (White Limestone Formation); Kirtlington Quarry, Oxfordshire (White Limestone Formation); Hornsleasow Quarry, Gloucestershire (Chipping Norton Limestone Formation) and Watton Cliff, Dorset (Forest Marble Formation). It was collected over a period of several decades by teams from different institutions, including the Natural History Museum (Woodeaton, Watton Cliff), University College London (Kirtlington, Watton Cliff) and the University of Bristol/Museum of Gloucester (Hornsleasow). Except for the Hornsleasow material, which is housed in the Museum of Gloucester, the specimens are held in the collections of the Natural History Museum, London. Some of the previously collected material had undergone an initial sort and was assigned either a general taxonomic identification (e.g. 'theropod') or a morphotype (e.g. 'morphotype A').

New material from Woodeaton was obtained by bulk sediment collection on site following initial fieldwork to identify productive horizons. Large bulk samples (often weighing several tonnes) were screen-washed using the methodology described by Ward (1981) to produce an initial concentrate of vertebrate material. This was split into four size fractions (500 μm–1 mm, 1–2 mm, 2–4 mm, >4 mm) to facilitate initial sorting and picking using a binocular microscope. We initially identified 164 isolated theropod teeth from the older collections and new Woodeaton material (Kirtlington, n = 49; Hornsleasow, n = 50; Watton Cliff, n = 4; Woodeaton, n = 61) of which 149 were sufficiently complete to warrant further investigation.

All teeth in the sample underwent a combination of optical imaging with a Dino-Lite AM 7915 MZTL microscope and scanning electron microscopy on a LEO 1455VP microscope. We also scanned each (complete) tooth using micro-computed tomography (μCT) with a Nikon Metrology HMX ST 225 μCT scanner and a Zeiss Versa μCT scanner at a range of voxel resolutions from 4 to 30 μm and created 3D models from the CT volumes using Avizo (v.8.1; ThermoFisher) (Appendix S1). Five morphometric variables were collected from each tooth, which were measured directly from the images using Fiji (Schindelin *et al.* 2012) and from the 3D models using Avizo. The measurements are simple 2D linear distances (Fig. 2) between landmarks on the tooth crown: crown height (CH), height of the crown measured from the tip of the tooth to the base of the enamel; crown base length (CBL), length of the base of the crown measured along its mesiodistal axis; crown base width (CBW), width of the base of the crown measured along its linguolabial axis perpendicular to the CBL; average number of denticles per millimetre along the mesial carina (MDM); and average number of denticles per millimetre along the distal carina (DDM). When a measurement could not be taken due to crown damage it was recorded as NA in the data, and carinae with no denticles were recorded as zero for either MDM or DDM variables. When required, the crown base ratio (CBR) is calculated as CBW/CBL and the denticle size density index (DSDI), a measure of the size difference between mesial and distal denticles (Rauhut & Werner 1995), as MDM/DDM.

Although other approaches, such as 3D data, are available (Hoyal Cuthill *et al.* 2019; Wills *et al.* 2021; MacLeod *et al.* 2022) we chose to use these 2D linear

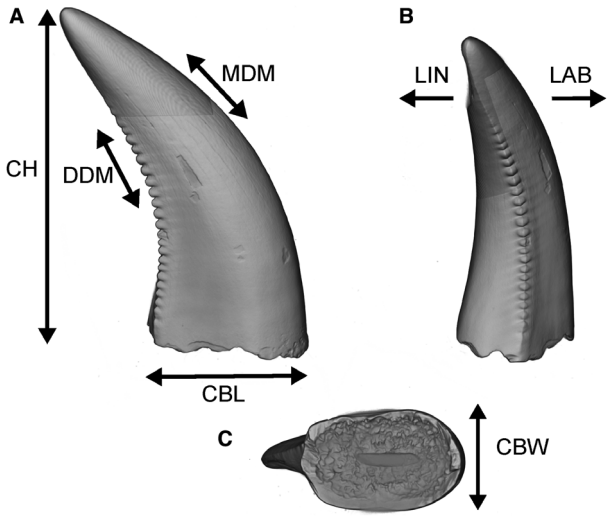

FIG. 2. Anatomical and morphometric terminology. Theropod tooth crown in: A, labial; B, distal; C, basal view. *Abbreviations*: CBL, crown base length; CBW, crown base width; CH, crown height; DDM, distal denticles per millimetre; LAB, labial; LIN, lingual; MDM, mesial denticles per millimetre. After Hendrickx *et al.* (2019) and Wills *et al.* (2021).

measurements because these variables are common to most published analyses of isolated theropod tooth datasets (e.g. Currie *et al.* 1990; Sankey *et al.* 2005; Smith *et al.* 2005; Larson 2008; Larson & Currie 2013; Williamson & Brusatte 2014; Hendrickx *et al.* 2015*a*, 2019, 2020; Larson *et al.* 2016; Noto *et al.* 2022), enabling direct comparisons with earlier work. Moreover, they have been shown to be useful taxonomic classifiers when used in both linear discriminant analysis (e.g. Larson & Currie 2013; Williamson & Brusatte 2014; Brusatte & Clark 2015; Gates *et al.* 2015; Hendrickx *et al.* 2020) and machine learning analysis (Wills *et al.* 2021). Given this and the lack of comparative digital image-based theropod tooth datasets we feel that the approach we have taken is appropriate.

To determine the taxonomic identifications of the teeth we undertook a quantitative analysis of morphometric data using a mixture of machine learning models following the methodology of Wills *et al.* (2021). We used three different machine learning techniques: mixture discriminant analysis (MDA), random forests (RF) and C5.0, and combined the classification results from all models to form an ensemble classifier. The three models differ in their approach to learning, enabling us to base the final classification prediction on the output of more than one technique. MDA is a non-linear extension of linear discriminant analysis whereby each class is modelled as a mixture of multiple multivariate normal subclass distributions, RF is an ensemble consisting of classification or regression trees (in this case classification trees) where the prediction from each individual tree is aggregated to form a final prediction, and C5.0 is a decision tree classifier based on information theory (Hastie & Tibshirani 1996; Breiman 2001; Kuhn *et al.* 2018; Wills *et al.* 2021). Models were combined into an ensemble classifier using both a simple majority voting rule and by combining the class prediction posterior probabilities for each tooth.

To build and train the models we combined several published datasets (Farlow *et al.* 1991; Sankey *et al.* 2002; Currie & Varrichio 2004; Smith *et al.* 2005; Larson 2008; Longrich 2008; Sankey 2008; Rauhut *et al.* 2010; Larson & Currie 2013; Hendrickx *et al.* 2015*a*; Gerke & Wings 2016; Larson *et al.* 2016; Young *et al.* 2019) that had been used for prior morphometric analysis with additional measurements taken as part of this study. The resultant dataset covers a wide range of theropod taxa with a broad geographical and temporal distribution, although there is some bias to North American Late Cretaceous taxa (Fig. 3). See the supporting data for a summary of the data used, taxonomic groups chosen and sample sizes used in our analysis.

Different definitions have been applied to these morphometric variables, with Smith *et al.* (2005) and Hendrickx *et al.* (2015*a*) differing in their methods for measuring CBL and CH (Fig. 2), and we used the corrected data provided by Gerke & Wings (2016) where possible. However, the difference in methodology has little overall effect on the reclassification rate, and the per-clade accuracies returned from the combined training dataset used here are similar to those reported by Wills *et al.* (2021). Prior to training these models the data were cleaned to improve model performance. First, we removed any outliers using a density-based spatial clustering algorithm (DBSCAN), which assumes that clusters of data form dense regions in space separated or surrounded by regions of lower density, with the outliers (or noise) falling in the lower density space (Ester *et al.* 1996). Outliers distort morphospace by shifting the mean centroid of a group to the direction of the outlier, which affects the model accuracy and the resultant classification. Second, we removed any classes with fewer members than the number of predictive variables (five), and last, we removed cases with missing data because this can have a detrimental effect on machine learning models; similarly, any unknown teeth with missing data were excluded from final classification. The data were log-transformed (adding a value of 1 to enable the transformation of zero values), scaled and centred prior to analysis. We made no attempt to directly address class imbalance by creating synthetic data (due the detrimental effect this has on model accuracy) and used equal prior probabilities in all models (Wills *et al.* 2021). From an initial dataset of 3886 specimens, data cleaning resulted in a final set of 1702 usable cases. We undertook an initial exploration of clade

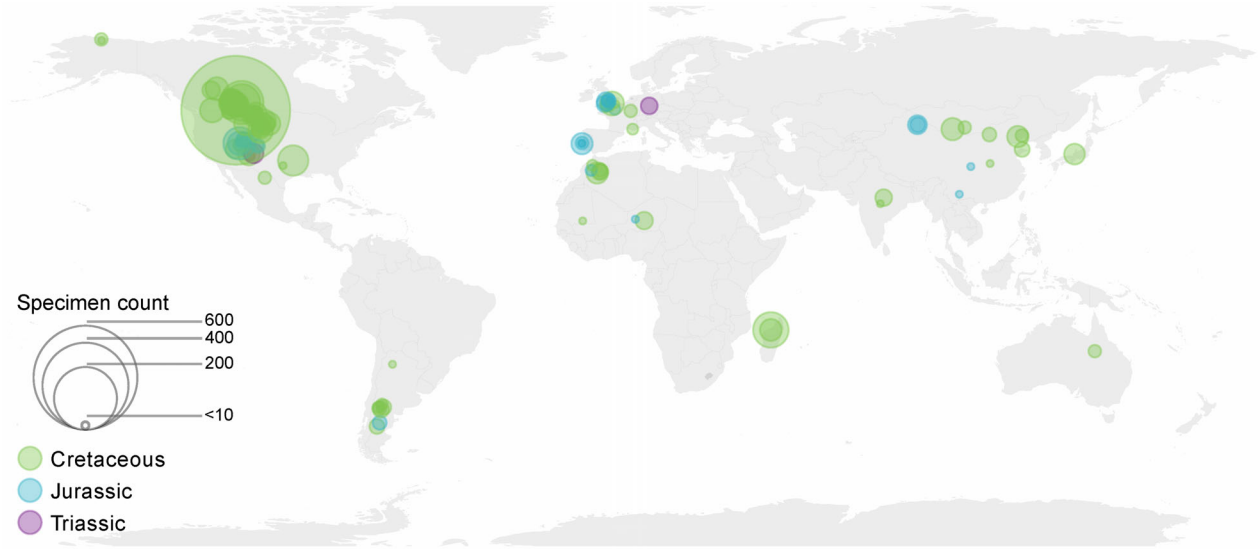

FIG. 3. Spatial distribution of training data samples, after removal of outliers and missing data, used for machine learning models.

feature space for the transformed morphometric variables using two different dimension reduction techniques to visualize the data, principal components analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE). We used both techniques given that PCA tries to preserve the global structure of the data whereas t-SNE looks to preserve local structure by keeping similar instances close to each other, potentially giving different insights into the data.

We undertook a series of non-parametric statistical analyses using permutational multivariate analysis of variance (PERMANOVA) with the Mahalanobis distance (Anderson & Walsh 2013; Anderson 2017), to obtain estimates of the statistical significance of training set group separations in feature space. PERMANOVA is used to compare groups of objects by testing for equivalence between the group centroids. The test works on the underlying distance matrix derived from the input variables rather than the raw or ordinated data. Given that PERMANOVA tests only whether all of the centroids in the data are equal, we performed post-hoc comparisons between the groups using a pairwise implementation of the PERMANOVA test with Bonferroni-corrected p-values. The PERMANOVA and pairwise-PERMANOVA tests were each performed with 10 000 replications.

For each model the cleaned data was split in an 80:20 ratio, preserving the overall class distribution of the data (Kuhn 2008), into a training dataset (1364 cases) and a testing dataset (338 cases). The models were developed on the training data and then assessed against the testing data. Testing data were not used in the initial model. The teeth to be classified were then run through each model in turn to provide independent classifications based on different techniques. We used k-fold cross-validation on the training set with k = 10 to give an overall model accuracy. We also ran each model permutation using a range of tuning parameters to obtain the highest accuracy. For MDA we modelled the response using a range of subclasses, from one to eight, for each taxonomic class; the RF model was tuned by varying the random subset of predictors that the model uses at each split in the tree (m_{try} parameter) from two to five and we grew the forest to 2000 trees; and for the C5.0 model we varied the number of model iterations from 1 to 100 and used both rule- and tree-based classifier models (Kuhn & Johnson 2013; Wills *et al.* 2021). In addition to the predicted class generated from the models we also calculated the posterior probability of the predicted class for each tooth. Training of the models relies on a random selection of teeth from the overall training data for each run, and indeed within each model there will be a degree of randomization input into the training. As a result, there may be slightly different results obtained from different training cycles of the models. For more details on the techniques involved and descriptions of the differences between the machine learning algorithms see Wills *et al.* (2021).

Dental terminology and nomenclature follows that outlined by Hendrickx *et al.* (2015*b*), and anatomical descriptions are based on morphological observations by one of the authors (SW). Geological, sedimentological and palaeoenvironmental observations are based on the study of published literature and field observations by one of the authors (SW).

All analyses were performed using R v.4.0.5 (R Core Team 2020) in RStudio (RStudio Team 2020). The following R packages were used for specific models or processes:

mda (Hastie *et al.* 2020), C5.0 (Kuhn *et al.* 2018), randomForest (Liaw & Wiener 2002, 2022), ranger (Wright & Ziegler 2017; Wright 2021) and caret (Kuhn 2008, 2022) for specific classification models; vegan (Oksanen *et al.* 2020) and RVAideMemoire (Hervé 2021) for PERMANOVA tests; ggplot (Wickham 2016) and gridextra (Auguie 2017) for plotting functions; and chronosphere (Kocsis & Raja 2019), divDyn (Kocsis *et al.* 2019, 2022) and rgplates (Kocsis & Raja 2021) for palaeogeographical reconstructions using the PALEOMAP plate model and data from Scotese (2016).

RESULTS

Machine learning models

The difficulties in providing accurate quantitative assessments of theropod tooth morphological discrimination are highlighted in Figure 4. Here, we show two different feature-space representations of the untrained morphological data, a PCA ordination and a t-SNE ordination, which clearly demonstrate the degree of overlap between numerous theropod clades. Non-parametric statistical tests on the t-SNE ordinated training data confirm this. The PERMANOVA test indicates that although the separation between groups is statistically significant overall ($F = 169.6$, $p < 0.01$), there is difficulty in resolving between-group structures for some group-pairs as demonstrated by the pairwise PERMANOVA tests (Fig. 5). This is consistent with previous reports in the literature in which attempts to distinguish theropod taxa using PCA or linear discriminant analysis have produced high degrees of feature-space overlap between some taxonomic groups (e.g. Hendrickx *et al.* 2019; Young *et al.* 2019; Noto *et al.* 2022). This result is unsurprising given that we are constrained in attempting to differentiate teeth with very similar gross morphology based on a small set of morphological measurements. As MacLeod *et al.* (2022) noted, however, that this does not preclude the possibility that different techniques may uncover significant between-group differences that can be used as the basis of a classification. In fact, when comparing the between-group structures for Maniraptora with other groups, the pairwise PERMANOVA tests (Table 1) suggest that these taxa are differentiable from most major theropod clades ($p < 0.01$).

We also conducted PERMANOVA tests on the trained MDA feature-space scores generated from the training data (Fig. 6). The overall test rejected the null hypothesis that there are no between-group differences ($p < 0.01$) but, as before, the post-hoc pairwise tests indicate that some group-pairs might be difficult to differentiate using this method, highlighting the importance of using multiple techniques to compare and classify isolated theropod teeth.

