Dinosaur micro-remains from the Middle Jurassic of Britain

Dinosaur micro-remains from the Middle Jurassic of Britain

Simon Wills

Thesis submitted for the degree of Doctor of Philosophy

Department of Earth and Planetary Sciences

Birkbeck, University of London

August 2023

In collaboration with: Natural History Museum, London

DECLARATION

I, Simon Wills, hereby declare that the research presented in this thesis is my original work. All work and materials that are drawn from others are always clearly attributed. The thesis has not been previously submitted to this or any other university for a degree.

ABSTRACT

In the Middle Jurassic, Great Britain was situated at ~30° north in an area of shallow seas with surrounding low-lying landmasses. Fluctuations in relative sea level resulted in emergent areas preserving snapshots of the terrestrial fauna in microvertebrate sites throughout southern England. Analysis of the dinosaur material, mostly isolated teeth, has resulted in a much more granular view of the taxa present including clades previously unknown or unconfirmed from this period.

I developed machine learning techniques, which combined with morphological-based approaches confirms the presence of at least three maniraptoran taxa in the assemblage: three dromaeosaur morphotypes; a troodontid; and a therizinosaur. These results provide the first quantitative support for the presence of maniraptoran theropods, including the oldest occurrences of troodontids and therizinosaurs worldwide, in the Middle Jurassic and are consistent with predictions made by phylogenetic analyses. There are at least six ornithischian taxa in the assemblage; a distinctive highly-ridged morphotype that cannot be referred with certainty to any known ornithischian taxa and therefore represents a new taxon; a number of small teeth with denticles restricted to the upper third of the crown which represent a hitherto unknown occurrence of heterodontosaurids in the Middle Jurassic of the UK; at least one morphotype of a basal thyreophoran; an indeterminate thyreophoran; a stegosaur, which represents one of the oldest stegosaurs worldwide; and a number of ankylosaur morphotypes which make up the vast majority of the isolated ornithischian teeth seen from these sites.

The application of machine learning, when combined with traditional morphological comparisons provides a powerful tool for the quantitative assessment of isolated teeth. This analysis increases the known diversity of Middle Jurassic dinosaur taxa in the UK and the confirmation of early maniraptorans, heterodontosaurids and stegosaurs highlights the importance of incorporating microvertebrate remains into faunal and evolutionary analyses.

Contents

Figures

Tables

Acknowledgements

Many people have aided and abetted me over the years, and I am grateful for all the help, guidance, support, and friendship they have given me.

I am eternally grateful to my supervisors Paul Barrett and Charlie Underwood for having faith and lots of patience in me. You have taught me so much and, editorial corrections from Paul notwithstanding, for that I thank you both.

Special thanks as well to the numerous other researchers and curators who have helped me throughout this PhD: Jerry Hooker, Susie Maidment, Andy Heckert, Pip Brewer, Emma Bernard, Charlotte Brassey, Natalie Cooper and Hilary Ketchum.

David Ward deserves a special mention for introducing me to the world of processing sediment for microvertebrates, having a never-ending capacity to process and pick fossils from tonnes of clay, and for going out of his way to be always supportive and massively helpful.

I am massively indebted to the wonderful people in the CT lab at the Natural History Museum, Brett Clark, Vincent Fernandez and Dan Sykes, who always managed to find time to help me whether it be the scanning of specimens or segmenting data.

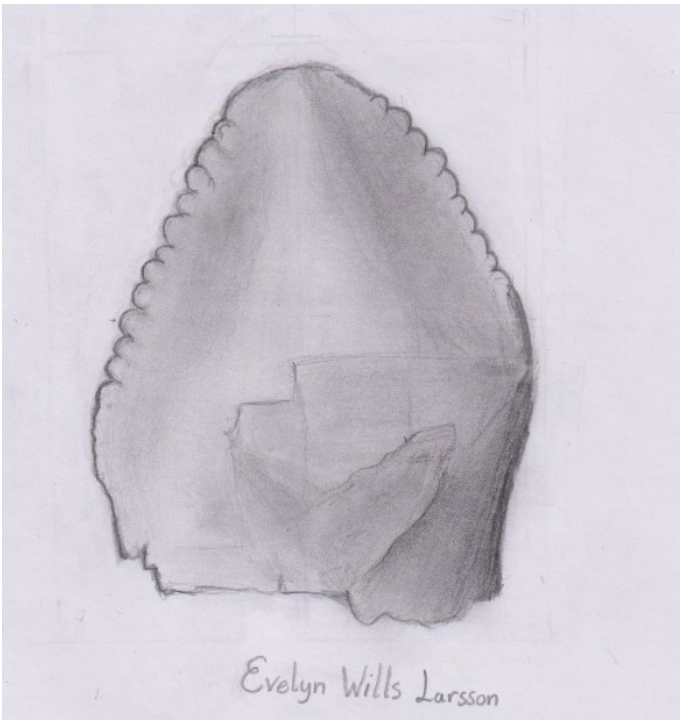
Norman MacLeod first introduced me to, and piqued my interest in, machine learning and I thank him for his patience as I slowly learnt what it all meant.

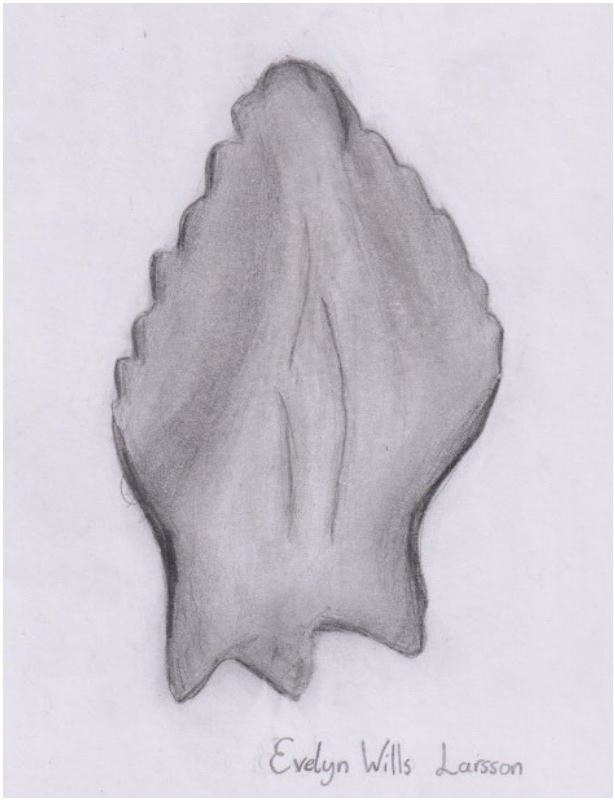
Fieldwork, often on a damp and rainy day, was an integral part of this thesis and I am grateful for everyone who helped out. Susie Maidment and Joe Bonsor who gave up their time on a cold January day to drag bags of heavy wet clay across a quarry floor in Gloucestershire. Many thanks to the large field crew without which the excavation at Woodeaton would not have been possible: Martha Richter, Tim Ewin, Mark Graham, Adrian Lister, Stephen Stukins, Andy Gale, Zoë Hughes, Alison Ward, Andrew Ward, Mike Smith, Karen Banton, Claire Bullar and Rebecca Schofield. Thanks to Breedon Group and Michael McKenna for allowing access to Hornsleasow and Woodeaton Quarries.

Thanks also to the many friends in academia I have made, my many office mates over the years and my research group at the Natural History Museum: Leila D'Souza, Tom

Raven, Joe Bonsor, João Leite, Terri Cleary, David Nicholson, David Button, Roger Benson, Richard Butler, Jonah Choiniere, Kimi Chapelle, David Ford and Kathleen Dollman.

Finally, to my wife Eva and my children Evelyn and Andrew who have put up with me vanishing for extended periods to dig holes in the ground.





To my parents, in loving memory, and my family Eva, Evelyn and Andrew

Chapter One: Introduction

This thesis investigates dinosaur micro-remains, namely isolated teeth, from four Middle Jurassic microvertebrate localities in southern Britain: Hornsleasow Quarry, Gloucestershire; Woodeaton and Kirtlington Quarries, Oxfordshire; and Watton Cliff, Dorset. Specimens from Hornsleasow, Kirtlington and Watton Cliff, collected in the 1970's to 1990's by the Museum of Gloucester and University College London respectively, have been supplemented by additional collecting during the course of this research at Hornsleasow and Watton Cliff and by the discovery and excavation of a new site at Woodeaton Quarry by the author and colleagues from the Natural History Museum in 2014 to 2017. The thesis synthesises the previously disparate taxonomic referrals of these specimens into a robust and quantifiable framework using modern methodologies such as CT scanning to develop 3D models and, for the first time, the application of machine learning to taxonomic classification of isolated dinosaur teeth.

In addition to the specimens from the UK, the taxonomic affinities of Middle Jurassic isolated theropod teeth from microvertebrate sites in Madagascar (Maganuco et al., 2005), India (Prasad & Parmar, 2020) and Kyrgyzstan (Averianov et al., 2005) are also reassessed, based on published morphometric measurements, and these results are documented in Chapter Five.

The Middle Jurassic is a critical stage in terrestrial vertebrate evolution, witnessing the onset of major radiations for many important clades, including squamates, lissamphibians, pterosaurs and dinosaurs (Barrett & Upchurch, 2005; Close et al., 2015; Evans, 2003; Jones et al., 2022; Raven et al., 2023; Upchurch et al., 2014). However, the poor terrestrial sedimentary record available for this period, combined with preservational and sampling biases, hampers our understanding of these events (Barrett et al., 2009; Ding et al., 2020; Evans, 2003; Yi et al., 2017). For example, 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. Similarly, Ornithischia, a speciose clade of mainly herbivorous dinosaurs, likely originated during the Middle to

Late Triassic (Baron et al., 2017a; Boyd, 2015; Dieudonné et al., 2021), nevertheless, representatives of this clade, although well-known from the Late Jurassic and Cretaceous (Weishampel et al., 2004), are poorly known from the Middle Jurassic. The lack of fossil evidence from the Middle Jurassic has similarly resulted in the creation of hypotheses to explain the difference between Jurassic and Cretaceous faunas in various geographical settings. The East Asian Isolation Hypothesis (Wilson & Upchurch, 2009), for example, postulates the geographical isolation of East Asia from the rest of Pangaea and the resultant evolution of groups endemic to East Asia. New evidence from the Middle Jurassic is now challenging this view suggesting that despite the poor fossil record many dinosaur groups had a global distribution prior to the breakup of Pangaea (Wills, Underwood, et al., 2023; Xu et al., 2018). Any records of Middle Jurassic dinosaurs are therefore of critical importance in untangling the evolutionary history of the group and microvertebrate sites, especially where temporally well constrained, have the potential to add to our knowledge of these important events.

Microvertebrate deposits

Microvertebrate sites are accumulations of small, usually disarticulated, vertebrate material in a geographically well-defined and stratigraphically limited unit. They contain the remains of multiple individuals and (usually) taxa resulting from processes that range from instantaneous mass death assemblages to time-averaged, dispersed occurrences of fragmentary bones and teeth (Behrensmeyer, 1991; Eberth et al., 2007). Microvertebrate sites occur in both marine and terrestrial environments throughout the Phanerozoic and have been found on all continents. They are usually exploited by a combination of bulk collecting and surface picking followed by breakdown in water or acid, sieving and then sorting (Baszio, 2008; Eberth et al., 2007; Ward, 1981). They generally yield a taxonomically diverse assemblage of small bones and teeth often including the remains of vertebrate taxa not known as macrofossils (e.g., Avrahami et al., 2018; Evans et al., 2006; Evans & Milner, 1994; Fanti & Miyashita, 2009; Foster & Heckert, 2011; Heckert et al., 2023; Panciroli, Benson, et al., 2020; Rauhut, 2000; Williams et al., 2022; Wills, Cavosie, et al., 2023; Wills, Underwood, et al., 2023; Zinke, 1998). Sample sizes are often large enough to allow detailed statistical and taphonomic studies to be undertaken (Avrahami et al., 2018; Foster, 2001; Wills, Cavosie, et al., 2023;

Wilson, 2008), and the range of taxa, and their restricted spatial extent and stratigraphy, makes them amenable to palaeoecological and palaeoenvironmental descriptions (Anderson et al., 2007; Carrano et al., 2016; Carrano & Velez-Juarbe, 2006; Csiki et al., 2008; Foster & Heckert, 2011; Gilbert et al., 2018). Consequently, they play an important role in community-level palaeoecological reconstructions and estimates of relative abundance and species-richness (Brinkman, 1990; Cullen & Evans, 2016; Cullen et al., 2016; Estes, 1964). In addition, and potentially of greater importance, they can contribute to the taxic diversity record, especially when sampling time periods with a poor fossil record, such as the terrestrial Middle Jurassic (Averianov et al., 2005; Barrett, 2006; Evans et al., 2006; Evans & Milner, 1994; Kriwet et al., 1997; Panciroli, Benson, et al., 2020; Wills, Underwood, et al., 2023; Young et al., 2019).

Dinosaur remains from microvertebrate sites are usually dominated by isolated teeth. Isolated tooth crowns have been utilised for a variety of purposes: taxonomic identification (e.g., Barrett et al., 2014; Becerra et al., 2013; Buckley et al., 2010; Currie et al., 1990; Evans et al., 2012; Evans & Milner, 1994; Farlow et al., 1991; Hendrickx & Mateus, 2014; Richter et al., 2012; Sankey et al., 2002; Smith, 2002; Smith, 2005; Smith et al., 2005; Thulborn, 1973; Zinke, 1998); palaeocology and palaeoenvironmental studies (e.g., Amiot et al., 2015; Baszio, 1997; Blob & Fiorillo, 1996; Brinkman, 1990; Brinkman et al., 2005; Carrano et al., 2016; D'Amore, 2009; Dodson, 1987; Sigogneau-Russell et al., 1998); population dynamics (e.g., Amiot et al., 2004; Fricke et al., 2009); physiology (e.g., Eagle et al., 2011; Fricke & Rogers, 2000); feeding behaviours (Fricke & Pearson, 2008); and macroevolutionary patterns (e.g., Larson et al., 2016; Williamson & Brusatte, 2014).

Evidence from microvertebrate sites in faunal and evolutionary analyses is often overlooked, however, because of uncertainties regarding the taxonomic affinities of specimens that might only be represented by isolated teeth or bones (e.g., Ding et al., 2020; Foth & Rauhut, 2017) and added to this is a degree of sampling bias towards larger skeletal remains. For example, the records of dinosaur occurrences in the Paleobiology Database are biased towards macro dinosaur remains with nearly 18,000 dinosaur records in total, of which only 512 refer to microvertebrates (Fig. 1.1).

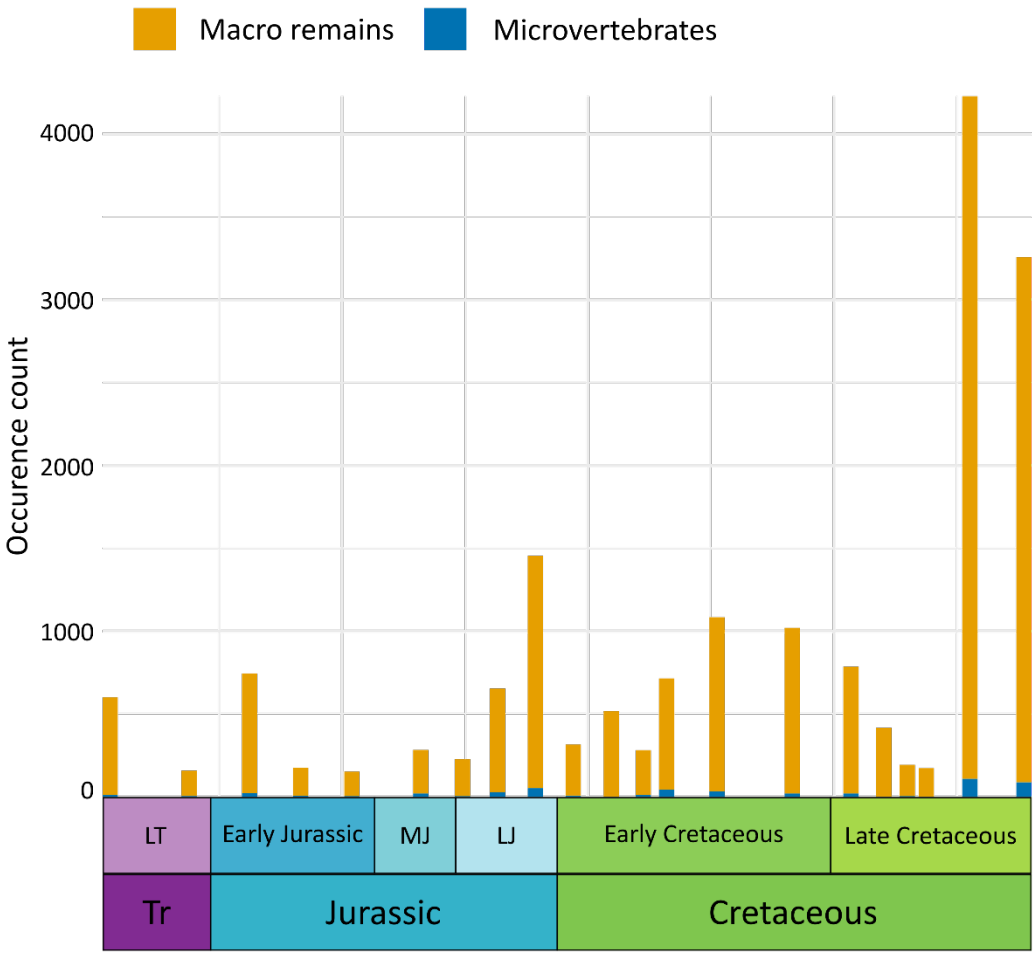


Fig. 1.1. Temporal distribution of Mesozoic dinosaur occurrence records in the Paleobiology Database for both macro remains and microvertebrates. The data were downloaded from the Paleobiology Database on 21 June 2023.

However, microvertebrate sites can offer important insights into major evolutionary questions. For example, confirming the presence of various dinosaur clades, such as maniraptoran theropods or ornithopod ornithischians, in Middle Jurassic microvertebrate deposits would be consistent with predictions made by phylogenetic analyses, whose ghost lineages posit the likely presence of these clades at this time (e.g., Carrano et al., 2012; Holtz, 2000; Rauhut & Foth, 2020). In this thesis I address the problem of taxonomic uncertainty in relation to isolated dinosaur teeth, with the development of a robust machine learning methodology to provide a quantifiable

classification framework. The application of machine learning, when combined with traditional morphological comparisons, now provides a powerful tool for the quantitative assessment of isolated teeth. Providing quantitative support for the taxonomic referral of isolated dinosaur teeth will allow for a more complete view of the evolutionary history, distribution and palaeoecology of the clade.

Fig. 1.2. The 'Mammal Pit' excavation at Durlston Bay, Swanage, Dorset. From Kingsley (1857).

Collecting small vertebrate remains, principally mammals, from Mesozoic microvertebrate sites in the UK has a history that stretches back over 200 years. Small mammalian remains from the British Middle Jurassic were known in the early part of the nineteenth century, with specimens collected from Stonesfield near Oxford by W.J. Broderip (in 1812 or 1814) and described by Buckland (1824), Broderip (1828) and Owen (1838). The earliest documented example of the systematic collection of microvertebrates is that of Wood (1839) who described collecting fish and mammal

teeth from a clay pit at Kingston in Suffolk. In 1856, Samuel Beckles undertook a huge excavation in the Lower Cretaceous Purbeck Limestone Group at Durlston Bay in Dorset. This excavation (Fig. 1.2), described as covering over 7,000 square feet by Kingsley (1857), was successful in recovering the remains of many small mammals and other vertebrates including dinosaurs such as the small heterodontosaurid *Echinodon becklesii* (Barrett & Maidment, 2011; Norman & Barrett, 2002; Owen, 1861). The Late Triassic and Early Jurassic sediments of southwest England and south Wales have yielded numerous microvertebrate assemblages, both marine and terrestrial, from a mixture of fissure fill deposits and bone beds which include mammals, small reptiles and dinosaurs (e.g., Kermack et al., 1968; Mears et al., 2016; Moore, 1867; Robinson, 1957; Savage, 1993; Slater et al., 2016; van den Berg et al., 2012; Whiteside et al., 2016; Williams et al., 2022).

The British Middle Jurassic, composed largely of marginal marine and marine strata, contains numerous terrestrial microvertebrate sites that were discovered in the 1970s (Freeman, 1976a, 1976b, 1979; Kermack et al., 1987; Waldman & Savage, 1972). The sites preserve a mixture of marine, freshwater and terrestrial taxa, including mammaliaforms, dinosaurs, pterosaurs, frogs, turtles, albanerpetontids, salamanders, lizards, choristoderes, crocodiles, bony fish, sharks and rays (Barrett, 2006; Evans et al., 2006; Evans & Milner, 1994; Evans & Waldman, 1996; Jones et al., 2022; Metcalf et al., 1992; Metcalf & Walker, 1994; Panciroli, Benson, et al., 2020; Tałanda et al., 2022; Underwood & Ward, 2004; Waldman & Evans, 1994; Wills et al., 2014; Wills et al., 2019; Wills, Underwood, et al., 2023; Young et al., 2019). These include the four Bathonian sites that form the basis of this thesis: Kirtlington Quarry, Oxfordshire (Evans & Milner, 1994; Freeman, 1976a, 1979; Kermack et al., 1987); Hornsleasow Quarry, Gloucestershire (Metcalf, 1995; Metcalf et al., 1992; Metcalf & Walker, 1994; Vaughan, 1989); Watton Cliff, Dorset (Evans, 1992; Evans & Milner, 1994; Freeman, 1976b) and Woodeaton Quarry, Oxfordshire (Wills et al., 2019; Wyatt, 2002). The geological settings of these sites are described in detail in Chapter Four.

Other important microvertebrate sites in Britain include: Chicksgrove Quarry, Wiltshire (Tithonian, Late Jurassic), which has produced a range of vertebrates including sauropod, theropod and ornithischian dinosaur material but remains undescribed (Benton et al., 2005); Sunnydown Farm, Dorset (Purbeck Group, Berriasian, Lower

Cretaceous) has a diverse vertebrate fauna including lizards, amphibians and dinosaurs (e.g., Barrett & Maidment, 2011; Benson & Barrett, 2009; Ensom, 1987, 1988; Ensom et al., 1991; Ensom et al., 1994; Evans & McGowan, 2002; Evans & Searle, 2002); sites in the Lower Cretaceous Wealden Group of both the Weald and Wessex Basins (Cook & Ross, 1996; Maidment et al., 2017; Sweetman, 2007; Sweetman, 2006); and the Eocene sites of Abbey Wood, Creechbarrow, Hordle Cliff, Headon Hill and the Isle of Wight (Benton et al., 2005; Hooker et al., 1980; Keeping, 1910; Wood, 1844).

In North America microvertebrate sites have been systematically sampled since the late 1950's, initially in the search for mammals (Clemens, 1960, 1966). The Upper Triassic Chinle Group (Heckert, 2004; Heckert et al., 2023; Heckert & Jenkins, 2005; Heckert et al., 2005) and Lower Jurassic Kayenta Formation (Curtis & Padian, 1999) in the southwestern USA, Upper Jurassic–Lower Cretaceous Morrison and Cloverly formations of the Midwest (Brett-Surman et al., 2005; Carrano et al., 2016; Carrano & Velez-Juarbe, 2006; Foster, 2001, 2006, 2007; Oreska et al., 2013), the Upper Cretaceous Oldman and Judith River formations (Campanian) of Dinosaur Provincial Park in Canada (Cullen et al., 2016; Currie et al., 1990; Eberth, 1990; Eberth et al., 2001; Peng et al., 2001; Rogers & Brady, 2010; Sankey et al., 2002), and the latest Cretaceous (Maastrichtian) Lance, Hell Creek and Prince Creek formations of the USA (Brown & Druckenmiller, 2011; Carpenter, 1982; Chiarenza et al., 2020; Estes, 1964; Fiorillo & Gangloff, 2001; Gates et al., 2015; Jackson & Varricchio, 2016; Sankey, 2008; Snyder et al., 2020) have all produced huge numbers of microvertebrate remains. Estes (1964) recorded over 30,000 specimens from the Lance Formation of Wyoming collected in a University of California Field project, and Dodson (1987) recorded over 16,000 specimens from the Judith River Formation of Alberta, mostly from Dinosaur Provincial Park. The richness of these sites, in terms of both pure numbers of specimens and the species richness, demonstrates the potential of microvertebrate deposits to contribute to our understanding of dinosaur evolution and some of the specimens recovered from these sites form part of the data used for the machine learning analysis described in Chapter Three and applied to isolated theropod teeth in Chapter Five.

Other sites include: the Bathonian of Guelb el Ahmar (Haddoumi et al., 2016) and nearby sites in the Upper Jurassic to Lower Cretaceous Ksar Metlili Formation of Morocco (Knoll

& Ruiz-Omeñaca, 2009; Lasseron, 2019; Sigogneau-Russell et al., 1998); a Bathonian microvertebrate site in Madagascar that has yielded a rare Gondwanan vertebrate fauna including isolated theropod teeth (Flynn et al., 2006; Flynn et al., 1998; Maganuco et al., 2005); Middle and Late Jurassic sites in Russia and Uzbekistan (e.g., Alifanov, 2014; Alifanov et al., 2001; Sues & Averianov, 2013); isolated theropod teeth from the Late Jurassic of Germany (Gerke & Wings, 2016); the Qigu and Shishugou formations (Middle and Upper Jurassic) of the Junggar Basin, China (Maisch & Matzke, 2003; Maisch et al., 2003; Wings et al., 2015); the Kimmeridgian (Upper Jurassic) at Guimorota, Portugal (Rauhut, 2000, 2001, 2003; Zinke, 1998); embryonic or hatchling sauropod teeth, theropods and ornithischians from the Berriasian (Lower Cretaceous) of Cherves-de-Cognac, France (Barrett et al., 2016; Mazin et al., 2006; Pouech et al., 2006); isolated theropod teeth from the Early Cretaceous of Bornholm in Denmark (Lindgren et al., 2008); sauropod and theropod teeth from the Cenomanian (Late Cretaceous) of Charentes, France (Vullo & Neraudeau, 2010; Vullo et al., 2007) and the rich Campanian – Maastrichtian (Late Cretaceous) microvertebrate site of Lo Hueco in Spain (Domingo et al., 2013, 2015; Ortega et al., 2015); the Middle Jurassic Kota Formation of India (Prasad & Parmar, 2020). The taxonomic affinities of many of the specimens recovered from these sites is still disputed (Ding et al., 2020) and as part of this thesis I reassess some of the published data (Averianov et al., 2005; Maganuco et al., 2005; Prasad & Parmar, 2020) using the machine learning techniques described in Chapter Three.

The study of isolated dinosaur teeth

Dinosaurs are polyphyodont and continually replaced their dentitions throughout their lifetime. Estimates for tooth replacement rates range from 14 – 98 days for some sauropodomorphs (D'Emic et al., 2013) to 290 – 777 days for coelurosaurian theropods (Erickson, 1996). The regular shedding of teeth, plus their resistance to wear and chemical alteration (Argast et al., 1987), ensures that isolated dinosaur teeth are common fossils in Mesozoic deposits, often making up the vast majority of dinosaur material recovered from microvertebrate sites, and representing the major source of information for interpretations of the taxa from such sites.

Teeth consist of an inner core of dentine surrounded by an outer shell of enamel. The principal component of both dentine and enamel is calcium phosphate (bioapatite)

which accounts for 70-75% and 96% (approximately) of their respective compositions by weight (Kohn et al., 1999). The phosphate crystals in tooth enamel are large (>1000 nm) when compared to those in dentine (<100 nm) and form a decussate texture that results in an extremely compact and strong structure which is resistant to diagenetic alteration and transport induced abrasion (Argast et al., 1987; Kohn et al., 1999). Experimental studies subjecting fossil vertebrate teeth to transport-induced abrasion and flume testing have concluded that whilst transport alone does not cause extensive changes in the morphology of enamel coated shed teeth (Argast et al., 1987), hydrodynamic sorting mechanisms that act during the accumulation of biological material in microvertebrate sites do have an effect on the faunal composition recovered, with tooth shape being linked to potential transport distance (Peterson et al., 2014).

Teeth have been used in dinosaur taxonomy since the beginning of vertebrate palaeontology, with the recognition that tooth morphology has the potential to be a useful diagnostic tool (e.g., Leidy, 1856; R. Lydekker, 1893; Owen, 1854, 1855, 1861). The first detailed monograph on the comparative anatomy of teeth, including those of dinosaurs, was published by Richard Owen in 1840. One of the earliest examples of the use of isolated teeth to name taxa was that of Leidy (1856) who described various teeth from the Late Cretaceous of the USA and established the taxa *Trachodon mirabilis*, *Troodon formosus* and *Deinodon horridus* on the basis of these. Both *Trachodon mirabilis* and *Deinodon horridus* eventually proved to be nomina dubia whist *Troodon* is now recognised as a valid taxon based on skeletal remains (Senter, 2007).

The use of isolated teeth as taxonomic tools has continued (e.g., Coombs, 1990; Currie et al., 1990; Fiorillo & Currie, 1994; Thulborn, 1973) although the emphasis has now shifted from qualitative measures of teeth to the development and use of quantitative frameworks and tooth datasets that can be analysed using multivariate statistical methods and morphometrics (e.g., Currie et al., 1990; Holtz et al., 1998; Larson & Currie, 2013; Smith, 2005; Smith et al., 2005; Williamson & Brusatte, 2014; Young et al., 2019). More recently, authors such as Hendrickx et al. (2019) have expanded both the taxonomic and spatiotemporal ranges of the taxa included to develop large morphometric datasets that can be adapted and reused. The vast majority of the published literature on this topic concentrates on the quantitative assessment of

isolated theropod teeth, and there has been little work on either ornithischian (Becerra et al., 2013) or sauropod (Barrett et al., 2016) teeth.

Research using isolated dinosaur teeth for taxonomic purposes initially concentrated on theropods from North America, such as those from the late Campanian Dinosaur Park Formation of Canada where skeletal remains could be used in association with the teeth to aid identification (Currie et al., 1990). Heckert (2002) noted that there are three positions that have been taken on the taxonomic utility of isolated dinosaur teeth: 1) that teeth are almost entirely non-diagnostic and have little or no taxonomic value (e.g., Charig & Crompton, 1974; Dodson & Dawson, 1991; Ostrom & Wellnhofer, 1990); 2) that teeth have some diagnostic value, but in the absence of comparable skeletal material the use of isolated teeth in diagnosing taxa at lower taxonomic levels is questionable (e.g., Larson & Currie, 2013; Padian, 1990; Sereno, 1991); and 3) that dinosaur teeth can be taxonomically diagnostic, with synapomorphies that can be used to erect valid taxa (e.g., Heckert, 2002, 2004; Hunt & Lucas, 1994; Thulborn, 1973, 1992).

Dinosaurs exhibit a diversity of tooth sizes and shapes within each of the major clades: theropod teeth are blade like, recurved, labio-lingually flattened and may have serrations (denticles) on either the mesial or distal carina; sauropod teeth show two generalised morphologies, either peg-like or spatulate; and ornithischian teeth have a low triangular crown with a well-developed constriction between the root and the crown, a crown often showing a degree of asymmetry along the linguolabial and mesiodistal axes, large, apically oriented denticles, and an asymmetrical swelling of the crown base (a cingulum). Most dinosaurs are considered to have relatively simple tooth forms with little variation in size, shape and function along the tooth row. Valkenburgh and Molnar (2002), in a comparison of dinosaurian and mammalian predators, noted that although teeth may vary in size along a theropod tooth row, there is little morphological and functional variation between the teeth.

However, dinosaurs exhibit a degree of morphological convergence in tooth shape between clades and, although generally homodont, often possess a degree of heterodonty along the tooth row. For example, early-diverging ornithischians, such as thyreophorans and early neornithischians, show little morphological variation along the tooth rows, with most changes related to size, but some of these taxa, such as

ornithopods and heterodontosaurids, possess premaxillary or specialised cheek teeth, adding to their diversity of form and function. In addition, many features of ornithischian teeth are shared with other Mesozoic archosaurs, such as the Late Triassic aetosaur *Revueltosaurus* (Boyd, 2015; Irmis et al., 2007; Nesbitt et al., 2007; Parker et al., 2005), and (with the exception of a basal cingulum) these cannot be considered as unique synapomorphies of ornithischians (Butler et al., 2008; Nesbitt et al., 2007).

Classification framework methods

To date two generic quantitative frameworks have been developed to assign isolated dinosaur teeth to taxa: linear morphometrics, using principal component analysis and linear discriminant analysis (Larson, 2008; Larson et al., 2016; Larson & Currie, 2013; Smith et al., 2005; Williamson & Brusatte, 2014); and a phylogenetic classification, using character-based descriptions in a phylogenetic framework (Hendrickx & Mateus, 2014; Hendrickx et al., 2019; Hendrickx et al., 2020). A third framework, developed in this thesis, explores the use of machine learning as an alternative approach to the problem.

Linear morphometric methods

The most common approaches applied to the quantitative analysis of isolated dinosaur teeth have been to use either principal component analysis (PCA) or linear discriminant analysis (LDA), or a combination of both (Larson, 2008; Smith et al., 2005; Williamson & Brusatte, 2014). However, both methods (Fig. 1.3), either individually or jointly, produce sub-optimal results. PCA is essentially a dimensionality reduction algorithm where the principal components are linear combinations of the original variables computed so that the first component explains the greatest amount of variance in the original data, the second component, calculated orthogonally to the first, the next highest amount of variance, and so on. If the requirement is to either visualise data for initial exploration, or to reduce high dimensional data to a lower number of dimensions for future calculations then PCA is a reasonable option. However, PCA is not designed to maximise separations between different group structures embedded within the data and therefore any classification based on PCA alone will be suspect (Jolliffe, 2002). LDA, although designed for classifying data and reasonably robust even when assumptions about the data are violated (Fisher, 1936; Hastie & Tibshirani, 1996; Hastie et al., 2009a), performs poorly on isolated dinosaur tooth data due to the degree of morphological

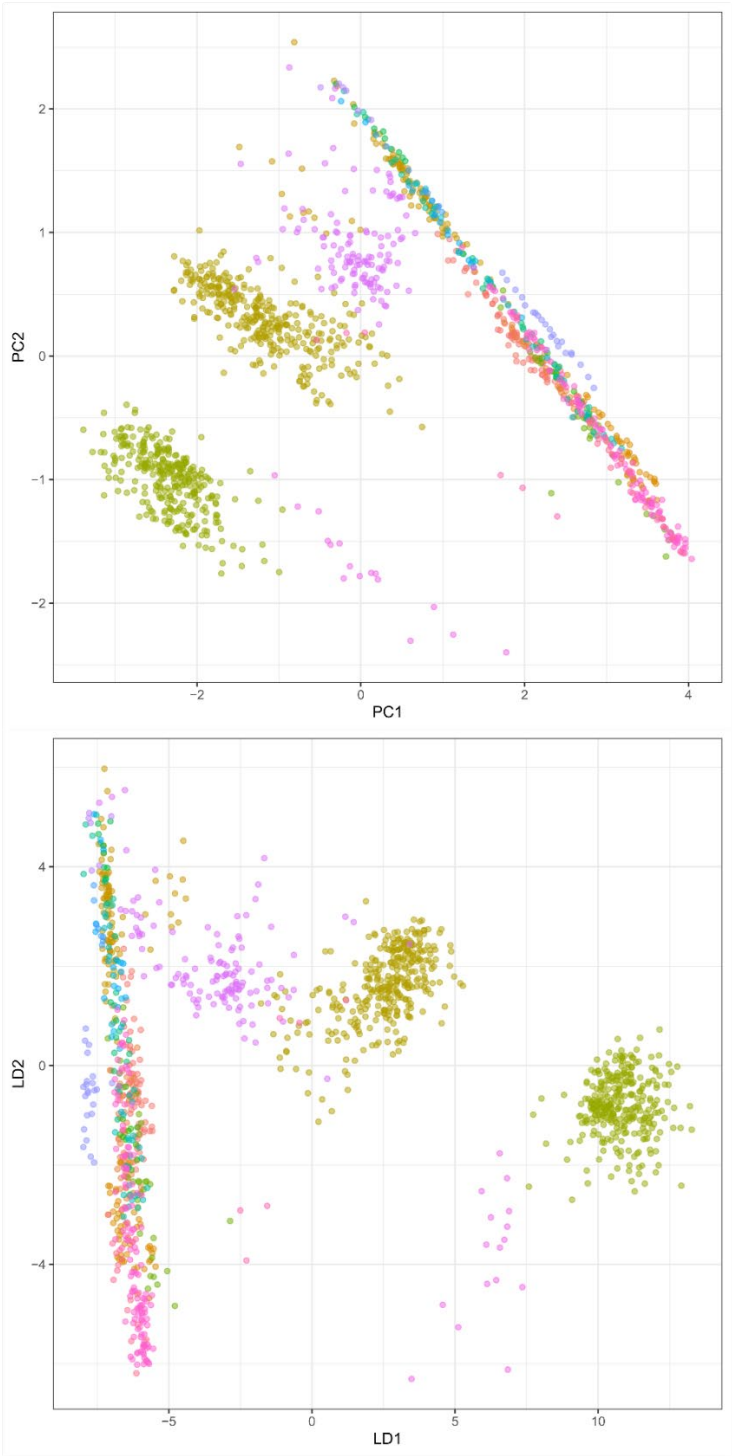


Fig. 1.3. Examples of two different ordinations of the theropod morphometric data used for machine learning analysis (Chapter Three) showing the morphospace overlap of different clades. A, PCA ordination and B, LDA ordination. Colours refer to different theropod groups and are equivalent across both ordinations. Group names are not shown. Each point represents a specimen in either the principal component or linear discriminant morphospace. Axes labels: PC1 and PC2 are the first two principle components, LD1 and LD2 are the first two discriminant functions.

convergence between different dinosaur clades. Figure 1.3 shows two different ordinations of the isolated theropod tooth data used in Chapter Three, a PCA ordination and a LDA ordination. Both ordinations show the inherent problem with large areas of morphospace overlap, apart from a few clades (such as some maniraptorans), ensuring that any resultant classification will be inaccurate. In Chapter Three, I explore the issues around PCA and LDA in more detail.