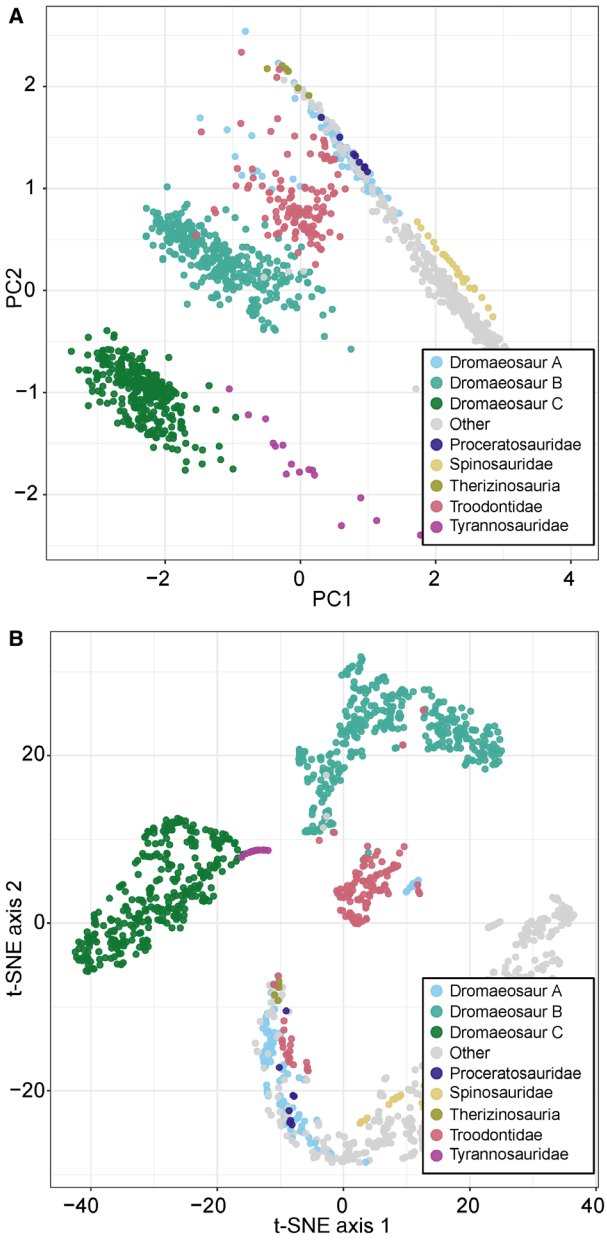

FIG. 4. Untrained ordinated feature-space occupation for teeth comprising the training data set formed by: A, the first two principal component (PC) axes; B, the first two t-distributed stochastic neighbour embedding (t-SNE) axes.

All three machine learning techniques have similar levels of accuracy (Table 2), with the overall accuracy of the machine learning models ranging from 82.4% (C5.0) to 85.6% (RF). When the models were run against the test dataset the two decision-tree algorithms, RF at 88.4% and C5.0 at 85.4%, slightly outperformed the MDA

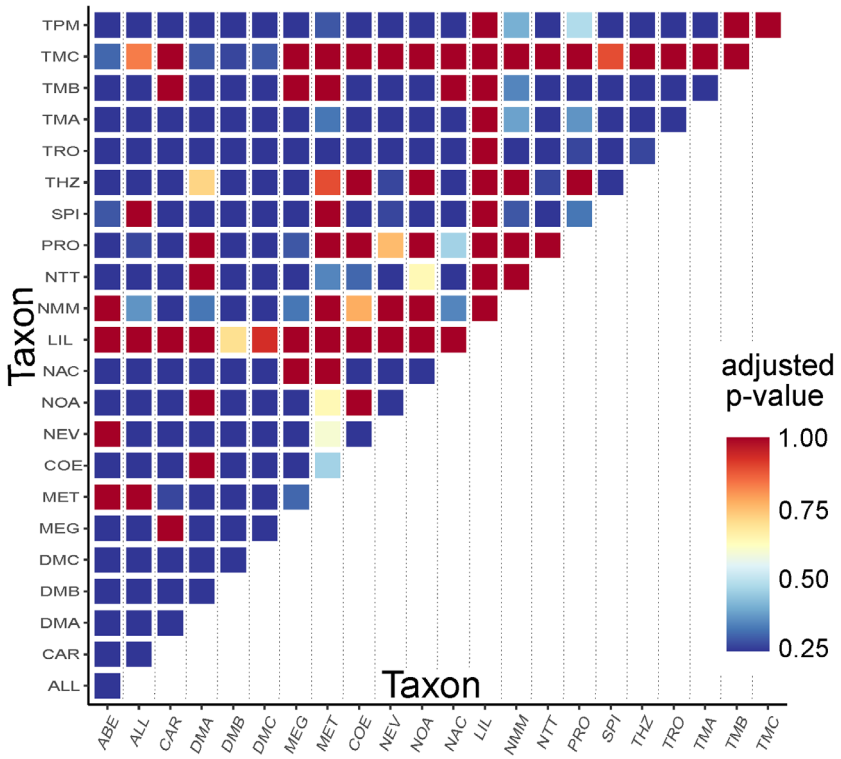

FIG. 5. Training data PERMANOVA Bonferroni adjusted p-values for pairwise clade groups using untrained t-distributed stochastic neighbour embedding (t-SNE) ordinated feature-space based on three t-SNE dimensions. *Taxon abbreviations*: ABE, Abelisauridae; ALL, Allosauridae; CAR, Carcharodontosauridae; COE, *Coelophysis*; DMA, Dromaeosaur morphotype A; DMB, Dromaeosaur morphotype B; DMC, Dromaeosaur morphotype C; LIL, *Liliensternus*; MEG, Megalosauridae; MET, Metriacanthosauridae; NAC, other Ceratosauria; NEV, Neovenatoridae; NOA, Noasauridae; NMM, other Megalosauroidea; NTT, other Tyrannosauroidea; PRO, Proceratosauridae; SPI, Spinosauridae; THZ, Therizinosauria; TMA, Tyrannosauridae morphotype A; TMB, Tyrannosauridae morphotype B; TMC, Tyrannosauridae morphotype C; TPM, Tyrannosauridae pre-maxillary; TRO, Troodontidae.

TABLE 1. By-group comparisons of maniraptoran clade-pairs.

Taxon pairs	Sum of squares	*F* model	R^2	p-value	p-value (adjusted)
Dromaeosaur A vs Dromaeosaur B	401.09	194.19	0.313	0.0001	0.0036
Dromaeosaur A vs Dromaeosaur C	394.49	195.40	0.329	0.0001	0.0036
Dromaeosaur A vs Therizinosauria	16.30	5.68	0.054	0.0012	0.0432
Dromaeosaur A vs Troodontidae	143.53	60.94	0.218	0.0001	0.0036
Dromaeosaur B vs Dromaeosaur C	621.39	306.68	0.326	0.0001	0.0036
Dromaeosaur B vs Therizinosauria	237.72	103.29	0.235	0.0001	0.0036
Dromaeosaur B vs Troodontidae	352.62	158.13	0.258	0.0001	0.0036
Dromaeosaur C vs Therizinosauria	299.81	147.01	0.322	0.0001	0.0036
Dromaeosaur C vs Troodontidae	411.89	201.67	0.321	0.0001	0.0036
Therizinosauria vs Troodontidae	53.61	20.58	0.139	0.0001	0.0036
Other vs Dromaeosaur A	248.56	97.40	0.151	0.0001	0.0036
Other vs Dromaeosaur B	735.38	355.97	0.312	0.0001	0.0036
Other vs Dromaeosaur C	719.66	350.49	0.316	0.0001	0.0036
Other vs Therizinosauria	67.75	23.70	0.049	0.0001	0.0036
Other vs Troodontidae	426.80	188.50	0.247	0.0001	0.0036

PERMANOVA Bonferroni-adjusted p-values on untrained t-distributed stochastic neighbour embedding (t-SNE) ordinated feature-space based on three t-SNE dimensions.

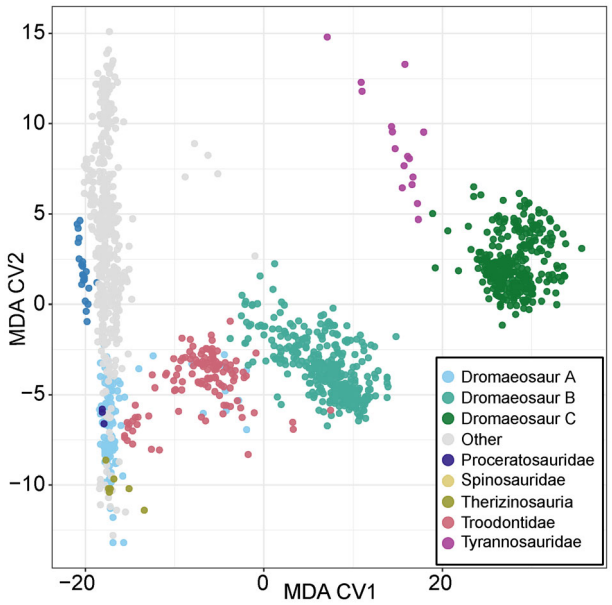

FIG. 6. Trained feature-space occupation of selected taxa from the training data based on two mixture discriminant analysis (MDA) dimensions. Total between-group variance explained 97.5% (CV1 = 90.4%, CV2 = 7.1%). CV, canonical variate.

TABLE 2. Machine learning model accuracies.

Machine learning model	Model accuracy (%)	Testing data accuracy (%)
RF	85.6	88.4
MDA	84.4	84.1
C5.0	82.4	85.4

RF, random forests: two randomly selected predictor variables at each tree node split and 2000 trees; MDA, mixture discriminant analysis: eight subclasses; C5.0: tree-based model with 40 boosting iterations.

model at 84.1%. We additionally assessed the RF model by calculating the out-of-bag (OOB) error, a subset of the original training data that the model uses to estimate the prediction error. In this case the overall OOB error is 0.15, meaning that 85% of the retained subset classify correctly, which corresponds well to the accuracy returned from the test data. RF prediction errors decrease as the forest is grown to its full extent of 2000 trees, with the overall OOB error and most individual clade OOB errors settling after around 200 trees. The accuracy of the RF model responses achieved by varying the m_{try} tuning parameter ranges from 84.8% to 85.6%, with the slightly higher accuracy achieved by using two randomly selected predictor variables at each tree node split.

The accuracy of the MDA model responses achieved by varying the number of potential subclasses in each taxonomic group ranges from 78.3% (one subclass) to 84.4% (eight subclasses), and the C5.0 model achieved the best response (82.4% accuracy) using a tree-based classifier with 40 boosting iterations.

At the individual clade level (Table 3; Fig. 7) the performance of both the ensemble model and the individual machine learning classifiers that make up this ensemble varies with classification accuracy, ranging from 50% to 100% (Fig. 7). Maniraptoran clades have a high level of classification accuracy regardless of the machine learning model used, ranging from 92.8% (Dromaeosaur morphotype A, RF model) to 100% (Dromaeosaur morphotype B, RF model; Dromaeosaur morphotype C, MDA model; and Therizinosauria, RF model). The variation in clade accuracy is driven by several factors, including the number of cases comprising the training group for that particular clade; morphological overlap with other clades; and the limited morphological measurements used to train the classifiers. The accuracy results reported here are derived from cross-tabulation tests on the classified testing data and confirm, as MacLeod *et al.* (2022) note, that good levels of discrimination for some clades can be achieved by machine learning even when group-level feature-spaces overlap.

UK Bathonian sites

The classification results from the UK Bathonian isolated teeth (Table 4) indicate the presence of three distinct dromaeosaur morphotypes. These morphotypes are strongly supported across all machine learning models and the ensemble classifier in either majority-vote or combined probability mode. Our confidence in the classifications is a combination of the machine learning results from three independent classifiers and our post-hoc morphological analysis. In all machine learning systems there is likely to be a degree of misclassification and in this case the models incorrectly classified GCLRM G8-23 as a dromaeosaur rather than a troodontid, NHMUK PV R37948 as a troodontid rather than a dromaeosaur and GCLRM G167-32 as a dromaeosaur rather than a therizinosaur (see Systematic palaeontology, below). The posterior probabilities from the ensemble classifier (Fig. 8) also add to our confidence in the machine learning prediction given that the majority of the teeth return high posteriors in favour of the assigned class, with the second-highest class posterior in each case also indicating maniraptoran affinities. In addition, it is clear from the trained MDA data (Fig. 9) that the small teeth from these sites occupy a segment of feature-space that is both congruent with a broad maniraptoran feature-space and distinct from that occupied by other Jurassic taxa.

TABLE 3. Classification accuracy by clade based on test data.

Taxon	Accuracy (%)†	RF (%)‡	C5.0 (%)‡	MDA (%)‡	No. cases§
Neotheropoda: *Coelophysis*	91.7	100	99.8	100	9
Non-averostran Neotheropoda: *Lilienstern*us	50.0	50	50	50	6
Ceratosauria: Abelisauridae	86.8	76.5	93.3	91.1	67
Ceratosauria: Noasauridae	55.6	66.7	50	73	13
Ceratosauria: other Ceratosauria	55.1	57.9	57.9	81.7	25
Megalosauroidea: Megalosauridae	71.3	78.6	78.3	82.5	29
Megalosauroidea: Spinosauridae	100.0	100	100	100	23
Megalosauroidea: other Megalosauroidea	50.0	50	50	49.9	8
Allosauroidea: Allosauridae	82.6	79.5	84.4	84.8	44
Allosauroidea: Carcharodontosauridae	95.5	99.4	91.8	93.4	56
Allosauroidea: Metriacanthosauridae	66.3	49.5	74.4	62.5	12
Allosauroidea: Neovenatoridae	60.6	49.5	49.7	68.5	16
Tyrannosauroidea: Tyrannosauridae morphotype A	91.6	99.8	62.5	100	16
Tyrannosauroidea: Tyrannosauridae morphotype B	86.5	83.9	85.5	91.5	132
Tyrannosauroidea: Tyrannosauridae morphotype C	50.0	50	50	58.3	6
Tyrannosauroidea: Tyrannosauridae premaxillary	66.5	66.5	66.4	87.4	16
Tyrannosauroidea: Proceratosauridae	99.8	99.8	100	68.8	8
Tyrannosauroidea: other Tyrannosauroidea	61.0	66.5	50	83.2	15
Therizinosauria	100.0	100	50	91.6	6
Dromaeosaur morphotype A	94.1	92.8	94.5	96.3	96
Dromaeosaur morphotype B	99.9	100	99.6	99.4	332
Dromaeosaur morphotype C	99.8	99.3	99.8	100	305
Troodontidae	94.6	95	93.4	97	124

†Ensemble model accuracy; ‡individual model accuracy; §no. cases per clade in the training data. RF, random forests: two randomly selected predictor variables at each tree node split and 2000 trees; MDA, mixture discriminant analysis: eight subclasses; C5.0: tree-based model with 40 boosting iterations.

SYSTEMATIC PALAEONTOLOGY

THEROPODA Marsh, 1881
MANIRAPTORA Gauthier, 1986
PARAVES Sereno, 1997
DROMAEOSAURIDAE Matthew & Brown, 1922
Gen. et sp. indet. Morphotype A
Figure 10

Referred specimens. GLRCM G100-14, G100-21, G100-9, G140-7, G14-27, G21-22, G38-10, G75704, G91702, G91705, G91706, NHMUK PV R37904, R37905, R37906, R37907, R37908, R37910, R37924, R37925, R37926, R37927, R37928, R37929, R37935, R37939, R37940, R37941, R37942, R37944, R37945, R37946, R37947, R37948, R37949, R37950, R37953.

Localities. Hornsleasow Quarry, Chipping Norton Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (14 teeth); Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (seven teeth); Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (15 teeth).

Description. Morphotype A tooth crowns (36 in total) are ziphodont, range in CH from 1.45 mm to 7.79 mm (Fig. 11) and have serrated distal carinae and unserrated mesial carinae. The distal crown margin is concave, the crowns are labiolingually compressed (CBR 0.36–0.76) and their lingual and labial surfaces possess centrally placed concave depressions that extend apically to the mid-height of the crown surface. These depressions, especially where strongly developed, result in a lemniscate (figure-of-eight) basal cross-section. Both mesial and distal carinae are well developed with the distal carina often deflected labially towards the crown base and the mesial carina twisted slightly and deflected lingually basally. The distal carina extends from the crown apex to the crown base and bears denticles that are generally restricted to the lower two-thirds of the carina, although occasionally reaching the apex. The distal denticles decrease in size both apically and distally from carina mid-length. Distal denticles are small, ranging in length from 0.05 mm to 0.27 mm (18.2 per mm to 3.6 per mm), are sub-rectangular in shape with a convex external margin, and are orientated perpendicular to the carina (except for a few teeth in which the denticles are slightly inclined apically). The mesial carina extends from the apex of the crown to a position approximately two-thirds down the crown and lacks denticles. The crown surface has a braided enamel texture consisting of sinuous grooves and ridges that are orientated apicobasally (Hendrickx *et al.* 2015*a*, 2019).

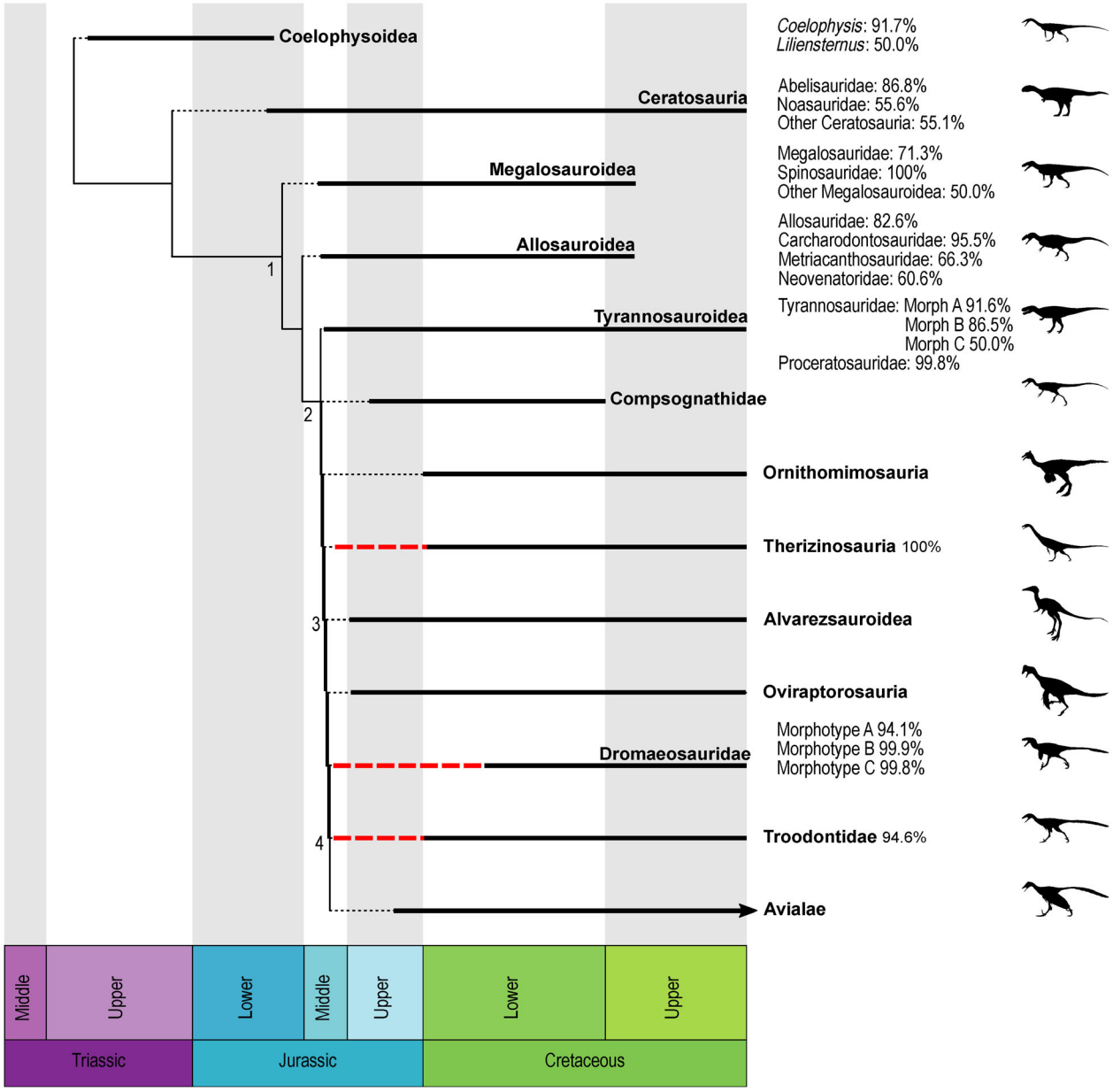

FIG. 7. Simplified time-calibrated theropod phylogeny showing the individual clade classification accuracies based on the machine learning ensemble and the range extensions (in red) implied by these results. *Nodes*: 1, Tetanurae; 2, Coelurosauria; 3, Maniraptora; 4, Paraves. For Therizinosauria and Troodontidae we have used the recent Berriasian age determination, rather than Barremian, for the Cedar Mountain Formation of Utah (Joeckel *et al.* 2020). Phylogeny modified after Rauhut & Foth (2020). All silhouettes taken from Phylopic (https://www.phylopic.org: Pranav Iyer: Dromaeosauridae (CC0 1.0); Michael Keesey: Abelisauridae, *Tyrannosaurus* (CC BY 3.0); Scott Harmann: *Coelophysis* (CC BY 3.0), Megalosauroidea, *Allosaurus*, Compsognathidae, Therizinosauria, *Troodon*, Alvialae (all CC BY-NC-SA 3.0); Nobu Tamura: Ornithomimosauria (CC BY 3.0); Funk Monk: Alvarezsauridae (CC0 1.0); Jaime Headden: Oviraptosauria (CC BY-NC-SA)).

Gen. et sp. indet. Morphotype B
Figure 12

Referred specimens. GCRLM GTUBE 67, G10022, G100-64, G10-37, G12-28, G14-22, G167-24, G68-1, G7.21-3, GHQ104 C-1, GTEMP3061, GX, NHMUK PV R36771, R36778, R37909, R37911, R37912, R37913, R37914, R37915, R37916, R37917, R37918, R37919, R37921, R37922, R37923, R37930, R37931, R37933, R37934, R37936, R37937, R37938, R37943, R37951, R37952.

Localities. Hornsleasow Quarry, Chipping Norton Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (12 teeth); Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Great Oolite Group, Bathonian, Middle

TABLE 4. Tooth morphotypes, UK Bathonian sites.

Specimen	Locality	Morphotype	Machine learning		
			Majority vote	Combined posterior probability	P
GLRCM G100-14	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
GLRCM G100-21	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
GLRCM G100-9	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
GLRCM G140-7	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
GLRCM G14-27	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
GLRCM G21-22	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
GLRCM G38-10	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
GLRCM G75704	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
GLRCM G91702	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.85
GLRCM G91705	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.86
GLRCM G91706	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
NHMUK PV R 37904	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37905	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.76
NHMUK PV R 37906	Hornsleasow	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.86
NHMUK PV R 37907	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37908	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37910	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.68
NHMUK PV R 37924	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.77
NHMUK PV R 37925	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.76
NHMUK PV R 37926	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.77
NHMUK PV R 37927	Kirtlington	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37928	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.88
NHMUK PV R 37929	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.75
NHMUK PV R 37935	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37939	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.7
NHMUK PV R 37940	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37941	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
NHMUK PV R 37942	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.83
NHMUK PV R 37944	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.66
NHMUK PV R 37945	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.85
NHMUK PV R 37946	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.84
NHMUK PV R 37947	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.64
NHMUK PV R 37948	Woodeaton	Dromaeosaur A	Troodontidae	Troodontidae	0.72
NHMUK PV R 37949	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.81
NHMUK PV R 37950	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.86
NHMUK PV R 37953	Woodeaton	Dromaeosaur A	Dromaeosaur A	Dromaeosaur A	0.87
GCRLM GTUBE 67	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G10022	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G100-64	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G10-37	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.98
GLRCM G12-28	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G14-22	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G167-24	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM G68-1	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.98
GLRCM G7.219–3	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM GHQ104 C -1	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM GTEMP3061	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
GLRCM GX	Hornsleasow	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 36771	Watton Cliff	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 36778	Watton Cliff	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.98

(continued)

TABLE 4. (Continued)

Specimen	Locality	Morphotype	Machine learning		
			Majority vote	Combined posterior probability	P
NHMUK PV R 37909	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.98
NHMUK PV R 37911	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.99
NHMUK PV R 37912	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.99
NHMUK PV R 37913	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.96
NHMUK PV R 37914	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.97
NHMUK PV R 37915	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37916	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.97
NHMUK PV R 37917	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37918	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37919	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37921	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37922	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.93
NHMUK PV R 37923	Kirtlington	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.83
NHMUK PV R 37930	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.97
NHMUK PV R 37931	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37933	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37934	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37936	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	1
NHMUK PV R 37937	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.92
NHMUK PV R 37938	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.59
NHMUK PV R 37943	Woodeaton	Dromaeosaur B	Troodontidae	Troodontidae	0.58
NHMUK PV R 37951	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.96
NHMUK PV R 37952	Woodeaton	Dromaeosaur B	Dromaeosaur B	Dromaeosaur B	0.73
NHMUK PV R 36779	Watton Cliff	Dromaeosaur C	Dromaeosaur B	Dromaeosaur B	0.63
NHMUK PV R 37920	Kirtlington	Dromaeosaur C	Dromaeosaur C	Dromaeosaur C	0.61
GLRCM G167-32	Hornsleasow	Therizinosauria	Dromaeosaur B	Dromaeosaur B	0.99
GLRCM G8-23	Hornsleasow	Troodontidae	Dromaeosaur B	Dromaeosaur B	0.98

Combined posterior probability, assigned tooth morphotype by combining posterior probabilities from three machine learning models; majority vote, assigned tooth morphotype following simple majority vote of three machine learning models; morphotype, assigned tooth morphotype following machine learning and visual description; P, combined posterior probability value.

Jurassic (10 teeth); Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (13 teeth); Watton Cliff, Forest Marble Formation, Great Oolite Group, Bathonian, Middle Jurassic (two teeth).

Description. Morphotype B crowns (37 in total) are grouped together by the machine learning analysis but show considerable variation in denticle size differences between carinae, hence might encompass several different subgroups with broadly similar morphology. Tooth crowns are ziphodont, slightly larger than morphotype A (CH ranging from 1.66 mm to 19 mm) and have a straight to concave distal margin. Most of the crowns are labiolingually narrow (CBR < 0.6) although four (NHMUK PV R37934, NHMUK PV R36778, NHMUK PV R37911 and NHMUK PV R37931) have a CBR of >0.8 (Fig. 11) and may represent more mesially positioned teeth (Hendrickx *et al.* 2019). In contrast to morphotype A, the depressions on the lingual and labial surfaces are less prominent. Consequently, the basal cross-section of morphotype B ranges from a weaker lemniscate outline to a more oval or lenticulate shape. The mesial and distal carinae are both well developed and extend from the crown apex to just above the crown base with the distal carina often exhibiting a labial deflection basally and the mesial carina (where preserved) twisted slightly lingually. In contrast to morphotype A, both mesial and distal carinae are denticulate. Mesial denticles are restricted to the apical region of the carina but distal denticles extend over the full length of the carina. The distal denticles are generally larger than the mesial denticles with DSDI > 1. However, in some smaller crowns (NHMUK PV R36778, NHMUK PV R37912, NHMUK PV R37913, NHMUK PV R37937, NHMUK PV R37911, NHMUK PV R37951) the DSDI is <1, indicating that the mesial denticles are larger than distal ones. In several crowns (NHMUK PV R37936, NHMUK PV R37943, NHMUK PV R37916, NHMUK PV R37931, GCLRM G167-24, GCLRM G10-37, GCLRM GTube 67, NHMUK PV R37923 and NHMUK PV R37938) the difference in size between mesial and distal denticles is exaggerated, with DSDI > 1.4, and it is possible that they may represent either a variation within this morphotype or a separate morphotype. However, in the absence of any other morphological

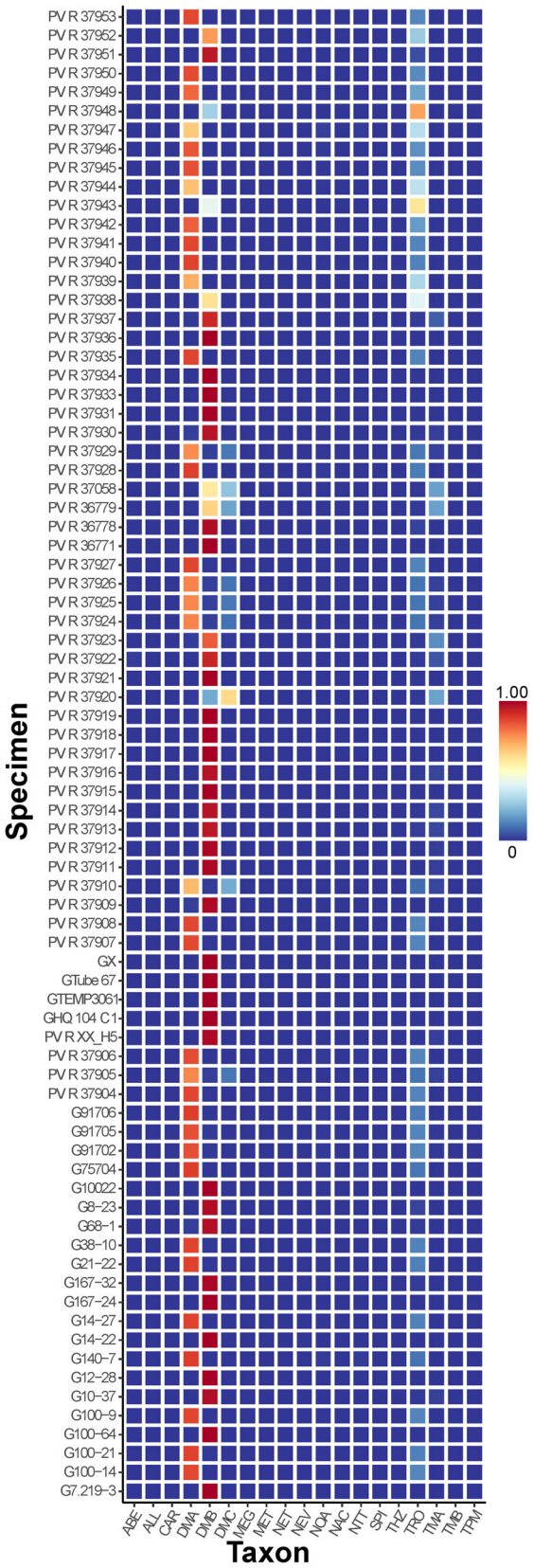

FIG. 8. Posterior probability of the final assigned taxon from the machine learning ensemble classifier for UK Bathonian teeth. Specimens with prefix PV R are Natural History Museum (NHMUK), those with prefix G are Museum of Gloucester (GLRCM). The colour scale represents the posterior probability. *Taxon abbreviations*: ABE, Abelisauridae; ALL, Allosauridae; CAR, Carcharodontosauridae; COE, *Coelophysis*; DMA, Dromaeosaur morphotype A; DMB, Dromaeosaur morphotype B; DMC, Dromaeosaur morphotype C; MEG, Megalosauridae; MET, Metriacanthosauridae; NAC, other Ceratosauria; NEV, Neovenatoridae; NOA, Noasauridae; NTT, other Tyrannosauroidea; SPI, Spinosauridae; THZ, Therizinosauria; TMA, Tyrannosauridae morphotype A; TMB, Tyrannosauridae morphotype B; TPM, Tyrannosauridae pre-maxillary; TRO, Troodontidae.

differences, and the machine learning support for this grouping, we have elected to keep these crowns in morphotype B. Mesial and distal denticles are all rectangular to subrectangular in shape with a convex external margin and are orientated perpendicular to the carina. The crown surface has a braided enamel texture consisting of sinuous grooves and ridges orientated apicobasally (Hendrickx *et al*. 2015*a*, 2019).

Gen. et sp. indet. Morphotype C
Figure 13

Referred specimens. NHMUK PV R36779, R37920.

Localities. Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic (one tooth); Watton Cliff, Forest Marble Formation, Great Oolite Group, Bathonian, Middle Jurassic (one tooth).