Phylogenetic methods

More recently, the use of a dentition-based phylogenetic method has been advocated to classify isolated theropod teeth (Hendrickx & Mateus, 2014; Hendrickx et al., 2019; Hendrickx et al., 2020) with the suggestion that this methodology provides superior results to LDA. This reasoning is based on perceived issues with the data quality used for LDA: missing measurements for damaged or incomplete tooth crowns; differences in measurement methodology between different workers; and the lack of classification resolution resulting from morphological convergence. However, much of this reasoning is flawed and arises from a lack of understanding of the LDA algorithm and the data used for the analysis. LDA does not accept cases with missing data unless those missing data values are somehow imputed: therefore, including such specimens in an analysis will degrade the results. Character-based approaches also suffer from missing data, and without an analysis of which characters are taxonomically informative it is difficult to understand how this would affect the results. In addition, a character-based methodology is, by its very nature, subjective with the potential for different researchers to code characters differently for the same specimen. The phylogenetic methodology necessitates the coding of up to 145 dentition-based characters and requires the use of a constrained backbone topology based on current theropod research to produce informative results. If the topological constraint is removed (Fig. 1.4) the resultant tree collapses into a polytomy. The phylogenetic method does have one potential advantage over other LDA or machine learning techniques in that it does not require a large training dataset with multiple examples of the same clade to function. However, this can also be seen as a disadvantage, as each clade is scored using a only single set of characters which then must represent an idealised view of the dental characters for that clade.

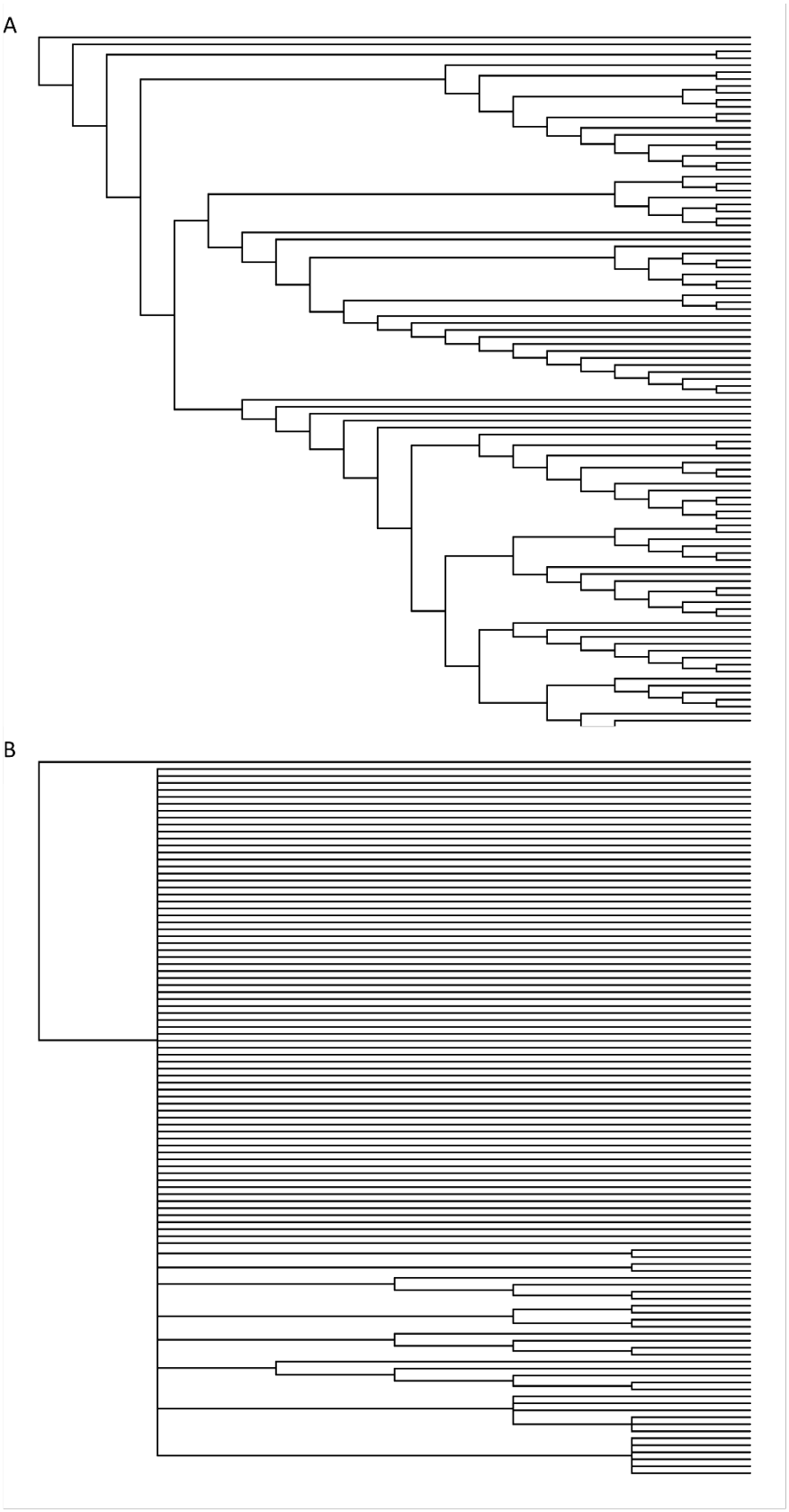


Fig. 1.4. Character based phylogenetic analysis of isolated theropod tooth crowns, based on data and methodology from Young et al. (2019). A, tree recovered when enforcing a backbone theropod topology. B, tree collapsed into a polytomy recovered with the backbone theropod topology removed. Taxon labels are not shown for clarity.

In Chapter Three, I discuss the issues around different measurement techniques of the morphometric data employed by different authors and conclude that the effects on the final classification result are negligible. I also discuss the negative effects of imputing missing data into a classifier and the more general limitations of linear discriminant analysis as a classifier.

Machine learning

Machine learning, the fitting of models to data and the identification of groupings of interest among those data, is a rapidly growing field that has been embraced by many disciplines. Machine learning is especially useful when attempting to analyse large and complex datasets in a reproducible fashion (Greener et al., 2022) and can operate in two modes: 1) supervised learning where the input (training data) has been labelled based on pre-existing knowledge of the data; and 2) unsupervised where the input is unlabelled and the algorithms are tasked to find patterns in the data with no prior knowledge of its structure. The use of machine learning in geology is ever increasing, with applications such as rock classification (Dawson et al., 2023), seismology (Mousavi et al., 2019), and lithofacies classifications (Halotel et al., 2020). In vertebrate palaeontology, however, the use of machine learning is still in its infancy (Monson et al., 2018). Machine learning has the potential to analyse datasets and find patterns in data where other methods have limitations (Hoyal Cuthill et al., 2019; MacLeod, 2017; Macleod et al., 2021; MacLeod & Kolska Horwitz, 2020; Wills et al., 2021) but the plethora of different machine learning techniques and algorithms – such as convolutional neural networks (CNNs), decision trees, support vector machines – and the requirement to have a robust training dataset is proving to be a barrier to uptake. In Chapter Three, I develop a machine learning methodology to classify isolated teeth, based on published linear morphometric data, which is then applied in Chapter Five to theropod teeth from British Middle Jurassic microvertebrate sites and to published data from other microvertebrate sites in India, Kyrgyzstan, and Madagascar.

Study aims

Microvertebrate dinosaur remains from the Middle Jurassic of Britain, predominantly isolated teeth, have remained understudied despite a long collecting history dating back to the 1970's. Taxonomic referrals of the teeth from four major sites (Hornsleasow

Quarry, Kirtlington Quarry, Woodeaton Quarry and Watton Cliff) remain in dispute masking the true picture of dinosaur diversity in this region during this important time period. In this thesis I aim to investigate the taxonomic affinities of the theropod and ornithischian teeth from these sites through the development of new machine learning methodologies in parallel with traditional morphology-based approaches.

Chapter Two: This chapter contains descriptions of the methods used for field site data collection, imaging, and modelling.

Chapter Three: This chapter describes a new machine learning methodology developed to classify isolated dinosaur teeth. Traditionally the classification of isolated dinosaur teeth, usually theropod teeth, has relied on a mixture of morphological description, phylogenetic analysis and standard linear models such as linear discriminant analysis. These techniques have had varying degrees of success, often with poor classification outcomes, due primarily to the high degree of morphological convergence in tooth shape between different dinosaur clades and a reliance on non-optimal modelling algorithms. Here, I develop an entirely new approach to the problem using multiple machine learning algorithms to build and combine classification models. This approach will then be used as the framework for the classification of isolated theropod teeth from the British Middle Jurassic and elsewhere (see Chapter Five).

Chapter Four: This chapter comprises a review of the regional geological setting during the Middle Jurassic of Britain and the local geological setting of the four sites under investigation. This includes new site descriptions of Woodeaton Quarry, Oxfordshire and Watton Cliff, Dorset that are based on new fieldwork carried out during 2014 to 2017.

Chapter Five: The first quantitative assessment of isolated theropod teeth, from the four study sites, using the new machine learning approaches described in Chapter Three. This work demonstrates the power of these techniques in providing quantitative assessments of large isolated tooth samples. I develop a robust, testable framework for taxonomic identifications, and use the results obtained to highlight the importance of assessing and including evidence from microvertebrate sites in faunal and evolutionary analyses. In addition to the four UK study sites, I also apply this methodology to

reappraisals of isolated theropod teeth from the Middle Jurassic of India, Morocco and Kyrgyzstan.

Chapter Six: An assessment of the ornithischian fauna from the four study sites using morphological descriptions based on 3D CT generated models of individual teeth. This chapter uses the CT techniques described in Chapter Two and discusses the implications of these identifications for dinosaur species-richness and evolution.

Chapter Seven: A discussion and synthesis of the key findings of this research, including possible future directions of this work. This includes the ongoing development of machine learning techniques and the use of other data collected from the UK sites.

Chapter Two: Material and Methods

Introduction

The microvertebrate material that forms the basis of this thesis comes from two sources: pre-existing collections made from Hornsleasow Quarry (Fig. 2.1), Kirtlington Quarry (Fig. 2.2) and Watton Cliff, and new material from Woodeaton Quarry (Figs. 2.3 and 2.4), Hornsleasow Quarry (Fig. 2.5) and Watton Cliff (Fig. 2.5). Original Hornsleasow material was sourced from the Museum of Gloucester (Metcalf et al., 1992; Metcalf & Walker, 1994; Vaughan, 1989) with Kirtlington and Watton Cliff material (Evans & Milner, 1994) coming from the collections of Susan Evans, University College London, which has now been transferred to the Natural History Museum. The new material was collected from the field by the author and, in the cases of Woodeaton and Hornsleasow Quarries, various teams from the Natural History Museum (Wills et al., 2019). New material from Hornsleasow Quarry was derived from microvertebrate-rich clay that was excavated during the late 1980's but left in the field (Darlington, 1988; Vaughan, 1989). The Museum of Gloucester held title to this material but transferred this to the Natural History Museum, London. The quarry owners, Breedon Group, generously gave permission for access to the quarry and collection of material in 2014 and 2022. A total of 6.4 tonnes of new sediment was collected and processed, see Table 2.1 for details.

Some of the existing collections of microvertebrate material were already sorted to some extent, with specimens often assigned to a general taxonomy or morphotype; in other cases this material consisted of a concentrate of mixed microvertebrate material. In preparation for analytical work this material was examined using optical microscopy, sorted, picked, and re-categorised as appropriate. New material collected in the field from Woodeaton (Wills et al., 2019), Hornsleasow and Watton Cliff was washed down to a concentrate prior to picking, sorting and imaging using standard processing techniques (Ward, 1981; Ward, 1984). The specimens from Kirtlington, Woodeaton, Watton Cliff and newly collected Hornsleasow material were given temporary (project-based) catalogue numbers and these will be accessioned into the main NHMUK collection.

Fig. 2.1. The original excavation at Hornsleasow Quarry in 1988 by teams from Gloucester County Museum and the University of Bristol. CNF, Chipping Norton Limestone Formation; HC, Hornsleasow Clay deposit, the microvertebrate rich horizon. Image taken from Metcalf et al. (1992).

Fig. 2.2. The original excavation at Kirtlington Quarry by University College London. A, view of the eastern quarry face at Kirtlington (date unknown) prior to excavation. The excavation site is lies at about three-quarters of the height of the quarry face, below the line of trees to the left of the image. B, excavation and removal of overburden in March 1981. C, the original site in April 1976. Images courtesy of Susan Evans, University College London.

All of the 3D models and CT scans that form part of the published works resulting from this research (Wills, Underwood, et al., 2023) are archived on Morphosource (www.morphosource.org, see the data archiving statement, Appendix One. As part of the fieldwork I updated geological sections and sedimentological descriptions for both Woodeaton Quarry and Watton Cliff. See Chapter Four, for the updated site descriptions and an overview of the sites including the history of collecting at each.

A total of 1,256 isolated teeth were picked from sample residues and imaged subsequently. These were supplemented by 253 specimens from other institutions and data repositories (see Material, below). For the statistical and machine learning analysis (Chapters Three and Five) of isolated theropod teeth, morphometric measurement data from 4,099 individual teeth were sourced from the literature. Individual teeth were scanned using micro-CT to develop 3D models from which morphometric measurements were taken. A bulk scanning methodology was developed to allow rapid CT data collection and segmentation.

Fieldwork

Fieldwork was undertaken at three of the sites: Woodeaton (Wills et al., 2019), in three field seasons between 2014 and 2016 with a large field-crew from the Natural History Museum (Fig. 2.3); Watton Cliff in 2012 and 2013 (Fig. 2.5); and Hornsleasow Quarry in 2014 and 2022 (Fig. 2.5). The aim of the fieldwork was to improve on existing site descriptions where possible and to obtain additional sediment for bulk processing and microvertebrate extraction. The section at Kirtlington Quarry in Oxfordshire, as observed in 2014, is currently overgrown and inaccessible.

Accessible sections were examined for each of the above mentioned quarries and cliff faces and detailed stratigraphic and sedimentological data were logged. All stratigraphic and sedimentological data was recorded on logging charts and later converted to digital format. Lithological descriptions consist of bed thickness, lithology, grain size, colour, trace and body fossils and sedimentary structures. Also captured were field interpretations of facies and palaeoenvironments (Coe et al., 2010; Miall, 1997). See Chapter Four for detailed geological descriptions of each site.

Fig. 2.3. Fieldwork at Woodeaton Quarry, Oxfordshire in 2014 by teams from the Natural History Museum, London. A and B, initial clearance of section with mechanical digger to expose the microvertebrate horizon. C, excavation of microvertebrate-rich clay by hand. The clay horizon is directly below the dark grey (palaeosol) towards the upper part of the exposed section. D, the quarry face after clearance and removal of vegetation. E, cleared section exposing (from top to bottom): flaggy bioclastic limestones of the Forest Marble Formation forming the overhang; dark grey palaeosol of the White Limestone Formation; brown microvertebrate rich marl. F, removal and transport of sediment from the quarry face.

A

B

Fig. 2.4. Team from University College London prospecting at Woodeaton Quarry in 1983. The sections exposed correspond to those excavated during the 2013 to 2016 fieldwork. A, section through the upper part of the White Limestone Formation and contact with the Forest Marble Formation. The bed labelled 'B' on the image corresponds to Bed 23 of the Bladon Member, White Limestone Formation which is the main microvertebrate-bearing horizon at Woodeaton Quarry. The bed labelled 'D' is the lowermost Forest Marble Formation. B, the quarry face at Woodeaton exposing the upper part of the White Limestone Formation and the Forest Marble Formation. The Forest Marble Formation forms the prominent band of limestone midway down the face. See Chapter Four, geological setting for more detail of the site geology. Images courtesy of Susan Evans, University College London

Fig. 2.5. New field collection at Watton Cliff, Dorset in 2015 and at Hornsleasow Quarry, Gloucestershire in 2022. A, excavating unconsolidated sediment in the Forest Marble Formation at Watton Cliff from a small invertebrate burrow. B, Watton Cliff looking west towards Eype Mouth. The slope in the foreground is composed of the Fullers Earth Formation littered with fallen blocks of bioclastic limestone of the Forest Marble Formation. C, collecting bags of sediment in 2022 at Hornsleasow Quarry. The sediment was left in the field after the initial excavation in 1989 by Gloucester City Museum.

Table 2.1. Summary of collection efforts at Hornsleasow Quarry, Kirtlington Quarry, Woodeaton Quarry and Watton Cliff.

Site	Initial sediment collection	New sediment collection	Teeth picked and imaged (*)
Hornsleasow	20 tonnes by Museum of Gloucester. 13 tonnes of which were processed by the museum and the University of Bristol (Darlington, 1988; Vaughan, 1989).	500 kg of material in 2014 and 1 tonne in 2022 by the author with assistance from S. Maidment and J. Bonsor from the Natural History Museum, London.	225
Kirtlington	7.3 tonnes by University College London between 1975 and 1980 (Evans & Milner, 1994).	-	733
Woodeaton	-	4.8 tonnes between 2013 and 2015 by the author and teams from the Natural History Museum, London (Wills et al., 2019).	216
Watton Cliff	510 kg by University College London between 1975 and 1976.	100 kg in 2012 and 2013 by the author.	9

(*) The numbers of teeth picked and imaged is not necessarily the total number of isolated teeth collected from each site. All dinosaur teeth were picked from the concentrates obtained at each site, with only representative samples of other taxa being picked.

Sediment collection and processing

For new material, specimens were obtained by bulk sediment collection on site following initial fieldwork to identify productive horizons and site clearance. Large bulk samples (often several weighing tonnes) were screen-washed using the methodology described by Ward (1981). Screen-washing removes the clay fraction (<500 µm) leaving an initial concentrate of material. This initial concentrate can be further processed in acid to remove the carbonate content if required. After screen-washing and acid processing the remaining concentrate was split into size fractions (< 1 mm, 1 mm - 2 mm, 2 mm - 4 mm, > 4 mm) to create convenient sample sizes for vertebrate picking (Fig. 2.6).

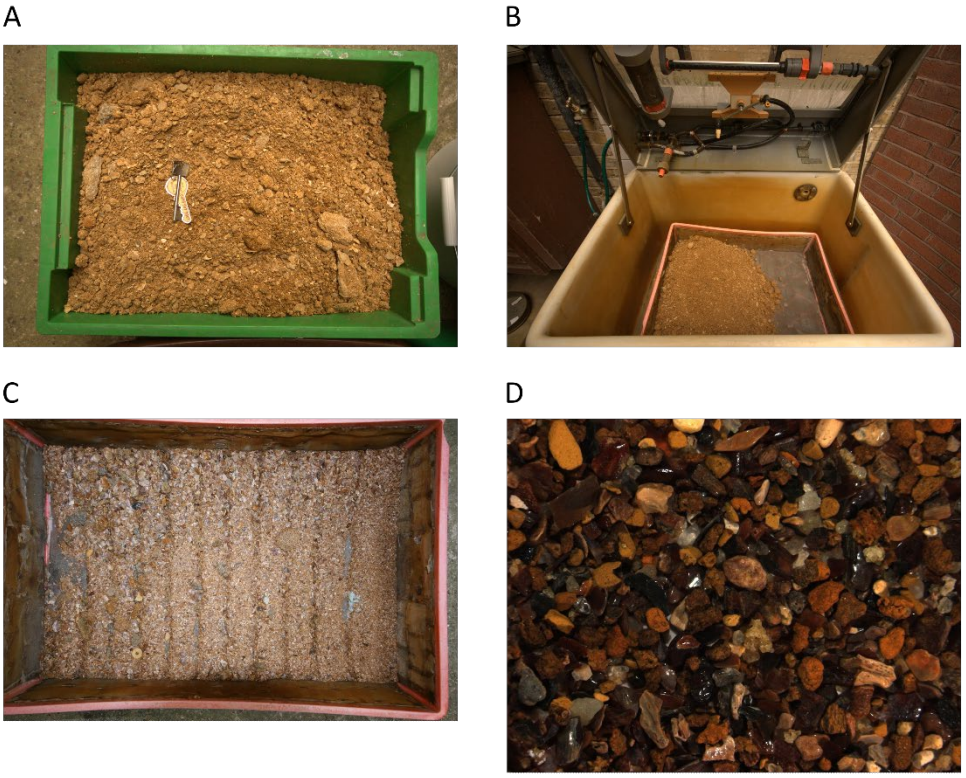


Fig. 2.6. Processing and screen-washing of bulk sediment from Watton Cliff at the Natural History Museum, London. A, initial microvertebrate-rich sediment collected from burrows and channels at Watton Cliff prior to processing. B, sediment awaiting processing in a screen-washing unit (Ward, 1981). C, washed and unsorted sediment after several hours of processing and prior to splitting into size fractions. D, 1–2 mm concentrate for picking under a binocular microscope.

At Woodeaton Quarry bulk sampling was carried out in both the Rutland Formation clays and in the clay and marl horizons of the Bladon Member (White Limestone Formation) over a period of three field seasons from 2013 to 2015. An initial 1.6 tonnes of rock were removed by hand, initially dried and then screen-washed. It was found that most of the dried sediments washed at a rate of between 3–10 kg per hour. The clay horizons in the Rutland Formation disaggregated readily, with the 500 µm residues comprising around 1–2% of the original sample weight. The oyster-rich horizons of the White Limestone Formation disaggregated more slowly and yielded large volumes of oyster shards, which represented approximately 10% of the original sample weight. The coarse fractions, between 6–12 mm, which were mainly oyster shell, were dried and sorted manually. The finer fractions, less than 6 mm, were decalcified in 7.5% formic acid buffered with 20% spent acid. A pH meter was used to ensure that the pH did not drop below 3.2. An untreated archive sample of the fine fraction was saved. The decalcified residue was washed several times to remove any calcium formate residues, dried, graded and picked under a binocular microscope. The sorted fraction was retained and archived. During the first stages of fieldwork a section along the western edge of the quarry was examined. This section exposed a sequence of limestones and clays of the Ardley and Bladon Members (Great Oolite Group, White Limestone Formation) through to the shelly detrital limestone of the Forest Marble Formation. This section was previously investigated by the team from University College London in 1983 (Fig. 2.4 and D. J. Ward, personal communication, May 2023). One bed, a variably lithified clay to impure limestone horizon with abundant plant material (Bed 23, see Chapter Four Geological Setting), produced substantial quantities of terrestrial microvertebrate remains. Bulk sampling concentrated on this unit thereafter, with 2.4 tonnes of sediment collected and an additional 800 kg sampled from the overlying bed. After screen-washing, drying, and splitting of the concentrate into size fractions, approximately 30 kg of concentrate from Bed 23 remained. Each size fraction was picked for microvertebrate remains under a binocular microscope and these were identified to the lowest taxonomic level possible and given temporary project numbers. Specimens recovered from the 2013 to 2015 field seasons are included in this thesis.

An additional 100 kg of material from Watton Cliff was collected during 2012 and 2013 to supplement the material previously collected by University College London. Unconsolidated sediment known to contain microvertebrates, found in channelised lenses and burrows within the main bioclastic limestone sequence of the Forest Marble Formation, was hand excavated. The screen-washing process followed the same methodology described above although different size fractions were chosen to split the screen-washed residue because of the coarse nature of the shelly concentrate (> 9.5 mm, 4–9.5 mm, 2–4 mm, 1–2 mm, 500 µm – 1 mm, < 500 µm). After removal of the clay fraction by washing, the residue still contained large amounts of carbonate material. This was mostly removed from the smaller size fractions by preparing the samples in 10% acetic acid to aid picking. The > 9.5 mm and 4–9.5 mm fractions were hand-picked for vertebrate material by eye. The smaller size fractions were picked using a binocular microscope. The initial 100 kg of sediment reduced in weight to 45 kg after screen-washing to remove the clay material. Acid preparation of 8.65 kg of the smaller size fractions resulted in a concentrate of 0.207 kg. In addition, some 10 kg of cemented limestone was dissolved in acetic acid and picked for vertebrate remains. Specimens recovered from the 2012 and 2013 trips are included in this thesis.

Initial excavations at Hornsleasow Quarry in the late 1980's recovered around 40 tonnes of microvertebrate-rich clay, the majority of which was transported and processed off-site (Darlington, 1988; Vaughan, 1989). After learning that material remained uncollected on site some 25 years post-excavation (D. J. Ward, personal communication, May 2013) permission was obtained from the quarry owners (initially Huntsmans Quarries, Cheltenham and later Breedon Group, Derby) to recover this material. Two collecting visits were made to Hornsleasow Quarry in 2014 and 2022. The 2014 visit identified that 3 tonnes of sediment remained from the original excavation. The sediment was still in the original bags although these were in poor condition, and sample labels were still present and legible in the majority of cases. Approximately 500 kg of sediment was re-bagged, labelled and transported to the Natural History Museum for screen-washing, sorting and picking. The remainder of the material was re-bagged, covered and left for later collection. The 2022 visit recovered an additional 1.1 tonnes

of sediment that is awaiting processing at the Natural History Museum. Specimens recovered from the 2014 trip are included in this thesis.

Material

For this thesis isolated theropod and ornithischian teeth were picked from concentrates housed at Gloucester City Museum and University College London, augmented with new samples collected in the field. A total of 1,256 isolated teeth were picked and imaged from the four sites under study, consisting of: 577 ornithischians; 164 theropods; seven sauropods; 284 crocodilians; 68 mammals; and 83 pterosaurs. The breakdown per-site is shown below.

Hornsleasow: 54 ornithischians, 50 theropods, three sauropods, 116 crocodilians, two pterosaurs.

Kirtlington: 489 ornithischians, 49 theropods, one sauropod, 113 crocodilians, 18 mammals, 63 pterosaurs.

Woodeaton: 29 ornithischians, 61 theropods, three sauropods, 55 crocodilians, 50 mammals, 18 pterosaurs.

Watton Cliff: five ornithischians, four theropods

Only the ornithischian and theropod teeth were analysed as part of this thesis, the remaining teeth were used for comparative morphology and will be used in future machine learning projects.

For morphological and statistical comparisons to known taxa, data were sourced from museum collections, data repositories and the literature. A summary of these sources is given below (see Appendix Two for a complete list).

Dorset County Museum, Dorchester, UK: 71 theropod specimens (isolated teeth).

Morphosource: five ornithischian taxa from various institutions.

Museo de Biología de la Universidad del Zulia, Venezuela: 10 ornithischian specimens (isolated teeth).

Natural History Museum, London, UK: 90 theropod, 50 ornithischian and 15 sauropod specimens.

Natural History Museum of Los Angeles County, Los Angeles, USA: one ornithischian specimen.

Oxford University Museum of Natural History, Oxford, UK: 11 theropod specimens (isolated teeth).

Published literature: 4,099 morphometric measurements of isolated theropod teeth covering a wide range of taxa, geographic and temporal distributions datasets (Currie & Varrichio, 2004; Farlow et al., 1991; Gerke & Wings, 2016; Hendrickx et al., 2015a; Larson, 2008; Larson et al., 2016; Larson & Currie, 2013; Longrich, 2008; Rauhut et al., 2010; Sankey, 2008; Sankey et al., 2002; Smith, 2005; Young et al., 2019).

Imaging and modelling

Digital images of specimens were initially obtained using either a Zeiss Axiocam HRc CCD digital camera mounted on a Leica stereo zoom dissecting microscope or a LEO 1455 variable pressure scanning electron microscope. This process is very time consuming and highlighted problems with the specimen handling and data capture with respect to microvertebrate specimens. These are invariably small, so difficult to handle and measure directly, are often fragile and generally occur in large quantities. Although many digital imaging techniques, such as the optical microscopy and SEM capture initially employed here, overcome the direct handling and measurement issues by allowing measurements to be taken from the image data rather than the specimen directly, the process is still time-consuming as images need to be taken in multiple orientations for measurements and morphological descriptions. To overcome these problems, I developed a bulk imaging technique using micro-CT, chosen because of its ability to bulk capture multiple specimens in a single scan and generate 3D models from the data. This technique allows direct measurement of morphometric data, multiple visualisation and manipulation options at different orientations, resolutions and scales, and the ability to use the 3D data for future work. The use of 3D models also negates the need to handle the original specimens. To reduce scan time, and to allow the simultaneous data capture of multiple specimens, a 4 x 4 cm bespoke specimen tray was designed and 3D printed using thermoplastic aliphatic polyester. Specimen locator boxes on the trays were printed at two sizes, 5 and 10 mm. This allows for the inclusion of up to 36 individual specimens per tray. Trays were stacked six high to produce a 4 x 4

x 4 cm cube containing a maximum of 216 specimens (Fig. 2.7). Each tray was labelled in order to be able to locate individual specimens consistently, e.g., tray 1, position A1.

Fig. 2.7. 3D printed specimen holder trays for CT scanning of microvertebrate specimens. Specimen trays are 4 x 4 cm in size with locator boxes either 0.5 x 0.5 cm (left tray) or 1 x 1 cm (right tray) allowing either 36 or nine individual specimens to be located per tray. Trays are stacked six high for scanning.

CT data was captured using both a Zeiss 520 Versa system (Carl Zeiss Microscopy GmbH, Jena, Germany) and Nikon HMX ST 225 (Nikon Metrology, Leuven, Belgium) at the Natural History Museum, London, UK. The full CT and post-processing workflow (Fig. 2.8) allows for the scanning and 3D model generation of multiple individual specimens in a few hours resulting in considerable time savings when compared to scanning and processing each specimen individually.

The full processing workflow is described below:

1. Create a spreadsheet with the following columns: *specimen number, tray, number, tray position, X_Max, X_Min, Y_Max, Y_Min, Z_Max, Z_Min, scan date*. These refer to the position of each specimen in the data cube once scanned.

2. Load all specimens to be scanned into series of specimen holders and stack to a maximum of six trays for a single scan. Update the spreadsheet with specimen number, tray number and tray position, e.g., NHMUK PV R35029, Tray 1, A3.

3. For the Versa, collection tomographic reconstructions were performed automatically using the Zeiss Scout-and-Scan scan software with the resulting 32-bit .xrm volume files converted to 16-bit tiff stacks using Zeiss XRM controller software. For the Nikon, collection tomographic reconstruction was performed using CT-agent software (Nikon Metrology GmbH, Alzenau, Germany) with a beam-hardening correction and a 3 x 3 median filter applied before saving the reconstructed dataset as a 16-bit tiff stack. Initial voxel resolutions range from 0.003 mm for the Versa system capture to 0.025 mm for the Nikon system capture.

4. After tomographic reconstruction and export of the data cube to a 16-bit tiff stack, Fiji v. 2.1 (Schindelin et al., 2012) was used to scan through the volume and locate each individual specimen using its range of x, y and z coordinates within the stack. The coordinates were then added to a .csv file (*X_Max, X_Min, Y_Max, Y_Min, Z_Max, Z_Min*) so that each line corresponds to an individual specimen location in the data cube (Fig. 2.9). Because of the limitations of 3D printing and the differing orientations of the specimen stacks (for different scans) in the CT scanners it was not possible to automate this process.

5. Several Python (v. 3.x) scripts and Fiji macros were developed that read the complete data stack, thresholding the data to automatically separate the specimen holder from the specimen, clipping out each individual element into its own stack based on the coordinates and labels in the .csv file, and generating a 3D surface mesh for each specimen from the resultant stacks labelled with the specimen number. In some cases additional processing of the image stacks was undertaken in Avizo v. 8.1 (ThermoFisher, 2014) to achieve better segmentation results.

6. For rapid visualisation and rendering, a combination of Matlab v. 2023a (The MathWorks Inc, 2023) and Python was used to re-orientate each model to its long axis, scale each to unit size, take a series of snapshots in xy, xz, yx, yz, zx and

zy views and take a series of snapshots at a user defined angle around each of the x, y and z axes in turn.

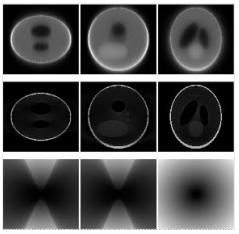
7. Linear morphometric measurements were taken directly from the 3D models using Avizo with the snapshots used to aid morphological descriptions. See the section on Machine Learning, below, for details of measurements taken.

All of the scripts used in this process are available in Appendix Three and at the following GitHub repository: https://github.com/simonwills/Bathonian_Dinosaurs

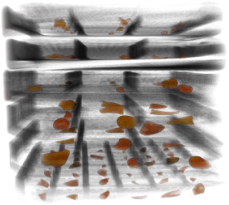
Loading of specimens into 6 x 6 or 3 x 3 trays for scanning.

Bulk scanning of multiple specimens with six specimen trays stacked together for one-pass scan.

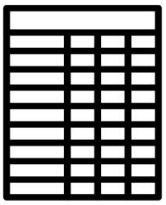
Tomographic reconstruction.



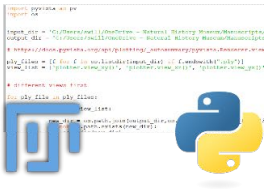
Export of entire multiple-specimen volume as single 16-bit tiff stack.



Completion of specimen locator spreadsheet containing specimen number and x, y, z position of each individual specimen in the full data volume.



Semi-automatic post-processing of entire stack using Python and Fiji to create individual tiff stacks and 3D meshes.