Description. Morphotype C includes two small, damaged crowns ranging in CH from 0.61 mm to 1.6 mm with a concave distal margin. The crowns are labiolingually narrow (CBR *c*. 0.5) and the depressions on the lingual and labial surfaces seen in morphotypes A and B are absent or weakly developed, resulting in a subcircular to oval basal cross-section. Both mesial and distal carinae are present, extending from the crown apex to just above the crown base, and are denticulate. Mesial denticles are restricted to the upper half of the carina, and distal denticles extend from the base to just below the crown apex. The mesial carina is extensively worn on both teeth. Denticles on the mesial and distal carinae are equal to subequal in size, with a DSDI of 1.1. The serration density on both the mesial and distal carinae is substantially greater than in morphotype B, with mesial denticles ranging from 15 per mm (NHMUK PV R36779) to 18 per mm (NHMUK PV R37920) and distal denticles from 13 per mm (NHMUK PV R36779) to 17.4 per mm (NHMUK PV R37920). By contrast, morphotype B mesial denticles average

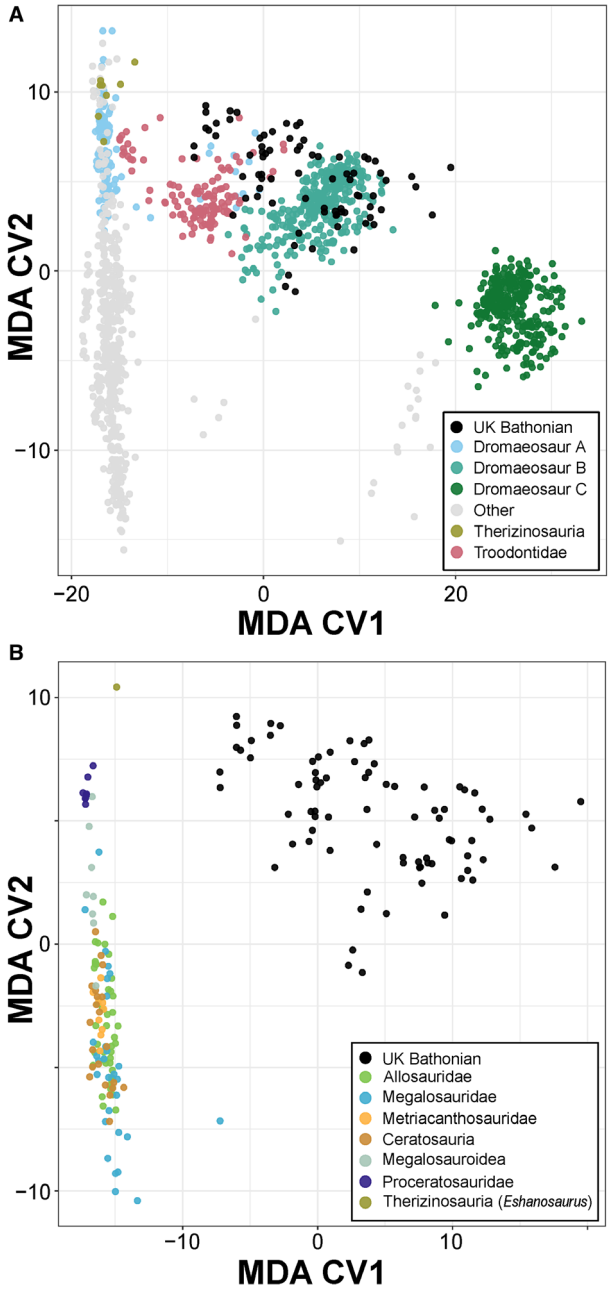

FIG. 9. Trained feature-space occupation of UK Bathonian teeth compared with training data based on two mixture discriminant analysis (MDA) dimensions. A, compared with all taxa in the training data with Maniraptoran clades highlighted. B, compared with Jurassic taxa. CV, canonical variate.

8.7 per mm and distal denticles, 7.0 per mm. Both mesial and distal denticles are rectangular to subrectangular in shape with a convex external margin and are orientated perpendicular to the carina. These small teeth, although damaged and worn in places, appear to represent a morphotype distinct from morphotype B based on their smaller size, and greater serration density on both carinae.

TROODONTIDAE Gilmore, 1924
Gen. et sp. indet.
Figure 14

Referred specimen. GCLRM G8-23.

Locality. Hornsleasow Quarry, Chipping Norton Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic.

Description. GLRCM 8-23 is a small, almost complete isolated tooth with a distinctive morphology. The tooth shows some damage at the base of the distal carina and at the crown apex where denticles are missing. The crown is small (CH 2.9 mm) and phylloform, with a slight lingual inclination. It is labiolingually compressed (CBR 0.53), lenticular in basal cross-section and has a weak constriction at the base. The distal margin of the crown is straight to weakly concave and the mesial margin is convex. The mesial and distal carinae are both denticulate with large, prominent and apically orientated denticles. The mesial carina reaches the base of the crown: however, due to damage it is not possible to confirm this for the distal carina. Distal denticles are both significantly larger and fewer in number than the mesial denticles with a DSDI of 1.43. Both mesial and distal denticles appear to extend from the base of the crown to the apex, although damage to the basal portion of the distal carina obscures this somewhat. Mesial denticles decrease in size both apically and basally from the crown midpoint whereas distal denticles increase in size slightly towards the apex. Distal denticles are subrectangular in shape, being slightly longer mesiodistally than apicobasally, and have convex external margins. The denticles are aligned perpendicular to the carina towards the base of the crown but become apically orientated and hooked midway along the carina. Mesial denticles have a parallelogram-shaped outline in labial view caused by the apical orientation of the denticles along the carina. Grooves are present between adjacent denticles on both carinae but do not extend to the crown surface.

THERIZINOSAUROIDEA Maleev, 1954
Gen. et sp. indet.
Figure 15

Referred specimen. GCLRM G167-32.

Locality. Hornsleasow Quarry, Chipping Norton Limestone Formation, Great Oolite Group, Bathonian, Middle Jurassic.

Description. GCLRM G167-32 is an isolated complete crown that is phylloform in shape, labiolingually compressed and subsymmetrical in both lingual and labial views with convex mesial and distal margins. The crown is small with a crown height of 3.5 mm, a maximum width of 2.8 mm (decreasing to 2.4 mm at the crown base: crown base occupying around 85% of the maximum crown width, CBR = 0.73), and has a small basal

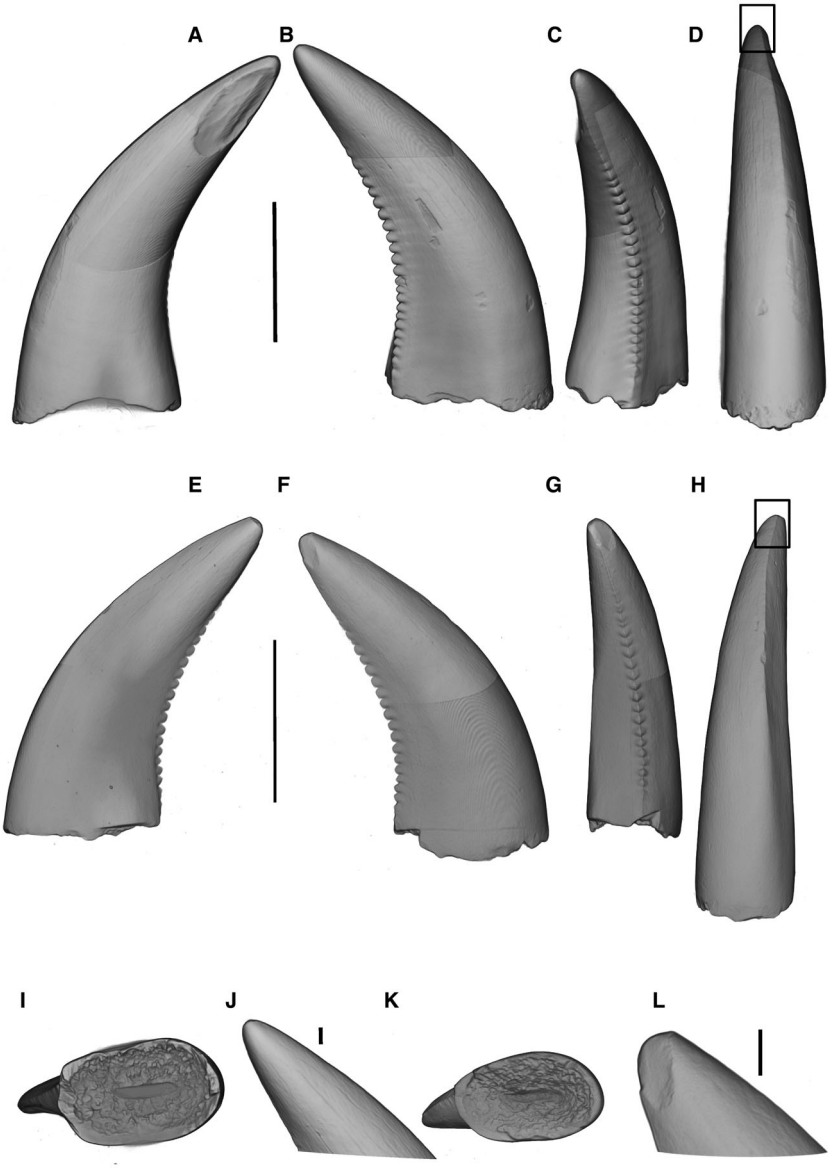

FIG. 10. Isolated crowns of indeterminate dromaeosaurs (morphotype A) from Woodeaton Quarry (NHMUK PV R37059; A–D, I–J) and Kirtlington Quarry (NHMUK PV R37925; E–H, K–L) in: A, E, lingual; B, F, labial; C, G, distal; D, H, mesial; I, K, basal view. J, L, magnification of apical regions boxed in D and H, respectively. Scale bars represent: 1 mm for general views (A–I, K); 0.1 mm for apical close-ups (J, L).

constriction. The labial surface is strongly convex. The lingual surface is dominated by a median ridge running from apex to base forming a slightly convex profile bounded by mesial and distal concave depressions adjacent to both carinae. Carinae are present on both margins of the teeth with the mesial carina restricted to the upper half of the crown and the distal carina extending toward, but not reaching, the crown base. Both carinae are denticulated with fewer, and larger, denticles towards the apex than at the mid-crown position. Average denticle sizes on both carinae are equal with the distal carina ranging from 6.3 per mm at mid-crown to 5.9 per mm apically and the mesial carina being 6.6 per mm at mid-crown to 5.7 per mm apically. Denticles appear to reach almost to the apex of the crown although slight damage and wear at the apex obscures this. The denticles are rectangular, being slightly longer apico-basally, have a convex exterior margin and are slightly inclined apically.

Morphological comparisons

Dromaeosaurid morphotypes. We interpret morphotypes A–C as dromaeosaurids based on the machine learning classification and several morphological characters of the teeth, which, in

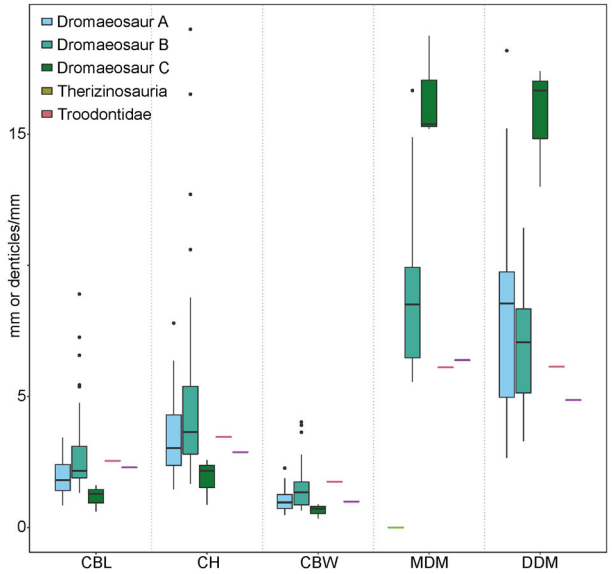

FIG. 11. Range of morphometric measurements across each maniraptoran morphotype. *Abbreviations*: CBL, crown base length; CBW, crown base width; CH, crown height; DDM, distal denticles per mm; MDM, mesial denticles per mm.

conjunction with their small size, preclude other taxa. The small crown size of these teeth tends to rule out taxa such as Tyrannosauroidea (excluding Proceratosauridae), Allosauroidea, Carcharodontosauridae, Ceratosauridae and Megalosauridae, although there is a possibility that these teeth may have come from juvenile individuals of larger adult taxa. Dromaeosaurid teeth have the following combination of features: relatively small size, with even larger-bodied taxa such as *Utahraptor* having teeth that are less than 5 cm in crown height (Hendrickx *et al.* 2019); ziphodonty; labiolingual compression; a twisted mesial carina; a distal carina that is deflected labially; and denticles present on either both the mesial and distal carinae or only the distal carina (Sankey *et al.* 2002; Hendrickx *et al.* 2019; Prasad & Parmar 2020). Dentitions with unserrated mesial carinae and denticulate distal carinae (as per morphotype A) are present in numerous clades in both the mesial and lateral dentitions, including a number of small dromaeosaurids (*Atrociraptor*, *Richardoestesia*, saurornitholestines, velociraptorines and dromaeosaurines) from the Late Cretaceous of North America (Larson 2008; Williamson & Brusatte 2014; Larson *et al.* 2016) and the Asian velociraptorine *Tsaagan* (Norell *et al.* 2006; Chiarenza *et al.* 2020). Larger distal denticles compared with mesial denticles has previously been used as a synapomorphy to identify velociraptorine teeth (e.g. Rauhut & Werner 1995; Sweetman 2004; van der Lubbe *et al.* 2009) however, this feature is found across a range of deinonychosaurian taxa such as saurornitholestines and *Richardoestesia* (Larson 2008; Larson & Currie 2013; Larson *et al.* 2016; Hendrickx *et al.* 2019; Chiarenza *et al.* 2020). A lemniscate cross-section is a feature present in many deinonychosaurs, including most dromaeosaurids, but with the exception of some metriacanthosaurids, megaraptorans and tyrannosauroids it is absent from non-maniraptoriform theropods (Hendrickx & Mateus 2014; Hendrickx *et al.* 2019).

Troodontid morphotype. We refer the single tooth GCLRM G8-23 to Troodontidae on both morphological-based considerations and machine learning morphospace position. GCLRM G8-23 resembles the teeth of troodontids based on its large, bulbous, widely spaced and apically inclined denticles on the distal carina, the overall phylloform shape of the crown, and the presence of a basal constriction. The presence of denticles on both the mesial and distal carinae is seen in derived troodontids (Makovicky *et al.* 2003; Goswami *et al.* 2013; Hendrickx *et al.* 2019), with only *Troodon formosus* (Leidy 1856), *Zanabazar junior* (Norell *et al.* 2009), *Saurornithoides mongoliensis* (Osborn 1924; Hendrickx *et al.* 2019) and a single isolated tooth from the Late Cretaceous of India (Goswami *et al.* 2013) having serrated mesial and distal carinae in at least part of the dentition. Abelisaurid lateral teeth also share this denticle morphology: however, the distal margins of most abelisaurid crowns, with a few exceptions, tend to be convex rather than straight to weakly concave and have a triangular crown shape rather than the phyllodont shape seen here (Hendrickx & Mateus 2014).

Therizinosauroid morphotype. We refer the single tooth GCLRM G167-32 to Therizinosauroidea on morphological-based considerations only, given that this tooth was incorrectly classified as a dromaeosaurid in the machine learning analysis. A subsymmetrical phylloform-shaped tooth with a basal constriction as seen in GCLRM G167-32 are features shared with therizinosauroids such as *Falcarius* (Kirkland *et al.* 2005; Zanno 2010*a*), *Erlikosaurus* (Barsbold & Perle 1980) and *Eshanosaurus* (Xu *et al.* 2001; Barrett 2009). The median ridge on the lingual surface running from crown apex to base resulting in concave surfaces adjacent to both carinae is consistent with that observed in the lateral teeth of therizinosaurs (Zanno 2010*a*; Hendrickx *et al.* 2019). The possession of a small number of large denticles on the carina is a feature shared between Therizinosauroidea and Troodontidae (Hendrickx *et al.* 2019): however, the sub-equal size of denticles on both carinae of GCLRM G167-32 and the regular, narrow spacing between denticles is in contrast to the large, bulbous and widely spaced denticles often seen in troodontids (Hendrickx *et al.* 2019). The convex distal margin of the crown in GCLRM G167-32 is a feature shared with a number of non-maniraptoran theropods (Abelisauridae, Spinosauridae and *Ceratosaurus*) and some maniraptorans (Ornithomimosauria, Alvarezsauroidea and Oviraptorosauria): however, the combination of crown shape and denticle morphology precludes these taxa. GCLRM G167-32 possesses a basal constriction, the crown base occupying *c.* 85% of maximum crown width, a feature assessed by Hendrickx *et al.* (2019) to represent an unambiguous dental synapomorphy of Therizinosauria.