Visualisation and measurement of 3D models in Avizo, Python and Matlab.

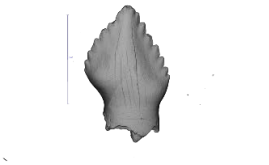


Fig. 2.8. Micro-CT workflow and post-processing steps to create 3D models from isolated microvertebrate specimens. The workflow covers from initial scanning to modelling and visualization.

Chapter Three: Machine Learning

Preface: This section has been published as:

Wills, S., Underwood, C. J., & 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(1), 75–99. https://doi.org/https://doi.org/10.1111/pala.12512

This paper is my original work. Barrett and Underwood contributed by editing the manuscript and providing discussion.

Abstract

The development of repeatable and quantifiable methodologies to deliver robust taxonomic identifications of isolated dinosaur teeth is crucial if such material from microvertebrate deposits is to be included in broader evolutionary studies and analyses. Taxonomic identification of fossils based on morphometric data traditionally relies on the use of standard linear models to classify such data. Machine learning and decision trees offer powerful alternative approaches to this problem but are not widely used in vertebrate palaeontology. Here, I apply these techniques to published morphometric data of isolated theropod teeth in order to explore their utility in tackling taxonomic problems. I chose two published datasets consisting of 886 teeth from 14 taxa and 3020 teeth from 17 taxa, respectively, each with five morphometric variables per tooth. I also explored the effects that missing data have on the final classification accuracy. My results suggest that machine learning and decision trees yield superior classification results over a wide range of data permutations, with decision trees achieving accuracies of 96% in classifying test data in some cases. Missing data or attempts to generate synthetic data to overcome missing data seriously degrade all classifiers predictive accuracy. The results of my analyses also indicate that using ensemble classifiers combining different classification techniques and the examination of posterior probabilities is a useful aid in checking final class assignments. The application of such techniques to isolated theropod teeth demonstrate that simple morphometric data can be used to yield statistically robust taxonomic classifications and that lower classification accuracy is more likely to reflect preservational limitations of the data or poor application of the methods.

Introduction

The use of non-linear analytical techniques (Table 3.1) that draw upon the rapidly expanding field of machine learning and decision trees has remained mostly unexplored with respect to characterizing fossil vertebrate morphology (Monson et al., 2018). By contrast, other disciplines have rapidly embraced machine learning techniques to undertake classification, prediction and various modelling tasks (Christin et al., 2019). Applications range from ecological modelling (Cutler et al., 2007; Džeroski, 2001), population monitoring (Britzke et al., 2011), automated taxonomic classification by phenotype (Hoyal Cuthill et al., 2019), medical image analysis (Ker et al., 2018), financial modelling and prediction (De Spiegeleer et al., 2018; Ma & Lv, 2019), psychology (Finch et al., 2014; Holden Bolin & Finch, 2014; Holden et al., 2011), and bioinformatics (Chen & Ishwaran, 2012; Couronné et al., 2018) to the digitisation of natural history collections (Schuettpelz et al., 2017). Automated and semi-automated approaches of data modelling have also been used for taxon identification, testing morphological variation and quantifying dietary inference from tooth surface morphology (Evans et al., 2007; MacLeod, 2007, 2017; Macleod et al., 2021; MacLeod & Kolska Horwitz, 2020; MacLeod et al., 2008; MacLeod & Steart, 2015; Melstrom & Irmis, 2019; Wilson et al., 2012) and are commonly used in the analysis of earth observation data (MacLeod, 2019; Onojeghuo et al., 2018; Son et al., 2018). Here I test the suitability of these methods for the taxonomic identificiation of fossils, using isolated non-avian theropod dinosaur teeth as a case study. Previously, standard linear classification models have been used to classify these specimens based on shape data (see below). Here I apply several alternative approaches to this problem and assess their comparative performance based on analysis of two datasets of isolated theropod tooth measurements.

Table 3.1. Glossary of terms used in machine learning and classification.

Term	Meaning	Reference(s)
Bagging	Also known as bootstrap aggregating. Used to reduce the variance of a decision tree classifier by creating training sample subsets on which to train the tree. A form of ensemble learning.	Kuhn and Johnson (2013b)
Boosting	A process whereby many weak classifiers are combined into a strong classifier.	Kuhn and Johnson (2013b), Valiant (1984)
C4.5	An algorithm used to create decision trees.	Quinlan (1993), Salzberg (1994)
C5.0	An algorithm used to create decision trees. The successor to C4.5.	Kuhn et al. (2018)
Decision trees	A supervised learning technique.	Kuhn and Johnson (2013b)
Ensemble learning	Combining a group of classifier models to produce a final prediction.	Hastie et al. (2009b)
Linear discriminant analysis (LDA)	A linear model for classification that seeks to find a combination of predictor values to categorise samples into groups. Also known as discriminant function analysis (DFA).	Fisher (1936), Welch (1939)
Linear model	A model in which the terms that describe the model form a linear equation.	Riffenburgh (2012)
Logistic regression (LR)	A linear model for regression and classification.	Finch et al. (2014)
Machine learning	A method of data analysis in which the model learns from new data.	MacLeod (2017), Mitchell (2006)
Mixture discriminant analysis (MDA)	A non-linear extension to linear discriminant analysis.	Hastie and Tibshirani (1996)
Naïve Bayes (NB)	A non-linear machine learning technique for group classification.	Russell and Norvig (2009)
Non-linear model	A model in which the terms that describe the model do not form a linear equation.	Riffenburgh (2012)
Posterior probability	The probability that a case can be assigned to a particular class after classification.	Kuhn and Johnson (2013c)
Principle component analysis (PCA)	A technique to reduce the dimensionality of data whilst minimizing information loss.	Jolliffe (2002)
Prior probability	In Bayesian statistics the prior distribution of the event i.e. the known or expected probability of an observation coming from a particular group before the classification is run.	Kuhn and Johnson (2013c)
Pruning (winnowing)	A process to reduce overfitting of a model generated using the C5.0 algorithm.	Kuhn (2008)
Random forests (RF)	An algorithm used to create a series of uncorrelated decision trees which are combined into one model.	Kuhn and Johnson (2013a)
Synthetic data	Data generated programmatically that does not exist in the original dataset.	Chawla et al. (2002)

The regular shedding of functional teeth (Currie et al., 1990; Farlow et al., 1991), plus their resistance to abrasion and chemical alteration (Argast et al., 1987; Peterson et al., 2014), results in the recovery of abundant, isolated dinosaur teeth in many Mesozoic terrestrial deposits (e.g., Evans & Milner, 1994; Fiorillo & Currie, 1994; Gates et al., 2015; Larson & Currie, 2013; Metcalf & Walker, 1994; Oliver W. M. Rauhut, 2002; Sankey et al., 2002; Sankey et al., 2005; Williamson & Brusatte, 2014). These teeth represent the vast majority of dinosaur material recovered from microvertebrate localities, and often represent the only source of information for interpretations of dinosaur species-richness and palaeoecology from such sites (e.g., Larson et al., 2016; Williamson & Brusatte, 2014; Wings et al., 2015). A reliable, repeatable framework for assessing the taxonomic identity of isolated teeth would therefore be useful in providing more accurate assessments of the faunal compositions of both microvertebrate localities and other localities where skeletal material is rare or uncommon. Historically, three positions have been taken on the taxonomic utility of isolated dinosaur teeth (Heckert, 2002): (1) that teeth are almost entirely non-diagnostic at generic or specific level and have little or no taxonomic value (e.g., Charig & Crompton, 1974; Dodson & Dawson, 1991; Ostrom & Wellnhofer, 1990); (2) that teeth have some diagnostic value, but in the absence of other skeletal material the use of isolated teeth in diagnosing taxa to higher taxonomic levels is questionable (e.g., Currie et al., 1990; Larson & Currie, 2013; Padian, 1990; Sereno, 1991); and (3) that dinosaur teeth can be taxonomically diagnostic and bear synapomorphies that can be used to erect valid taxa or assign isolated teeth to known existing taxa (e.g., Heckert, 2002, 2004; Hendrickx et al., 2020; Hunt & Lucas, 1994; Thulborn, 1973, 1992). Recent work based on detailed character descriptions, morphometric analyses, or a combination of these approaches indicates that at least some diagnostic value can be extracted from dinosaur teeth (Barrett et al., 2014; Boyd, 2015; Hendrickx & Mateus, 2014; Hendrickx et al., 2015a; Hendrickx et al., 2019; Hendrickx et al., 2020; Larson & Currie, 2013; Ősi et al., 2016; Smith, 2002; Smith, 2005; Smith et al., 2005; Strickson et al., 2016). Nevertheless, as tooth morphology can vary ontogenetically, positionally (within the jaws of the same animal) and between individuals, as well as taxonomically (Coombs, 1990; Hendrickx et al., 2019), there is still disagreement regarding the most appropriate method for assigning isolated teeth to defensible, recognizable morphotaxa, which could then form a basis for further

investigation. Indeed, Hendrickx et al. (2015, 2020) have suggested that morphometric data alone are sub-optimal for classification and that far better results can be obtained using detailed descriptions of morphological characters and cladistic analyses based on a dentition-based data matrix.

Currie et al. (1990) and Farlow et al. (1991) were the first to apply a morphometric approach to isolated dinosaur teeth in a systematic fashion to aid taxonomic identification and examine the functional significance of different tooth crown morphologies. Smith (2005) and Smith et al. (2005), building on previous work (Baszio, 1997; Chandler, 1990; Currie et al., 1990; Farlow et al., 1991), provided a preliminary framework for the taxonomic identification of theropod dinosaur teeth by applying multivariate statistical methods to standard morphometric measurements. Following this work a generic approach applying principle component analysis (PCA) and linear discriminant analysis (LDA) has become the 'standard' quantitative methodology for the identification of isolated theropod teeth (e.g., Buckley et al., 2010; Fanti & Therrien, 2007; Gerke & Wings, 2016; Larson, 2008; Larson & Currie, 2013; Samman et al., 2005; Torices et al., 2014; Williamson & Brusatte, 2014; Young et al., 2019). Similar methodologies have been applied to ornithischian dinosaurs (Becerra et al., 2013) and isolated teeth from other extinct taxa, such as sharks (Marramà & Kriwet, 2017) and archosauriforms (Hoffman et al., 2019).

However, caution is warranted when applying this methodology. The use of PCA alone is not suitable to assess between-group differences and can mask differences when the group structure is embedded within variables exhibiting lower variances (MacLeod, 2018), or when group differences are assessed on a limited number of principle components by simply plotting PC1 against PC2. It is, however, useful as a dimensionality reduction transformation where there is a requirement to reduce the number of predictor variables while retaining the embedded information content, or as an investigative tool to explore data structures (Jolliffe, 2002; MacLeod, 2018). LDA is commonly used as either a follow-on classifier from PCA – by submitting the retained PCA eigenvectors to the LDA model – or as a classifier applied directly to the raw data. Most applications of LDA assume that the data under investigation meets the requirements of the technique, but do not always check that this is the case. This is

important, as LDA can be adversely affected by small or widely unequal group sizes, data outliers, unequal covariance matrices and non-Gaussian distributions, and the method works more effectively when the smallest group has significantly more cases than predictor variables. The effects of these caveats may be marginal in practice (Feldesman, 2002) but thus far these issues have not received detailed discussion in this context. If the data under consideration do violate these assumptions it calls into question the results obtained from such analyses, especially in the absence of verification by other methods (e.g., Corentin & Salvador, 2018; Fraser & Theodor, 2011; Hendrickx et al., 2015a; Hendrickx et al., 2019; Milla Carmona et al., 2016; Whitenack & Gottfried, 2010).

The algorithms employed in these analyses (Table 3.2) belong to a category of supervised classifiers known as 'eager-learners', where a model is generated from a set of training data before being applied to an 'unknown' dataset. The function of a supervised classifier is to build a model that then enables correct assignment of a future object described by predictor variables to a known class (Maugis et al., 2011; Rausch & Kelley, 2009). Eager-learners often take a long time to construct a model but can make predictions quickly. It is also possible to use some of these techniques, such as random forests, in unsupervised mode to assess and detect meaningful structures in a dataset and to classify objects to groups that are not known *a priori* (Afanador et al., 2016; Criminisi et al., 2012; Shi & Horvath, 2006). Although I have employed these techniques on fairly simple morphometric measurements, there is no reason why the techniques discussed below could not be employed on more complex morphological datasets such as 3D-shape data or digital images. Below I include a short introduction to the techniques I applied, including the use of ensemble model classifiers.

Linear models

Linear discriminant analysis. Linear discriminant analysis (LDA), a technique that identifies linear combinations of predictor variables to maximise the multivariate distance between groups (Fisher, 1936; Welch, 1939), is perhaps the most widely used method for classification. The functions are calculated in such a way that the first function captures as much of the group differences as possible, with subsequent functions each representing group differences not captured by previous functions. The combinations of predictors and prior probabilities are then used to calculate the

posterior probability distribution for each case. Group membership is assigned by selecting the group with the highest posterior probability for each case. For LDA to function appropriately two underlying assumptions regarding the data are made: that the data is multivariate normal; and that the group covariance matrices for the predictor variables are equal (Feldesman, 2002; Hastie et al., 2009a). LDA is also sensitive to highly-correlated predictors and is dependent on the ability to invert the covariance matrix, requiring more samples than predictors per group.

Logistic Regression. Logistic regression (LR), although commonly used to solve two-class problems, can be extended to a multi-class scenario and uses a linear predictor function to assess the likelihood of a particular class outcome. LR uses the log of the odds of being in one group compared to the others as the basis of its prediction. No assumptions are made regarding the distribution of the predictor variables entered into the model, nor does it assume equal covariance matrices and therefore no additional data pre-processing is required (Finch et al., 2014; Kuhn & Johnson, 2013a; Rausch & Kelley, 2009).

Non-linear models

Mixture Discriminant Analysis. Mixture discriminant analysis (MDA) is a non-linear extension of LDA whereby each class is modelled as a mixture of multiple multivariate normal distributions, i.e., each class can contain an unobserved number of sub-classes (Finch et al., 2014; Hastie & Tibshirani, 1996; Kuhn & Johnson, 2013a). Unlike LDA, there is no assumption of equal covariance matrices across groups for MDA. In a biological classification of taxa such sub-classes are particularly relevant, especially when classifying data to higher taxonomic levels. MDA has been applied with some success in other fields and often exhibits high predictive accuracy (Britzke et al., 2011; Finch et al., 2014; Rausch & Kelley, 2009).

Naïve Bayes. Naïve Bayes (NB) is a non-linear machine learning approach to group classification (Marsland, 2015; Russell & Norvig, 2009) that is known to work well with small group sample sizes (MacLeod, 2018). The model assumes that all the predictors are independent of each other which results in relatively quick computational times (Kuhn & Johnson, 2013a).

Decision trees

The final methodologies I explore are a departure from the standard linear or non-linear families of classification models. Both random forests and C5.0 are decision tree-based techniques that expand on the seminal work of Breiman et al. (1984), which introduced classification and regression trees.

Before exploring the detail of the two techniques it is useful to understand the basics of a decision tree. Decision trees are used in everyday life to make decisions based on a series of criteria. A simple example would be to decide on which train to catch to reach a certain destination at a preferred time without changing stations. In order to reach this decision I effectively run through a series of steps, each step is a question and the answer to the question dictates a path that the decision can follow. A suitable decision tree for such a choice is shown in Fig. 3.1. Every decision tree is a nested hierarchy of questions and answers (or if/then statements). For the example of catching a train to London Victoria station, the following hypothetical decision tree (one of many possible trees) might be followed:

If the final destination of the train is Brighton, then it is the wrong train

or

If the final destination of the train is London, and the station is London Bridge, then it is the wrong train

or

If the final destination of the train is London, and the station is Victoria, and it is a not direct train, then it is the wrong train

or

If then final destination of the train is London, and the station is Victoria, and it is a direct train, and it arrives between 08:00 and 08:15, then it is the correct train.

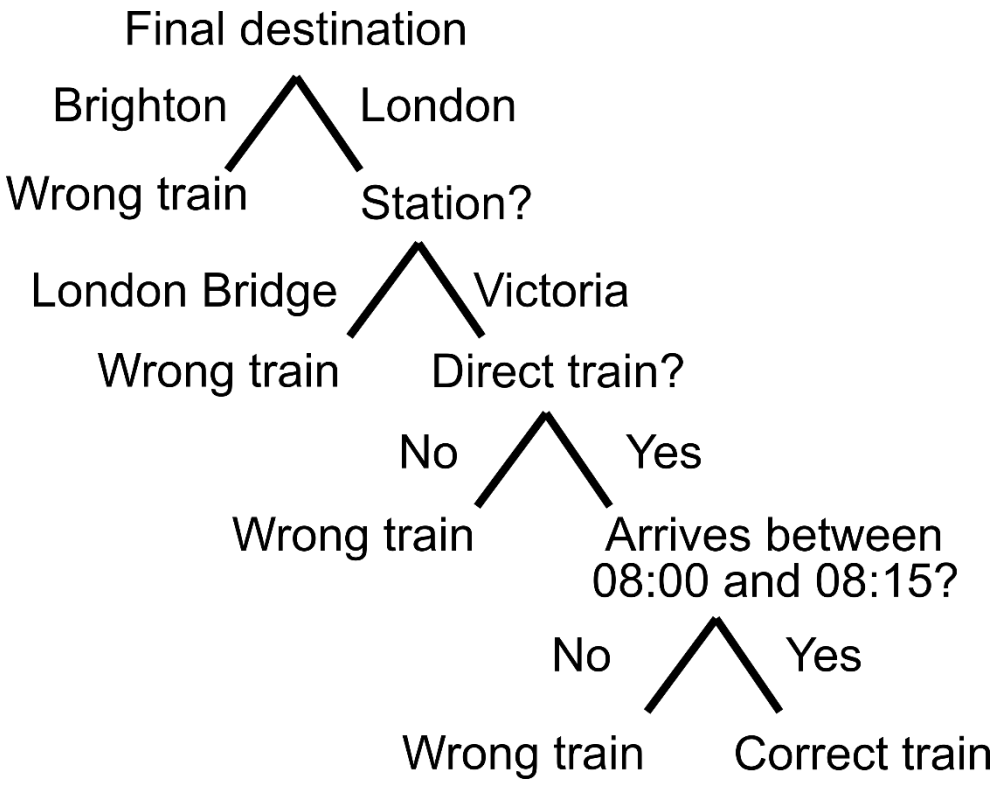


Fig. 3.1. Hypothetical categorical decision tree for the example of catching a train to London Victoria station.

A decision tree is essentially a flowchart of questions or rules that leads down a path to a prediction. Data is inputted into the root node of the tree. The decision tree algorithm then progressively divides the data into smaller and smaller groups based on the splitting criteria until the point at which the dataset can either be split no more or it reaches a rule that orders the splitting to finish. Decision trees can either be regression trees where the predicted outcome is a value (e.g., a house price) or classification trees where the predicted outcome is categorical (e.g., a taxon). The concepts of decision trees and random forests are similar. A decision tree is effectively built upon the entire dataset to produce one tree. A random forest combines many decision trees into a single model, where each of the trees in the model is generated on random subsets of observations and variables.

The major advantages of decision trees over techniques such as LDA or logistic regression are that: 1) they can accommodate missing data; 2) there is no need for the data to conform to a normal distribution, as the techniques are non-parametric; 3) outliers have little effect on the final classification as they will rarely define a splitting

node; 4) they can use both categorical and numerical data as predictor variables; and 5) transformed predictor variables (e.g., log transforms) have no effect on the tree structure (Feldesman, 2002). A drawback with decision tree methods is that of overfitting the data. This occurs when a tree is grown that perfectly predicts the classification pattern of the training data by defining terminal nodes (or leaves) that fit particular idiosyncrasies of the training process, i.e. that are relevant to that particular dataset only. This can have the effect of producing a classifier that can correctly predict > 99.9% of cases in the training data but cannot generalise to other observations. In these cases the model is picking up not only the patterns in the data, but also the 'noise' within the dataset. Tree-based methods are also prone to bias if some classes dominate the data and care needs to be taken to account for this prior to fitting.

Random Forests. Random forests (RF) is an ensemble learning method where a large number of uncorrelated decision trees are aggregated to form a final classification (Breiman, 2001; Breiman et al., 1984). This final classification is based on either an average of all the individual tree estimates (for regression trees) or a simple majority vote (for classification trees). The decision trees are built by randomly selecting predictors and observations to create individual trees. This random selection process increases the diversity in the forest and leads to a more robust prediction. Random predictors (i.e., variables) are used at each split in the tree which de-correlate the trees forming the forest. The number of predictors used is controlled by a parameter setting (m_{try}) which Kuhn and Johnson (2013a) and Breiman (2001) recommend setting to the square root of the number of predictors. RF classifications are sensitive to the number of trees used to build the forest with error rates reducing with increasing numbers of trees. Random forests tend to be stable and produce good predictive performance. However, they do have a number of disadvantages: even though some parameters are controllable, such as the number of trees or the number of predictors available at each split, the actual make up of each tree and therefore the forest is random and the forest itself (not the prediction) is less easy to interpret than a single decision tree; training a large number of trees can have higher computational overhead than a simple single decision tree.

C5.0. The C5.0 rule-based decision tree classifier is an updated version of the C4.5 model of Quinlan (1993) where the splitting criterion is based on information theory to choose the most informative variables for classifying the training set (Kuhn & Johnson, 2013a; Mehta & Shukla, 2015). As with decision trees in general, each sub-set resulting from the initial split is then re-split (usually on a different field) with the process repeating until no more splitting is possible. Each split can be either binary or multi-branched. C5.0 then tries to reduce the effects of overfitting by undertaking a pruning (winnowing) process on the lower level splits to remove those that do not contribute to the final model and produce simpler and more accurate trees. Unlike random forests the C5.0 tree is built by default on the entire dataset using all the variables and cases. The winnowing process attempts to uncover predictor variables that have a relationship to the desired model outcome with the final model only built using those variables. The C5.0 algorithm also allows for the implementation of adaptive boosting, which generates multiple classifiers rather than one with the final prediction resulting from majority voting across the classifiers. Unlike random forests, which creates multiple random trees, the C5.0 adaptive boosting trees are linked back to the classification errors generated from the first tree or ruleset. The first classifier will usually make mistakes on some groups. A second classifier is then generated that focusses on the misclassified data from the first tree in an attempt to improve the misclassification rate. Errors from the second tree are passed to a third and so on. The process continues for a user pre-defined number of iterations (trials). For a more detailed description of both C4.5 and C5.0 methods see Kuhn and Johnson (2013a).

Both random forests and C5.0 classifiers use either entropy or the Gini index for the node splitting criterion at each point in the tree (Afanador et al., 2016). Entropy is essentially a measure of purity in the tree node with a mixed class node being more impure than a node with a single class. The splitting decision will be based on calculating the entropy of both the parent node and all possible variations of child nodes. The child node that produces the biggest decrease in entropy between parent and child is the split that is chosen. The Gini index (or Gini impurity) is a calculation of the probability of a random instance being misclassified. The process is essentially the same as for an entropy split, with the Gini index of the parent node and all variations of child nodes

being calculated in order to choose a split which lowers the Gini index (Afanador et al., 2016; Shi & Horvath, 2006).

Ensemble models

Ensemble learning methods take a series of classifier models and combine the predictions to produce a final classification (Dietterich, 2001; Roli et al., 2001). A key to a good ensemble model is that the individual classifier techniques should be diverse to create a stronger overall prediction. There are a number of different methods to combine the results of the models making up the ensemble such as bagging, boosting and stacking (Dietterich 2001), here I use majority-voting and model stacking to arrive at the final classification. Majority voting simply takes the majority rule of the predictions from each classifier as the final classification result, for example if two classifiers predict a case to be 'class 1' and one classifier predicts the case as 'class 2' then the ensemble classification for that case is 'class 1'. Model stacking is where a single training dataset is run through multiple models. The predictions from these models are then used as the input to a second-level model from which the final classification is drawn.

Materials and Methods

Here I describe the datasets used for the analysis and the data preparation steps involved. I also discuss how I dealt with common issues found in multivariate datasets used for classification models, such as class balancing and missing data. In addition, I examine how the choice of prior probabilities and the resultant classification posterior probabilities affect the models. The methods and models developed herein are applied in Chapter Five to isolated theropod teeth from the Middle Jurassic of Britain (Wills, Underwood, et al. (2023) and Chapter Four), India (Prasad & Parmar, 2020), Kyrgyzstan (Averianov et al., 2005) and Madagascar (Maganuco et al., 2005).

I used two published datasets that include multiple linear measurements for isolated theropod teeth and that were used as the basis for prior morphometric analyses (Hendrickx et al., 2015; Larson et al., 2016). These each include a wide range of theropod taxa, with broad spatial and temporal distributions. Each specimen has five measured morphometric variables that are simple 2D linear distances (or representations thereof) between repeatable landmarks on the tooth crowns (Fig. 3.2): 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; crown height (CH), height of the crown measured from the tip of the tooth to the base of the enamel; number of denticles per millimetre along the midpoint of the anterior carina (ADM); and number of denticles per millimetre along the midpoint of the posterior carina (PDM) (Currie et al., 1990; Smith et al., 2005; Larson and Currie 2013). These datasets comprise human-selected and hand measured morphometric data rather than measurements derived from photographic or other digital sources of information (such as CT-data) that have also been used in machine learning classifications (e.g. Hoyal Cuthill et al., 2019). As such, it is inevitable that some degree of error will be introduced into the measurement process. However, given that the classification of isolated theropod teeth is a common requirement in vertebrate palaeontology, and the currently available datasets are all hand measured morphometric data, I feel there is value in applying such techniques to this data.

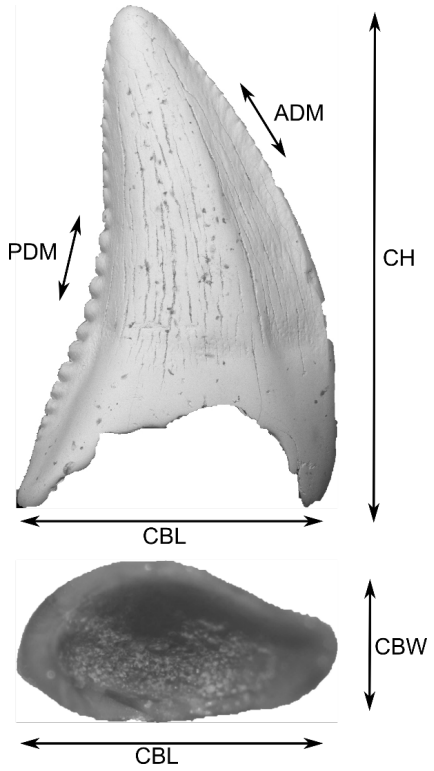


Fig. 3.2. Theropod tooth measurements used for machine learning analysis. ADM, anterior denticles per millimetre; CBL, crown base length; CBW, crown base width; CH, crown height; PDM, posterior denticles per millimetre.

The Hendrickx et al. (2015a) dataset consists of 995 individual cases belonging to 62 taxa from 19 major theropod clades (e.g., Megalosauridae, Tyrannosauridae, Dromaeosauridae, Abelisauridae) ranging in age from the Pliensbachian to the Maastrichtian with a global distribution. I analysed the data at two different taxonomic levels: a genus-level grouping of 680 cases and 32 classes and a higher-level clade aggregation comprising 886 cases and 14 classes. The dataset of Larson et al. (2016) comprises 3,104 maniraptoran theropod teeth from 18 lithostratigraphic units ranging in age from the uppermost Santonian (Milk River Formation) through to the Maastrichtian (Hell Creek Formation) of western North America. I analysed these data at two different taxonomic levels: a generic-level grouping containing 3020 cases and 17 classes; and a higher-level aggregation containing 3020 cases and four classes (Dromaeosauridae, Troodontidae, *Richardoestesia* and cf. Aves). I did not undertake a species level analysis due to the lack of species-level data with enough complete cases. See Appendix Two for a complete description of the data used in the models and Appendix Four for a list of institutional abbreviations.

Data preparation

Prior to analysis I undertook a series of data exploration and general preparation steps. Each published dataset reports individual specimens at different taxonomic levels. For example, Hendrickx et al. (2015a) list specimens at the generic level, whereas Larson et al. (2016) use species, with some of the latter split into stratigraphic units. To compare different models across both datasets, I aggregated groups of specimens to increasingly higher taxonomic levels. I removed any cases where it was unclear from the literature that a zero value in the data indicated a true zero (e.g., no anterior denticles) or represented missing data and, as some of the techniques applied require no missing data in the predictor variables, I removed all incomplete cases. Some classification techniques, such as LDA, are sensitive to the number of cases comprising individual groups in relation to the total number of predictor variables (Kuhn & Johnson, 2013a; Zavorka & Perrett, 2014) and require more cases per group than predictor variables. In addition, MacLeod (2018) noted that true group structures can be masked when the number of variables is greater than the number of cases. This is caused by having insufficient numbers of data points per group to describe the group structures correctly. At each taxonomic level tested, I therefore removed entire groups where the total

number of group members was less than or equal to the number of predictor variables. As no dataset exhibited a multivariate normal distribution, the predictor variables were log-transformed with a constant value of one added prior to transformation to allow the log of true zero values.

For each taxonomic level tested I split the data into training and testing samples with a 80:20 ratio using the R package Caret (Kuhn, 2008) which attempts to balance the class distributions within the training and testing sets. To optimise the models I undertook k-fold cross validation on the training set. Cross validation reduces the problems of underfitting, not capturing enough information in the model to accurately predict new data, and overfitting where the model performs well on the training set but does not generalise enough to perform well on new data (Hastie et al., 2009a) K-fold cross validation randomly divides the original data into k equally-sized subsamples. In this case I used a k-value of 10, so that the original training dataset is randomly divided into 10 subsamples. Nine subsamples are used as the training set and one as the testing set. This is then repeated 10 times such that each case forms part of a training set k-1 times and a testing set once. The model effectiveness is then averaged over each repeat to give a single overall model accuracy. I additionally ran the subsequent models on the retained testing samples, i.e., the samples not used to create the classification model, to provide more accurate assessments of the predictive accuracy of each model on unknown data.

Some of the models require specific parameters or preparation: for Naïve Bayes, in order to compensate for the non-independence of variables in the test data, I used PCA scores as input into the model rather than the original data; for random forests, the models used 2000 trees (to ensure model stability) and a range of m_{try} values from two to five; for C5.0, I ran models both with and without winnowing and set the model to stop the boosting process at 100 trials. I also generated a classifier ruleset for each model comprising simple if-then rules for the predicted class based on the input predictor variables.

All analyses were performed using R version 3.6.0 (R Core Team, 2019) in R Studio (RStudio Team, 2020) with the Caret package (Kuhn 2008) used for model generation. The following R packages were used for specific models or processes: UBL for synthetic

data generation (Branco et al., 2016); missForest to introduce random missing data (Stekhoven, 2013; Stekhoven & Buehlmann, 2012); mice for data imputation (van Buuren & Groothuis-Oudshoorn, 2011); MASS, C5.0 and randomForest for specific classification models (Kuhn et al., 2018; Liaw & Wiener, 2002; Venables & Ripley, 2002); and ggplot, gridextra, cowplot and ggalluvial for plotting functions (Brunson, 2019; Wickham, 2016; Wilke, 2019).

Data balancing

A common issue with published datasets on tooth linear measurements is the unequal distribution of group members between distinct groups within the dataset. For example, the Larson et al. (2016) dataset contains 3020 specimens broken down into 17 generic groups. The distribution of group membership within these data ranges from 1176 individual cases to only six cases. As previously noted, groups defined by small numbers of cases suffer from the inability for the cases to correctly define the group structure. This imbalance also causes the performance of machine learning classifiers to be degraded as there is a bias towards the majority classes in an attempt to reduce the overall classification error. There are various methods that can be used to balance a dataset, all of which involve either the addition or removal of data points. Undersampling works on the majority classes, reducing the number of cases in each class in turn to create a more balanced dataset. This has the negative effect of removing informative data about these classes. Oversampling works on the minority classes by increasing the number of observations by replication. Whilst this does not result in information loss the implicit assumption is that the minority class structures are adequate to define those classes. I employed a methodology, Synthetic minority oversampling technique (SMOTE), that shifts the learning bias towards minority classes by generating synthetic data in these classes (Chawla et al., 2002). SMOTE oversamples the minority classes by creating new data points in feature-space randomly along a line joining an existing point to its nearest neighbours. I tested two scenarios to balance the training dataset to see if this resulted in a more accurate classification overall. First, random undersampling (i.e., removal) of the most populated classes combined with oversampling (by synthetic data generation) of the least populated classes to create a new dataset containing approximately the same number of overall cases as the original. Second, oversampling of the least populated classes to create an enlarged dataset with

no undersampling of the most populated classes. I created these synthetic datasets based on the Larson et al. (2016) data at two different taxonomic levels running a number of different classifier models across the synthetic data to compare results to the original.

Dealing with missing data

Fossil datasets commonly contain incomplete morphometric information due to the nature of their preservation. Parts of a specimen may be missing due to breakage or wear, distortion as a result of geological processes may result in a measurement being suspect and therefore excluded, and the presence of host matrix can obscure particular features. The problem of missing data can be overcome either by deleting cases with missing values, using a variety of techniques to predict missing values based on the overall dataset, or by using a technique that is not reliant on complete cases. The first two techniques are problematic: deleting cases can remove useful information from the dataset, and replacing values with either mean substitution or values imputed from multiple regression has a tendency to distort the dataset and therefore the resultant classification (Feldesman, 2002; Schafer, 1997). Here I test different scenarios using the C5.0 tree-based classifier, which is not reliant on complete data. To look at the effects of missing data I used the Larson et al. (2016) dataset, which was edited to contain only complete cases. I then generated five new training datasets (Fig. 3.3) from this where I introduced increasing proportions of randomly generated missing data into the predictor variables (at 5, 10, 20, 30 and 50% levels) using the missForest package (Stekhoven and Buehlmann 2012; Stekhoven 2013). C5.0 classification models were then built for each of these new training datasets and applied to the retained testing data each time, allowing us to model changes in classification accuracy as the amount of missing data in the training set varied. I examined the effect of predicting missing values for each of the new training sets where I had previously introduced missing data using the MICE package (van Buuren and Groothuis-Oudshoorn 2011). For each training set containing missing data I created five imputed data sets that differ only in imputed missing values. I then built C5.0 classification models for each of these imputed datasets and stacked the results together to generate a training set containing imputed data. The imputed training set was fed into a secondary C5.0 model to provide the final

classification (Fig. 3.3). Finally, I generated a C5.0 model using the original, complete Larson et al. (2016) dataset where I retained cases with missing data.

Prior and posterior probabilities

Bayesian classifiers use a prior probability distribution of group membership to calculate the posterior probability distribution, i.e., the resultant classification. The prior is essentially the probability that an observation comes from a particular group. There are three ways of defining prior probabilities: the prior probabilities are equal for all the groups, such that there is an equal chance that an observation can come from any group; the probabilities of group membership are proportional to the training dataset group observations; or the true group distribution is known (irrespective of the current dataset) and the priors can be defined explicitly to match this. The choice of prior will affect the outcome of classifications, especially when some group populations may be rare due to either unequal sampling or are a true reflection of the population under study (Zavorka and Perrett 2014). I modelled the effects of defining both equal and proportional priors on the final classification result.

Understanding how the final class assignment is made by a classifier is also important before any value can be attached to the result. Classifiers base their decisions on final class values on the calculated posterior probability for each class on a case-by-case basis. Classifiers that use ensemble techniques to arrive at a final result will still use posterior probability to assign classes within each of the models before creating the ensemble. The class assigned to a particular case is simply the class with the highest posterior probability. In some cases the results are fairly unequivocal, but in others a degree of caution is required. Take a simple example of a three class problem and two cases. Case one reports posterior probabilities of: Class A = 0.8, Class B= 0.1 and Class C = 0.1. Case two reports posterior probabilities of: Class A = 0.34, Class B = 0.33 and Class C = 0.33. Both cases are assigned to class A on the basis of the highest posterior probability, but it is clear from the results that the strength of the classification in case two is weak. Here I look at how the posterior probability varies on a case-by-case basis for a classification derived from an MDA model (see Results, below).

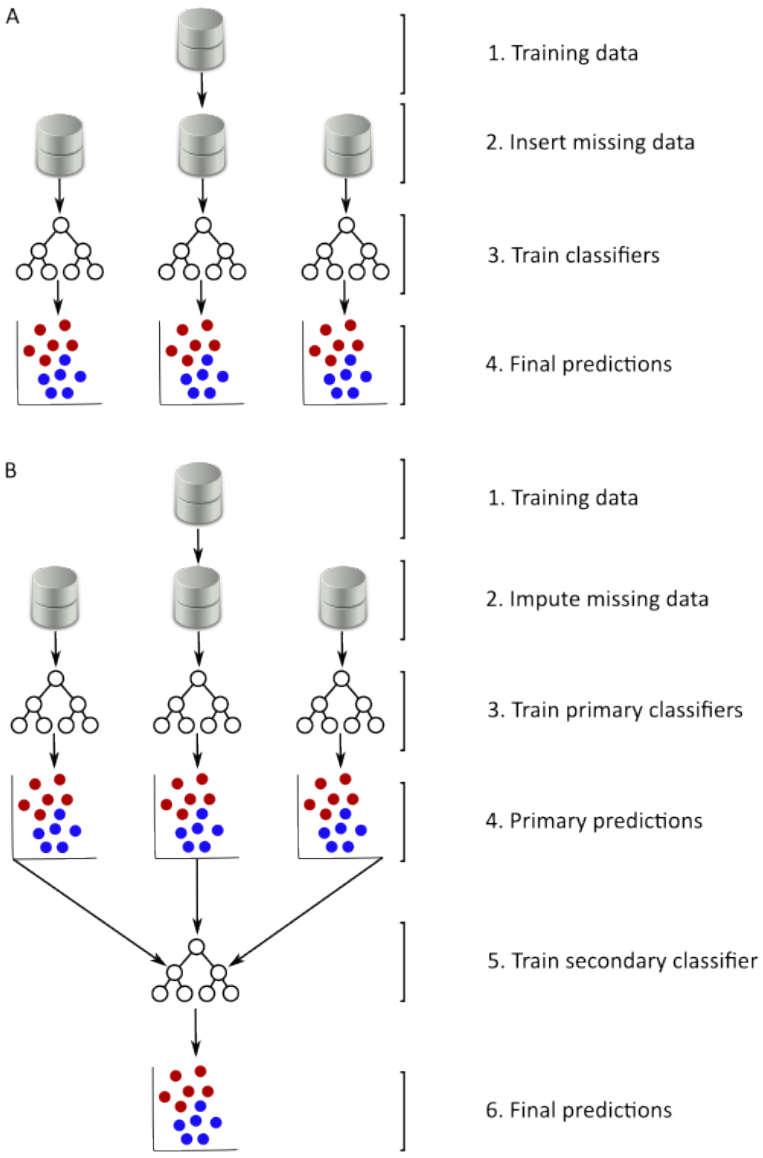


Fig. 3.3. Workflow for looking at the effect of missing data on predictive accuracy. A, Generating new datasets with missing data inserted at random. For this exercise I added missing data into the predictor variables at 5, 10, 20, 30 and 50% levels. A1, original training data with no missing data; A2, New data sets with missing data inserted by random replacement of original training data; A3, C5.0 classifiers trained on each new data set; A4, Final predictions, each classifier produces an individual prediction. B, Replacing missing data with substituted values. For the sake of clarity I have only shown the workflow for one of the training datasets containing missing data. This dataset was derived from workflow A. B1, Training data, data set with randomly inserted missing data; B2, New data sets with missing data replaced by imputed data. Here we imputed five times to create five new data sets from the original training data; B3, Primary classification, C5.0 classifiers trained on each imputed data set; B4, Primary predictions. Each classifier produces a prediction based on its imputed training data set; B5, Aggregation stage. All the primary model predictions are combined. A secondary C5.0 classifier is trained using the aggregated data as input; B6, Final predictions.

Ensemble classifier

For the ensemble classifier I combined the logistic regression, MDA and RF models as these employ differing techniques, with MDA and RF generally achieving the highest individual model accuracy (see Results, below). I used majority-voting and model stacking to combine the individual classification results and generate the final classification.

Results

Comparison of classification models

Table 3.2 shows the overall accuracies of the models as applied to both the Hendrickx et al. (2015) and Larson et al. (2016) datasets. The top performers in each case are the non-linear MDA model and the decision tree based random forests and C5.0 models. Linear models (LDA and LR) perform poorly across both datasets as does the non-linear naïve Bayes model. Overall classification accuracy, irrespective of the model employed, increases as the number of classes decreases (Table 3.2). This increase in accuracy is as a result of true group structures being correctly described by having sufficient numbers of datapoints per group.

Using the Hendrickx et al. (2015) dataset I ran the classifiers at two taxonomic levels, the first a genus level with 32 classes and 680 cases and the second at a higher (family) taxonomic level with 14 classes and 886 cases. The 32-class model accuracies range from 59.2% (naïve Bayes) to 77.4% (MDA) accuracy. The 14-class models show an overall increase in classification accuracy with accuracies for the two highest performing models (random forests and C5.0) at around 80%. Compared to the equivalent Hendrickx et al. (2015) classification using LDA, the tree based models increase the overall accuracy of the prediction by around 10% from 70.9% as reported by Hendrickx et al. (2015) to 80.4% from the RF classifier. The LDA model based on these data similarly boosts the overall accuracy to 76.7%. Fig. 3.4 depicts the normalised confusion matrix for the Hendrickx et al. (2015) 14-class dataset from LDA, MDA, RF and C5.0 classifiers showing per-clade accuracies for each model.

Table 3.2. Classification results for different models using the Hendrickx et al. (2015) and Larson et al. (2016) datasets.

	Hendrickx, et al. (2015)				Larson, et al. (2016)			
	680 cases, 32 classes		886 cases, 14 classes		3020 cases, 17 classes		3020 cases, 4 classes	
	Accuracy							
	Model	Testing data	Model	Testing data	Model	Testing data	Model	Testing data
LDA	0.645	0.690	0.752	0.767	0.705	0.697	0.958	0.942
LR	0.687	0.730	0.753	0.759	0.721	0.726	0.962	0.951
MDA	0.745	0.774	0.803	0.796	0.732	0.734	0.965	0.963
NB	0.647	0.592	0.755	0.750	0.698	0.697	0.930	0.933
RF	0.742	0.758	0.786	0.804	0.748	0.750	0.965	0.962
C5.0	0.710	0.749	0.775	0.802	0.741	0.746	0.962	0.957

Accuracies are shown for both the classification model and the testing data. LDA, linear discriminant analysis; LR, logistic regression; MDA, mixture discriminant analysis; NB, naïve Bayes; RF, random forests; C5.0, rule-based decision tree.

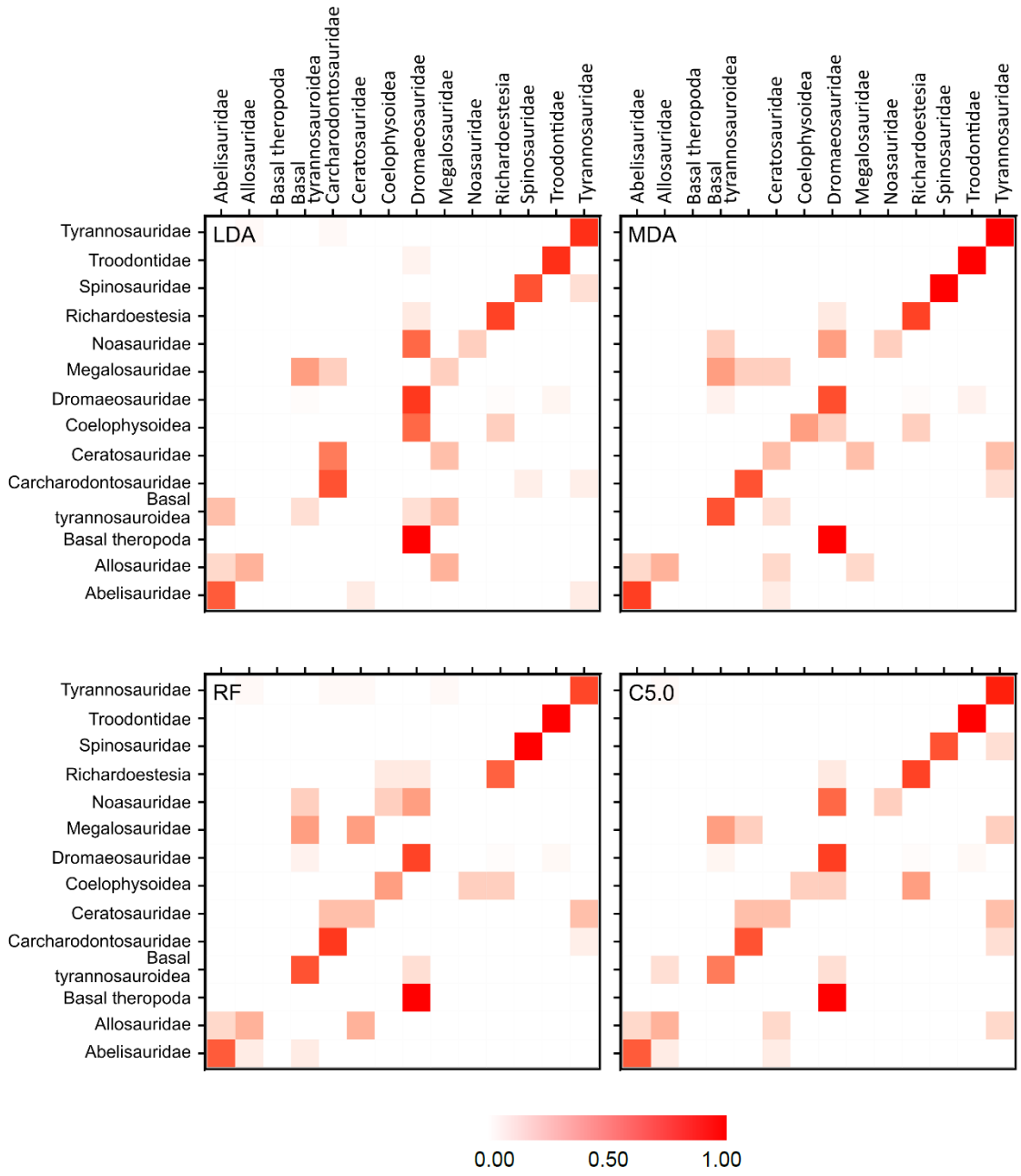


Fig. 3.4. Normalized confusion matrices for LDA, MDA, RF and C5.0 classification models based on the Hendrickx et al. (2015) 14-class dataset. Reference classes are plotted on the x-axis and predicted classes on the y-axis. Red scale refers to the percentage classified for each taxon-pair with 1 equalling 100%.

Two dimensional scatterplots of canonical variates obtained from MDA for the Hendrickx et al. (2015) dataset are shown in Fig. 3.5A, which visually depict the group separations in discriminant space. The random forest classifier (Fig. 3.5B) demonstrates the decrease in error rates both overall and for most individual clades as the number of trees in the model increases. I ran all random forest models with 2,000 trees: however, the results indicate that little improvement in model performance is reached after 1,000 trees. The models used three randomly selected predictors (m_{try} value) for the 32-class dataset and two for the 14-class dataset. Fig. 3.6 depicts the overall C5.0 model accuracies for the Hendrickx et al. (2015) dataset at a range of boosting iterations and using both winnowing and no winnowing. Across both taxonomic levels tested the overall accuracy settles down at around 25–30 boosting iterations. For the 32-class dataset the rules-based model using no winnowing improves the predictive accuracy slightly, for the 14-class dataset the rules-based model again shows a slight improvement in predictive accuracy irrespective of the use of winnowing.

Results from analysis of the Larson et al. (2016) dataset, again at two different taxonomic levels, broadly reinforce the previous analysis (Table 3.2). Decision trees and MDA return the highest classification accuracies with LDA performing relatively poorly. The difference between accuracies narrows as the number of groups in the data decreases and the numbers of cases making up each class increases. Accuracy for the 17-class dataset models ranges from 69.7% (LDA and NB) to 75% (RF) when applied to the testing data, with the 4-class dataset accuracies ranging from 93.3% (NB) to 96.3% (MDA). As with the previous dataset, the accuracy of classification increases as data is aggregated to higher and higher taxonomic levels. This increase in accuracy is reflective of the increasing certainty of the taxonomy, an increase in the number of cases making up the training groups and the removal of misclassification errors between closely related clades such as *Richardoestesia gilmorei* and *R. isosceles*, which have a tendency to classify to each other.

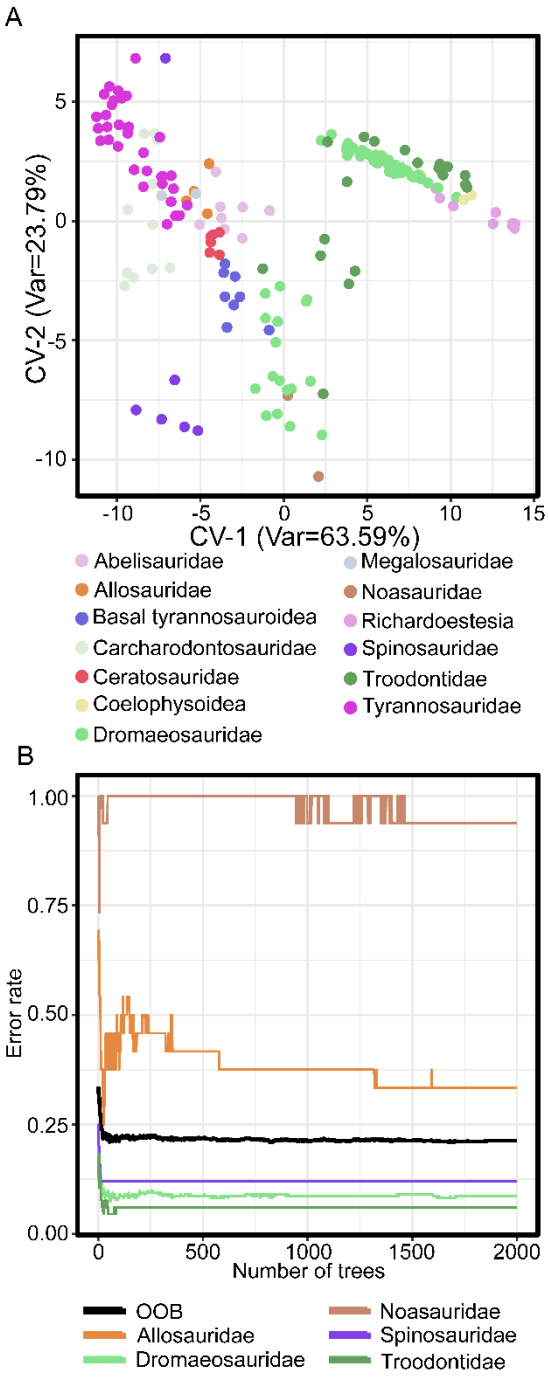


Fig. 3.5. Hendrickx et al. (2015) 14-class dataset machine learning analysis. A, MDA canonical variates showing group separations in discriminant space. B, random forest error rate per taxon and overall (OOB) classification error rate; for the sake of clarity only five taxa are shown.

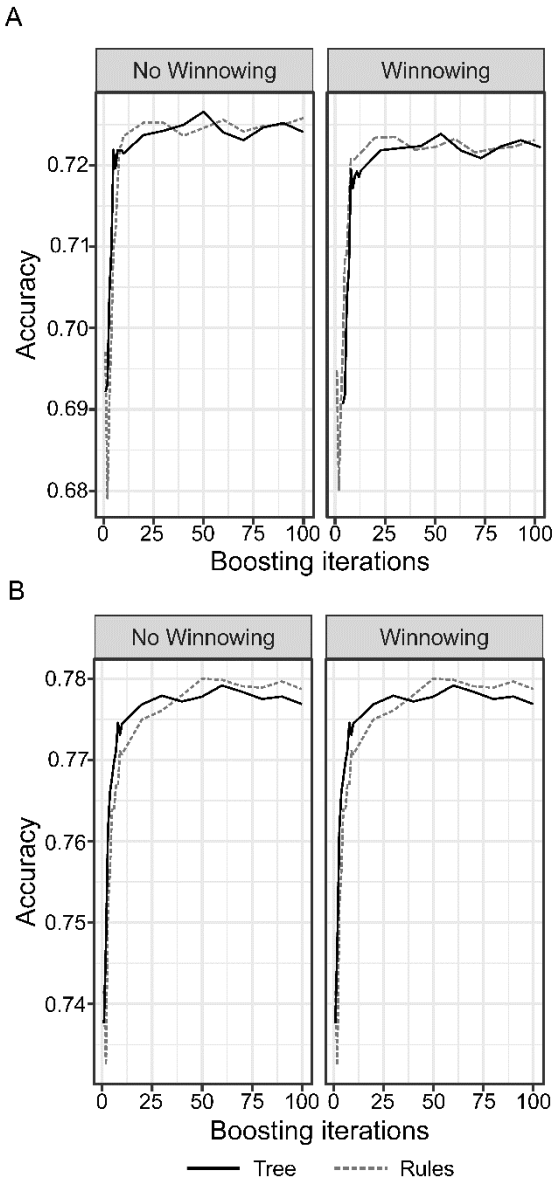


Fig. 3.6. C5.0 accuracy plots for Hendrickx et al. (2015) data showing the effects of winnowing predictor variables and the rules vs tree based models at different boosting iterations. A, results for the 32-class model. B, results for the 14-class model.

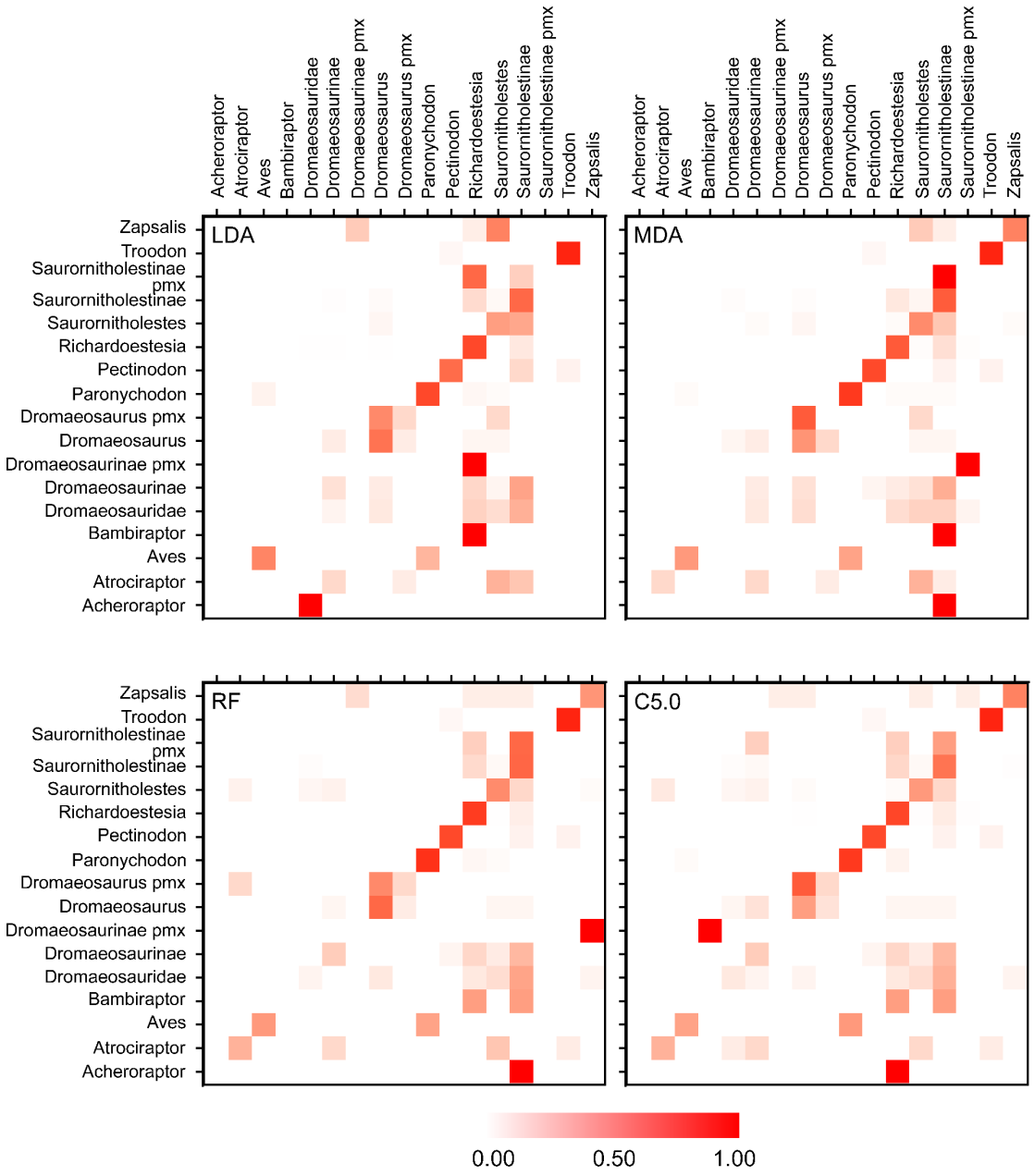


Fig. 3.7. Normalized confusion matrices for Linear Discriminant Analysis, Mixture Discriminant Analysis, Random Forest and C5.0 classification models based on the Larson et al. (2016) 17-class dataset. Reference classes are plotted on the x-axis and predicted classes on the y-axis. Abbreviation: pmx, pre-maxillary tooth. Red scale refers to the percentage classified for each taxon-pair with 1 equalling 100%.

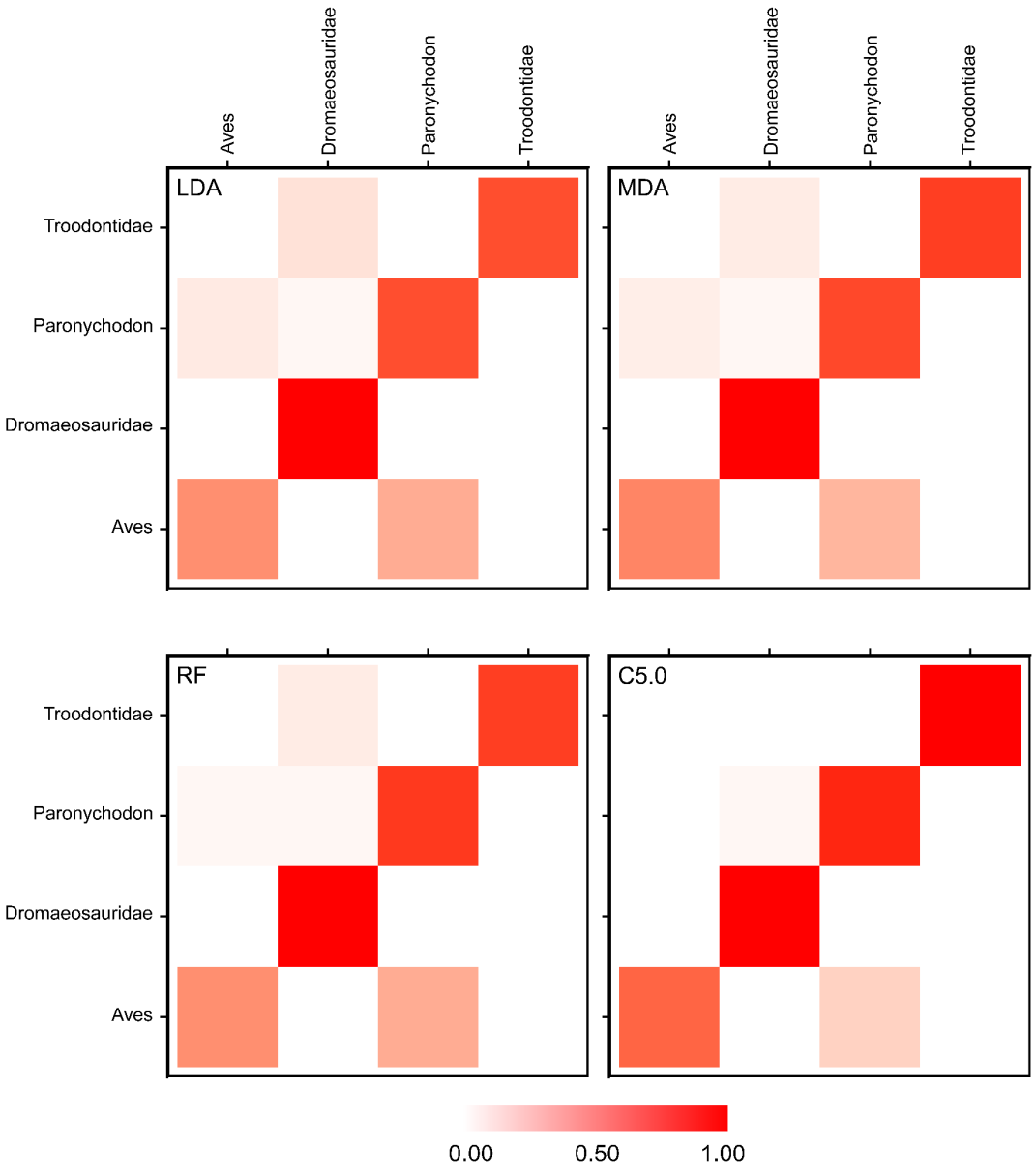


Fig. 3.8. Normalized confusion matrices for the Linear Discriminant Analysis, Mixture Discriminant Analysis, Random Forest and C5.0 classification models based on the Larson et al. (2016) 4-class dataset. Reference classes are plotted on the x-axis and predicted classes on the y-axis. Red scale refers to the percentage classified for each taxon-pair with 1 equalling 100%.

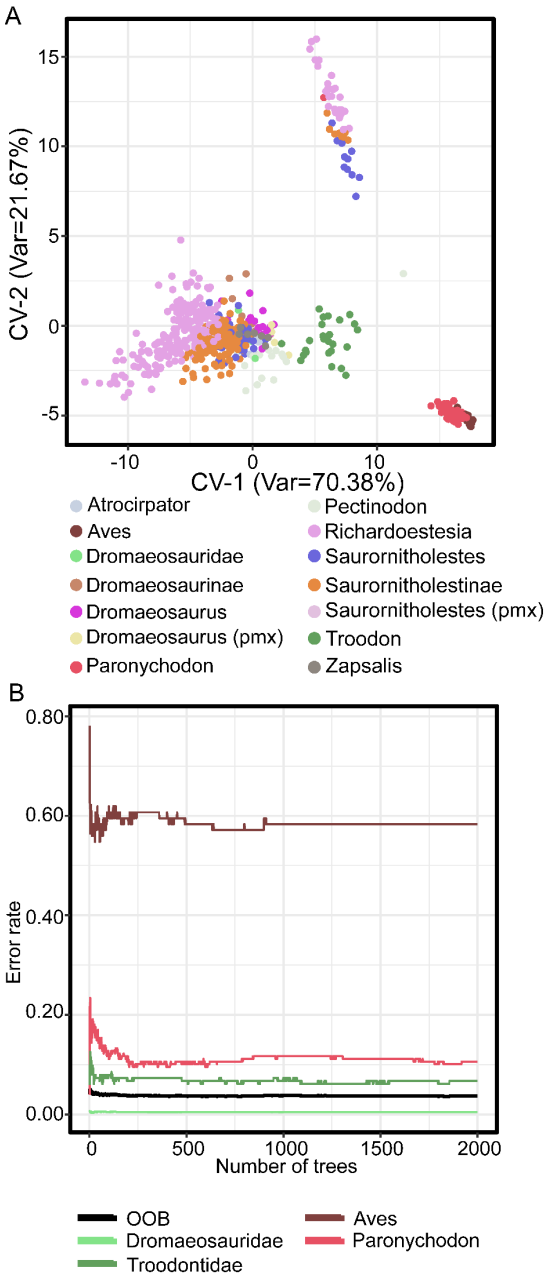


Fig. 3.9. Machine learning visualisations for Larson et al. (2016) data showing A, group separations in MDA discriminant space. B, random forest error rate for Larson et al. (2016) 4-class model. Abbreviation: pmx, pre-maxillary tooth.

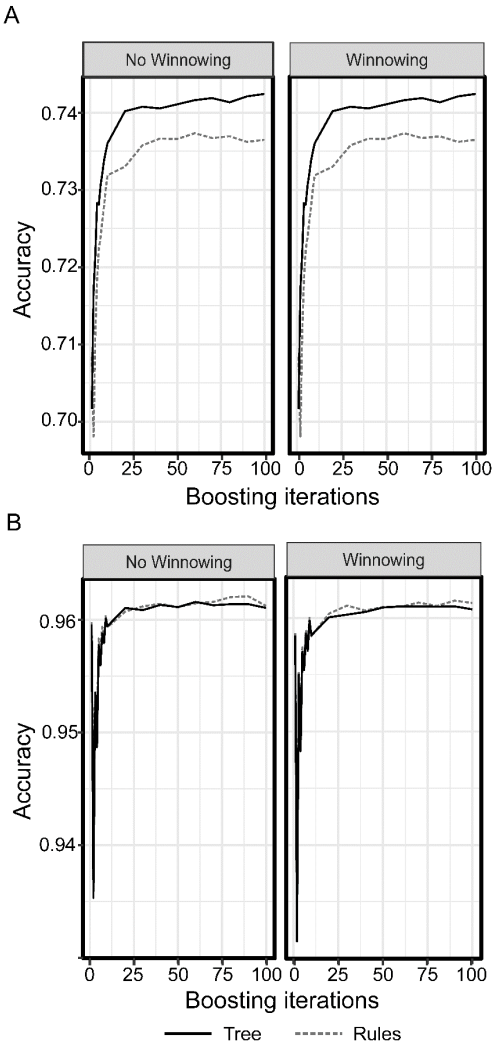


Fig. 3.10. C5.0 models for Larson et al. (2016) data showing the effects of winnowing predictor variables and the rules vs tree based models at different boosting iterations. A, 17-class model. B, 4-class model.

Figures 3.7 and 3.8 depict the normalised confusion matrices for the 17- and 4-class Larson et al. (2016) datasets from the LDA, MDA, RF and C5.0 classifiers. Group separations in discriminant space obtained from the MDA classifier are shown in Fig. 3.9A, the first two canonical variates are plotted that together account for around 93% of the total variation in each case. The random forest classifiers (Fig. 3.9B) again demonstrate the decrease in error rates as the number of trees in the model increases.

The 4-class model overall accuracy and the accuracy of Troodontidae and Dromaeosauridae show little change after 250 trees but Aves and *Paronychodon* are unstable to around 1,000 trees. The 17-class model is noisier but again settles down at around 1,000 trees. Fig. 3.10 depicts the overall C5.0 model accuracies and Fig. 3.11 visualises one of the decision trees for the 4-class model. Across both taxonomic levels tested the tree-based model outperforms the rules based model although the difference between the two is minimal especially at the 4-class level. Winnowing of the predictor variables has a negative impact on the accuracy at 17 classes but little if any effect at the 4-class level. Boosting iterations settle at around 25 for the 4-class model and 50 for the 17-class model.

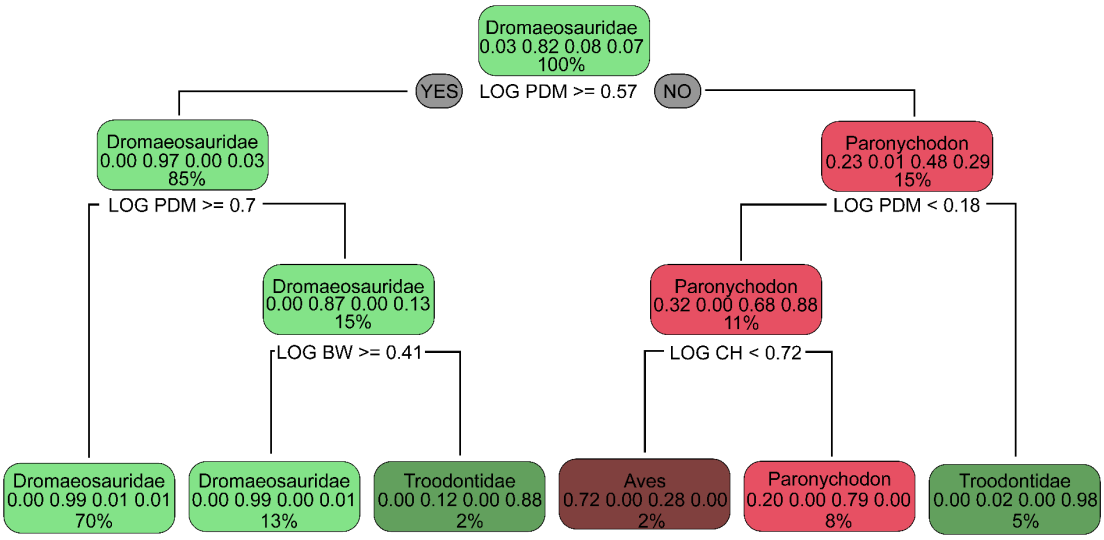


Fig. 3.11. Extract from the 4-class decision tree classifier (C5.0) using the Larson et al. (2016) data. Each node shows: the predicted class; the predicted probability of each class; the percentage of observations in each node.

Data balancing

Fig. 3.12A and Table 3.3 depict the changes in classification accuracy for LDA, MDA, RF and C5.0 models as I generated synthetic data in an attempt to balance the number of cases per class. The results show that attempting to balance class membership by either a combination of undersampling and oversampling (balanced results) or by oversampling alone produces significantly worse accuracy than no balancing.

Table 3.3. Classification accuracy results for synthetic data generation (SMOTE) compared to unbalanced data for LDA, MDA, RF and C5.0 classifiers.

	17-class			4-class		
	Accuracy					
	Balanced	Oversampled	None	Balanced	Oversampled	None
LDA	0.588	0.614	0.697	0.925	0.934	0.942
MDA	0.599	0.624	0.734	0.930	0.958	0.963
RF	0.654	0.681	0.750	0.942	0.960	0.962
C5.0	0.621	0.686	0.746	0.952	0.963	0.957

Accuracy based on Larson et al. (2016) data.
LDA, linear discriminant analysis; MDA, mixture discriminant analysis; RF, random forests; C5.0, rule-based decision tree.

Missing data

Table 3.4 summarises the results of introducing missing data at various percentage levels into the Larson et al. (2016) dataset and then using imputation to replace missing values. The classification accuracies decrease as the amount of missing data increases, with the 17-class model accuracy dropping off at a sharper rate than the 4-class model. The results indicate that the C5.0 classifier copes reasonably well with up to 20% missing data in some scenarios (Fig. 3.12B). The 4-class model accuracy decreases from 96.2% with no missing data to 93.9% with 20% missing data. Data imputation has a positive effect on the classification accuracies in the 4-class scenario with imputation at the 5% level slightly outperforming the original (no missing data) classifier. In the case of the 17-class models imputation has little effect on the classification accuracy with most imputed models showing a slightly lower accuracy rate than the models developed with missing data.