Comparisons with other UK Middle Jurassic theropod taxa

Of the theropod dinosaurs known from the Middle Jurassic of the UK we can exclude larger non-maniraptoran taxa such as the megalosaurids *Megalosaurus* (Buckland 1824; Benson 2010*a*), *Magnosaurus* (Huene 1932; Benson 2010*b*), *Duriavenator* (Owen 1883; Waldman 1974; Benson 2008) and *Eustreptospondylus* (Walker 1964; Sadleir *et al.* 2008), and the basal tetanuran

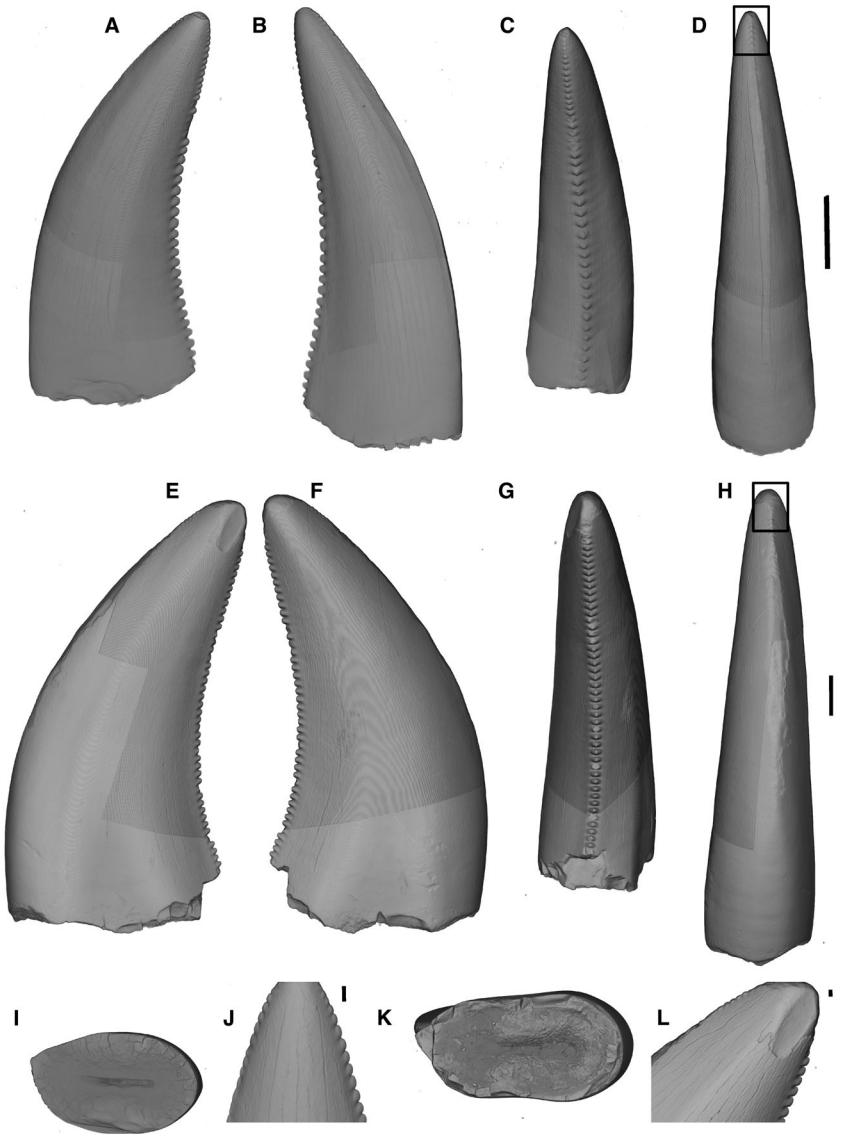

FIG. 12. Isolated crowns of indeterminate dromaeosaurs (morphotype B) from Woodeaton Quarry (NHMUK PV R37916; A–D, I–J) and Hornsleasow Quarry (GCLRM G167-24; E–H, K–L) in: A, E, lingual; B, F, labial; C, G, distal; D, H, mesial; I, K, basal view. J, L, magnification of apical regions boxed in D and H, respectively. Scale bars represent: 1 mm for general views (A–I, K); 0.1 mm for apical close-ups (J, L).

Cruxicheiros (Benson & Radley 2010) on the basis of morphospace position (Fig. 9), size and overall morphology (Hendrickx *et al.* 2015*a*). The teeth of an earlier-branching coelurosaur, the tyrannosauroid *Proceratosaurus bradleyi* (Rauhut *et al.* 2010), from the Bathonian of England, bear a superficial similarity to dromaeosaur morphotype B. However, the overall crown shape of *Proceratosaurus* maxillary and dentary teeth differs in that the basal part of the crown is almost straight with only the apical part strongly recurved. In addition, the basal longitudinal depressions on both lingual and labial surfaces are strongly developed, giving a clear lemniscate basal cross-section in contrast to that present in morphotype B (Rauhut *et al.* 2010). Moreover, *Proceratosaurus* denticles are chisel-shaped with flattened exterior margins in contrast to the convex margin seen in morphotype B (Woodward 1910; Rauhut *et al.* 2010).

DISCUSSION

The application of machine learning techniques, combined with morphological-based approaches, to isolated teeth from Bathonian microvertebrate sites confirms the presence of at least three maniraptoran taxa in the assemblage: three dromaeosaur morphotypes (which might indicate multiple dromaeosaur taxa); a troodontid; and a

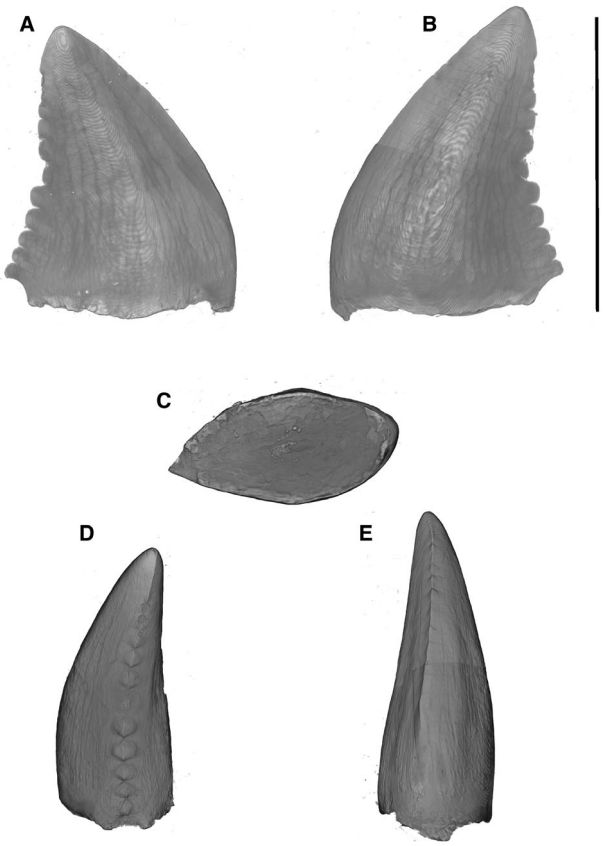

FIG. 13. Isolated crown of indeterminate dromaeosaur (morphotype C) from Woodeaton Quarry (NHMUK PV R37920), in: A, lingual; B, labial; C, basal; D, distal; E, mesial view. Scale bar represents 1 mm.

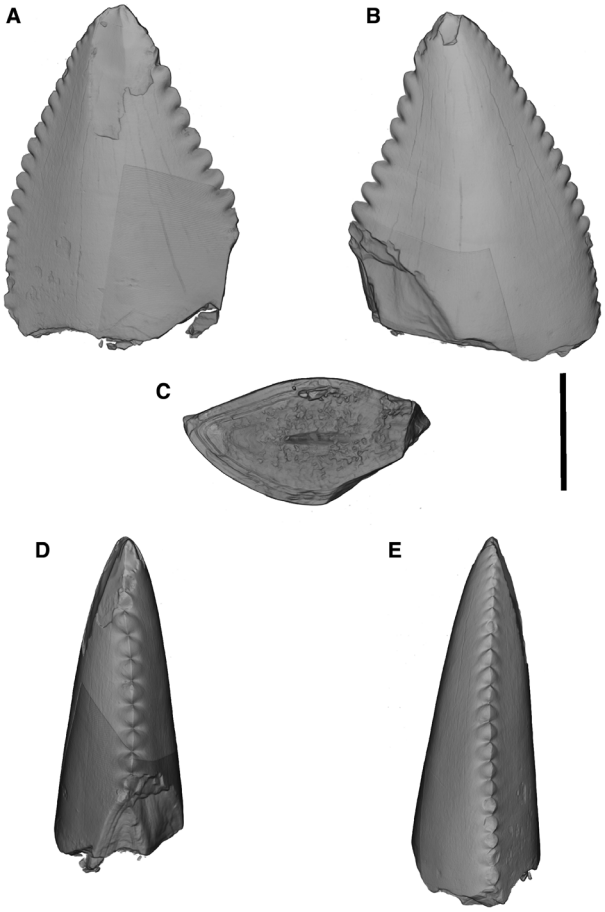

FIG. 14. Isolated crown of an indeterminate troodontid (GCLRM G8-23) from Hornsleasow Quarry, Gloucestershire in: A, lingual; B, labial; C, basal; D, distal; E, mesial view. Scale bar represents 1 mm.

therizinosaur. Apart from morphological changes due to taxonomy, there are a number of other sources of possible variation in tooth morphology including those potentially introduced by positional or ontogenetic changes. Unfortunately there are few relevant datasets or previous studies on theropod tooth variation on which to rigorously test these hypotheses. Buckley *et al.* (2010) and Buckley & Currie (2014) examined tooth variation in single populations of the tyrannosaurid *Albertosaurus sarcophagus* (Buckley *et al.* 2010) and the Late Triassic theropod *Coelophysis bauri* (Buckley & Currie 2014). The analysis of *A. sarcophagus* teeth suggests that strongly heterodont dentitions can influence morphospace occupation, with premaxillary teeth quantifiably different to maxillary and dentary teeth but with no quantifiable difference between maxillary and dentary teeth. Analysis of 848 teeth from 23 skulls of *C. bauri* using both discriminant analysis and canonical variate analysis shows that positional variation does not influence morphospace occupation but that it can be influenced by ontogeny. This does suggest that a degree of caution is warranted when ascribing morphotypes of isolated theropod teeth to different taxa; hence here we distinguish the teeth only as morphotypes within a broader taxonomic framework.

These results provide the first quantitative support for the presence of maniraptoran theropods in the Middle Jurassic, from sites that are well constrained biostratigraphically in Bathonian ammonite zones, increase the known diversity of Middle Jurassic theropods from the UK, and provide the oldest occurrences of troodontids and therizinosaurs worldwide (Fig. 7). These identifications provide the first definitive body-fossils consistent with predictions made by phylogenetic analyses, which posited the likely presence of these clades at this time (Holtz 2000; Rauhut 2003; Xu *et al.* 2010; Carrano *et al.* 2012; Rauhut & Foth 2020). Previous reports of Middle Jurassic maniraptoran occurrences have been disputed (Foth & Rauhut 2017; Ding *et al.* 2020) or have considerable temporal and stratigraphic confusion (Sullivan *et al.* 2014; Xu *et al.* 2016). The age of the paravians from the Middle to Upper Jurassic Daohugou Beds (Yanliao biota) in northeastern China is controversial because

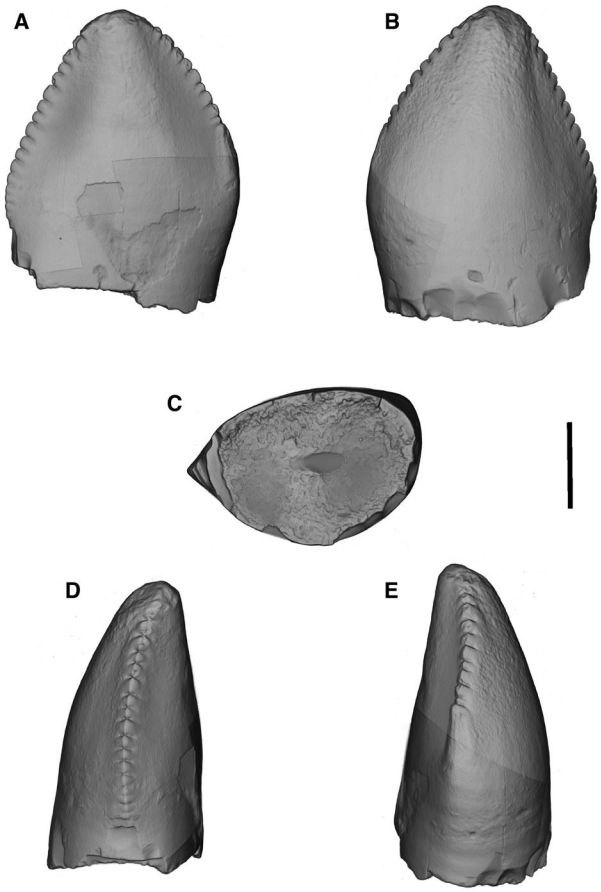

FIG. 15. Isolated crown of an indeterminate therizinosauroid (GCLRM G167-32) from Hornsleasow Quarry, Gloucestershire in: A, lingual; B, labial; C, basal; D, distal; E, mesial view. Scale bar represents 1 mm.

of stratigraphic uncertainties surrounding the placement of volcanic rocks in the sequence used to obtain radiometric dates (Sullivan *et al.* 2014; Xu *et al.* 2016). The beds are close to the Middle–Upper Jurassic boundary and have been referred to either the upper part of the Jiulongshan (Haifanggou) Formation and/or the lower part of the overlying Tiaojishan Formation (or both). Recent radiometric dating suggests that the *Anchiornis*-bearing bed is Oxfordian in age with most anchiornithines coming from the Tiaojishan Formation and scansoriopterygids from the underlying Jiulongshan Formation. Notwithstanding this, the uncontroversial acceptance of avialians such as Archaeopteryx (Meyer 1861; Owen 1864; Huxley 1868) from the Late Jurassic of Germany, as well as the Yanliao biota maniraptorans (Godefroit *et al.* 2013*a*, 2013*b*; Sullivan *et al.* 2014; Xu *et al.* 2015; Foth & Rauhut 2017) and probable maniraptorans from the Middle to Upper Jurassic Shishugou Formation (Han *et al.* 2011), clearly indicate that the clade should have been established by the late Middle Jurassic (Choiniere *et al.* 2012).

Our results show that Maniraptora was not only established by the Bathonian but was already diverse at this time, at least in Laurasia, and also extend significantly the known temporal ranges of all major maniraptoran clades. Therizinosaurians, excluding the controversial occurrence of *Eshanosaurus* (Xu *et al.* 2001; Barrett 2009), are currently known mainly from the Cretaceous of Asia apart from the basal, and oldest, therizinosauroids *Falcarius* and *Martharaptor* from the Berriasian Cedar Mountain Formation of Utah (Kirkland *et al.* 2005; Senter *et al.* 2012; Joeckel *et al.* 2020) and the Turonian taxon *Nothronychus* from New Mexico and Utah (Kirkland & Wolfe 2001). The occurrence of a therizinosaur in the Bathonian of the UK extends the temporal range of this clade by *c.* 27 myr (Fig. 7). Dromaeosaurs had an almost pan-global distribution during the Late Cretaceous, although they are best known from Asia and North America. The earliest definitive dromaeosaurs, excluding records of referred isolated teeth, are from the Barremian Jehol biota of China (Xu *et al.* 2000; Zheng *et al.* 2009). Isolated teeth from the Middle and Late Jurassic of Laurasia and Gondwana have been assigned to the clade previously (Zinke 1998; Hendrickx & Mateus 2014; Vullo *et al.* 2014; Prasad & Parmar 2020) but their identifications have not been widely accepted (Foth & Rauhut 2017; Ding *et al.* 2020; Sellés *et al.* 2021). Our results, however, offer the first quantitative assessment of potential dromaeosaur teeth from the Middle Jurassic, confirming the existence of the clade by the Bathonian and a confirmed range extension of some 38 myr (Fig. 7). Based on comparisons with our data, it seems likely that some other published Jurassic records also represent this clade, although rigorous analysis will be needed to confirm this suggestion. Troodontids are known primarily from the Cretaceous of Asia, Europe and North America (Brown & Schlaikjer 1943; Russell 1946; Barsbold *et al.* 1987; Currie 1987; Sellés *et al.* 2021) and possibly the Late Jurassic of China (Hu *et al.* 2009; Turner *et al.* 2012; Brusatte *et al.* 2014), although more recent analyses consider these Late Jurassic taxa to be basal avialians (Foth & Rauhut 2017; Pei *et al.* 2017). Isolated teeth from the Late Jurassic of Portugal and North America and the Late Cretaceous of India have been assigned to the clade (Chure 1994; Zinke 1998; Goswami *et al.* 2013) although many of these identifications have been questioned (Ding *et al.* 2020). Thus, our confirmed Middle Jurassic European troodontid pushes back the origin of this clade by 27 myr (Fig. 7) from the Berriasian (*Geminiraptor*, Utah; Senter *et al.* 2010) to the Bathonian.

The presence of this diverse Middle Jurassic biota also suggests we need to re-visit the biogeographical scenarios that have been proposed to account for patterns in maniraptoran faunal distributions (Case *et al.* 2007; Rauhut *et al.* 2010; Zanno 2010*b*; Ding *et al.* 2020). Two non-

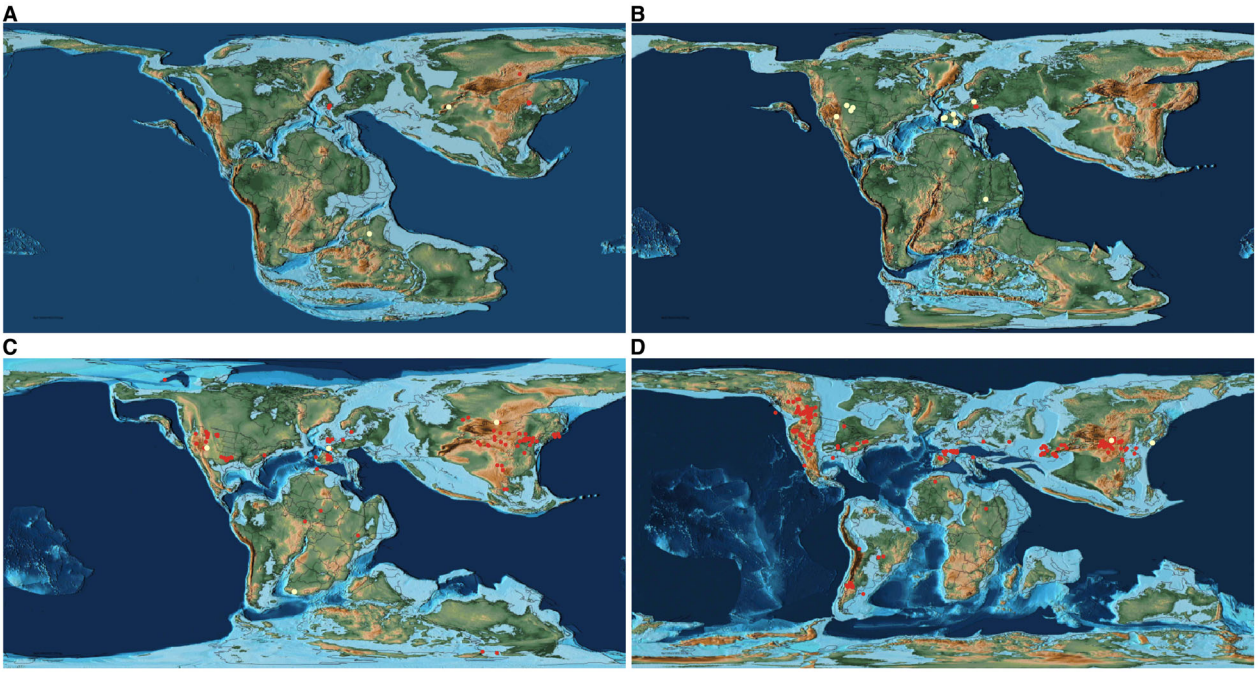

FIG. 16. Biogeographical history of Maniraptora from the Middle Jurassic to the Late Cretaceous. A, Middle Jurassic. B, Late Jurassic. C, Early Cretaceous. D, Late Cretaceous. Maniraptoran occurrences shown in red circles and unconfirmed occurrences in yellow circles. Middle Jurassic occurrences include the results from this study, dromaeosaurs reported from the Kota Formation of India (Prasad & Parmar 2020) and possible occurrences in Ethiopia (Goodwin *et al.* 1999) and Kyrgyzstan (Averianov *et al.* 2005). Palaeogeographical maps from Scotese (2021).

mutually exclusive scenarios are widely accepted as having major impacts on maniraptoran biogeographical distributions: vicariance from a widespread initial distribution, driven by continental break-up and fragmentation (Fastovsky & Weishampel 1996; Upchurch *et al.* 2002; Zanno 2010*b*; Ding *et al.* 2020), and faunal dispersal with dispersal routes shaped by the establishment of land bridges between continental masses (Upchurch *et al.* 2002; Dunhill *et al.* 2016; Ding *et al.* 2020). It is also likely that regional extinction events played a part in shaping biogeographical distributions (Sereno 1997; Barrett *et al.* 2011; Benson *et al.* 2012). The presence of Middle Jurassic Laurasian proceratosaurids and earliest Cretaceous Gondwanan ornithomimosaurs suggests that coelurosaurs were widespread before the break-up of Pangaea (Rauhut *et al.* 2010; Choiniere *et al.* 2012), with a recent analysis by Ding *et al.* (2020) suggesting that continental-scale vicariance was an important factor in accounting for coelurosaurian biogeographical distributions. Due to the uncertainty created by the absence of definitive and temporally well constrained pre-Cretaceous maniraptorans (Zanno 2010*b*; Foth & Rauhut 2017; Sellés *et al.* 2021), several different scenarios have been put forward to account for maniraptoran distributions while accepting that more fossil evidence would be needed in order to test these. For example, Foth & Rauhut (2017) suggested that all maniraptoran clades more derived than *Ornitholestes* originated and diversified in eastern Asia, followed by dispersal from this area to Europe and North America by the Late Jurassic. By contrast, the pan-Laurasian distribution of Early Cretaceous therizinosaurs has been taken to indicate either a vicariance event, with therizinosaurs present in Asia and North America prior to major rifting and the opening of the North Atlantic, or a dispersal of basal therizinosaurs between North America and Asia via land bridges after the rifting event (Zanno 2010*b*; Ding *et al.* 2020; Scotese 2021). Dromaeosaur biogeography has been suggested to indicate a vicariance event driven by the break-up of Pangaea and subsequent continental separation (Ding *et al.* 2020), implying a widespread distribution before the break-up. Troodontids are common across Asia and North America by the Campanian and Maastrichtian, and a tooth attributed to the clade has been reported from the Late Cretaceous of India, the first Gondwanan representative of the clade (Goswami *et al.* 2013), although Ding *et al.* (2020) suggested that this identification should be provisional. Possible scenarios to account for troodontid biogeography include multiple Laurasian dispersal events, a dispersal event from Laurasia to Gondwana or a wider clade distribution prior to the break-up of Pangaea (Senter *et al.* 2010; Goswami *et al.* 2013; Ding *et al.* 2020).

The presence of maniraptorans in the Middle Jurassic (Fig. 16) suggests that a pan-Pangaean distribution was established before continental separation began at *c.* 170 Ma (Scotese 2021). A combination of vicariance events driven by continental separation, regional extinctions and later dispersal events can be invoked that then lead to the later Mesozoic distributions.

Machine learning provides a powerful new tool that can provide quantitative assessments of isolated theropod tooth identifications and has been shown to outperform other analytical methods (Wills *et al.* 2021). The use of multiple machine learning algorithms. as applied here, enables the corroboration of results by checking predictions derived from another technique. It is also important to note the limitations of any technique and our study was constrained (due to the nature of the training datasets available) to a small number of morphometric variables. Moreover, data availability was too poor to accurately describe a model in some cases. However, we expect the ability to classify isolated teeth in this manner to improve with the collection of more data (including 3D data) to train the classifiers. For now, we emphasize the importance of cross-checking results from machine learning analyses with more traditional morphological-based approaches.

CONCLUSION

The use of machine learning algorithms has enabled us to confirm, in a quantifiable framework, the presence of a diverse maniraptoran theropod fauna in the Middle Jurassic (Bathonian) of the UK. Our sample includes the oldest-known occurrences of Troodontidae and Therizinosauroidea. This confirms a Middle Jurassic (or earlier) origin for Maniraptora and suggests that the clade had a pan-Pangaean distribution prior to continental break-up. The presence of these early maniraptorans, currently known only from isolated teeth, highlights the importance of incorporating microvertebrate remains into faunal and evolutionary analyses. The accuracy of machine learning results is hampered by the quality of the data used to train the models, and larger datasets will be required to improve model performance, but the combination of these results with morphological-based identifications can overcome this issue to provide a robust, testable framework for taxonomic identifications.

Acknowledgements. We would like to thank Susan Evans for access to previously collected material from Kirtlington and Watton Cliff, and David Rice and the Museum of Gloucester for access to the Hornsleasow material. Michael McKenna provided access to Woodeaton quarry, and Breedon Aggregates provided access to Hornsleasow quarry. We also thank Vincent Fernandez, Brett Clarke and Alex Ball at the Imaging and Analysis Centre, Natural History Museum for CT and SEM access and scanning; Christophe Hendrickx for assistance with identifications and useful discussions; and Susannah Maidment for registering the new specimens in the collections of the Natural History Museum and for assistance in the field. We are grateful to all of the field crews at Woodeaton and Hornsleasow quarries: David Ward, Emma Bernard, Pip Brewer, Martha Richter, Tim Ewin, Mark Graham, Adrian Lister, Stephen Stukins, Andy Gale, Zoë Hughes, Alison Ward, Andrew Ward, Mike Smith, Karen Banton, Claire Bullar, Rebecca Schofield and Joe Bonsor. Additional thanks go to Jerry Hooker and Natalie Cooper for useful discussions and to the reviewers Elisabete Malafaia and Jonah Choiniere for their useful comments that helped to improve this contribution.

Author contributions. **Conceptualization** S Wills (SW), CJ Underwood (CJU), PM Barrett (PMB); **Data Curation** SW; **Formal Analysis** SW; **Funding Acquisition** SW, CJU, PMB; **Investigation** SW; **Methodology** SW; **Project Administration** SW, CJU, PMB; **Resources** SW, CJU, PMB; **Software** SW; **Supervision** CJU, PMB; **Validation** SW, CJU, PMB; **Visualization** SW; **Writing – Original Draft Preparation** SW; **Writing – Review & Editing** SW, CJU, PMB.

DATA ARCHIVING STATEMENT

Data for this study (including R scripts and morphometric data) are available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.6q573n61w). Image data are available in MorphoSource (https://www.morphosource.org/projects/000457042).

Editor. Philip Mannion

SUPPORTING INFORMATION

Additional Supporting Information can be found online (https://doi.org/10.1002/spp2.1487):

Appendix S1. List of DOIs for CT image series available in MorphoSource (specimens GLRCM G167-24, G167-32; NHMUK PV R30759, R37925, R37916, 37 920).

REFERENCES

ANDERSON, M. J. 2017. Permutational multivariate analysis of variance (PERMANOVA). 1–15. *In* Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F. and Teugels, J. L. (eds) *Wiley StatsRef: Statistics Reference Online*, https://doi.org/10.1002/9781118445112.stat07841

ANDERSON, M. J. and WALSH, D. C. I. 2013. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? *Ecological Monographs*, **83**, 557–574.

ARGAST, S., FARLOW, J. O., GABET, R. M. and BRINKMAN, D. L. 1987. Transport-induced abrasion of fossil reptilian teeth: implications for the existence of Tertiary dinosaurs in the Hell Creek Formation, Montana. *Geology*, **15**, 927–930.

AUGUIE, B. 2017. gridExtra: Miscellaneous Functions for "Grid" Graphics. R package v.2.3. https://cran.r-project.org/web/packages/gridExtra/index.html

AVERIANOV, A. O., MARTIN, T. and BAKIROV, A. A. 2005. Pterosaur and dinosaur remains from the Middle Jurassic Balabansai Svita in the northern Fergana Depression, Kyrgyzstan (Central Asia). *Palaeontology*, **48**, 135–155.

BARRETT, P. M. 2009. The affinities of the enigmatic dinosaur *Eshanosaurus deguchiianus* from the Early Jurassic of Yunnan Province, People's Republic of China. *Palaeontology*, **52**, 681–688.

BARRETT, P. M., BENSON, R. B. J., RICH, T. H. and VICKERS-RICH, P. 2011. First spinosaurid dinosaur from Australia and the cosmopolitanism of Cretaceous dinosaur faunas. *Biology Letters*, **7**, 933–936.

BARRON, A. J. M., LOTT, G. K. and RIDING, J. B. 2012. Stratigraphical framework for the Middle Jurassic strata of Great Britain and the adjoining continental shelf, British Geological Survey Research Report, RR/11/06, 187 pp.

BARSBOLD, R. and PERLE, A. 1980. Segnosauria, a new infraorder of carnivorous dinosaurs. *Acta Palaeontologica Polonica*, **25**, 187–195.

BARSBOLD, R., OSMÓLSKA, H. and KURZANOV, S. M. 1987. On a new troodontid (Dinosauria, Theropoda) from the Early Cretaceous of Mongolia. *Acta Palaeontologica Polonica*, **32**, 121–132.

BENSON, R. B. J. 2008. A redescription of '*Megalosaurus*' *hesperis* (Dinosauria: Theropoda) from the Inferior Oolite (Bajocian, Middle Jurassic) of Dorset, United Kingdom. *Zootaxa*, **1931**, 57–67.

BENSON, R. B. J. 2010*a*. A description of *Megalosaurus bucklandii* (Dinosauria: Theropoda) from the Bathonian of the UK and the relationships of Middle Jurassic theropods. *Zoological Journal of the Linnean Society*, **158**, 882–935.

BENSON, R. B. J. 2010*b*. The osteology of *Magnosaurus nethercombensis* (Dinosauria, Theropoda) from the Bajocian (Middle Jurassic) of the United Kingdom and a re-examination of the oldest records of tetanurans. *Journal of Systematic Palaeontology*, **8**, 131–146.

BENSON, R. B. J. and RADLEY, J. D. 2010. A new large-bodied theropod dinosaur from the Middle Jurassic of Warwickshire, United Kingdom. *Acta Palaeontologica Polonica*, **55**, 35–42.

BENSON, R. B. J., RICH, T. H., VICKERS-RICH, P. and HALL, M. 2012. Theropod fauna from southern Australia indicates high polar diversity and climate-driven dinosaur provinciality. *PLoS One*, **7**, e37122.

BENTON, M. J., COOK, E. and HOOKER, J. J. 2005. *Mesozoic and Tertiary fossil mammals and birds of Great Britain*. Geological Conservation Review Series No. 32, Joint Nature Conservation Committee, 215 pp.

BREIMAN, L. 2001. Random forests. *Machine Learning*, **45**, 5–32.

BROWN, B. and SCHLAIKJER, E. M. 1943. A study of the troodont dinosaurs with the description of a new genus and four new species. *Bulletin of the American Museum of Natural History*, **82**, 115–150.

BRUSATTE, S. L. and CLARK, N. D. L. 2015. Theropod dinosaurs from the Middle Jurassic (Bajocian–Bathonian) of Skye, Scotland. *Scottish Journal of Geology*, **51**, 157–164.

BRUSATTE, S. L., LLOYD, G. T., WANG, S. C. and NORELL, M. A. 2014. Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur–bird transition. *Current Biology*, **24**, 2386–2392.

BUCKLAND, W. 1824. Notice on *Megalosaurus*. *Transactions of the Geological Society of London*, **1**, 390–396.

BUCKLEY, L. G. and CURRIE, P. J. 2014. Analysis of intraspecific and ontogenetic variation in the dentition of *Coelophysis bauri* (Late Triassic), and implications for the systematics of isolated theropod teeth. *New Mexico Museum of Natural History & Science Bulletin*, **63**, 1–72.

BUCKLEY, L. G., LARSON, D. W., REICHEL, M. and SAMMAN, T. 2010. Quantifying tooth variation within a single population of *Albertosaurus sarcophagus* (Theropoda: Tyrannosauridae) and implications for identifying isolated teeth of tyrannosaurids. *Canadian Journal of Earth Sciences*, **47**, 1227–1251.

CARRANO, M. T., BENSON, R. B. J. and SAMPSON, S. D. 2012. The phylogeny of Tetanurae (Dinosauria: Theropoda). *Journal of Systematic Palaeontology*, **10**, 211–300.

CASE, J. A., MARTIN, J. E. and REGUERO, M. A. 2007. A dromaeosaur from the Maastrichtian of James Ross Island and the Late Cretaceous Antarctic dinosaur fauna. 1–4. *In* COOPER, A. K., RAYMOND, C. R. and ISAES EDITORIAL TEAM (eds) *Antarctica: A keystone in a changing world*. Online proceedings of the 10th ISAES. US Geological Survey and the National Academies. USGS Open-File Report 2007-1047, Short Research Paper, **083**.

CHANNON, P. J. 1950. New and enlarged Jurassic sections in the Cotswolds. *Proceedings of the Geologists' Association*, **61**, 242–260.

CHIARENZA, A. A., FIORILLO, A. R., TYKOSKI, R. S., McCarthy, P. J., FLAIG, P. P. and CONTRERAS, D. L. 2020. The first juvenile dromaeosaurid (Dinosauria: Theropoda) from Arctic Alaska. *PLoS One*, **15**, e0235078.