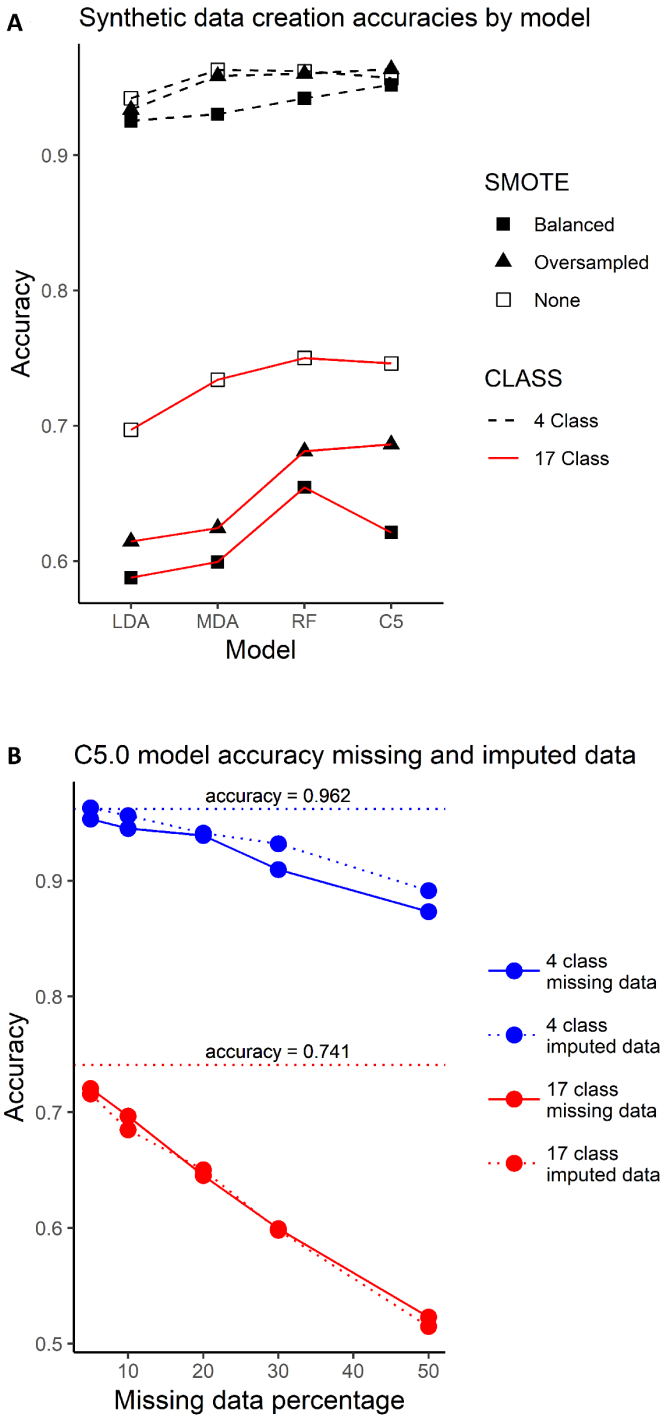


Fig. 3.12. The effects of synthetic data creation and missing data on machine learning model accuracies. A, C5.0 classifier accuracy for synthetically generated class balanced datasets B, C5.0 classifier accuracies for missing and imputed data at different levels. Horizontal dotted lines show the C5.0 model accuracy with no missing or imputed data.

Table 3.4. C5.0 decision tree classifier results on missing and imputed data for Larson et al. (2016) dataset.

	17-class		4-class	
	Accuracy			
	Missing data	Imputed data	Missing data	Imputed data
Percentage data missing / imputed				
0	0.741	0.741	0.962	0.962
5	0.721	0.716	0.953	0.963
10	0.696	0.685	0.945	0.956
20	0.645	0.650	0.939	0.941
30	0.599	0.598	0.909	0.932
50	0.523	0.515	0.873	0.891

Prior and posterior probabilities

The effects of changing prior probabilities are summarised in Table 3.5 for LDA and MDA classifiers. Equal prior probabilities have the effect of increasing the bias towards smaller and potentially unstable groups reducing the overall accuracy of the model when compared to proportional priors. This is seen most markedly for the MDA classifier. Posterior probabilities from the MDA classifier for 10 cases of the Larson et al. (2016) dataset are shown in Table 3.6. For most of the cases the classifier results in unambiguous predicted classes such as for cases 2–4 where the probability of the case classifying to Dromaeosauridae is 1.0. In other cases there is a degree of ambiguity as to the final class prediction. This is demonstrated by cases 1 and 8 where the final class prediction is only weakly supported (probabilities of 0.57 and 0.55, respectively). Fig. 3.13 shows the posterior probability mapping for the Larson et al. (2016) 17-class dataset. It is apparent from the overall map that clades such as *Richardoestesia* and *Troodon* have well supported final class prediction compared to *Acheroraptor* and *Bambiraptor*.

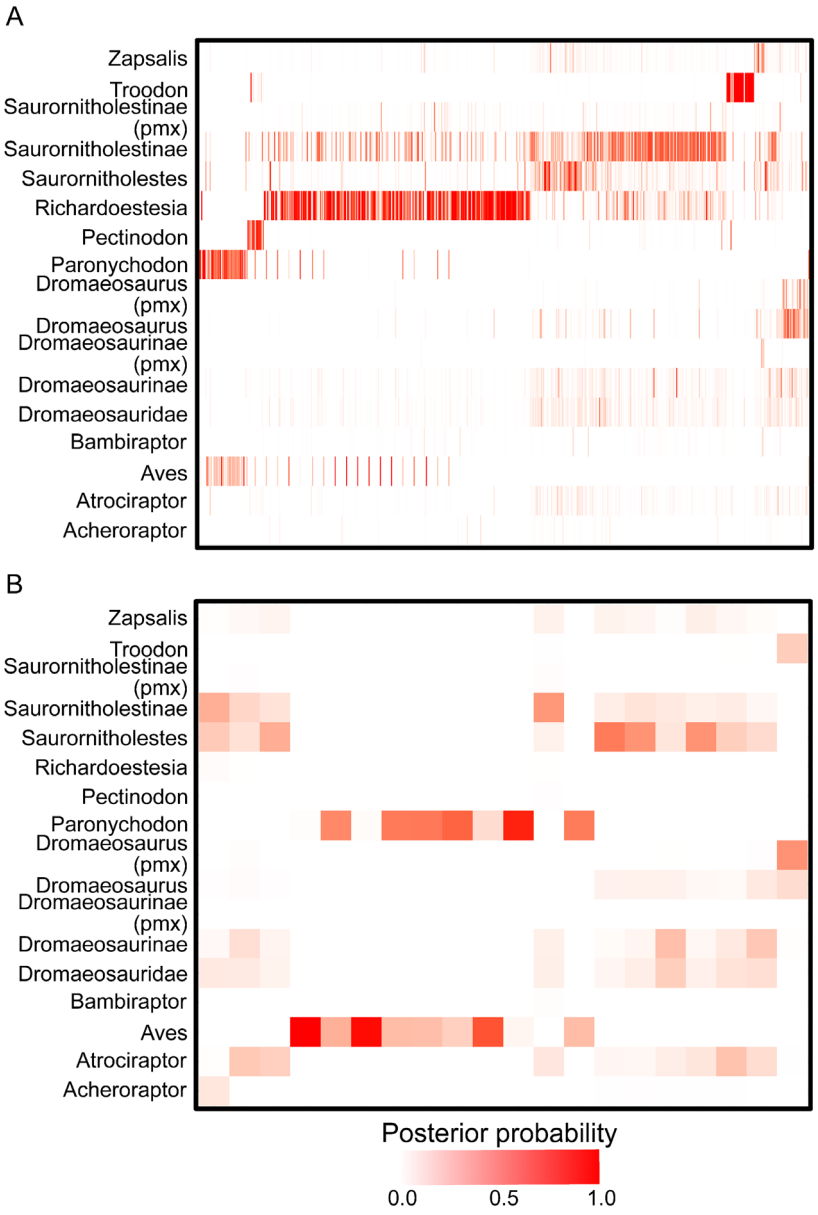


Fig. 3.13. Classification posterior probability heatmap for the Mixture Discriminant Analysis model using the Larson et al. (2016) 17-class dataset. A, entire test dataset. B, first 30 cases. Each block on the x-axis represents one case. Abbreviation: pmx, premaxillary tooth.

Table 3.5. Effect of different prior probabilities on linear discriminant (LDA) and mixture discriminant (MDA) model accuracy.

	Hendrickx, et al. (2015) 14-class model		Larson, et al. (2016) 17-class model	
	Accuracy			
Model	equal priors	proportional priors	equal priors	proportional priors
LDA	0.767	0.774	0.697	0.708
MDA	0.796	0.841	0.734	0.746

Table 3.6. Posterior probabilities for 10 cases selected at random from the MDA classifier using the Larson et al. (2016) 4-class dataset.

Taxon / Case	1	2	3	4	5	6	7	8	9	10
Aves	0.57	0.00	0.00	0.00	0.98	0.02	0.00	0.00	0.00	0.00
Dromaeosaruidae	0.00	1.00	1.00	1.00	0.00	0.00	1.00	0.55	1.00	0.02
Paronychodon	0.42	0.00	0.00	0.00	0.02	0.98	0.00	0.00	0.00	0.00
Troodontidae	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.45	0.00	0.98

Ensemble classifiers

Table 3.7 summarises the accuracy achieved by stacking three different models to create an ensemble classifier and the accuracy of the majority vote ensemble. The stacking ensemble increases the overall classification accuracy in all cases with the exception of the Hendrickx et al. (2015) 32-class model. The increase in accuracy ranges from 0.5% for the Hendrickx et al. (2015) 14-class model to 1.1% for the Larson et al. (2016) 4-class data. The majority voting ensemble increases the overall model accuracy for the Larson et al. (2016) 4-class data to 97.5% (a similar level to the stacked ensemble) but is less successful for the other data analysed with either the individual classifiers or the stacked ensemble outperforming. Fig. 3.14 shows how the classification of the Larson et al. (2016) dataset changes as a result of using different classifiers (LR, MDA, RF) and a majority vote ensemble classification based on all three individual classifiers. Clades

such as *Pectinodon*, *Zapsalis*, *Paronychodon* and Aves have a relatively consistent classification outcome across all classifiers. This contrasts with many of the other dromaeosaurids which cross-classify depending on the chosen classification algorithm. Fig. 3.14 also depicts an 'unknown' group in the final majority voting ensemble. This is where none of the constituent classifiers agreed on a final class and is an indication that there may be a sub-group present in the data that was incorrectly assigned a class in the training data.

Table 3.7. Machine learning ensemble model accuracy using model stacking of three different models.

	Hendrickx, et al. (2015)		Larson, et al. (2016)	
	Accuracy			
	32 class	14 class	17 class	4 class
LR	0.685	0.733	0.731	0.958
MDA	0.749	0.785	0.733	0.963
RF	0.745	0.791	0.751	0.960
Ensemble stack	0.620	0.796	0.759	0.974
Majority vote	0.743	0.779	0.737	0.975

Accuracies are shown for both the individual models that make up the ensembles and the stacked and majority vote ensembles.
LDA, linear discriminant analysis; MDA, mixture discriminant analysis; RF, random forests; C5.0, rule-based decision tree.

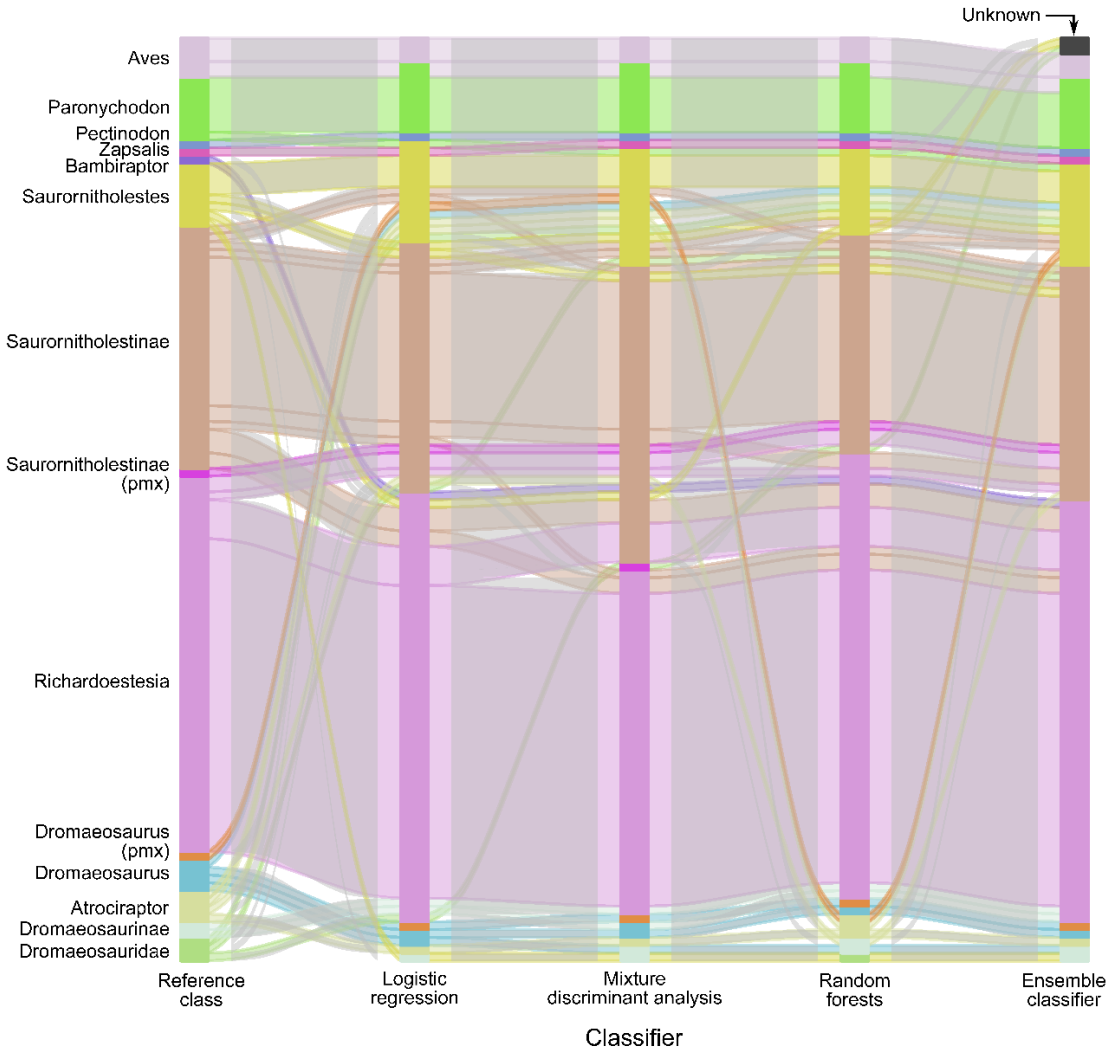


Fig. 3.14. Ensemble classifier showing classification changes at the clade level using Logistic Regression, Mixture Discriminant Analysis and Random Forest classifiers combined using a majority vote rule to form an ensemble classifier for the Larson et al. (2016) 17-class data. Vertical bars represent the clade predictions for each classifier, flows between the bars represent changes in prediction between the different classifiers. The ensemble classifier has an additional 'unknown' class where none of the individual classifiers were in agreement with a prediction. Abbreviation: pmx, pre-maxillary tooth.

Discussion

These results demonstrate that the non-linear and machine learning techniques I applied to hand-measured morphometric data derived from isolated theropod teeth consistently outperform LDA. When applying similar tests to anthropological data, Feldesman (2002) found that there was little difference between LDA and classification trees with LDA outperforming tree-based methods in some cases, whereas other authors (e.g., Finch et al., 2014; Holden et al., 2011) found LDA (and LR) to be the worst performers across a range of scenarios. This obviously raises the question of how to

choose the most appropriate classifier to apply to a dataset. As pointed out by Feldesman (2002), unless the data meet all of the theoretical conditions of the technique in question then there must be a lack of confidence in the predictions delivered. At a minimum, therefore, I would stress the importance of applying more than one technique to test the classification. In most studies, decision trees such as RF and C5.0 have been shown to be among the best performers and have few (if any) prior assumptions regarding data structures. I therefore recommend that a decision tree approach (or MDA, another strong classifier) be either the primary classifier or at least used to test the classification returned from the chosen primary classifier. Ensemble classifiers can increase the predictive power over a single classifier and also offer the opportunity to reduce the risk of choosing the 'wrong' classifier and, where possible, I advocate their usage also (Dietterich 2001).

I also demonstrate that the choice of prior probability can affect the outcome of the classification. As the true population distributions of fossil taxa are unknown, and sampling of taxa is essentially opportunistic, a reasonable assumption is that the probability of a random observation coming from a particular group is equal across the groups under investigation. I accept that a choice of equal prior probabilities can increase the bias towards smaller and potentially unstable groups and reduce the overall accuracy of the model (Table 3.5). Nonetheless, I would recommend using equal priors, as with fossil taxa the true population is unknown and therefore the sample population cannot reflect reality. Rigorous data preparation to reduce the number of small unstable groups can help, but there is then a trade-off between overall model accuracy and the potential that a group may need to be excluded from the model. Datasets that contain missing data within the predictor variables complicate matters, as traditional LDA algorithms will not use incomplete cases. My results indicate that imputing data as an alternative to deleting incomplete cases degrades the classifier accuracy substantially (Table 3.4; Fig. 3.12B). As decision trees can handle missing data I would recommend them over other alternatives as a first choice where the analysis of cases with missing data is a requirement. Class-imbalanced data biases the prediction towards majority groups and some techniques such as LDA perform badly with class imbalances. My results suggest that using methods such as SMOTE to address this, by balancing class

ratios via either synthetic case generation or under-sampling, degrades the classifier accuracy substantially (Fig. 3.12A). Blagus and Lusa (2013), however, concluded that whereas SMOTE was ineffective for discriminant analysis classifiers it may be of some benefit for other classifiers, such as decision trees. Although I would not rule out using synthetic data generation to balance classes, the effects of doing so need to be clearly understood (for example driving a bias towards the original minority classes) and the results tested against other classifiers using the imbalanced data. I would strongly recommend that posterior probabilities are checked as part of the process to verify the final classification.

Recent studies, such as Hendrickx et al. (2019), suggest that apomorphic character-based morphological data is potentially a more useful tool for distinguishing isolated theropod tooth crowns than morphometric data. However, I show that the careful application of machine learning techniques using the frameworks discussed in this study demonstrate that continuous quantitative morphometric data can also discriminate isolated theropod teeth with taxonomic accuracy of up to 96% in the specific datasets I used. The use of appropriate multiple classifiers coupled with a considered approach and understanding of the effects of missing data, initial group sizes and class imbalances are an improvement on the current commonly used techniques and yield rapid and statistically robust group predictions. Classification of isolated teeth in this manner will improve with better data, namely more cases per clade, to train the classifiers on. The careful addition of new measurement variables may also improve classification accuracies. As machine learning techniques have already been shown to be able to successfully classify taxa even with evolutionary convergence (e.g., Hoyal Cuthill et al., 2019) it is likely that even highly heterodont theropod clades and clades exhibiting dental morphological convergence could be accurately distinguished given the right amount of data and careful pre-processing of the data. It is probable that in some circumstances a combination of a dentition-based cladistic analysis and morphometric analysis may achieve the best results. The taxon-level grouping that is chosen will have an impact on the overall accuracy of the model simply because this controls the number of cases per group which in turn impacts on the ability of the classifier to accurately describe that group. An attempt to classify at a species level where each species is

described by, for example, four individual teeth will be less accurate than a genus level classification where each genus is represented by several hundred teeth.

Conclusions

In order to assess the performance of machine learning techniques on basic morphometric data derived from isolated theropod dinosaur tooth crowns a comparative study was undertaken using two published datasets. Various machine learning procedures were applied to each dataset in order to test the predictive accuracy under a range of different conditions. The results presented here, although specific to the tested datasets, demonstrate several important points:

Although LDA was generally the poorest performer in terms of accuracy, its predictive capability improved with larger class sizes.

Data subjected to predictive classification techniques needs to be rigorously assessed prior to classification for normality, missing data, class imbalances and class size. If data fail these tests then alternatives to LDA need to be considered.

Decision tree techniques such as random forest and C5.0 consistently outperformed other methods and I would advocate their usage for such classification problems.

Attempts to balance classes either by synthetic data generation, or by over- or undersampling of classes, significantly degraded the classification accuracy and care must be taken before employing these techniques.

Increasing percentages of missing data and the use of imputation to correct for this caused steep decreases in the predictive accuracy of those classifiers designed to handle such data (e.g., C5.0).

Different classifiers will assign the same case to different classes. The use of ensemble classifiers and an assessment of the resultant posterior probabilities helps to reduce the possibility of the 'wrong' technique being chosen.

As a result of this study I would recommend the use of decision trees as an alternative approach to LDA. The final aim of the analysis should guide the choice of random forest or C5.0. If the goal is to predict the taxon that a tooth falls into then random forests are a good choice. If the aim is to classify and to be able to see how the classification is built within the tree structure then C5.0 should be used. In practice I would recommend

corroboration of any results by checking predictions with another technique, preferably via the use of ensemble classifiers. The use of such techniques on isolated theropod teeth demonstrates that high levels of predictive taxonomic accuracy are possible from simple morphometric data as long as care is taken to understand the structure of the data in question and the assumptions that various techniques require.

Chapter Four: Geological Setting

Abstract

The Middle Jurassic of Britian was dominated by areas of low-lying land surrounded by shallow seas with a mixture of fully marine to marginal marine sedimentation. Scattered throughout the thick sequence of carbonates and mudstone that make up the Bathonian Great Oolite Group are impersistent horizons of terrestrial sediment formed during periods of regional uplift and emergence which preserve elements of terrestrial, semi-aquatic and marine vertebrate faunas. These horizons are near-shore or shoreline environments and represent either small ponds and lakes formed on the exposed carbonate platforms or, as at Watton Cliff, a tidal dominated shoreline system. The limited spatial extent of these deposits captures and preserves a very localised fauna. Detailed geological fieldwork at Woodeaton Quarry has clarified the regional lithostratigraphic framework of the White Limestone Formation (Great Oolite Group) and allowed both Woodeaton and Kirtlington Quarries to be correlated accurately.

Regional Geology

In the Middle Jurassic, Great Britain was situated at ~30° north and east of what would eventually become the North Atlantic rift (Fig. 4.1). The area was dominated by a series of structural highs and rifted basins forming areas of low-lying land surrounded by shallow epicontinental seas with a mixture of fully marine, shelf and lagoonal sedimentation (Fig. 4.2). Fluctuations in relative sea-level during the Bathonian resulted in a series of regressive units and locally emergent areas on the carbonate shelves surrounding the landmasses (Barron et al., 2012; Hesselbo, 2008; Horton et al., 1995; Palmer, 1979; Palmer & Jenkyns, 1975; Underwood, 2004; Wills et al., 2019; Wyatt, 1996). These periods of emergence preserved snapshots of the terrestrial fauna in several microvertebrate sites that are located stratigraphically in the Bathonian Great Oolite Group (Dineley & Metcalf, 1999; Evans & Milner, 1994; Metcalf et al., 1992; Metcalf & Walker, 1994; Wills et al., 2019; Wills, Underwood, et al., 2023).

The onshore Aalenian, Bajocian and Bathonian strata form a sequence of rocks up to 1300 m in thickness comprising mostly marine mudstones and carbonates in southern England, becoming increasingly arenaceous and non-marine towards the East Midlands, Yorkshire and Scotland and are interspersed with thin and impersistent horizons of

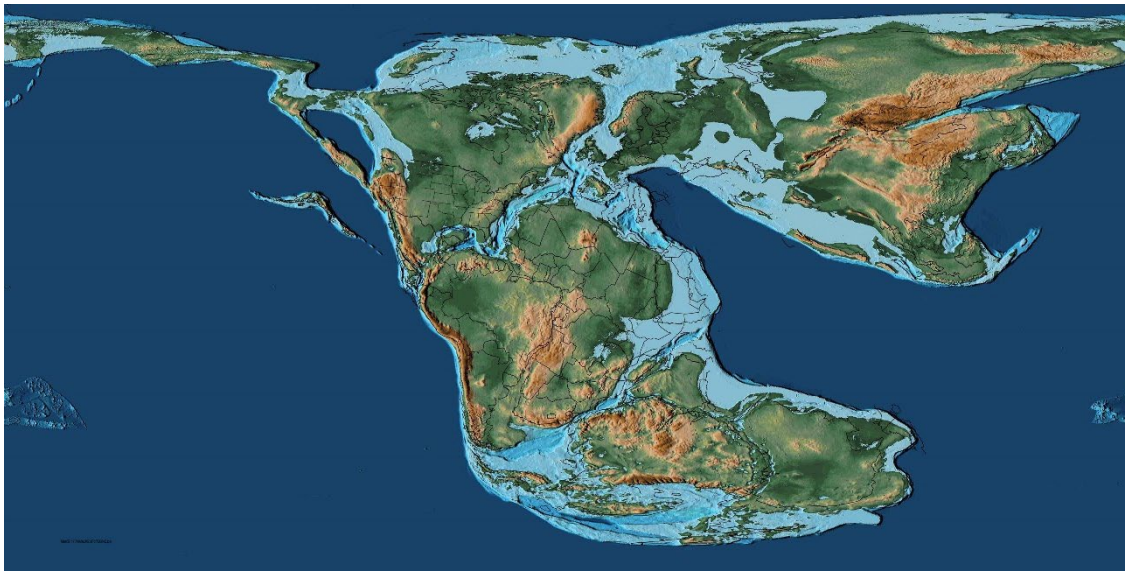


Fig. 4.1 Global palaeogeography in the Bathonian (Middle Jurassic). Great Britain was situated at ~30° north and east of what would eventually become the North Atlantic rift in an area of shallow seas with surrounding low-lying landmasses. After Scotese (2016).

terrestrial sediments. Hornsleasow, Kirtlington and Woodeaton Quarries are situated on the carbonates of the Cotswold Shelf (Fig. 4.3), and lay on, or slightly to the north of, the oolite barrier system formed during the Bathonian whereas Watton Cliff is situated to the south of the barrier in the open marine environment of the Wessex Basin (Fig. 4.3). The succession is characterised by a series of facies belts, representing increasing marine influence that prograded basinwards with time (Sumbler & Wyatt, 1999; Wyatt, 1996; Wyatt, 2011) and included an oolite barrier which played an important role in controlling regional depositional environments. Open marine conditions with deeper water sedimentation prevailed to the south of the oolites with nearshore marine, lagoonal and coastal plain environments to the north (Barron et al., 2012; Hesselbo, 2008; Palmer, 1974; Palmer, 1979; Palmer & Jenkyns, 1975; Sumbler & Wyatt, 1999; Underwood, 2004; Wyatt, 1996).

Localised fluctuations in relative sea level during the Bathonian resulted in a number of thin regressive units occurring throughout the succession. The units are represented by shallowing-up depositional sequences of sediment, often capped by hardgrounds or the development of karsts in exposed marine limestone sequences (Bradshaw, 1978; Cripps, 1986; Horton et al., 1995; Palmer, 1979; Palmer & Jenkyns, 1975; Sumbler, 1984; Sumbler et al., 2000; Wyatt, 1996).

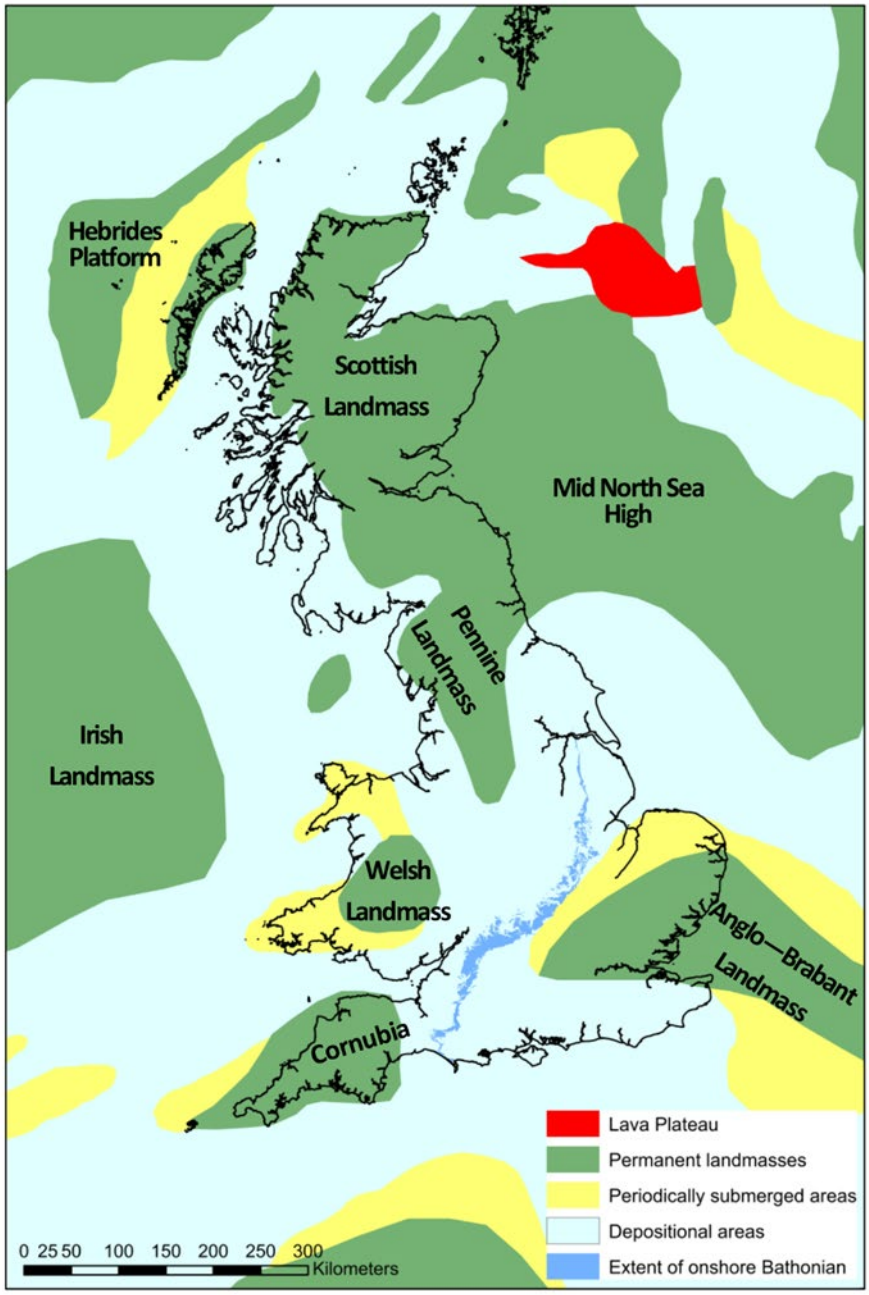


Fig. 4.2 Middle Jurassic palaeogeography of the British Isles. The major land masses and surrounding depositional areas are shown along with the extent of the onshore Bathonian strata. After Bradshaw et al. (1992).

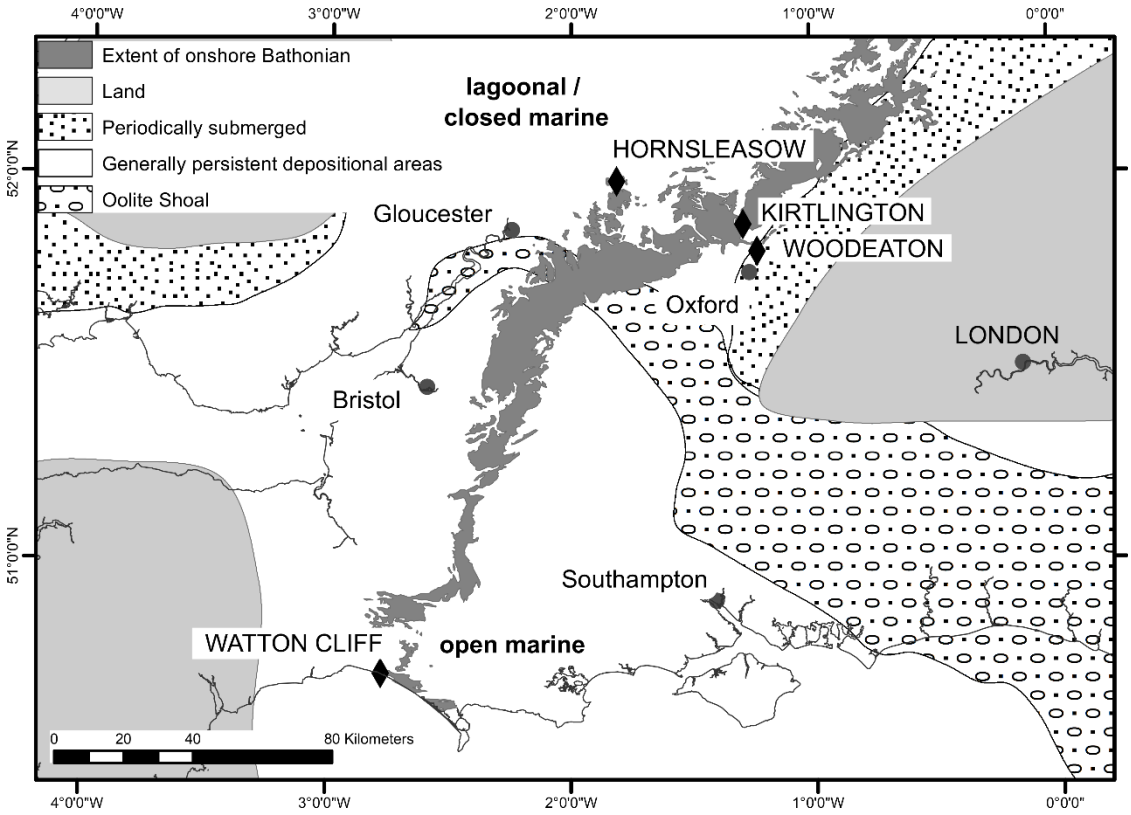


Fig. 4.3. Palaeogeography and depositional regimes of southern England during the Bathonian showing the location of the four microvertebrate sites: Hornsleasow, Kirtlington and Woodeaton Quarries in the northerly restricted marine environment and Watton Cliff in the southerly open marine environment. Site localities and Middle Jurassic palaeogeography and depositional regimes of southern England. After Underwood (2004) and Wills et al. (2019).

Lithostratigraphic framework

The Middle Jurassic in Britain has been the subject of detailed research, stretching back well into the 19th Century, with an extensive literature, resulting in a detailed lithostratigraphic framework for the group and a biozonation based principally on ammonites (Arkell, 1931, 1933a, 1933b, 1942, 1947; Arkell et al., 1933; Barron et al., 2012; Bradshaw, 1978; Bradshaw et al., 1992; Cope et al., 1992; Cope, 2012; Cox & Sumbler, 2002; Cripps, 1986; Hull, 1857, 1859; Judd, 1875; Murchison, 1834; Palmer, 1979; Phillips, 1871; Richardson, 1929, 1933; Sumbler et al., 2000; Sumbler & Wyatt, 1999; Torrens, 1969a, 1969b; Woodward, 1894; Wyatt, 1996; Wyatt, 2002; Wyatt, 2011). All four of the microvertebrate sites described in this thesis fall within the Great Oolite Group. They range from the basal Chipping Norton Formation (Hornsleasow Quarry), through the White Limestone Formation (Woodeaton and Kirtlington) to the Forest Marble Formation (Watton Cliff).