CHOINIERE, J. N., FORSTER, C. A. and DE KLERK, W. J. 2012. New information on *Nqwebasaurus thwazi*, a coelurosaurian theropod from the Early Cretaceous Kirkwood Formation in South Africa. *Journal of African Earth Sciences*, **71–72**, 1–17.

CHURE, D. C. 1994. *Koparion douglassi*, a new dinosaur from the Morrison Formation (Upper Jurassic) of Dinosaur National Monument; the oldest troodontid (Theropoda: Maniraptora). *Brigham Young University, Geology Studies*, **40**, 11–15.

COPE, J. C. W., DUFF, K. L., PARSONS, C. F., TORRENS, H., WIMBLEDON, W. A. and WRIGHT, J. K. 1980. *A correlation of Jurassic rocks in the British Isles. Part Two: Middle and Upper Jurassic*. Geological Society of London, Special Report No. 15, 109 pp.

COX, B. M. and SUMBLER, M. G. 2002. *British Middle Jurassic stratigraphy*. Geological Conservation Review Series, Joint Nature Conservation Committee, Peterborough, 508 pp.

CURRIE, P. J. 1987. Bird-like characteristics of the jaws and teeth of troodontid theropods (Dinosauria, Saurischia). *Journal of Vertebrate Paleontology*, **7**, 72–81.

CURRIE, P. J. and VARRICHIO, D. J. 2004. A new dromaeosaurid from the Horseshoe Canyon Formation (Upper Cretaceous) of Alberta, Canada. 112–132. *In* CURRIE, P. J., KOPPELHUS, E. B., SHUGAR, M. A. and WRIGHT, J. L. (eds) *Feathered dragons: studies on the transition from dinosaurs to birds*. Indiana University Press.

CURRIE, P. J., RIGBY, J. K. and SLOAN, R. E. 1990. Theropod teeth from the Judith River Formation of southern Alberta, Canada. 107–125. *In* CARPENTER, K. and CURRIE, P. J. (eds) *Dinosaur systematics: Perspectives and approaches*. Cambridge University Press.

DINELEY, D. and METCALF, S. J. 1999. British Jurassic fossil fishes sites. 355–415. *In* DINELEY, D. and METCALF, S. J. (eds) *Fossil fishes of Great Britain*. Geological Conservation Review Series No. 16. Joint Nature Conservation Committee.

DING, A., PITTMAN, M., UPCHURCH, P., O'CONNOR, J., FIELD, D. J. and XU, X. 2020. The biogeography of coelurosaurian theropods and its impact on their evolutionary history. 117–158. *In* PITTMAN, M. and XU, X. (eds) *Pennaraptoran theropod dinosaurs: Past progress and new frontiers*. Bulletin of the American Museum of Natural History, 440.

DUNHILL, A. M., BESTWICK, J., NAREY, H. and SCIBERRAS, J. 2016. Dinosaur biogeographical structure and Mesozoic continental fragmentation: a network-based approach. *Journal of Biogeography*, **43**, 1691–1704.

ESTER, M., KRIEGEL, H.-P., SANDER, J. and XU, X. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. 226–231. *In KDD-96: Proceedings of the second international conference on knowledge discovery and data mining*. AAAI.

EVANS, S. E. and MILNER, A. R. 1994. Middle Jurassic microvertebrate assemblages from the British Isles. 303–321. *In* FRASER, N. C. and SUES, H.-D. (eds) *In the shadow of the dinosaurs: Early Mesozoic tetrapods*. Cambridge University Press.

FARLOW, J. O., BRINKMAN, D. B., ABLER, W. L. and CURRIE, P. J. 1991. Size, shape and serration density of theropod dinosaur lateral teeth. *Modern Geology*, **16**, 161–198.

FASTOVSKY, D. E. and WEISHAMPEL, D. B. 1996. *The evolution and extinction of the dinosaurs*. Cambridge University Press, 479 pp.

FIORILLO, A. R. and CURRIE, P. J. 1994. Theropod teeth from the Judith River Formation (Upper Cretaceous) of South-Central Montana. *Journal of Vertebrate Paleontology*, **14**, 74–80.

FOTH, C. and RAUHUT, O. W. M. 2017. Re-evaluation of the Haarlem *Archaeopteryx* and the radiation of maniraptoran theropod dinosaurs. *BMC Evolutionary Biology*, **17**, 236.

FREEMAN, E. F. 1976*a*. Mammal teeth from the Forest Marble (Middle Jurassic) of Oxfordshire, England. *Science*, **194**, 1053–1055.

FREEMAN, E. F. 1976*b*. A mammalian fossil from the Forest Marble (Middle Jurassic) of Dorset. *Proceedings of the Geologists' Association*, **87**, 231–236.

FREEMAN, E. F. 1979. A Middle Jurassic mammal bed from Oxfordshire. *Palaeontology*, **22**, 135–166.

GATES, T. A., ZANNO, L. E. and MAKOVICKY, P. J. 2015. Theropod teeth from the upper Maastrichtian Hell Creek Formation "Sue" Quarry: new morphotypes and faunal comparisons. *Acta Palaeontologica Polonica*, **60**, 131–139.

GAUTHIER, J. A. 1986. Saurischian monophyly and the origin of birds. *Memoirs of the California Academy of Sciences*, **8**, 1–55.

GERKE, O. and WINGS, O. 2016. Multivariate and cladistic analyses of isolated teeth reveal sympatry of theropod dinosaurs in the Late Jurassic of northern Germany. *PLoS One*, **11**, e0158334.

GILMORE, C. W. 1924. On *Troodon validus*, an orthopodous dinosaur from the Belly River Cretaceous of Alberta, Canada. *Department of Geology, University of Alberta Bulletin*, **1**, 1–43.

GODEFROIT, P., CAU, A., HU, D.-Y. H., ESCUILLIE, F., WU, W. and DYKE, G. 2013*a*. A Jurassic avialan dinosaur from China resolves the early phylogenetic history of birds. *Nature*, **498**, 359–362.

GODEFROIT, P., DEMUYNCK, H., DYKE, G., HU, D., ESCUILLIÉ, F. and CLAEYS, P. 2013*b*. Reduced plumage and flight ability of a new Jurassic paravian theropod from China. *Nature Communications*, **4**, 1394.

GOODWIN, M. B., CLEMENS, W. A., HUTCHISON, J. H., WOOD, C. B., ZAVADA, M. S., KEMP, A., DUFFIN, C. J. and SCHAFF, C. R. 1999. Mesozoic continental vertebrates with associated palynostratigraphic dates from the northwestern Ethiopian Plateau. *Journal of Vertebrate Paleontology*, **19**, 728–741.

GOSWAMI, A., PRASAD, G. V. R., VERMA, O., FLYNN, J. J. and BENSON, R. B. J. 2013. A troodontid dinosaur from the latest Cretaceous of India. *Nature Communications*, **4**, 1703.

HAN, F., CLARK, J. M., XU, X., SULLIVAN, C., CHOINIERE, J. and HONE, D. W. E. 2011. Theropod teeth from the Middle–Upper Jurassic Shishugou Formation of Northwest Xinjiang, China. *Journal of Vertebrate Paleontology*, **31**, 111–126.

HASTIE, T. and TIBSHIRANI, R. 1996. Discriminant analysis by Gaussian mixtures. *Journal of the Royal Statistical Society: Series B: Methodological*, **58**, 155–176.

HASTIE, T., TIBSHIRANI, R., LEISCH, F., HORNIK, K., RIPLEY, B. D. and NARASIMHAN, B. 2020. mda: Mixture and flexible discriminant analysis. R package v.0.5-2. https://rdrr.io/cran/mda/

HENDRICKX, C. and MATEUS, O. 2014. Abelisauridae (Dinosauria: Theropoda) from the Late Jurassic of Portugal and dentition-based phylogeny as a contribution for the identification of isolated theropod teeth. *Zootaxa*, **3759**, 1–74.

HENDRICKX, C., MATEUS, O. and ARAÚJO, R. 2015*a*. The dentition of megalosaurid theropods. *Acta Palaeontologica Polonica*, **60**, 627–642.

HENDRICKX, C., MATEUS, O. and ARAÚJO, R. 2015*b*. A proposed terminology of theropod teeth (Dinosauria, Saurischia). *Journal of Vertebrate Paleontology*, **35**, e982797.

HENDRICKX, C., MATEUS, O., ARAÚJO, R. and CHOINIERE, J. 2019. The distribution of dental features in non-avian theropod dinosaurs: taxonomic potential, degree of

homoplasy, and major evolutionary trends. *Palaeontologia Electronica*, **22** (3), 74 (110 pp).

HENDRICKX, C., TSCHOPP, E. and EZCURRA, M. D. 2020. Taxonomic identification of isolated theropod teeth: the case of the shed tooth crown associated with *Aerosteon* (Theropoda: Megaraptora) and the dentition of Abelisauridae. *Cretaceous Research*, **108**, 104312.

HERVÉ, M. 2021. RVAideMemoire: Testing and plotting procedures for biostatistics. R package v.0.9-80. https://cran.r-project.org/web/packages/RVAideMemoire/index.html

HESSELBO, S. P. 2008. Sequence stratigraphy and inferred relative sea-level change from the onshore British Jurassic. *Proceedings of the Geologists' Association*, **119**, 19–34.

HOLLOWAY, S. 1983. The shell-detrital calcirudites of the Forest Marble Formation (Bathonian) of southwest England. *Proceedings of the Geologists' Association*, **94**, 259–266.

HOLTZ, T. R. 2000. A new phylogeny of the carnivorous dinosaurs. *Gaia*, **15**, 5–61.

HORTON, A., SUMBLER, M. G., COX, B. M. and AMBROSE, K. 1995. Geology of the country around Thame. Memoir of the British Geological Survey, Sheet 237 (England & Wales).

HOYAL CUTHILL, J. F., GUTTENBERG, N., LEDGER, S., CROWTHER, R. and HUERTAS, B. 2019. Deep learning on butterfly phenotypes tests evolution's oldest mathematical model. *Science Advances*, **5**, eaaw4967.

HU, D., HOU, L., ZHANG, L. and XU, X. 2009. A pre-*Archaeopteryx* troodontid theropod from China with long feathers on the metatarsus. *Nature*, **461**, 640–643.

HUENE, F. VON. 1932. Die fossile Reptil-Ordnung Saurischia, ihre Entwicklung und Geschichte. *Monographien zur Geologie und Palaontologie, Series 1*, **4**, 1–361.

HUNTER, A. W. and UNDERWOOD, C. J. 2009. Palaeoenvironmental control on distribution of crinoids in the Bathonian (Middle Jurassic) of England and France. *Acta Palaeontologica Polonica*, **54**, 77–98.

HUXLEY, T. H. 1868. Remarks upon *Archaeopteryx lithographica*. *Proceedings of the Royal Society of London*, **98**, 1–5.

JOECKEL, R. M., LUDVIGSON, G. A., MÖLLER, A., HOTTON, C. L., SUAREZ, M. B., SUAREZ, C. A., SAMES, B., KIRKLAND, J. I. and HENDRIX, B. 2020. Chronostratigraphy and terrestrial palaeoclimatology of Berriasian–Hauterivian strata of the Cedar Mountain Formation, Utah, USA. *Geological Society of London, Special Publication*, **498**, 75–100.

KIRKLAND, J. I. and WOLFE, D. G. 2001. First definitive therizinosaurid (Dinosauria; Theropoda) from North America. *Journal of Vertebrate Paleontology*, **21**, 410–414.

KIRKLAND, J. I., ZANNO, L. E., SAMPSON, S. D., CLARK, J. M. and DEBLIEUX, D. D. 2005. A primitive therizinosauroid dinosaur from the Early Cretaceous of Utah. *Nature*, **435**, 84–87.

KOCSIS, Á. T. and RAJA, N. B. 2019. chronosphere: Earth system history variables. R package v.0.2.0. https://cran.r-project.org/web/packages/chronosphere/index.html

KOCSIS, Á. T. and RAJA, N. B. 2021. rgplates: R interface for the GPlates web service and desktop application. R package v.0.1.0. https://cran.r-project.org/web/packages/rgplates/index.html

KOCSIS, Á. T., REDDIN, C. J., ALROY, J. and KIESSLING, W. 2019. The R package divDyn for quantifying diversity dynamics using fossil sampling data. *Methods in Ecology & Evolution*, **10**, 735–743.

KOCSIS, Á. T., REDDIN, C. J., ALROY, J. and KIESSLING, W. 2022. divDyn: Diversity dynamics using fossil sampling data. R Package v.0.8.1. https://cran.r-project.org/web/packages/divDyn/index.html

KUHN, M. 2008. Building predictive models in R using the caret Package. *Journal of Statistical Software*, **1**, 1–26.

KUHN, M. 2022. Caret: classification and regression training. R package v.6.0-88. https://cran.r-project.org/web/packages/caret/index.html

KUHN, M. and JOHNSON, K. 2013. Classification trees and rule-based models. 369–413. *In* KUHN, M. and JOHNSON, K. (eds) *Applied predictive modelling*. Springer.

KUHN, M., WESTON, S., CULP, M., COULTER, N. and QUINLAN, R. 2018. C5.0 decision trees and rule-based model. RuleQuest Research Pty Ltd. https://cran.r-project.org/web/packages/C50/index.html

LARSON, D. W. 2008. Diversity and variation of theropod dinosaur teeth from the uppermost Santonian Milk River Formation (Upper Cretaceous), Alberta: a quantitative method supporting identification of the oldest dinosaur tooth assemblage in Canada. *Canadian Journal of Earth Sciences*, **45**, 1455–1468.

LARSON, D. W. and CURRIE, P. J. 2013. Multivariate analyses of small theropod dinosaur teeth and implications for paleoecological turnover through time. *PLoS One*, **8**, e54329.

LARSON, D. W., BROWN, C. M. and EVANS, D. C. 2016. Dental disparity and ecological stability in bird-like dinosaurs prior to the end-Cretaceous mass extinction. *Current Biology*, **26**, 1325–1333.

LEIDY, J. 1856. Notices of remains of extinct reptiles and fishes, discovered by Dr. F. V. Hayden in the bad lands of the Judith River, Nebraska Territory. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **8**, 72–73.

LIAW, A. and WIENER, M. 2002. Classification and regression by randomForest. *R News*, **2**, 18–22.

LIAW, A. and WIENER, M. 2022. randomForest: Breiman and Cutler's random forests for classification and regression. R package v.4.7-1.1. https://cran.r-project.org/web/packages/randomForest/index.html

LONGRICH, N. R. 2008. Small theropod teeth from the Lance Formation of Wyoming, USA. 135–158. *In* SANKEY, J. T. and BASZIO, S. (eds) *Vertebrate microfossil assemblages: Their role in paleoecology and paleobiogeography*. Indiana University Press.

MacLEOD, N. and KOLSKA HORWITZ, L. 2020. Machine-learning strategies for testing patterns of morphological variation in small samples: sexual dimorphism in gray wolf (*Canis lupus*) crania. *BMC Biology*, **18**, 113.

MacLEOD, N., CANTY, R. J. and POLASZEK, A. 2022. Morphology-based identification of *Bemisia tabaci* cryptic species puparia via embedded group-contrast convolution neural network analysis. *Systematic Biology*, **71**, 1095–1109.

MAKOVICKY, P. J., NORELL, M. A., CLARKE, J. M. and ROWE, T. 2003. Osteology and relationships of *Byronosaurus*

jaffei (Theropoda: Troodontidae). *American Museum Novitates*, **2003**, 1–32.

MALEEV, E. A. 1954. Noviy cherepachoobrazhniy yashcher v Mongolii [New tortoise-like saurian from Mongolia]. *Priroda*, **1954**, 106–108. [in Russian]

MARSH, O. C. 1881. Principal characters of American Jurassic dinosaurs. Part V. *American Journal of Science, Series 3*, **21**, 417–423.

MATTHEW, W. D. and BROWN, B. 1922. The family Deinodontidae, with notice of a new genus from the Cretaceous of Alberta. *Bulletin of the American Museum of Natural History*, **46**, 367–385.

McKERROW, W. S., JOHNSON, R. T. and JAKOBSON, M. E. 1969. Palaeoecological studies in the Great Oolite at Kirtlington, Oxfordshire. *Palaeontology*, **12**, 56–83.

MELVILLE, R. V. and FRESHNEY, E. C. 1982. *The Hampshire Basin and adjoining areas*. HMSO, London, 146 pp.

METCALF, S. J. 1995. The palaeoenvironment and palaeoecology of a Middle Jurassic vertebrate-bearing fen-type paleosol in a coastal carbonate regime. Unpub. PhD thesis. University of Bristol.

METCALF, S. J. and WALKER, R. J. 1994. A new Bathonian microvertebrate locality in the English Midlands. 322–331. *In* FRASER, N. C. and SUES, H.-D. (eds) *In the shadow of the dinosaurs: Early Mesozoic tetrapods*. Cambridge University Press.

METCALF, S. J., VAUGHAN, R. F., BENTON, M. J., COLE, J., SIMMS, M. J. and DARTNELL, D. L. 1992. A new Bathonian (Middle Jurassic) microvertebrate site, within the Chipping Norton Limestone Formation at Hornsleasow Quarry, Gloucestershire. *Proceedings of the Geologists' Association*, **103**, 321–342.

MEYER, H. V. 1861. Die Feder von Solenhofen. *Neues Jahrbuch für Mineralogie, Geognosie, Geologie und Petrefaktenkunde*, **1861**, 678–679.

NORELL, M. A., CLARK, J. M., TURNER, A. H., MAKOVICKY, P. J., BARSBOLD, R. and ROWE, T. 2006. A new dromaeosaurid theropod from Ukhaa Tolgod (Ömnögov, Mongolia). *American Museum Novitates*, **3545**, 1–51.

NORELL, M. A., MAKOVICKY, P. J., BEVER, G. S., BALANOFF, A. M., CLARK, J. M., BARSBOLD, R. and ROWE, T. 2009. A review of the Mongolian Cretaceous dinosaur *Saurornithoides* (Troodontidae: Theropoda). *American Museum Novitates*, **3654**, 1–63.

NOTO, C. R., D'AMORE, D. C., DRUMHELLER, S. K. and ADAMS, T. L. 2022. A newly recognized theropod assemblage from the Lewisville Formation (Woodbine Group; Cenomanian) and its implications for understanding Late Cretaceous Appalachian terrestrial ecosystems. *PeerJ*, **10**, e12782.

OKSANEN, J., GUILLAUME, B. F., FRIENDLY, M., KINDT, R., LEGENDRE, P., McGLINN, D., MINCHIN, P. R., O'HARA, R. B., SIMPSON, G. L., SOLYMOS, P., HENRY, M., STEVENS, H., SZOECS, E. and WAGNER, H. 2020. vegan: Community ecology package. R package version 2.5-7. https://cran.r-project.org/web/packages/vegan/index.html

OSBORN, H. F. 1924. Three new Theropoda, *Protoceratops* zone, central Mongolia. *American Museum Novitates*, **144**, 1–12.

OWEN, R. 1864. On the *Archeopteryx* of von Meyer, with a description of the fossil remains of a long-tailed species, from the lithographic stone of Solnhofen. *Philosophical Transactions of the Royal Society*, **153**, 33–47.

OWEN, R. 1883. On the skull of *Megalosaurus*. *Quarterly Journal of the Geological Society of London*, **39**, 334–347.

PALMER, T. J. 1973. Field meeting in the Great Oolite of Oxfordshire: 5–7 May 1972. *Proceedings of the Geologists' Association*, **84**, 53–64.

PALMER, T. J. 1974. *Some palaeoecological studies of the middle and upper Bathonian of central England and northern France*. Unpub. PhD thesis, University of Oxford.

PALMER, T. J. 1979. The Hampen Marly and White Limestone Formations: Florida-type carbonate lagoons in the Jurassic of Central England. *Palaeontology*, **22**, 189–228.

PALMER, T. J. and JENKYNS, H. C. 1975. A carbonate island barrier in the Great Oolite (Middle Jurassic) of Central England. *Sedimentology*, **22**, 125–135.

PEI, R., LI, Q., MENG, Q., NORELL, M. A. and GAO, K.-Q. 2017. New specimens of *Anchiornis huxleyi* (Theropoda, Paraves) from the Late Jurassic of northeastern China. *Bulletin of the American Museum of Natural History*, **411**, 1–66.

PRASAD, G. V. R. and PARMAR, V. 2020. First ornithischian and theropod dinosaur teeth from the Middle Jurassic Kota Formation of India: paleobiogeographic relationships. 1–30. *In* PRASAD, G. V. R. and PATNAIK, R. (eds) *Biological consequences of plate tectonics: New perspectives on post-Gondwana break-up. A tribute to Ashok Sahni*. Springer.

R CORE TEAM. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

RAUHUT, O. W. M. 2003. The interrelationships and evolution of basal theropod dinosaurs. *Special Papers in Palaeontology*, **69**, 1–213.

RAUHUT, O. W. M. and FOTH, C. 2020. The origin of birds: current consensus, controversy, and the occurrence of feathers. *In* RAUHUT, O. W. M. and FOTH, C. (eds) *The evolution of feathers*. Springer.

RAUHUT, O. W. M. and WERNER, C. 1995. First record of the family Dromaeosauridae (Dinosauria: Theropoda) in the Cretaceous of Gondwana (Wadi Milk Formation, northern Sudan). *Paläontologische Zeitschrift*, **69**, 475–489.

RAUHUT, O. W. M., MILNER, A. C. and MOORE-FAY, S. 2010. Cranial osteology and phylogenetic position of the theropod dinosaur *Proceratosaurus bradleyi* (Woodward, 1910) from the Middle Jurassic of England. *Zoological Journal of the Linnean Society*, **158**, 155–195.

RICHARDSON, L. 1929. The country around Moreton-in-the-Marsh. *Memoirs. Geological Survey of Great Britain*, **217**, 1–162.

RSTUDIO TEAM. 2020. RStudio: Integrated Development Environment for R. https://www.rstudio.com

RUSSELL, L. S. 1946. The lower jaw of the theropod dinosaur *Troodon*. *Transactions of the Royal Society of Canada, Series 3*, **60**, 171.

SADLEIR, R., BARRETT, P. M. and POWELL, H. P. 2008. The anatomy and systematics of *Eustreptospondylus oxoniensis*, a theropod dinosaur from the Middle Jurassic of

Oxfordshire, England. *Monographs of the Palaeontographical Society*, **160**, 1–82.

SANKEY, J. T. 2008. Diversity of latest Cretaceous (Late Maastrichtian) small theropods and birds: teeth from the Lance and Hell Creek Formations. 117–137. *In* SANKEY, J. T. and BASZIO, S. (eds) *Vertebrate microfossil assemblages: Their role in paleoecology and paleobiogeography*. Indiana University Press.

SANKEY, J. T., BRINKMAN, D. B., GUENTHER, M. and CURRIE, P. J. 2002. Small theropod and bird teeth from the Late Cretaceous (late Campanian) Judith River Group, Alberta. *Journal of Paleontology*, **76**, 751–763.

SANKEY, J. T., STANDHARDT, B. R. and SCHIEBOUT, J. A. 2005. Theropod teeth from the Upper Cretaceous (Campanian–Maastrichtian), Big Bend National Park, Texas. 127–152. *In* CARPENTER, K. (ed.) *The carnivorous dinosaurs*. Indiana University Press.

SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J.-Y., WHITE, D. J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. and CARDONA, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, **9**, 676–682.

SCOTESE, C. R. 2016. Tutorial: PALEOMAP PaleoAtlas for GPlates and the PaleoData plotter program. https://doi.org/10.13140/RG.2.2.34367.00166

SCOTESE, C. R. 2021. An atlas of Phanerozoic paleogeographic maps: the seas come in and the seas go out. *Annual Review of Earth & Planetary Sciences*, **49**, 679–728.

SELLÉS, A. G., VILA, B., BRUSATTE, S. L., CURRIE, P. J. and GALOBART, À. 2021. A fast-growing basal troodontid (Dinosauria: Theropoda) from the latest Cretaceous of Europe. *Scientific Reports*, **11**, 4855.

SELLWOOD, B. W. and McKERROW, W. S. 1974. Depositional environments in the lower part of the Great Oolite Group of Oxfordshire and North Gloucestershire. *Proceedings of the Geologists' Association*, **85**, 189–210.

SENTER, P., KIRKLAND, J. I., BIRD, J. and BARTLETT, J. A. 2010. A new troodontid theropod dinosaur from the Lower Cretaceous of Utah. *PLoS One*, **5**, e14329.

SENTER, P., KIRKLAND, J. I. and DEBLIEUX, D. D. 2012. *Martharaptor greenriverensis*, a new theropod dinosaur from the Lower Cretaceous of Utah. *PLoS One*, **7**, e43911.

SERENO, P. C. 1997. The origin and evolution of dinosaurs. *Annual Review of Earth & Planetary Sciences*, **25**, 435–489.

SMITH, J. B., VANN, D. R. and DODSON, P. 2005. Dental morphology and variation in theropod dinosaurs: implications for the taxonomic identification of isolated teeth. *The Anatomical Record (Part A)*, **285**, 699–736.

STRAHAN, A. 1898. The Geology of the Isle of Purbeck and Weymouth. *Memoirs of the Geological Survey, England and Wales*, 278 pp.

SULLIVAN, C., WANG, Y., HONE, D. W. E., WANG, Y., XU, X. and ZHANG, F. 2014. The vertebrates of the Jurassic Daohugou Biota of northeastern China. *Journal of Vertebrate Paleontology*, **34**, 243–280.

SWEETMAN, S. C. 2004. The first record of velociraptorine dinosaurs (Saurischia, Theropoda) from the Wealden (Early Cretaceous, Barremian) of southern England. *Cretaceous Research*, **25**, 353–364.

TORRENS, H. 1969*a*. Guide for Dorset and South Somerset. Excursion No. 1. *International field symposium on the British Jurassic*. Department of Geology, University of Keele, A21–A40.

TORRENS, H. 1969*b*. Excursion to the Cotswolds and Oxfordshire. *International field symposium on the British Jurassic*. Department of Geology, University of Keele, B8–B23.

TURNER, A. H., MAKOVICKY, P. J. and NORELL, M. A. 2012. A review of dromaeosaurid systematics and paravian phylogeny. *Bulletin of the American Museum of Natural History*, **371**, 1–206.

UNDERWOOD, C. J. 2004. Environmental controls on the distribution of neoselachian sharks and rays within the British Bathonian (Middle Jurassic). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **203**, 107–126.

UPCHURCH, P., HUNN, C. A. and NORMAN, D. B. 2002. An analysis of dinosaurian biogeography: evidence for the existence of vicariance and dispersal patterns caused by geological events. *Proceedings of the Royal Society B*, **269**, 613–621.

VAN DER LUBBE, T., RICHTER, U. and KNÖTSCHKE, N. 2009. Velociraptorine dromaeosaurid teeth from the Kimmeridgian (Late Jurassic) of Germany. *Acta Palaeontologica Polonica*, **54**, 401–408.

VAUGHAN, R. F. 1989. The excavation at Hornsleasow Quarry. Interim report for the Nature Conservancy Council, City of Gloucester Museums, 68 pp.

VULLO, R., ABIT, D., BALLÈVRE, M. P., BILLON-BRUYAT, J., BOURGEAIS, R., BUFFETAUT, É., DAVIERO-GOMEZ, V., GARCIA, G., GOMEZ, B. M., MAZIN, J., MOREL, S., NÉRAUDEAU, D., POUECH, J. C., RAGE, J., SCHNYDER, J. and TONG, H. 2014. Palaeontology of the Purbeck-type (Tithonian, Late Jurassic) bonebeds of Chassiron (Oléron Island, western France). *Comptes Rendus Palevol*, **13**, 421–441.

WALDMAN, M. 1974. Megalosaurids from the Bajocian (Middle Jurassic) of Dorset. *Palaeontology*, **17**, 325–340.

WALKER, A. D. 1964. Triassic reptiles from the Elgin area: *Ornithosuchus* and the origin of carnosaurs. *Philosophical Transactions of the Royal Society B*, **248**, 53–134.

WARD, D. 1981. A simple machine for bulk processing of clays and silts. *Tertiary Research*, **3**, 121–124.

WICKHAM, H. 2016. *ggplot2: Elegant graphics for data analysis*. Springer, 260 pp. https://ggplot2.tidyverse.org/

WILLIAMSON, T. E. and BRUSATTE, S. L. 2014. Small theropod teeth from the Late Cretaceous of the San Juan Basin, northwestern New Mexico and their implications for understanding latest Cretaceous dinosaur evolution. *PLoS One*, **9**, e93190.

WILLS, S., BARRETT, P. M. and WALKER, A. 2014. New dinosaur and crocodilian material from the Middle Jurassic (Bathonian) Kilmaluag Formation, Skye, Scotland. *Scottish Journal of Geology*, **50**, 183–190.

WILLS, S., BERNARD, E. L., BREWER, P., UNDERWOOD, C. J. and WARD, D. J. 2019. Palaeontology, stratigraphy and sedimentology of Woodeaton Quarry

(Oxfordshire) and a new microvertebrate site from the White Limestone Formation (Bathonian, Jurassic). *Proceedings of the Geologists' Association*, **130**, 170–186.

WILLS, S., UNDERWOOD, C. J. and BARRETT, P. M. 2021. Learning to see the wood for the trees: machine learning, decision trees, and the classification of isolated theropod teeth. *Palaeontology*, **64**, 75–99.

WOODWARD, A. S. 1910. On a skull of *Megalosaurus* from the Great Oolite of Minchinhampton (Gloucestershire). *Quarterly Journal of the Geological Society*, **66**, 111–115.

WOODWARD, H. B. 1894. The Jurassic Rocks of Britain, Vol. 4. The Lower Oolitic Rocks of England (Yorkshire Excepted). *Memoir of the Geological Survey of Great Britain*, 628 pp.

WRIGHT, M. N. 2021. ranger: A fast implementation of random forests. R package v.0.13.1. https://cran.r-project.org/web/packages/ranger/index.html

WRIGHT, M. N. and ZIEGLER, A. 2017. ranger: a fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, **77**, 1–17.

WYATT, R. J. 1996. A correlation of the Bathonian (Middle Jurassic) succession between Bath and Burford, and its relation to that near Oxford. *Proceedings of the Geologists' Association*, **107**, 299–322.

WYATT, R. J. 2002. Woodeaton, Oxfordshire. 234–239. *In* COX, B. M. and SUMBLER, M. G. (eds) *British Middle Jurassic stratigraphy, Geological Conservation Review Series, No. 26*. Joint Nature Conservation Committee, Peterborough.

XU, X., ZHOU, Z. and WANG, X. 2000. The smallest known non-avian theropod dinosaur. *Nature*, **408**, 705–708.

XU, X., ZHAO, X. and CLARK, J. M. 2001. A new therizinosaur from the Lower Jurassic Lower Lufeng Formation of Yunnan, China. *Journal of Vertebrate Paleontology*, **21**, 477–483.

XU, X., MA, Q.-Y. and HU, D.-Y. 2010. Pre-*Archaeopteryx* coelurosaurian dinosaurs and their implications for understanding avian origins. *Chinese Science Bulletin*, **55**, 3971–3977.

XU, X., ZHENG, X., SULLIVAN, C., WANG, X., XING, L., WANG, Y., ZHANG, X., O'CONNOR, J. K., ZHANG, F. and PAN, Y. 2015. A bizarre Jurassic maniraptoran theropod with preserved evidence of membranous wings. *Nature*, **521**, 70–73.

XU, X., ZHOU, Z., SULLIVAN, C., WANG, Y. and REN, D. 2016. An updated review of the Middle–Late Jurassic Yanliao Biota: chronology, taphonomy, paleontology and paleoecology. *Acta Geologica Sinica (English Edition)*, **90**, 2229–2243.

YOUNG, C. M. E., HENDRICKX, C., CHALLANDS, T. J., FOFFA, D., ROSS, D. A., BUTLER, I. B. and BRUSATTE, S. L. 2019. New theropod dinosaur teeth from the Middle Jurassic of the Isle of Skye, Scotland. *Scottish Journal of Geology*, **55**, 7–19.

ZANNO, L. E. 2010*a*. Osteology of *Falcarius utahensis* (Dinosauria: Theropoda): characterizing the anatomy of basal therizinosaurs. *Zoological Journal of the Linnean Society*, **158**, 196–230.

ZANNO, L. E. 2010*b*. A taxonomic and phylogenetic re-evaluation of Therizinosauria (Dinosauria: Maniraptora). *Journal of Systematic Palaeontology*, **8**, 503–543.

ZHENG, X., XU, X., YOU, H., ZHAO, Q. and DONG, Z. 2009. A short-armed dromaeosaurid from the Jehol Group of China with implications for early dromaeosaurid evolution. *Proceedings of the Royal Society B*, **277**, 211–217.

ZINKE, J. 1998. Small theropod teeth from the Upper Jurassic coal mine of Guimarota (Portugal). *Paläontologische Zeitschrift*, **72**, 179–189.