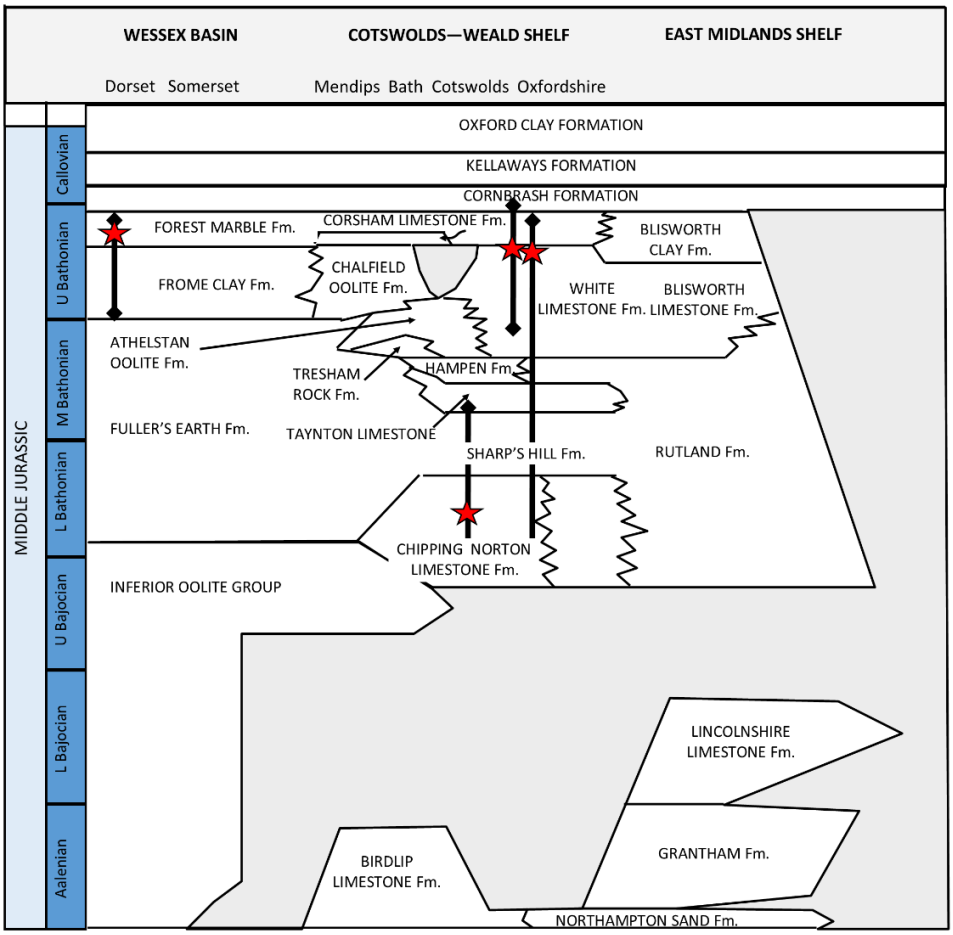


Fig. 4.4. Lithostratigraphic framework of the onshore British Middle Jurassic after Barron et al. (2012) and Wills et al. (2019). The vertical bars show the approximate vertical extent of sections exposed at (from left to right) Watton Cliff, Dorset; Hornsleasow Quarry, Gloucestershire; and Kirtlington and Woodeaton Quarries, Oxfordshire. The position of the microvertebrate horizon sampled for this thesis is shown in red.

The Great Oolite Group forms a complex sequence of mudstones and limestones, up to 200 m thick onshore (Norton et al., 2004). The gently shelving nature of the structural highs, coupled with minor fluctuations in sea level, caused major changes in shoreline positions and local depositional environments. As a result of this, many of the units are thin, diachronous, laterally impersistent and show rapid lateral and vertical facies changes and non-sequences (Barron et al., 2012; Cox & Sumbler, 2002). Depositional types range from clastic muds of deep water marine conditions, micrites on the shelf foreslope, oolitic and peletic carbonate sands on the very shallow marine shelf and mixed carbonate deposition in protected shallow marine lagoons (Barron et al., 2012; Palmer, 1979; Palmer & Jenkyns, 1975). Non-sequences and hardgrounds are relatively

common with karstic weathering surfaces, palaeosol development and influxes of terrigenous sediment containing abundant plant and woody material indicating pauses in marine sedimentation and occasional local emergence prior to the onset of fully marine conditions in the overlying Forest Marble Formation.

The nomenclature of the lithostratigraphy of the Great Oolite Group (Fig. 4.4) has grown in complexity over the years with many named units often having no formal boundary definitions and lacking correlation with their lateral equivalents. The description of the White Limestone Formation sequence at Woodeaton Quarry (in this chapter) helps to clarify the stratigraphy and relationships within this part of the Bathonian.

Hornsleasow Quarry

Hornsleasow Quarry (National grid reference: SP 131322, Fig. 4.5) is a designated Site of Special Scientific Interest (SSSI) located to the west of Moreton-in-Marsh in the Cotswolds of Gloucestershire. The quarry exposes a complete section through the Bajocian–Bathonian Chipping Norton Limestone Formation and the overlying Bathonian Sharp's Hill Formation (Fig. 4.6) and has been described in detail by several authors (Channon, 1950; Cox & Sumbler, 2002; Richardson, 1929; Sellwood & McKerrow, 1974; Torrens, 1969b). The microvertebrate horizon occurs in the Chipping Norton Limestone Formation as an 11 x 1 m clay lens lying on a palaeokarst surface (Fig. 4.7) that can be traced throughout the quarry (Metcalf, 1995; Metcalf et al., 1992; Metcalf & Walker, 1994; Vaughan, 1989).The succession at Hornsleasow is placed in the *Zigzagiceras zigzag* zone of the lower Bathonian (Cope et al., 1980) and represents the oldest of the British Bathonian microsites.

The clay lens was exposed in 1987 by blasting during quarrying operations. Channon (1950) previously noted a clay lens exposed at the same level in Hornsleasow Quarry in the section exposed at the time of his description. An amateur geologist, Mr Kevin Gardner, discovered pelvic bones and vertebrae that were referred to *Cetiosaurus* plus a large theropod tooth within the clay. The finds were reported to Gloucester City Museum who subsequently obtained the services of the Crickley Hill Archaeological Trust to excavate the site with the help of the quarry owners, Huntsmans Quarries Limited (Darlington, 1988; Vaughan, 1989). A sample of clay sieved for Gloucester City Museum using techniques developed by Ward (1981) produced a substantial

microvertebrate fauna following which the entire clay lens was excavated and removed for processing (see Chapter Two). The lens is made up of a lower grey smectite clay with limestone clasts and an upper green illite-chlorite clay (Metcalf, 1995; Metcalf & Walker, 1994). The larger vertebrate remains were recovered from the base of the grey clay unit and the microvertebrates were preferentially found in this layer (Metcalf, 1995). Palmer (in Vaughan, 1989) suggested a palaeokarst origin for the underlying surface developed during the emergence of marine limestones during the early Bathonian. The clay unit developed following a flooding event which introduced the initial sediment into the karstic hollow, subsequently a coastal marsh pond supporting a wide range of freshwater aquatic life became established (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.

The microvertebrate fauna of Hornsleasow comprises marine, freshwater and terrestrial elements with the marine taxa (shark and crocodile teeth) being reworked from the underlying limestones. The autochthonous fauna includes chondrichthyan and osteichthyan fishes, amphibians, turtles, small crocodiles and choristoderes, with terrestrial taxa represented by dinosaurs, mammals, pterosaurs and lepidosauromorphs (Evans & Kermack, 1994; Metcalf, 1995; Vaughan, 1989).

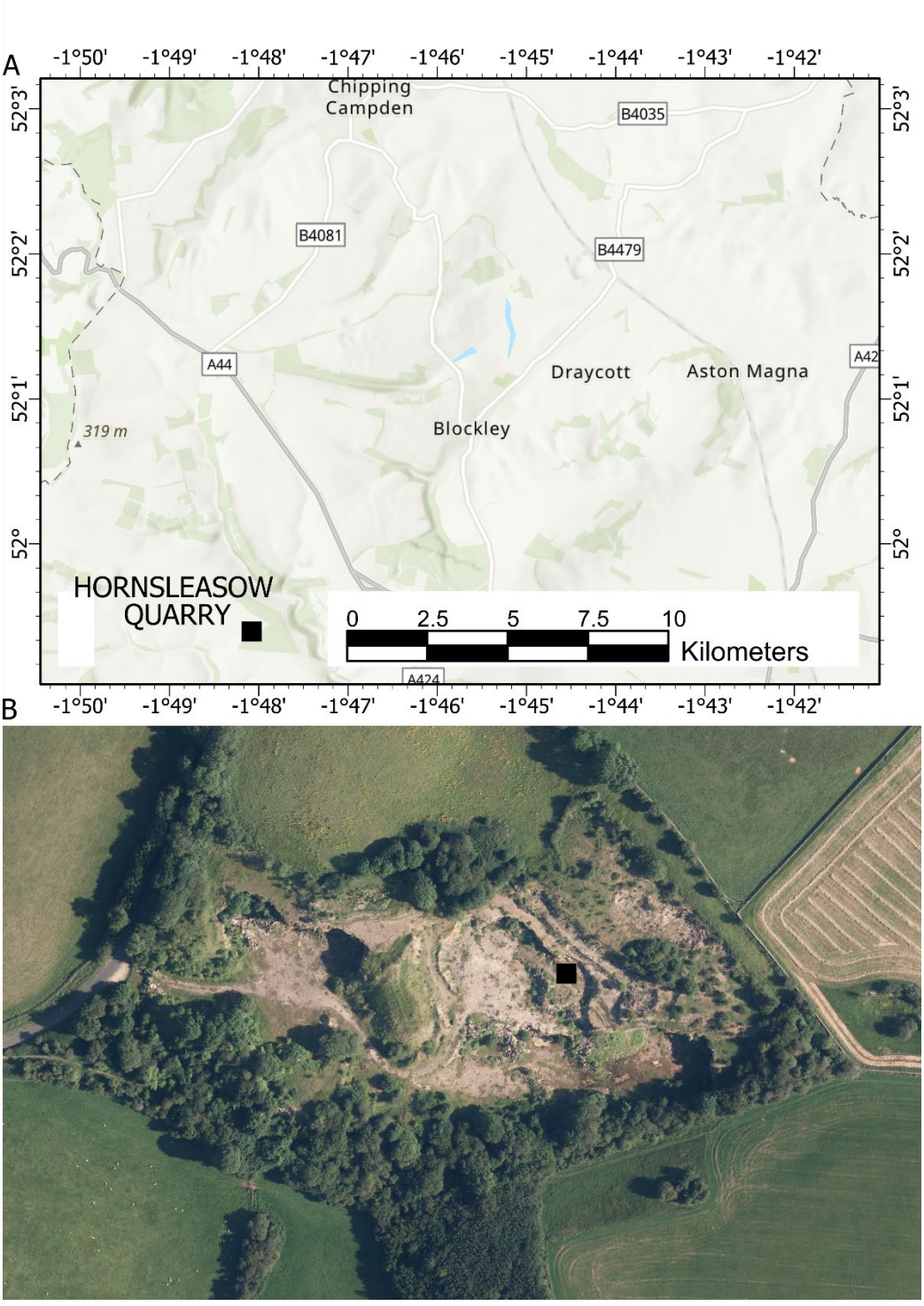


Fig. 4.5. The microvertebrate site at Hornsleasow Quarry, Gloucestershire. A, locality map. B, aerial photograph of the quarry in 2016 showing the location of the original excavation.

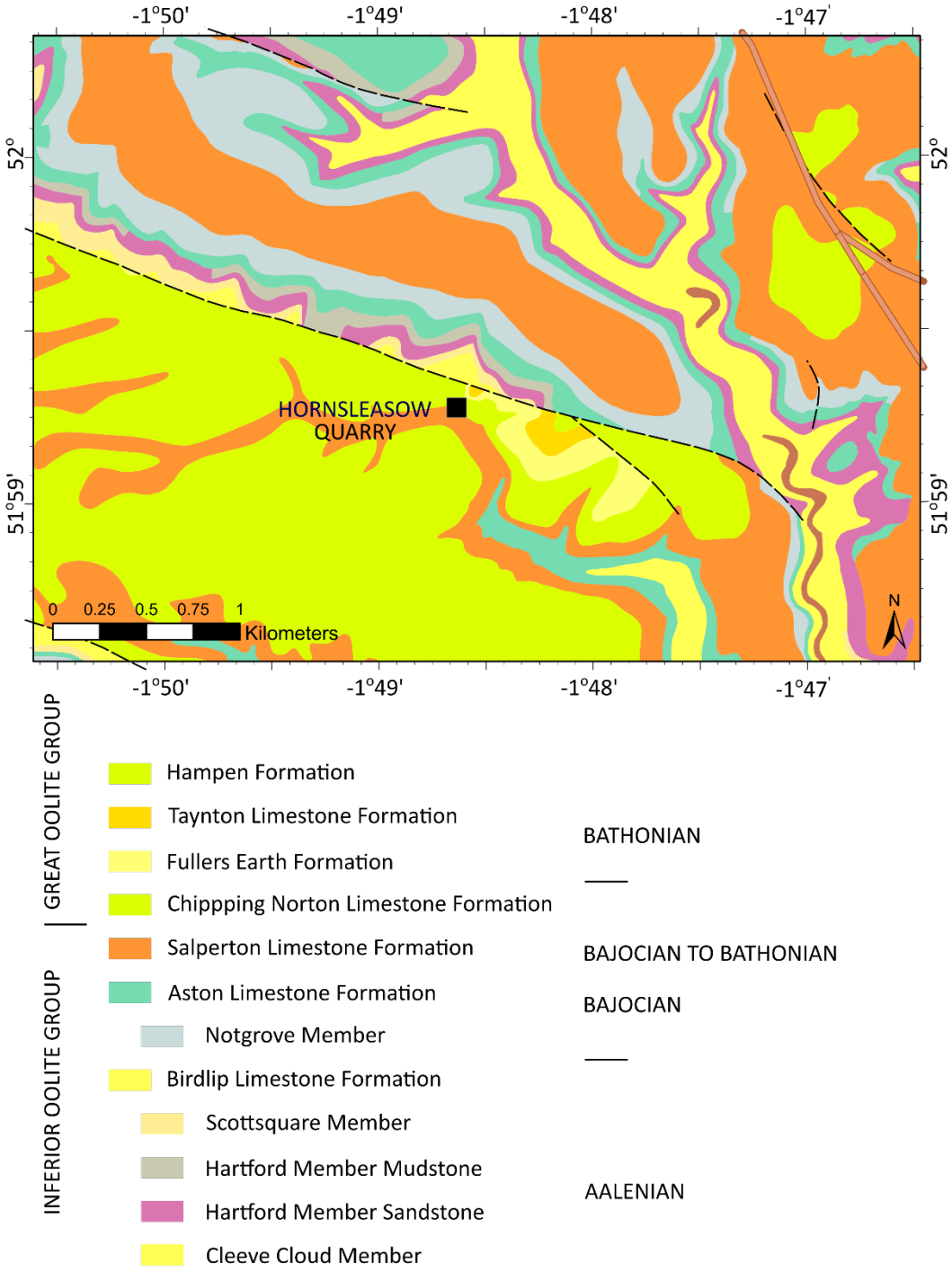


Fig. 4.6. Geological setting of Hornsleasow Quarry, Gloucestershire. Licence number 2017/024 ED British Geological Survey (c) NERC. All rights reserved.

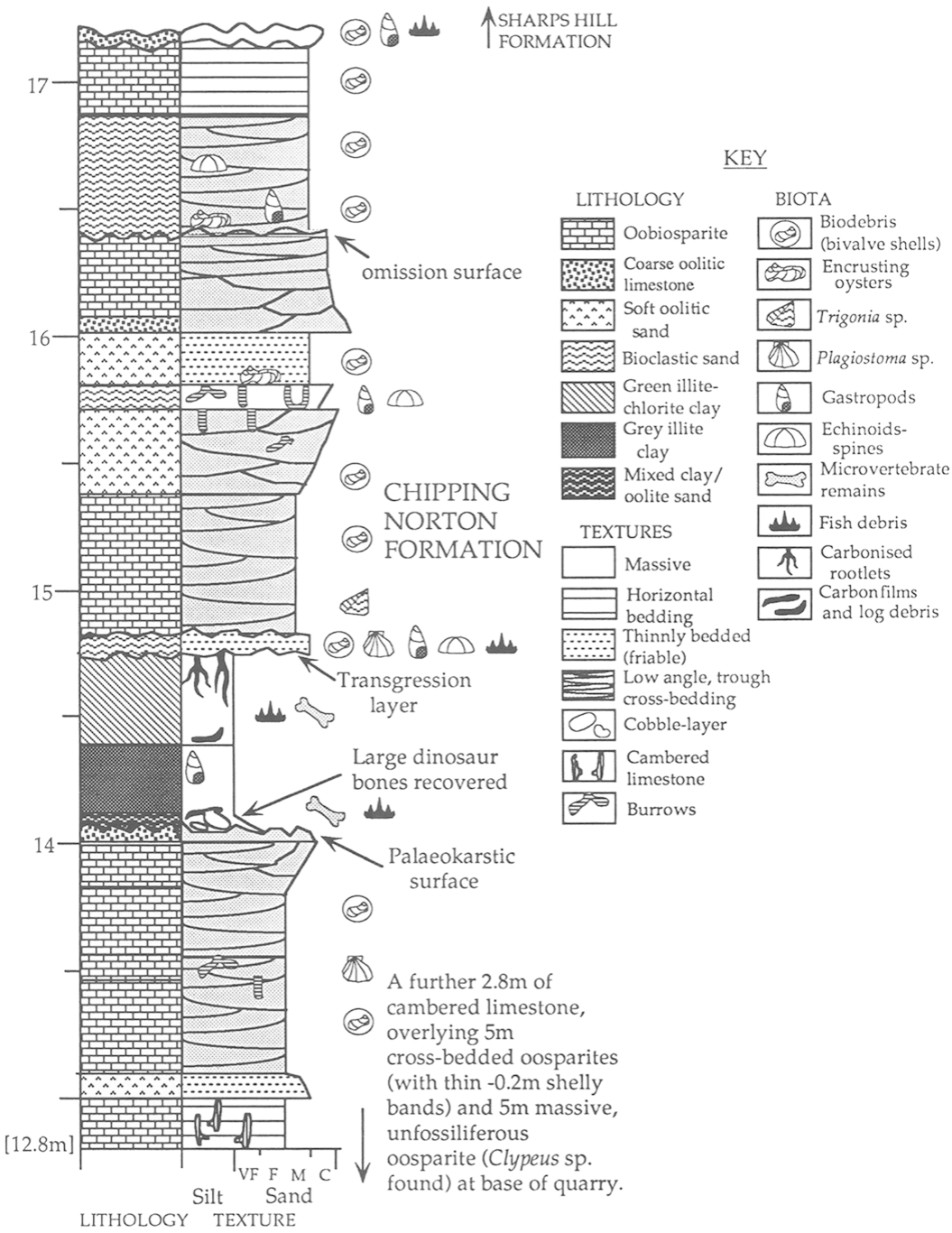


Fig. 4.7. Lithological log of Hornsleasow Quarry, from Metcalf et al. (1992), through the original exposed section at Hornsleasow Quarry in 1988. The clay lens ("Hornsleasow Clay") containing both large sauropod remains and microvertebrates rests on the karstic weathering surface at 14 m in the section.

Kirtlington Quarry

Kirtlington Cement Quarry (National grid reference: SP 494199, Fig. 4.8) was worked commercially in the early part of the 20th century. It is renowned for a rich Bathonian vertebrate fauna including important mammalian remains, for which it is a designated SSSI and a Geological Conservation Review site. The quarry exposes sections through the White Limestone, Forest Marble and Cornbrash formations of the Great Oolite Group (Fig. 4.9) and has been described in detail by many authors (Arkell, 1931; Benton et al., 2005; McKerrow et al., 1969; Richardson et al., 1946). The microvertebrate fauna of Kirtlington was discovered in 1974 by an amateur palaeontologist, Eric Freeman, who began prospecting for mammal teeth in what is now known as the 'Mammal Bed' (Freeman, 1976a) and subsequently worked by a team from University College London (Evans and Kermack (1994) and Fig. 4.10). The microvertebrates were collected by bulk sampling of the unconsolidated marl, and later processing produced a wide range of fauna including fish, frogs, salamanders, turtles, sphenodontians, lizards, pterosaurs, dinosaurs, crocodiles and mammals (including early multituberculates) (Benton et al., 2005; Benton & Spencer, 1995; Butler & Hooker, 2005; Evans, 1990; Evans, 1991; Evans & Kermack, 1994; Evans et al., 1988; Freeman, 1976a, 1979; Kermack et al., 1998; Kermack et al., 1987; Sigogneau-Russell, 2001, 2003a, 2003b).

The main microvertebrate horizon, the 'Mammal Bed', occurs in a bed of unconsolidated clay, which had been assigned to the Forest Marble Formation (Bed 3p of McKerrow et al. (1969) and Benton et al. (2005)) above an oolitic limestone that grades down into a more massive coralline limestone (Coral–*Epithyris* limestone of the Baldon Member, White Limestone Formation). The Forest Marble Formation is considered to be upper Bathonian in age ranging from the *Retrocostatum* to *Discus* Zones with the base of the formation placed immediately above the coralline limestone bed (Barron et al., 2012; Torrens, 1969b), implying a *Retrocostatum* zone age for the 'Mammal Bed' (Evans & Milner, 1994; Freeman, 1979). However, the presence of an ostracod in the coralline limestone could indicate a slightly younger *Discus* zone age (Benton et al., 2005).

More recent site observations, carried out as part of this work, show that the microvertebrate horizon occurs in the upper part of the White Limestone Formation, around or just above the level of the *Fimbriata-waltoni* Bed of previous studies (Horton

et al., 1995), rather than in the Forest Marble Formation, and is slightly younger than the approximately coeval deposit at Woodeaton Quarry.

The 'Mammal Bed' occurs in the north eastern corner of the quarry (see Text-Fig. 2 of McKerrow et al. (1969)) and comprises a thin and impersistent lens of unconsolidated brown marl some 4–25 cm thick, extending over a section of 21.5 m and resting on an erosional surface with the underlying limestones (Freeman, 1979). The palaeoenvironment of the Kirtlington 'Mammal Bed' has been interpreted as a pool formed during a period of marine regression along a shallow coastal plain region characterised by coastal lakes, swamps and lagoons (Evans & Milner, 1994; Freeman, 1979; Palmer, 1979). The bed contains both indigenous and derived faunal assemblages with the latter comprising reworked marine forms from the underlying limestone and the former representing a non-marine aquatic to semi-aquatic and terrestrial fauna (Benton et al., 2005; Evans & Milner, 1994; Freeman, 1979). A hard-pan at the top of bed 3p (the 'Mammal Bed') indicates a period of sub-aerial exposure of the sediments before the onset a marine influx associated with the overlying cross-bedded limestone of the Forest Marble Formation.

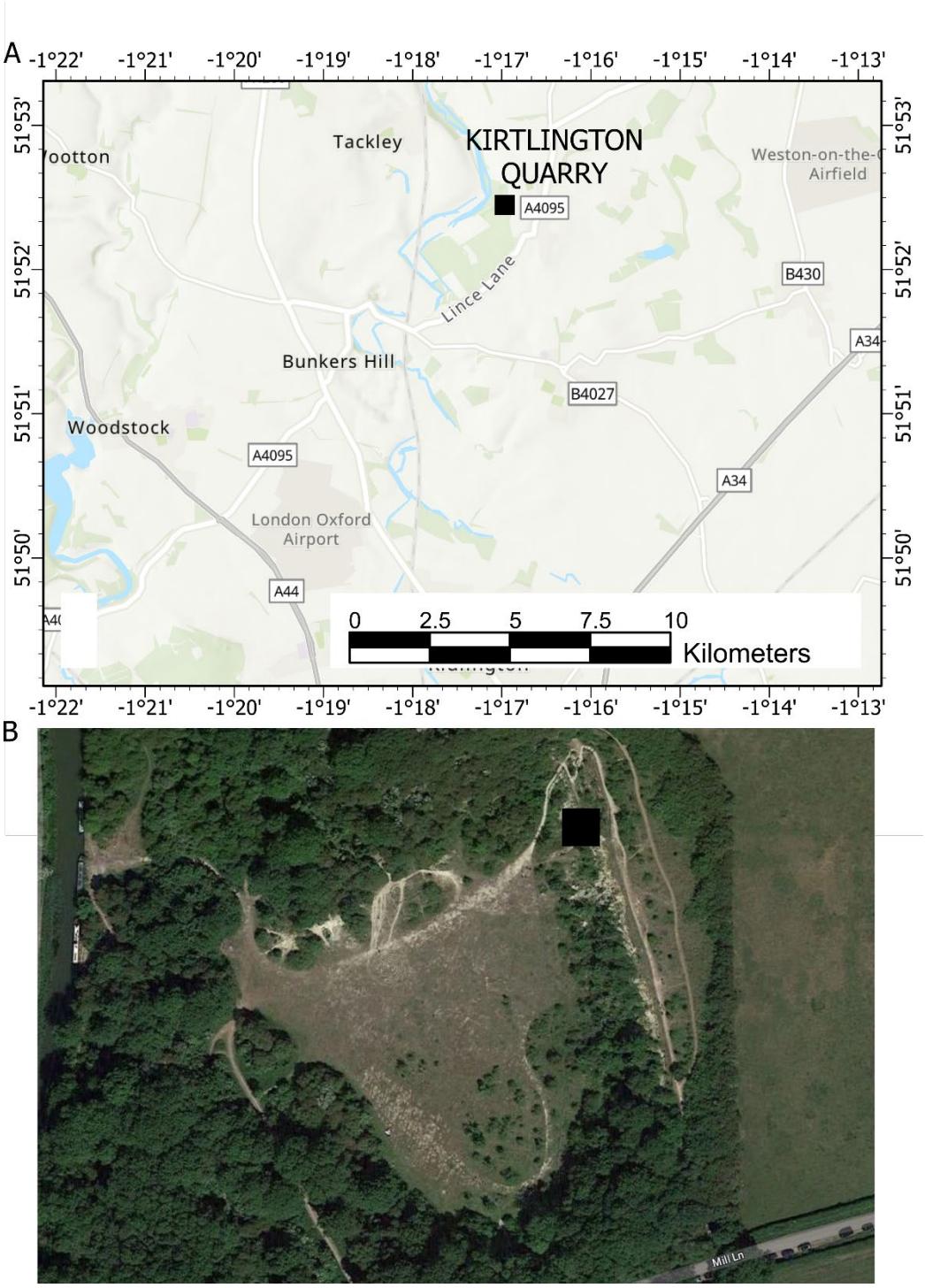


Fig. 4.8. The microvertebrate site at Kirtlington Quarry, Oxfordshire. A, locality map. B, aerial photograph of the quarry in 2023 showing the location of the original excavation of the microvertebrate rich “mammal bed”.

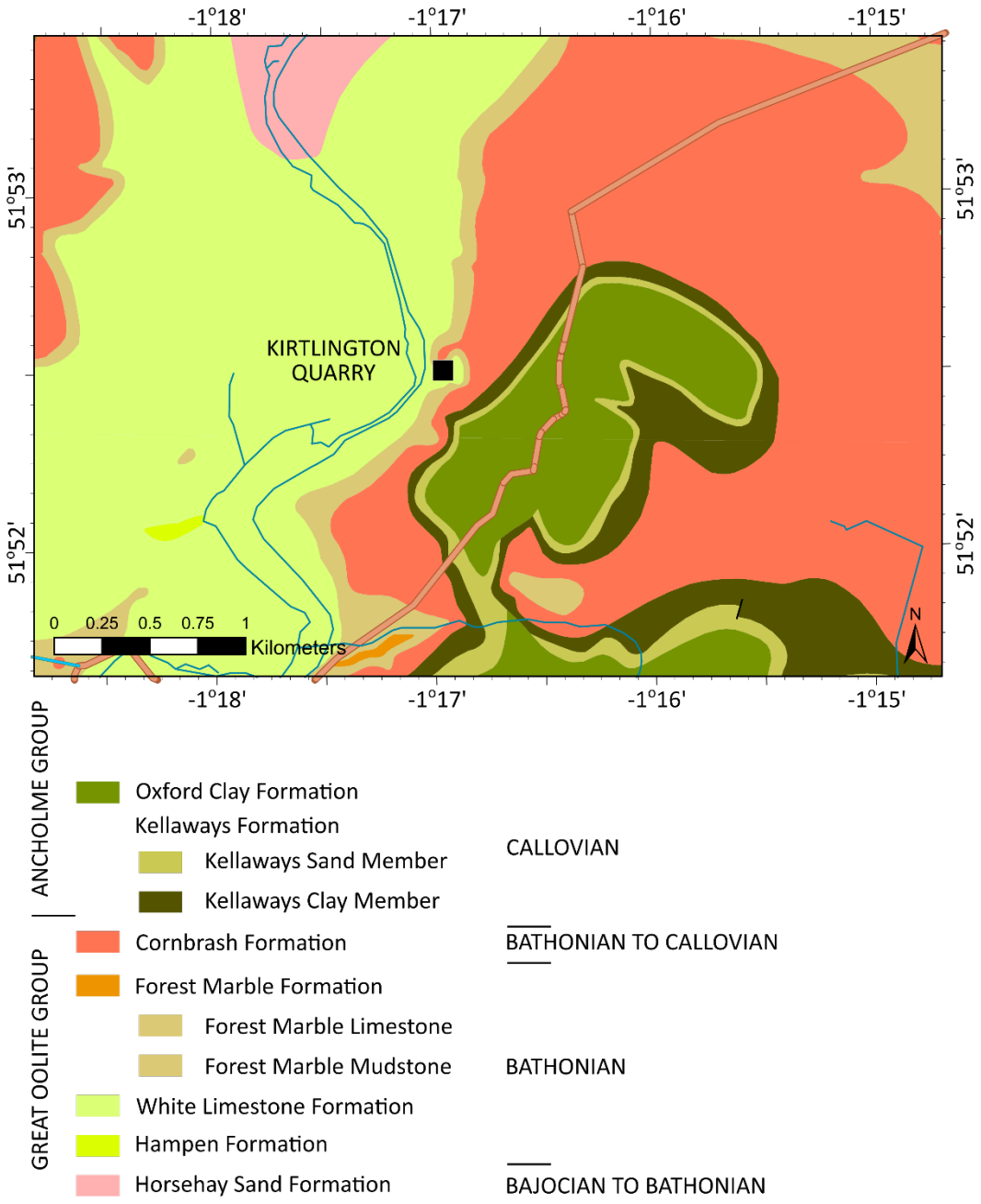


Fig. 4.9. Geological setting of Kirtlington Quarry, Oxfordshire. Licence number 2017/024 ED British Geological Survey (c) NERC. All rights reserved.

Fig. 4.10. Excavations by University College London at Kirtlington Quarry in 1980 and 1981. The bed numbers in the images refer to those from McKerrow et al. (1969). A, the main microvertebrate horizon at Kirtlington (Bed 3p) at the base of the section exposed in the north-east corner of the quarry. B, excavation at Kirtlington in 1981 showing the relative position of bed 3p and the *Fimbriata-waltoni* bed. C, development of a hard-pan at the top of bed 3p. Similar hard-pans are seen at Woodeaton Quarry. Images courtesy of Susan Evans, University College London.

Woodeaton Quarry

Preface: This section has been published as;

Wills, S., Bernard, E. L., Brewer, P., Underwood, C. J., & 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(2), 170-186. https://doi.org/https://doi.org/10.1016/j.pgeola.2019.02.003

The majority of this paper was my original work including the drafting of the manuscript. Brewer provided additional information on the mammal fauna and Bernard on the fish fauna. Underwood and Wills undertook logging of the measured sections and Underwood prepared the main log which is reproduced here. All authors contributed equally to the fieldwork and Ward undertook a substantial part of the initial screen-washing of sediment.

Here I review and update the stratigraphy, sedimentology and geological setting of Woodeaton Quarry near Oxford (Fig. 4.11, National Grid Reference SP533123) (Bathonian, Great Oolite Group), including the discovery of a new horizon rich in microvertebrate remains. The material described herein was collected between 2013 and 2016 by the Natural History Museum, London and Birkbeck College, University of London. The initial work was undertaken to recover representative samples from each bed and to verify the stratigraphy before these sections became inaccessible, as the quarry was scheduled to be infilled and partly made into a nature reserve. A previous visit by the Natural History Museum in 2002 recovered the partial remains of a small sauropod from the Rutland Formation (Great Oolite Group), which will be described elsewhere (P. M. Barrett, pers. comm.).

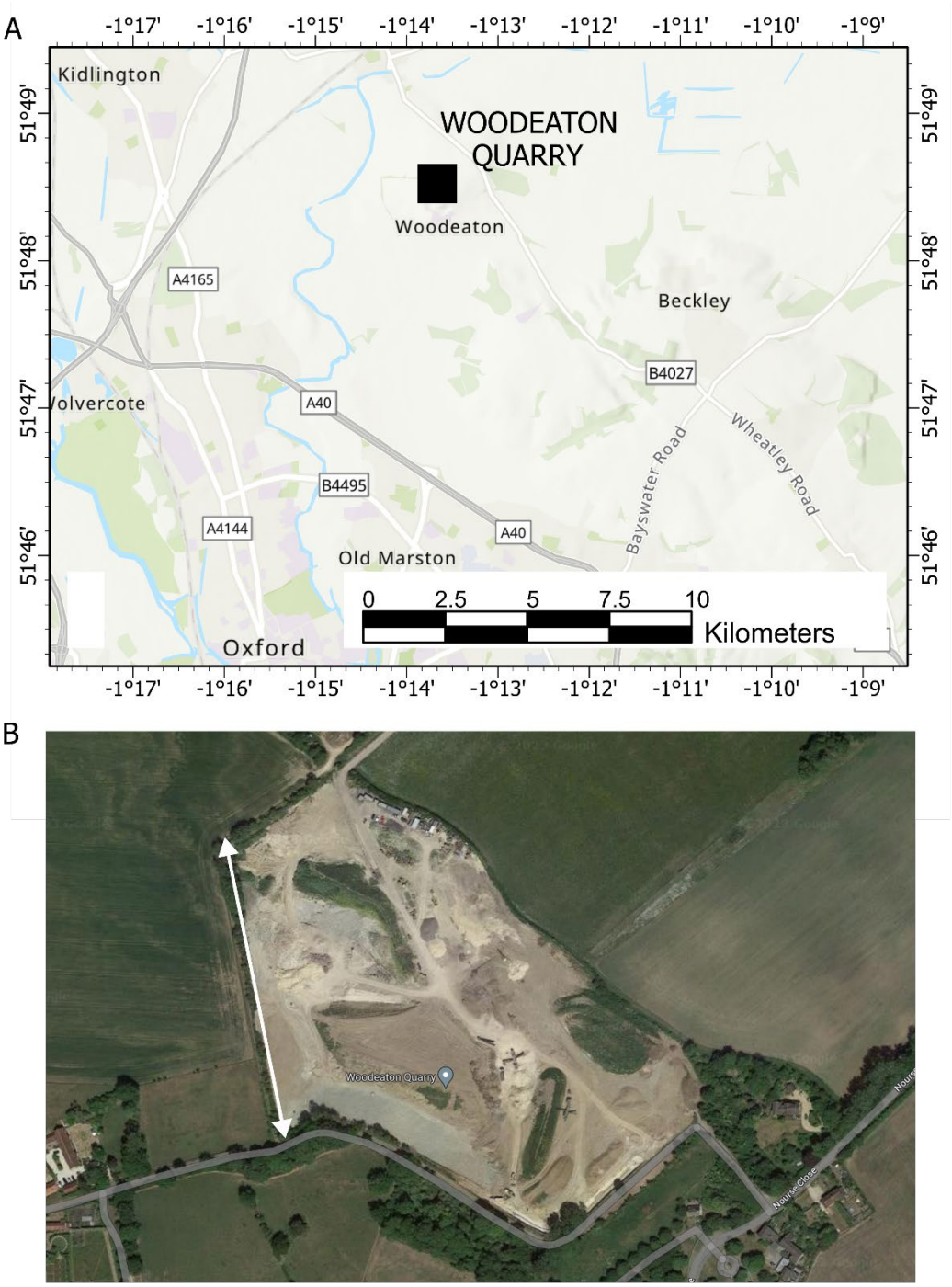


Fig. 4.11. The microvertebrate site at Woodeaton Quarry, Oxfordshire. A, locality map. B, aerial photograph of the quarry in 2023 showing the location of the excavation of the microvertebrate rich horizon along the western quarry face.

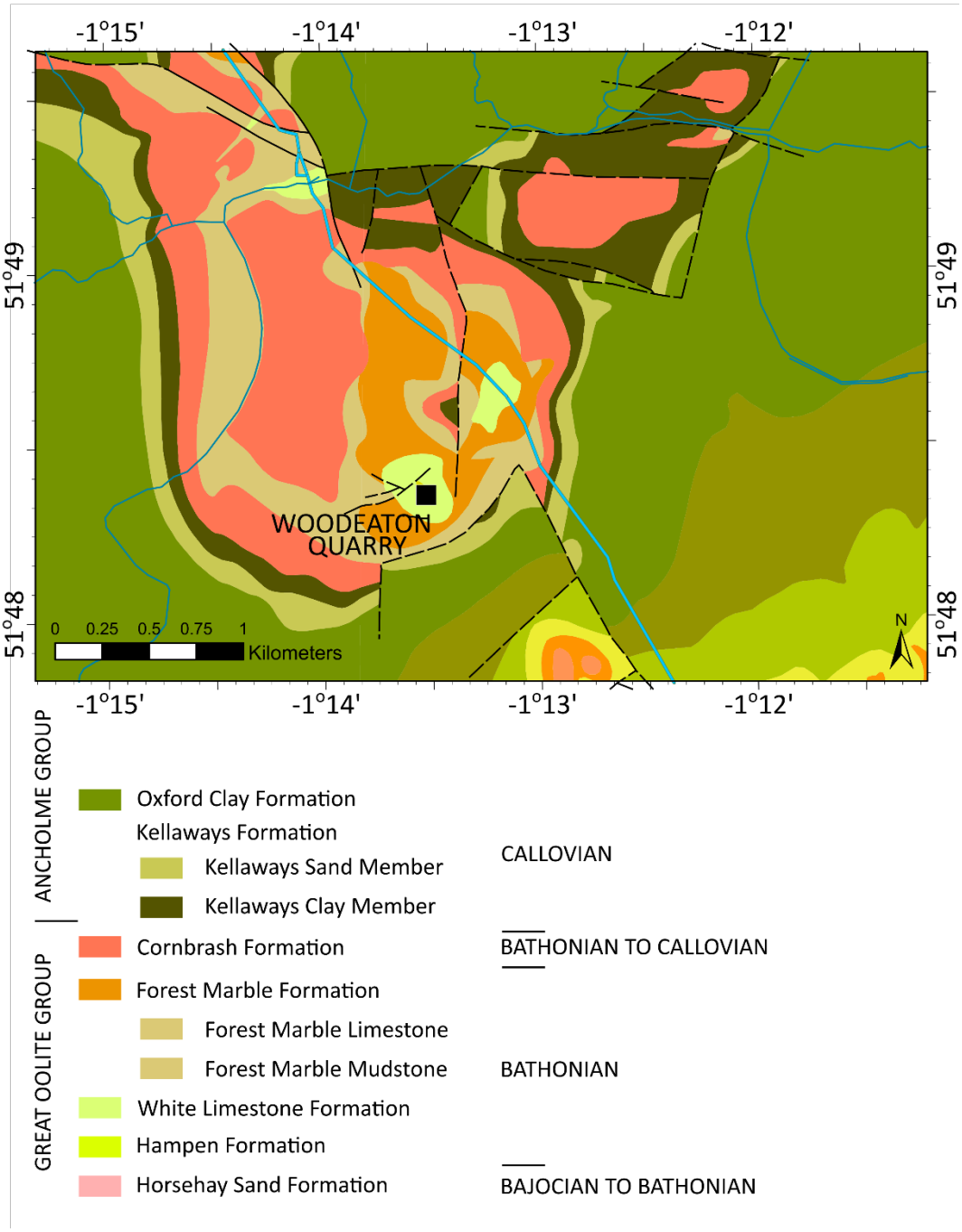


Fig. 4.12. Geological setting of Woodeaton Quarry, Oxfordshire. Licence number 2017/024 ED British Geological Survey (c) NERC. All rights reserved.

Woodeaton Quarry in Oxfordshire had previously yielded several large sauropod vertebrae and other (unpublished) dinosaur remains from a horizon in the Rutland Formation. Fieldwork at Woodeaton was conducted in 1983 by the same team from University College London as was previously involved in excavations at Kirtlington Quarry. That fieldwork concentrated on the Rutland Formation, although the section through the Bladon Member of the White Limestone Formation from which microvertebrates were collected as part of this thesis was exposed and examined at the time. No review of the wider terrestrial fauna from Woodeaton has been published to date. Here I present an overview of new material recovered from a microvertebrate site at the top of the White Limestone Formation (Middle Jurassic, Bathonian, Great Oolite Group, *Retrocostatum* Zone) and review the stratigraphy to provide a comprehensive local stratigraphic framework and place the quarry in the correct regional context. The terrestrial fauna is similar to that found from other British Middle Jurassic microvertebrate sites and includes theropods, ornithischians, tritylodontids, and early mammals, including amphitheriids, docodonts, “eutriconodonts”, “haramyids”, and multituberculates (Evans & Milner, 1994; Freeman, 1976a; Kermack et al., 1998; Kermack et al., 1987; Metcalf & Walker, 1994; Wills et al., 2019). Placement of the White Limestone Formation boundaries is clarified with respect to the Rutland and Forest Marble Formations. This indicates that the microvertebrate site fauna from Woodeaton Quarry is slightly older than that of the well-known ‘Mammal Bed’ from nearby Kirtlington Quarry. A detailed description of the theropod and ornithischian dinosaur material is presented in Chapters Five and Six respectively.

Woodeaton Quarry (Fig. 4.12) is situated on a periclinal inlier of Great Oolite Group surrounded by the Kellaways and Oxford Clay formations at the intersection of the Islip anticline with the Wheatley fault zone (Arkell, 1944; Horton et al., 1995; Wyatt, 2002). The area around Woodeaton has been extensively quarried in the past and Arkell (1947) noted that there were at least 28 quarries in the Noke Hill area. Woodeaton Quarry, also known as Grove Quarry, is one of two original quarries situated in the vicinity of Woodeaton Village with Hope Farm Quarry long since closed (Cripps, 1986). The quarry formerly exposed one of the most complete sequences of the middle to late Bathonian

in the UK from the Chipping Norton Limestone Formation to the lower part of the Forest Marble Formation (Fig. 4.13).

Palmer (1973) provided the first detailed description of Woodeaton Quarry, with subsequent descriptions in Cripps (1986), Horton et al. (1995), Palmer (1974) Palmer (1979), Palmer and Jenkyns (1975) and Wyatt (2002). The Rutland Formation, White Limestone Formation and lower part of the Forest Marble Formation are well-known from long-term exposures (e.g. Horton et al., 1995). Of these units, the Rutland Formation was least exposed and commonly rather degraded, to the extent that nearly half of its thickness, now known to contain three palaeosols, was grouped by Horton et al. (1995) into a single bed. Considerable lateral variation is seen in some of the units, but this is not mentioned in publications on the stratigraphy of the site. Renewed quarrying in 2002 greatly expanded the exposed succession, with units referred to the Chipping Norton Limestone, Sharps Hill, Charlbury and Taynton Limestone formations being exposed. These lower units were rapidly obscured and no mention was made of them in later publications (Guthrie et al., 2014). Of these lower units, only the uppermost part of the Taynton Limestone Formation was still exposed in 2014–2016, although grey marl with rhynchonellids, originating from the Charlbury Formation, was being used as a quarry lining in several places. The quarry is currently being used as a landfill site and as of 2017 the only remaining extant sections visible are from the White Limestone Formation (Ardley Member) through to the Forest Marble Formation.

During the first stages of fieldwork, a section along the western edge of the quarry was examined. This section exposed a sequence of limestones and clays of the Ardley and Bladon members (Great Oolite Group, White Limestone Formation) through to the shelly detrital limestone of the Forest Marble Formation. One bed, a variably lithified clay to impure limestone horizon with abundant plant material (Bed 23, see section below), produced substantial quantities of terrestrial microvertebrates. Bulk sampling concentrated on this unit thereafter.

The geology and stratigraphy of Woodeaton Quarry is outlined below, and sedimentary logs resulting from this study are presented in Fig. 4.13.

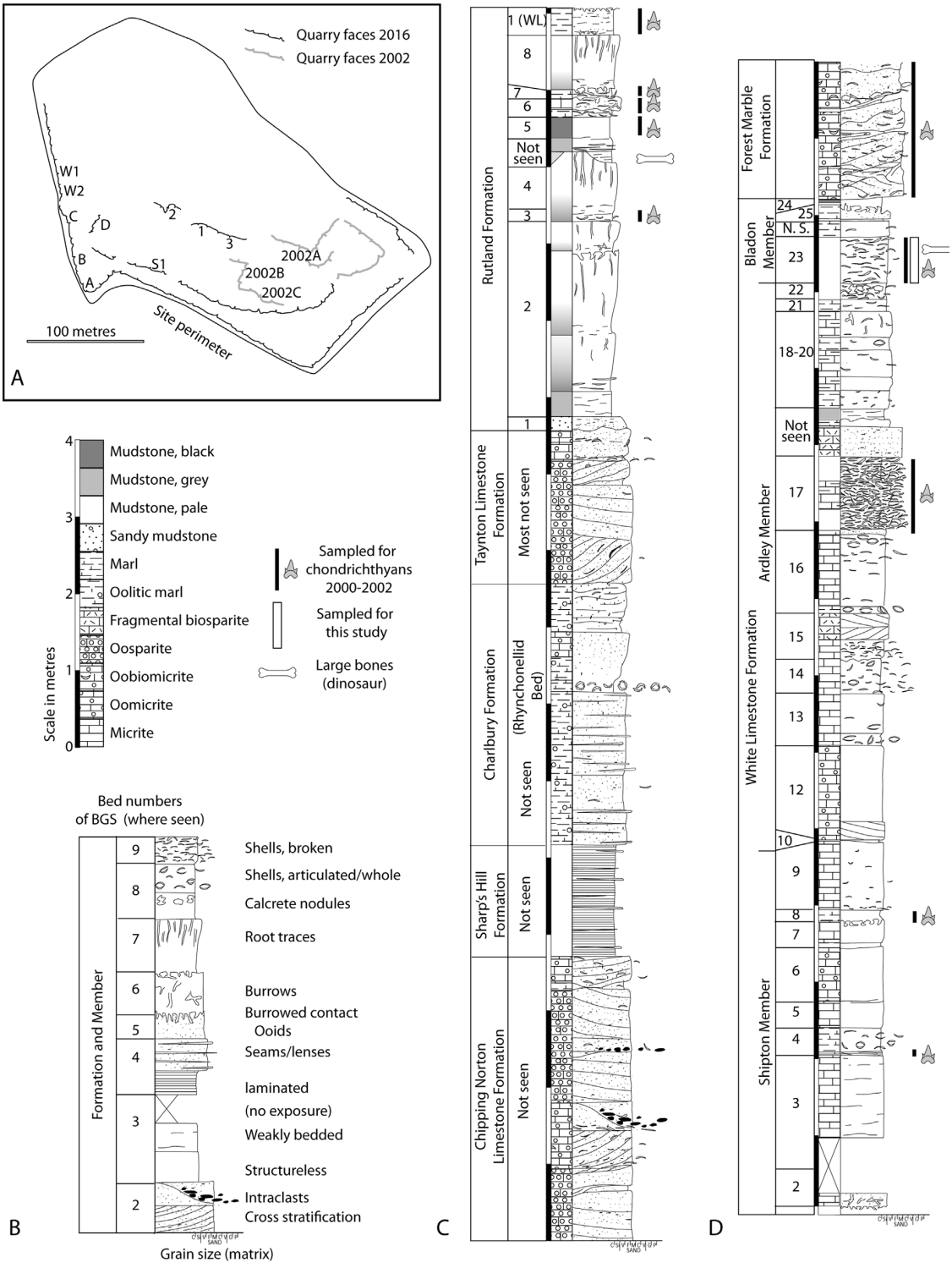


Fig. 4.13. Stratigraphic logs through the section at Woodeaton Quarry, Oxfordshire. A, location of described sections at Woodeaton Quarry. B to C, composite stratigraphic logs from the Chipping Norton Formation to the Forest Marble Formation showing sampling undertaken in 2000–2002 and for this study. Bed numbers referred to British Geological Survey numbering from Horton et al. (1995). Logs produced by Underwood in Wills et al. (2019). Bed 23 is the microvertebrate horizon sampled as part of this study.

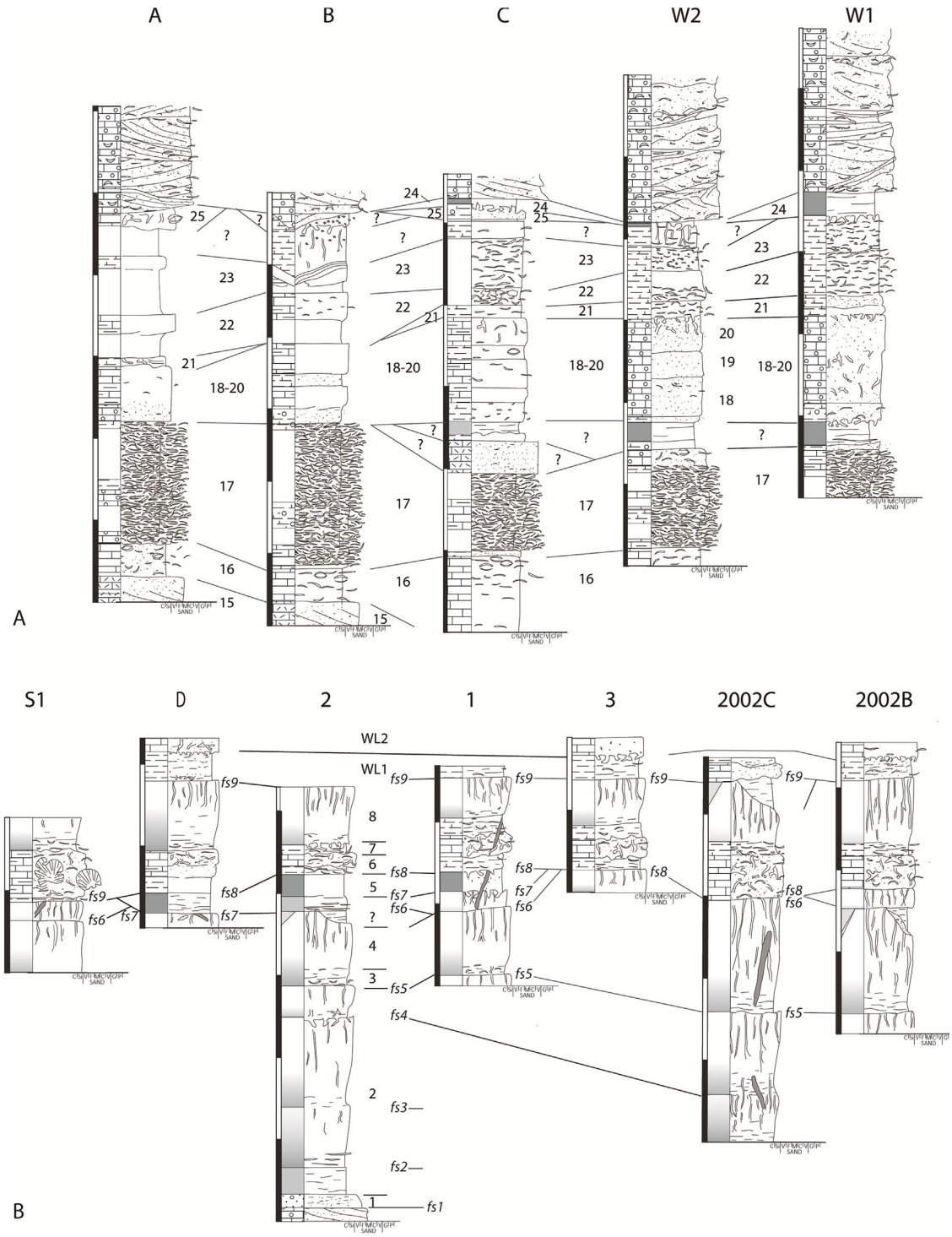


Fig. 4.14. Correlated stratigraphic logs at Woodeaton Quarry, Oxfordshire. Section names refer to localities on Fig. 4.13A, bed numbers refer to the section depicted in Fig. 4.13B, fs – flooding surface. A, north – south section along eastern perimeter of quarry. B, east – west section along western perimeter of quarry (Fig. 4.11). Horton et al. (1995). Bed 23 is the microvertebrate horizon. Logs produced by Underwood in Wills et al. (2019).

Chipping Norton Limestone Formation

Up to 3.8 m of this unit was previously exposed during the working life of the quarry. The base was not seen, but a spring line near the exposed base suggests that a contact with an underlying impervious unit (presumably Toarcian mudstones) is present not far below the exposure. The Chipping Norton Limestone Formation comprises yellow-orange, strongly cross-stratified oolitic and bioclastic packstones and some grainstones. The cross-stratification appears to be mostly in the form of trough sets, some reaching 1 m high. The majority of the unit is oolitic with only rare bioclasts, but some beds are composed primarily of small and fragmented bioclasts. One small channel contained large numbers of oval intraclasts, with rare intraclasts being present elsewhere in the unit.

Sharp's Hill Formation

This muddy unit was never well exposed, but where exposure occurs it reaches about 1.3 m in thickness. It comprises pale grey silty claystones with thin lenses and seams of calcareous siltstone. Lamination is apparent throughout, and the unit is virtually unfossiliferous.

Charlbury Formation

The Charlbury Formation is about 3.5 m thick and is composed largely of grey to yellowish impure limestone. The dominant lithology is a soft micrite with matrix supported ooids throughout. Thin lenses and seams of calcareous siltstone are present. Fossils are generally uncommon with scattered oysters and brachiopods, with the exception of a bed of marl crowded with well-preserved rhynchonellid brachiopods (*Kallirhynchia*). The latter represents the so-called Rhynchonella Bed as recognized at other sites (e.g. Boneham & Wyatt, 1993).

Taynton Limestone Formation

The majority of the 2 m thick Taynton Limestone Formation comprises strongly cross-stratified cream coloured oolitic grainstones. The tabular cross-sets are typically at a low angle and bioclasts are small and uncommon throughout. Towards the top of the formation, there is more matrix and bioclasts, typically broken oysters, are common. The top surface of the unit is irregular and orange-stained and appears karstic.

Rutland Formation

The Rutland Formation here corresponds to the Hampen Marly Beds of Arkell et al. (1933) and Arkell (1947), the Hampen Marly Formation of Palmer (1979) and the Upper Estuarine Series of Bradshaw (1978). The term Hampen Marly Formation is now restricted to the more marine clay and limestone facies that occur to the west of the area (Horton et al., 1995). The Rutland Formation is a variable package of lagoonal to subaerial lagoon margin facies totalling about 5.2 m in thickness. Whilst the lower parts were only seen in two excavated sections in 2002 and 2014, the uppermost 2 m of the section was well exposed and shows considerable lateral variation. The bed numbering scheme of Horton et al. (1995) works for the upper part of the formation, but it does not take lateral variation into account and groups several distinct units together in the lower part. In addition, the uppermost unit of the Rutland Formation was included in the White Limestone Formation by Horton et al. (1995) despite being identical lithologically to some underlying beds.

The base of the Rutland Formation is an oolitic clay with clasts of limestone from the underlying Taynton Limestone Formation. Above this, the Rutland Formation comprises a sequence of shallowing upwards cycles terminating in rootlet-bearing palaeosols, with flooding surfaces terminating the rootlet horizons. Above the basal flooding surface, eight additional surfaces were recognized (fs 2-8, see Fig. 4.14), but not all were seen in all sections due to erosional downcutting at some of the flooding levels. Where complete, each of the cycles in the lower to middle part of the formation, comprises a basal dark mudstone with small aragonitic shells and thin burrows, grading up to a paler grey mudstone devoid of obvious fossils and finally to a pale greenish marl with vertical, carbonaceous roots. In addition to the typical small root traces, large subvertical roots, seen to penetrate at least 1.5 m and reach 0.1 m in diameter, are present, being initiated from several different palaeosol surfaces. In places, below flooding surfaces fs7 and fs9, small, incised depressions are filled with laminated pale mudstone containing (where studied) non-marine ostracods and charophytes. In some sections, a level above surface fs7 comprises a dark green and brown mudstone with greenish slickensides. This presumably represents a subaqueous rooted horizon, in contrast to the other rooted levels that were subaerial. Above surfaces fs8 and fs9 are black, gritty (bioclast-rich) mudstones with a restricted marine biota of oysters and other bivalves, echinoids

(*Hemicidaris*) and diverse neoselachian sharks. One interval, Bed 6 of Horton et al. (1995) of this more marine interval forms a dark impure limestone with a similar restricted fauna, but at site S1 also contains large numbers of highly bio-eroded coral colonies. It is likely that older records of a "Monster Bed", Bed 5 of Palmer (1973), refer to the succession between surfaces fs5 and fs6, which yielded sauropod dinosaur remains within a palaeosol in 2002.

White Limestone Formation

The White Limestone Formation is approximately 13 m thick, although some of the lowest part was not exposed. It is typically divided into the Shipton, Ardley and Bladon members, but at Woodeaton Quarry the Shipton Member and lower part of the Ardley Member cannot be readily separated on lithological grounds. The lower part of the formation, up to Bed 16 of Horton et al. (1995), comprises massive pale limestones with several thin seams of marl (Beds 8, 10 and the base of Bed 7). The limestones are typically micrites or oolitic wackestones although a finely biocalstic level (Bed 15) is also present. Sedimentary structures are limited to low angle tabular cross-stratification in Beds 12 and 15. Fossils are rare in some beds, but common in others where a low diversity marine biota is dominated by oysters, modiolids and the brachiopod *Epithyris*. The upper part of the White Limestone Formation is lithologically very variable with considerable lateral variation.

The first major lithological deviation from the pale limestones of below comes with Bed 17 (Horton et al., 1995). This varies from 1.1–2 m in thickness and comprises a dark grey to orange clay, marl or rubbly muddy limestone matrix with matrix to clast supported oyster shells. There are ooids in places in addition to an associated fauna including echinoderms, brachiopods, pectinid bivalves and neoselachian sharks. Above this are two laterally impersistent units that were not noted in previous studies. The lower of these is a very hard grainstone composed of finely comminuted shell fragments but with no recognizable fossils. There is also a laterally discontinuous black clay with thin partings of ooids but no obvious fossils. Above these units the remainder of the Ardley Member comprises pale soft limestones and marls, which are variably bioturbated and oolitic. The lithology is rather variable laterally and the beds numbered by Horton et al. (1995) cannot often be recognized with certainty.

The Bladon Member (Fig. 4.15) is thin, usually less than 1 m thick, and the base is sharply defined in some places but gradational in others. The lower part of this, Bed 23 of Horton et al. (1995), is laterally persistent but overlying beds are not. Bed 23 comprises a pale grey to almost white, massive clay, marl or impure limestone within which the degree of lithification is highly variable. This is the *Fimbriata-waltoni* Bed of previous studies (Horton et al., 1995). Small aragonitic bivalves (mostly *Corbula*) as well as naticid, cerithiid and planorbid gastropods are common in the more calcareous parts, but are present as moulds elsewhere. Small dark fragments of both fusinite and lignite are common throughout. In most places this is overlain by a hard pale, marl with numerous small rootlets. The contact between these is in places seen to be highly undulose and there is often a thin heterolithic or lamninated band at the contact. This upper unit was apparently not recognised by Horton et al. (1995). Above these two beds are two additional units that are discontinuously present. These appear to correspond to Beds 24 and 25 of Horton et al. (1995), although their descriptions only loosely match the observed lithologies. One unit comprises a dark grey to black clay. This has an irregular base and appears to be associated with the rootlets seen below it. In places a thin limonitic crust is developed at the base with small irregular quartz fragments present. The other unit is a pale nodular micrite with *Epithyris*. The latter appears to correspond to eroded remnants of the Coral-*Epithyris* Limestone of other Bathonian sites in the region (McKerrow et al., 1969). These uppermost units are not seen in contact (although it is inferred that they do by Horton et al., 1995), but one section showed the black rooted clay overlying a very thin weathered limestone containing small calcrete nodules. It therefore appears that the *Epithyris*-bearing unit (Bed 25) is overlain by the dark clay (Bed 24).

Forest Marble Formation

The Forest Marble Formation overlies the White Limestone Formation with a clearly undulose, erosive base. The exposed part of the Forest Marble Formation is composed almost entirely of limestone facies, although mudstone-dominated facies have been mapped very close to Woodeaton Quarry (Fig. 4.12). The limestones are strongly cross-stratified and bed thicknesses and geometries are extremely variable. The limestone lithology varies from packstone to grainstone and the grains vary from largely ooids with rare bioclasts to almost entirely comminuted shell fragments. Mudstone intraclasts are

present in places. Some bedding surfaces have diverse assemblages of bivalves and echinoderms but generally the diversity appears low. Thin drapes of yellow to grey clay are present between some of the limestone units but remain volumetrically insignificant. Acid digestion of the limestones yielded a diverse vertebrate assemblage including neoselachian and hybodont sharks, osteichthyan fish teeth and rare tetrapod fragments.

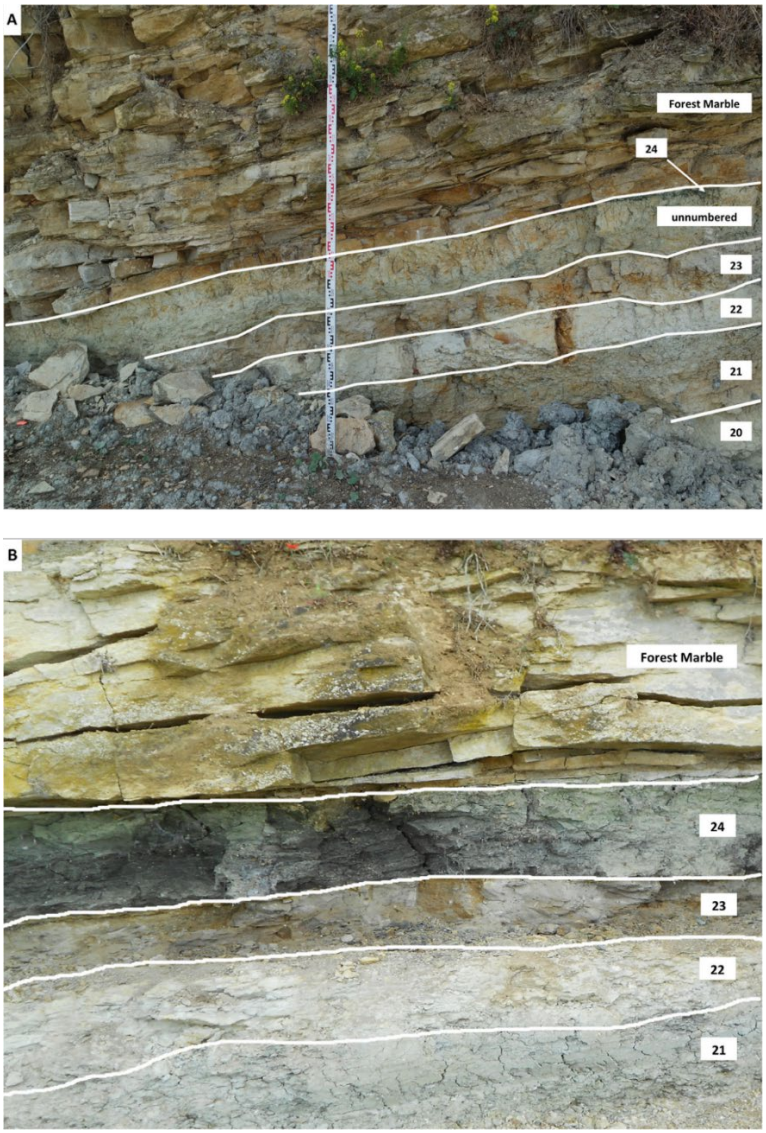


Fig. 4.15. Sections through the Bladon Member of the White Limestone Formation (Bathonian, Middle Jurassic) to the Forest Marble Formation (Bathonian, Middle Jurassic) at Woodeaton Quarry, Oxfordshire. Bed numbers correspond to those shown on Fig. 4.13. A. Section W2 showing the almost complete removal of bed 24 by the erosion surface at the base of the Forest Marble Formation, and the addition of a small channel deposit between beds 23 and 24 penetrated by roots from the overlying bed 24. B, section W1 beds 21 to 24 of the Bladon Member overlain by flaggy cross stratified limestone of the Forest Marble Formation.

Vertebrate palaeontology

Vertebrate remains are present in various horizons at Woodeaton Quarry. Large dinosaur bones have previously been recovered from both the Rutland Formation and White Limestone Formation. Palmer (1973) mentions two large dorsal vertebrae from Bed 5 of the Rutland Formation (the "Monster Bed", Fig. 4.13), referred to the sauropod *Cetiosaurus*. This material is probably in the Oxford University Museum collections (Palmer, 1973) but could not be located by the authors. Previous work conducted by the Natural History Museum recovered the partial remains of a sauropod (currently awaiting formal description) at a similar level in the Rutland Formation. Fragmentary large bones of indeterminate dinosaurs are also present in Bed 23 of the White Limestone Formation (Fig. 4.13) along with possible ornithischian remains collected from approximately the same horizon whose current location is unknown. A partial incisor from Bed 5 (the "Monster Bed", Fig. 4.13) of the Rutland Formation was tentatively identified as a mammal (Eric Freeman pers. comm. in Clemens et al. (1979), Evans and Milner (1994) and pers. comm. to H. Ketchum (2017)); however, this cannot be confirmed. Sampling of beds above and below Bed 5 by researchers at University College London in 1983 (Frances Mussett pers. comm. in Evans and Milner (1994) and Underwood in 2002 (C. J. Underwood, pers. comm. 2018), produced no evidence of mammals. Extensive sampling of Bed 23 of the Bladon Member, White Limestone Formation, *Retrocostatum* Zone (Barron et al., 2012; Cox & Sumbler, 2002) as part of the present study (Fig. 4.13) yielded a diverse microvertebrate assemblage including mammals (teeth and edentulous jaws), tritylodontids (teeth and possible vertebrae), dinosaurs (teeth), pterosaurs (teeth), crocodiles (teeth and osteoderms), turtles (carapace fragments), lizards (jaw fragments), albanerpetontids (jaw fragments), salamanders (jaw fragments), frogs (limb elements) and fish (teeth, scales and jaw fragments). In addition, numerous small fragmentary pieces of reptilian eggshell (including dinosaur) have been recovered from this bed and awaiting description. The dinosaur fauna from this bed is described in Chapter Five (theropods) and Chapter Six (ornithischians).

Watton Cliff

Preface: This section is partially based on new field observations made between 2014 and 2016 (Chapter Two).

Watton Cliff lies on the Dorset coast between West Bay and Eype Mouth (Fig. 4.16) and includes the most complete section through the Forest Marble Formation seen in Dorset. The cliff exposes strata from the marls of the Frome Clay Formation at the base of the section, the lateral equivalent of the White Limestone Formation on the Cotswold Shelf (Fig. 4.4), through to the Forest Marble Formation at the cliff top (Fig. 4.16). The section is found on the downthrown side of a fault (the Eype Mouth Fault) bringing Middle Jurassic sediments to the east of the fault into contact with the Lower Jurassic Lias Group (Fig. 4.17). Microvertebrates have been recovered from channelized lenses of sediment within a bioclastic limestone sequence exposed at the site. The initial discovery of mammalian microvertebrates at Watton Cliff was made in the early 1970's by Freeman (1976b) who described the allotherian *Eleutherodon oxfrdensis*. Ensom (1977) also reported a single tritylodontid cynodont tooth from the site. The vertebrate fauna includes cynodont, mammalian, dinosaurian, amphibian, crocodile, fish and shark remains and is slightly younger than the microvertebrate fauna of Kirtlington, Oxfordshire (Benton et al., 2005; Callomon & Cope, 1995; Dineley & Metcalf, 1999; Ensom, 1977; Evans, 1992; Evans & Milner, 1994; Freeman, 1976a; Kermack, 1988; Kermack et al., 1987; Underwood & Ward, 2004).

Watton Cliff is one of the type sections for the Forest Marble Formation, exposing around 25 m of the formation including the basal *Boueti* Bed overlying the Frome Clay Formation (Barron et al. 2012). The section (Fig. 4.18 and Fig. 4.19) comprises a 10 m thick lower sequence of clays, argillaceous limestones and calcareous grits, overlain by 3–5 m of cross-bedded lenticular bioclastic limestones (the microvertebrate horizon), 9 m of interbedded clays and shaley limestones and 3 m of flaggy blue limestone. The sequence at Watton Cliff is well-known and has been described by a number of authors (Callomon & Cope, 1995; Cope et al., 1980; Cope, 2012; Holloway, 1983; House, 1989; Melville & Freshney, 1982; Strahan, 1898; Torrens, 1969a; Wilson et al., 1958; Woodward, 1894). The Forest Marble bioclastic limestones have been placed in the *Clydoniceras discus* Zone of the upper Bathonian (Barton et al., 2011; Bristow et al.,

1995; Hunter & Underwood, 2009; Penn, 1982), with similar lenticular limestone bodies known from the Shaftesbury district (Bristow et al., 1995), Bath (Woodward, 1894) and Yeovil (Kellaway & Wilson, 1941).

The limestone comprises broken and poorly sorted shell debris with ooids, dominated by bivalve remains (*Camptonectes, Plagiostoma* and *Praeexogyra*), and crinoid elements (*Isocrinus nicoleti, Millericrinus cf. exilis, Apiocrinites sp.,* and *Solanocrinites ooliticus*) forming a sequence of sheets and lenses with shallow cross-cutting lenses of uncemented sediment (Fig. 4.20) (Bristow et al., 1995; Cope et al., 1980; Cope, 2012; Holloway, 1983; Hunter & Underwood, 2009). Plant debris is relatively common ranging in size from small fragments up to logs several meters in length. Individual limestone beds generally show a fining upward sequence of shell debris and are often separated by thin clay drapes. The strong influence of tidal action is indicated by herringbone cross bedding with associated mud-drapes formed during tidal ebb/flood cycles. Ripple marked bedding surfaces suggest at times locally emergent or near emergent conditions (Fig. 4.20 and Fig. 4.21) with the bivalve and crinoid faunas typical of hardgrounds (Hunter & Underwood, 2009). The unconsolidated patches in the bioclastic limestone unit lack cement to form a bioclastic gravel, which can often be extracted by hand, and seem to represent either channels or invertebrate burrows. These patches commonly contain water-worn vertebrate material (Benton et al., 2005; Dineley & Metcalf, 1999) 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 and Underwood (2009).

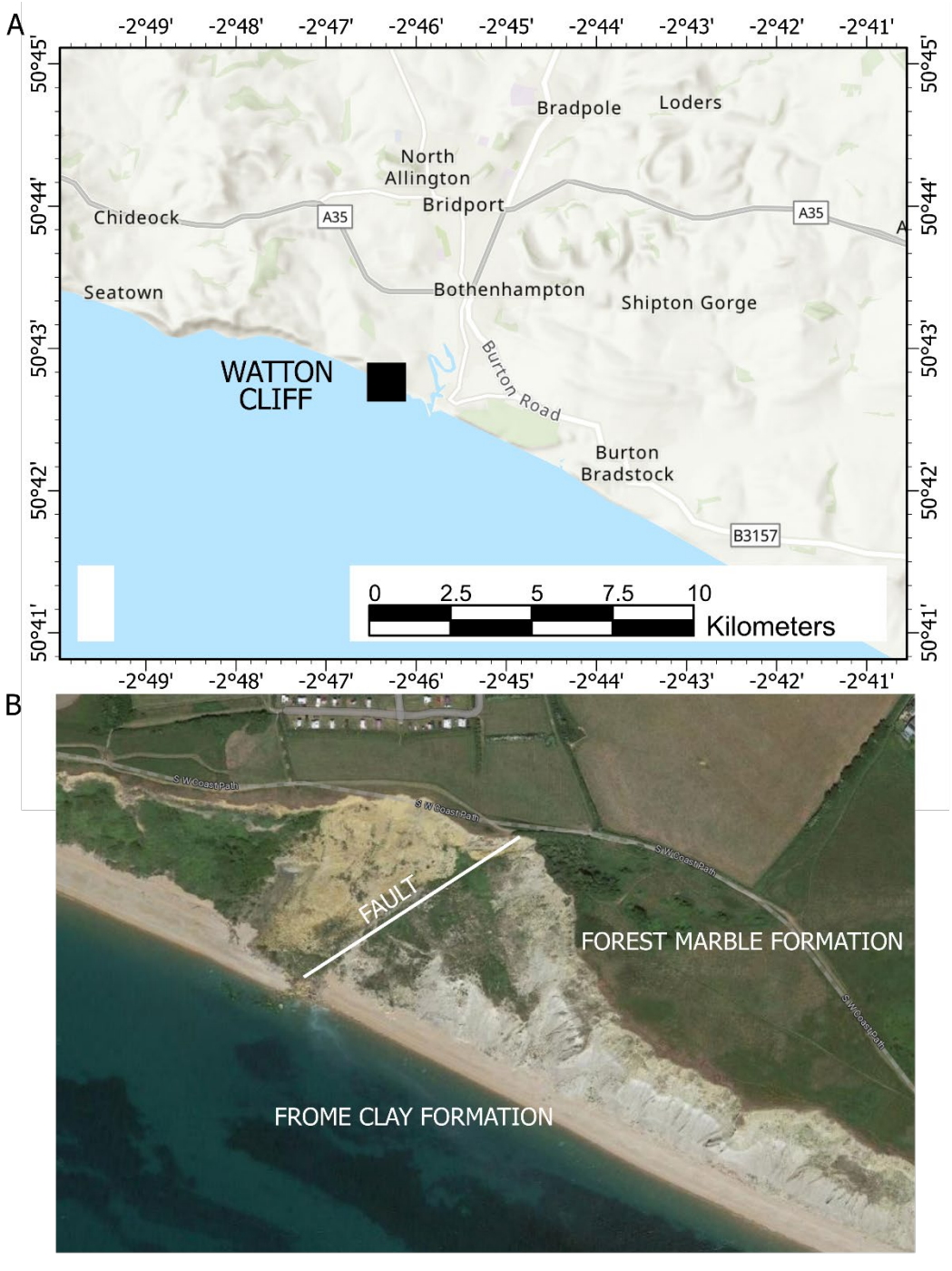


Fig. 4.16. The microvertebrate site at Watton Cliff, Dorset. A, locality map. B, aerial photograph of the site in 2023. Microvertebrates are recovered from fallen blocks of the Forest Marble Formation, seen at the cliff top, on the cliff slope forming the downthrow of the Eype Mouth Fault (eastern side).

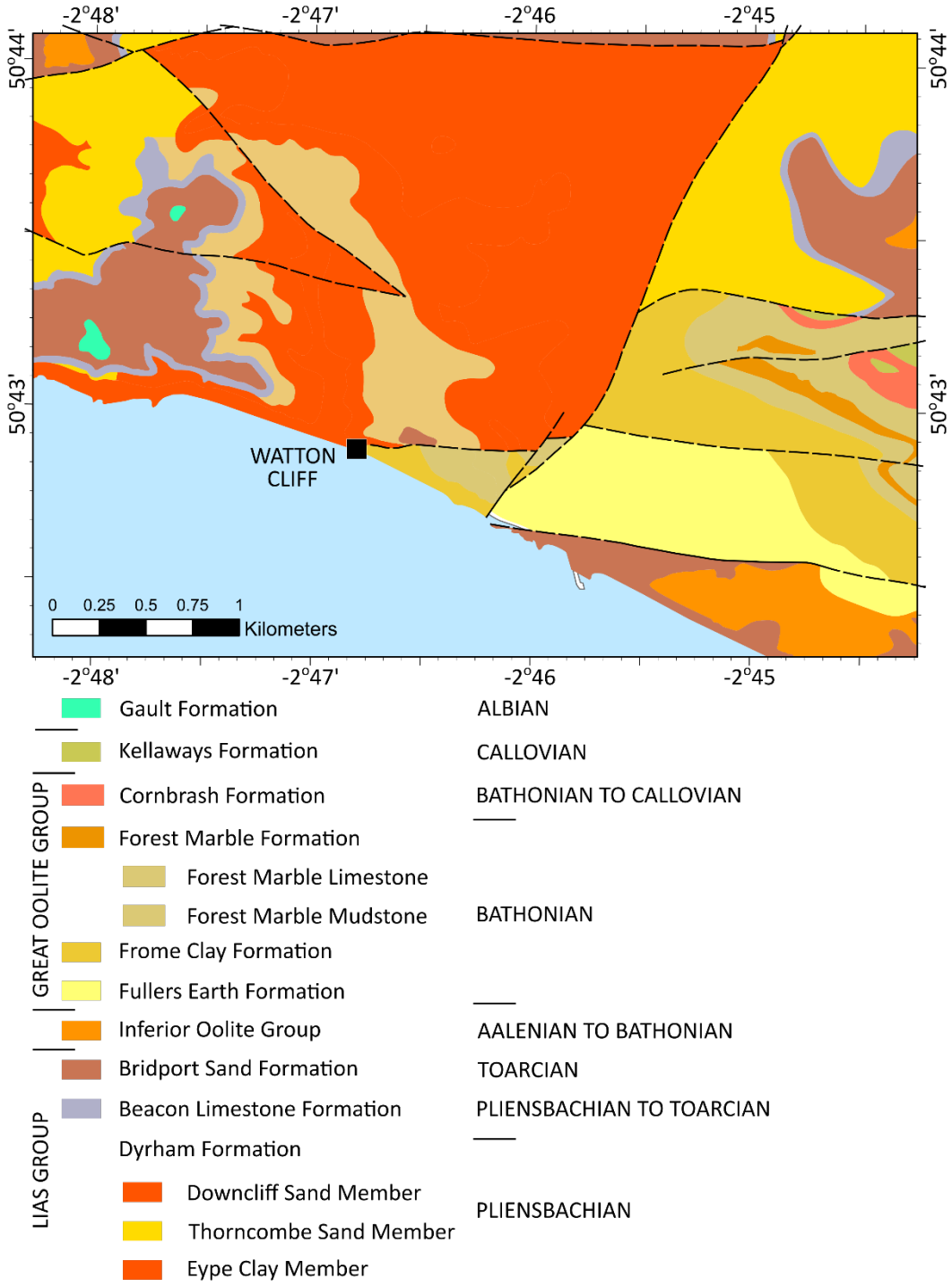


Fig. 4.17. Geological setting of Watton Cliff, West Bay, Dorset. Licence number 2017/024 ED British Geological Survey (c) NERC. All rights reserved.

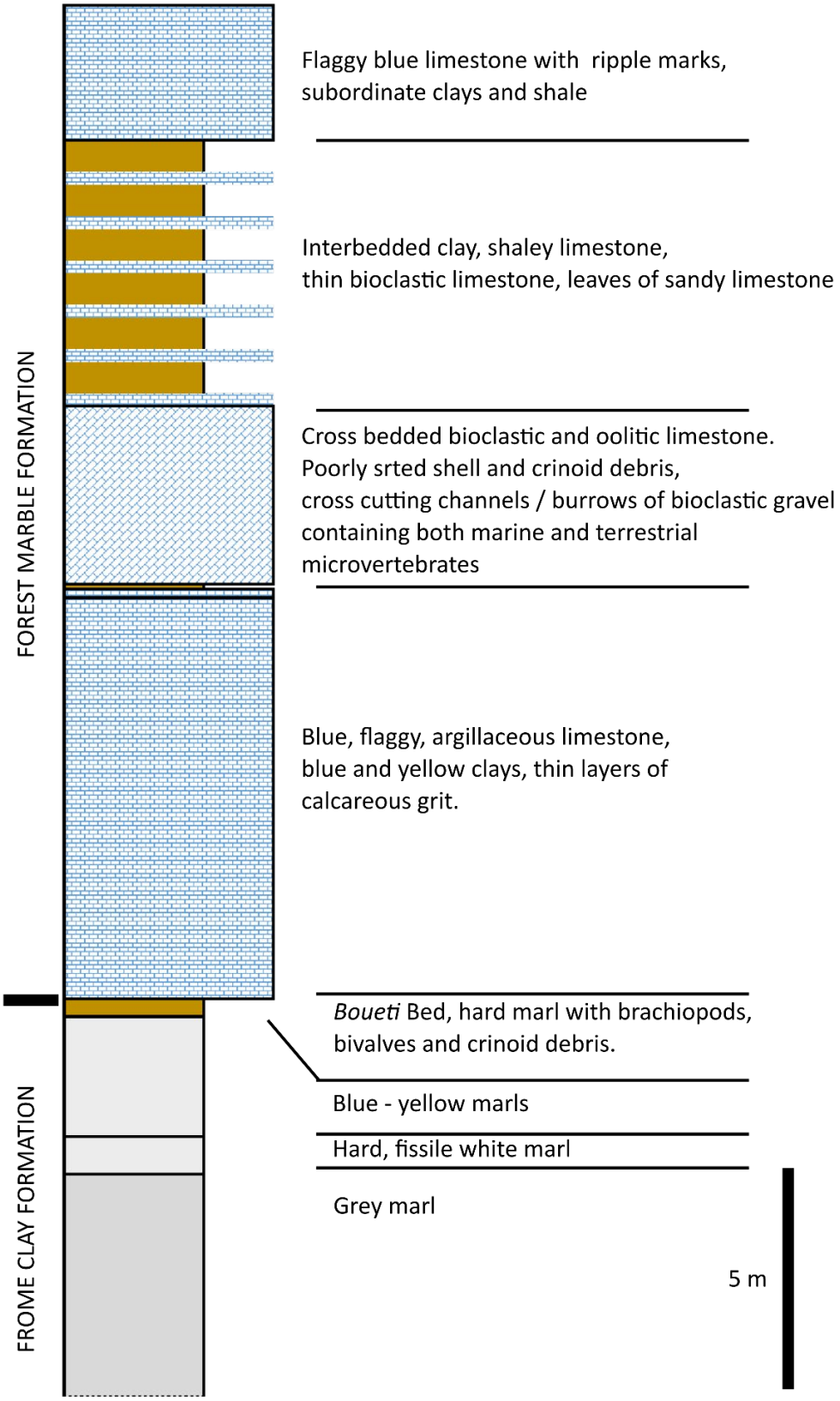


Fig. 4.18. Generalised section through the Middle Jurassic Frome Clay Formation and Forest Marble Formation (Great Oolite Group, Bathonian, Middle Jurassic) at Watton Cliff, Dorset. The section is based on field observation undertaken in 2014 to 2016 and published data (Dineley & Metcalf, 1999; Hunter & Underwood, 2009; West, 2012)

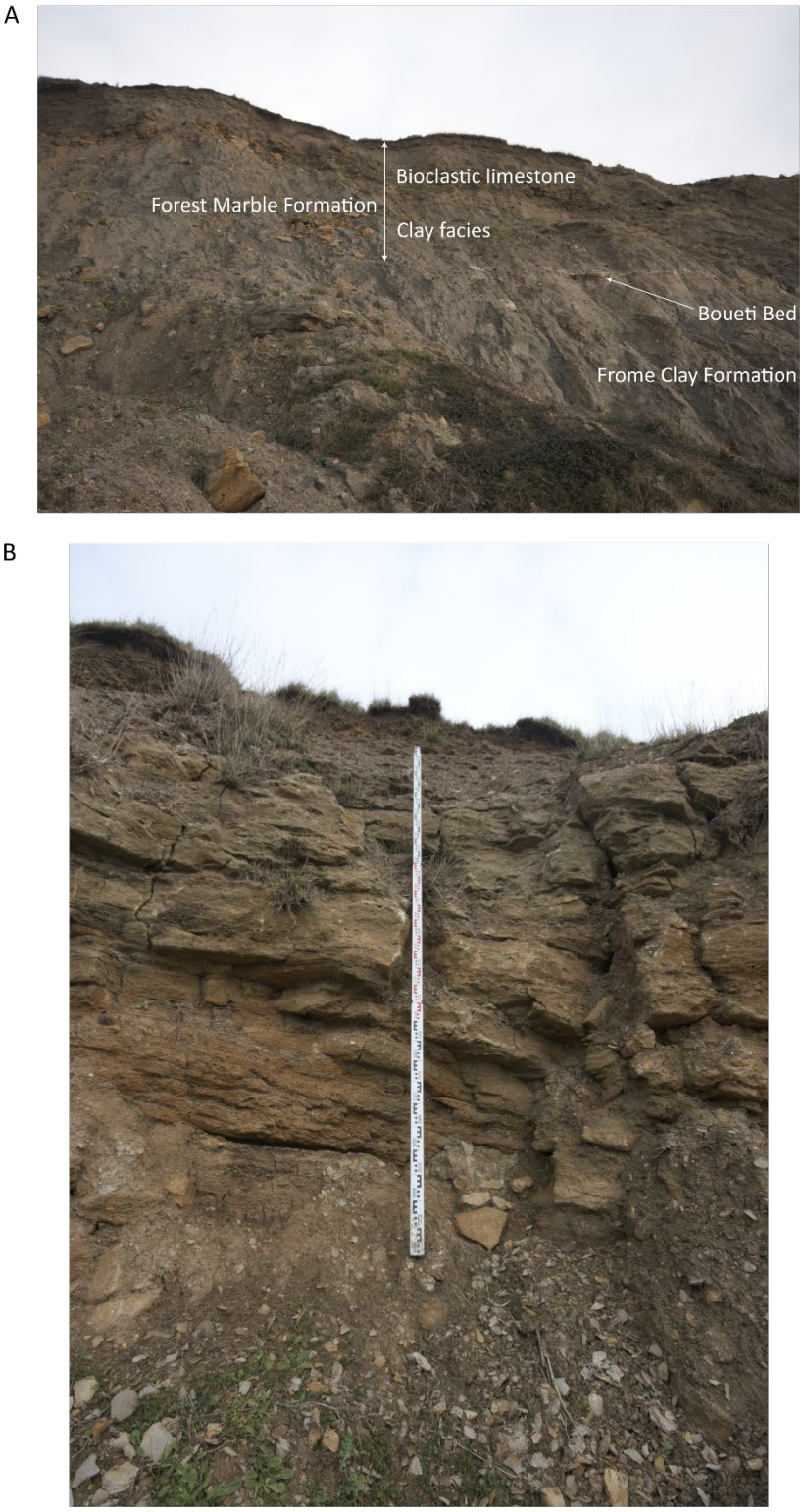


Fig. 4.19. Cliff sections at Watton Cliff, West Bay, Dorset. A, cliff section through the Great Oolite Group with the grey calcareous mudstones of the Frome Clay Formation making up the majority of the cliff with the Forest Marble Formation exposed above the Boueti Bed towards the top of the cliff. B, close-up of the bioclastic limestones of the Forest Marble Formation, containing microvertebrates, exposed at the top of the cliff.

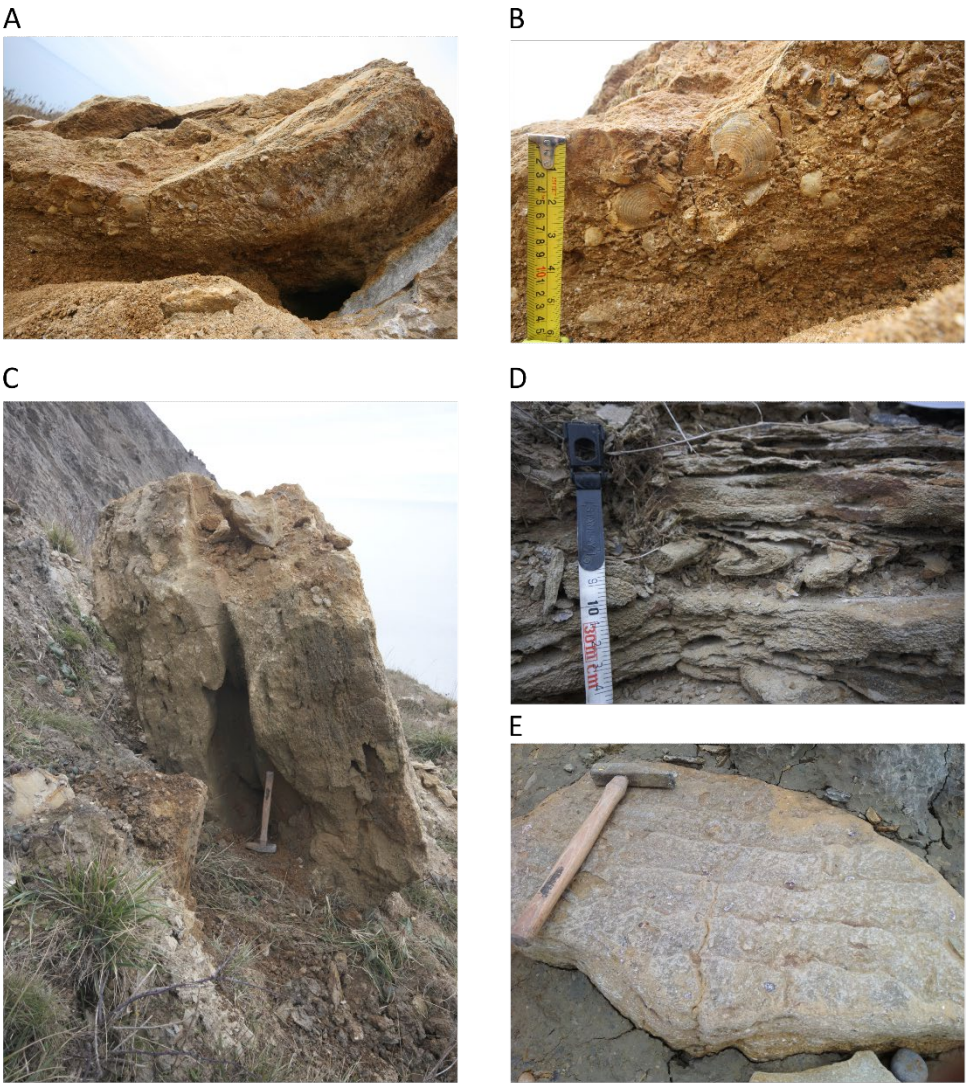


Fig. 4.20. Sedimentary structures in the upper bioclastic limestones of the Forest Marble Formation (Bathonian, Middle Jurassic) at Watton Cliff, Dorset. A, lens of unconsolidated bioclastic gravel within an otherwise cemented unit of bioclastic limestone. B, close-up of the lens in A showing the bivalve and crinoid faunas developed on a hardground. C, larger lens of bioclastic gravel (excavated) possibly representing a large invertebrate burrow. D, mud drapes and rip-up clasts in the bioclastic limestone. E, fallen block of Forest Marble with a ripple marked surface.

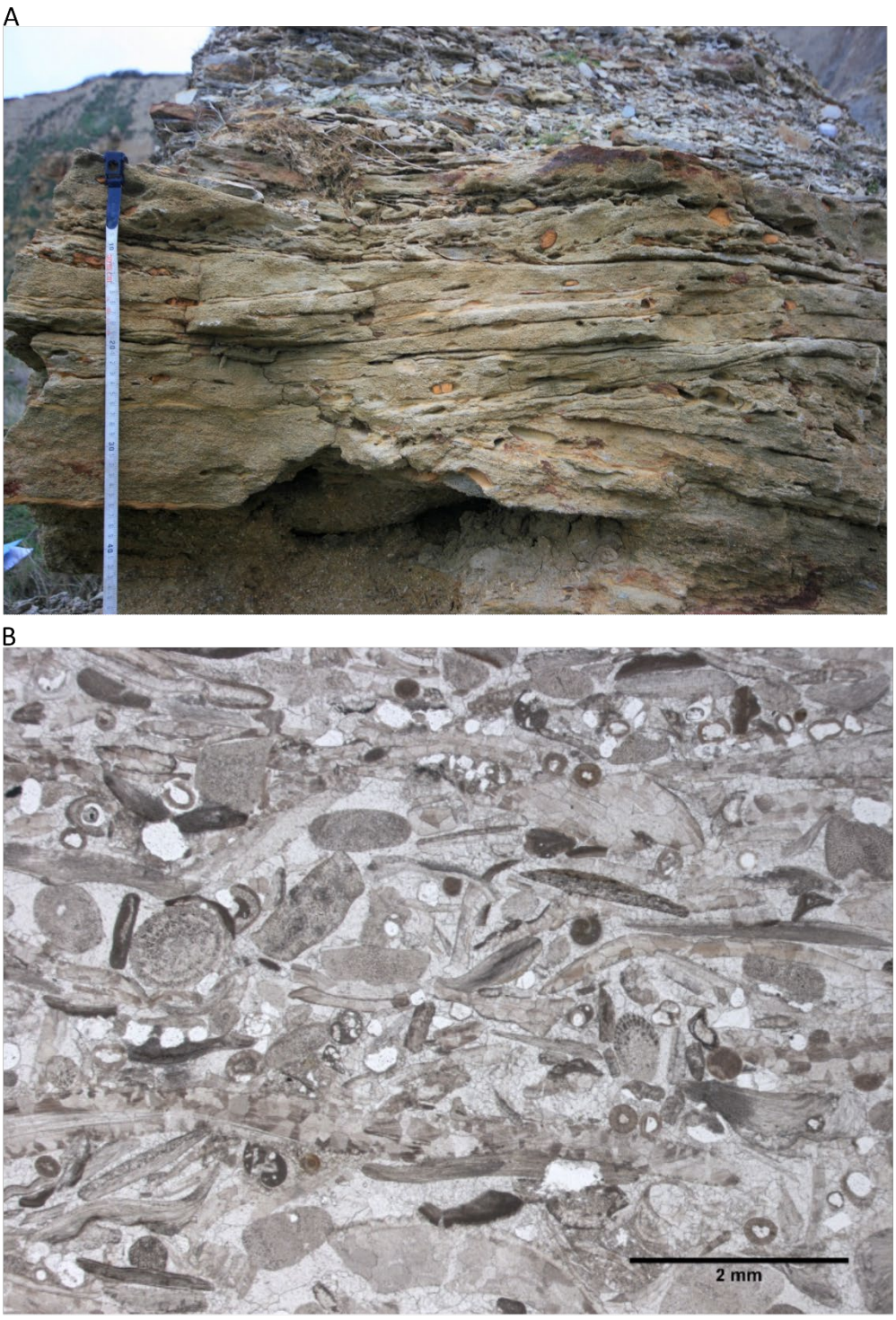


Fig. 4.21. Sedimentary structures and lithology of the Forest Marble Formation (Bathonian, Middle Jurassic), Watton Cliff, Dorset. A, herringbone cross-bedding and mud rip-up clasts in the Forest Marble Formation. B, thin-section of cemented Forest Marble bioclastic limestone in plane polarised light with bivalve and crinoid debris.

Discussion

The sections described herein demonstrate the extreme lateral and vertical variability of the Rutland and White Limestone Formations. Historically, placing the boundaries between the Rutland, White Limestone and Forest Marble formations has been a matter of considerable debate. Horton et al. (1995) placed the boundary between the Rutland and White Limestone formations at an erosive surface between a mudstone unit (Fig. 4.13, Bed 8 of the Rutland Formation) and an overlying thin shelly detrital marl, presumably reflecting the gradual change towards the more marine conditions of the White Limestone Formation. It is often difficult to distinguish between these two units in the field due to lateral variation across the section and the similarity in lithology between the two units. Consequently, I concur with Palmer (1973) and include the lowermost Shipton Member unit of Horton et al. (1995) in the Rutland Formation rather than the White Limestone Formation (Fig. 4.13).

The boundary placement between the White Limestone and Forest Marble formations has had a similarly tortuous history. This is in part due to the variability of the formations both vertically and laterally across the area. Horton et al. (1995) place the boundary above the 'Upper *Epithyris* Bed' (Fig. 4.13, Bed 25) agreeing with many earlier authors (Hull, 1859; Odling, 1913; Palmer, 1973; Woodward, 1894) whereas McKerrow et al. (1969) and Benton et al. (2005) include most of the Bladon Member (White Limestone Formation) in the Forest Marble Formation. I note that Beds 24 and 25 (Fig. 4.13) of Horton et al. (1995) appear to be positionally reversed in the section measured at Woodeaton with Bed 24 overlying Bed 25 (the 'Upper *Epithyris* Bed'). This appears to correspond to the section measured at Kirtlington Quarry by Arkell (1931) who notes a grey clay of varying thickness overlying the 'Upper Epithyris Bed'. Consequently, I concur with Arkell (1931) in placing the White Limestone and Forest Marble formation boundary at the sharply erosive and undulose contact at the base of the strongly cross stratified, oolitic and bioclastic limestones (Fig. 4.13), with the Forest Marble Formation limited on lithostratigraphic grounds to this facies above the contact.

The microvertebrate fauna recovered from Bed 23 (Bladon Member, White Limestone Formation, Fig. 4.13) is comparable to that known from the approximately coeval section at Kirtlington Quarry, Oxfordshire (Evans & Milner, 1994). McKerrow et al. (1969) placed the microvertebrate horizon at Kirtlington (the 'Mammal Bed', Bed 3p of

McKerrow et al. (1969) in the upper part of the 'Upper *Epithyris* Bed' (the equivalent to Bed 25 at Woodeaton Quarry, Fig. 4.13), whereas Freeman (1979) and Cripps (1986) suggested that it forms a distinct deposit lying above this unit, directly below the overlying Forest Marble Formation. At Kirtlington Quarry, Benton et al. (2005) place the 'Middle *Epithyris* Bed' of Arkell (1931) and overlying beds, including the 'Mammal Bed' in the Forest Marble Formation. The definition of the White Limestone and Forest Marble formation boundary adopted above places these beds in the White Limestone Formation. This implies that the Bed 23 microvertebrate horizon at Woodeaton is slightly older than Kirtlington.

Most British Middle Jurassic microvertebrate deposits represent shallow brackish to freshwater ponds or lakes, or marginal marine environments, formed on emergent carbonate platforms or restricted shallow lagoons and are geographically restricted with a limited aerial extent (Evans & Milner, 1994; Metcalf et al., 1992; Metcalf & Walker, 1994). For example, the microvertebrate rich clay at Hornsleasow Quarry formed in a small spatially restricted pond on a palaeokarst surface with subsequent soil development and drying out of the deposit (Metcalf, 1995). The presence of hardgrounds, calcrete nodules, palaeosols and karstic topography formed at the terrestrial sedimentary horizons suggests aerial exposure at these sites in the order of tens of thousands of years (Stockmann et al., 2014). The main microvertebrate bearing unit at Woodeaton (*fimbriata-waltoni* Bed 23) appears to be an exception to the geographically restricted extent of similar deposits in the Bathonian of the UK. This unit is traceable across the entire quarry face (where exposed) at Woodeaton Quarry and is also known regionally, although to date terrestrial microvertebrates have only been recovered from this horizon at Woodeaton (e.g. Horton et al., 1995; McKerrow et al., 1969; Palmer, 1979). The facies of this unit suggests that it represents a larger scale, brackish water lagoon hemmed in by some form of barrier. Invertebrate and vertebrate biotas suggest a fluctuating salinity so it is likely that periodic influxes of seawater during tides flooded into the area that otherwise had constant supply of freshwater re-supply from the land. Poorly developed calcrete nodules elsewhere in the succession and abundant evidence of biomass burning may suggest at least seasonal aridity. A detailed taphonomic analysis is required to confirm these findings and to underpin any further analyses (other than taxonomic) on the fossils recovered from this bed.

Conclusions

The sections described herein clarify the boundaries between the Rutland, White Limestone and Forest Marble formations. I place the lowermost Shipton Member unit of Horton et al. (1995) in the Rutland Formation rather than the White Limestone Formation. The boundary between the White Limestone Formation and overlying Forest Marble Formation is likewise clarified in placing the White Limestone and Forest Marble formation boundary at the sharply erosive and undulose contact at the base of the strongly cross stratified, oolitic and bioclastic limestones, with the Forest Marble Formation limited on lithostratigraphic grounds to this facies above the contact. This places the 'Mammal Bed' at Kirtlington Quarry at the top of the White Limestone Formation rather than in the Forest Marble Formation contra Benton et al. (2005). The extent of the microvertebrate horizon at Woodeaton Quarry suggests that it represents a large scale depositional environment, unlike the more spatially restricted deposits at Hornsleasow and Kirtlington, and is positioned stratigraphically lower than the approximately coeval 'Mammal Bed' at Kirtlington Quarry.

Chapter Five: Theropoda

Preface: This chapter has been published as:

Wills, S., Underwood, C. J., & Barrett, P. M. (2023). Machine learning confirms new records of maniraptoran theropods in Middle Jurassic UK microvertebrate faunas. Papers in Palaeontology, 9(2), e1487. https://doi.org/https://doi.org/10.1002/spp2.1487

I conducted all of the data collection and analyses presented in this paper and the majority of it is my own original work. Barrett and Underwood provided discussion and edited the manuscript.

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, I apply different machine learning models, in conjunction with morphological comparisons, to a suite of isolated theropod teeth from Bathonian microvertebrate sites in the UK, in order to determine if any of these can be confidently assigned to Maniraptora, as had previously been suggested based on morphology alone. I generated three independent models developed on a training dataset with a wide range of theropod taxa and broad geographical and temporal coverage. Classifying the Middle Jurassic teeth in my sample against these models, and undertaking morphological comparisons, 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 at least some 27 million years. In addition to testing the suite of theropod teeth from the UK I also applied the same models to potential Middle Jurassic maniraptoran teeth described from India, Kyrgyzstan and Madagascar, the results of which confirm their maniraptoran affinities. These results indicate that not only were maniraptorans present in the Middle Jurassic, as predicted by previous phylogenetic analyses, but had already radiated into a diverse global 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.

Introduction

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 their initial radiation likely occurred in the Middle Jurassic (Carrano et al., 2012; Holtz, 2000; Rauhut, 2003; Rauhut & Foth, 2020; Xu, Choiniere, et al., 2010). 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*, a potential therizinosaur whose identification and dating remains contentious (Barrett, 2009; Kirkland & Wolfe, 2001; Xu et al., 2001). Nevertheless, fragmentary, generically indeterminate remains of some maniraptoran sub-clades, such as possible dromaeosaurs, have been reported from Middle Jurassic microvertebrate sites in Europe, Asia and Africa (Averianov et al., 2005; Evans & Milner, 1994; Maganuco et al., 2005; Metcalf & Walker, 1994; Prasad & Parmar, 2020), including the four British Bathonian sites under study in this thesis. 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, 2010a; Ding et al., 2020; Foth &

Rauhut, 2017; 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, 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; Peterson et al., 2014). 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; Gates et al., 2015; Larson & Currie, 2013). The comparatively simple structure of theropod teeth has made identifications difficult historically, as 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 potential solutions to this problem (e.g., Chiarenza et al., 2020; Currie et al., 1990; Farlow et al., 1991; Gerke & Wings, 2016; Hendrickx & Mateus, 2014; Hendrickx et al., 2019; Hendrickx et al., 2020; Larson, 2008; Larson & Currie, 2013; Smith et al., 2005; Williamson & Brusatte, 2014; Young et al., 2019). In Chapter Three, I show how the use of machine learning procedures can produce accurate group-discrimination when applied to morphological data (Hoyal Cuthill et al., 2019; MacLeod & Kolska Horwitz, 2020). I applied this technique to a diverse test sample of theropod teeth and demonstrated that these methods lead to higher classification accuracies than more traditional statistical analyses (Wills et al., 2021). Here, I apply these new machine learning methods to a large sample of isolated theropod teeth from the series of UK Middle Jurassic microvertebrate sites described in Chapter Four, and In addition I also re-evaluate possible maniraptoran theropod teeth from the Middle Jurassic of India (Prasad & Parmar, 2020), Madagascar (Maganuco et al., 2005) and Kyrgyzstan (Averianov et al., 2005).

Using these machine learning and morphological-based approaches I 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 globally 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.

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. 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 comprised open marine conditions in the South, a lagoon and mudflat complex in the North and a series of oolitic shoals separating these. Sea-level fluctuations throughout the Bathonian often caused pauses in marine sedimentation with occasional localised emergence accompanied by the development of hardgrounds, palaeosols, and terrigenous sediment influxes (Barron et al., 2012; Hesselbo, 2008; Horton et al., 1995; Palmer, 1979; Palmer & Jenkyns, 1975; Underwood, 2004; 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, the focus of this study, have yielded large numbers of small vertebrate remains, including sharks, bony fish, mammals, turtles, crocodilians, choristoderes, pterosaurs, squamates and amphibians (Evans & Milner, 1994; Freeman, 1976a, 1976b, 1979; Metcalf et al., 1992; Metcalf & Walker, 1994; Wills et al., 2014;

Wills et al., 2019). Dinosaur teeth are common and some of these were referred tentatively to various coelurosaurian theropod clades (Evans & Milner, 1994; Freeman, 1976a, 1976b, 1979; Metcalf et al., 1992; Wills et al., 2014; Wills et al., 2019).

For a full description of the local geological setting of Hornsleasow, Kirtlington and Woodeaton Quarries and Watton Cliff see Chapter Four. The geological settings of the teeth described from India, Kyrgyzstan and Madagascar are outlined briefly, below.

India: Prasad and Parmar (2020) described isolated theropod and ornithischian teeth from the upper part of the Jurassic Kota Formation of the Pranhita-Godavari Valley in the Adilabad district, Telangana, India. Before the breakup of Gondwana, a series of sedimentary basins formed in intra-continental rift structures across what is now the Indian subcontinent. The Kota Formation, in the fluvial dominated Pranhita-Godavari Basin (late Carboniferous to Cretaceous), is a series of terrestrial sediments including sandstones, clays and limestones that has been variously dated from Early Jurassic to Early Cretaceous (Dasgupta, 2021; Prasad & Parmar, 2020). The age is controversial as there are no radiometric dates associated with the formation with its lower and upper boundaries being constrained by the Upper Dharmaram Formation (Lower Jurassic) and the Gangapur Formation (Lower Cretaceous). Various authors have attempted to date the Kota Formation by using faunal and floral comparisons with other terrestrial Jurassic deposits. Estimates range from the Early Jurassic based on the fish and squamate fauna, charophytes and palynofauna (Bhattacharya et al., 1994; Evans et al., 2002; Jain, 1980; Prabhakar, 1989) to Middle Jurassic based on freshwater ostracods and the dinosaur microvertebrate fauna (Govindan, 1975; Prasad & Parmar, 2020). A recent review indicates a latest Early Jurassic to Middle Jurassic age for the Kota Formation (Dasgupta, 2021) suggesting that the teeth described by Prasad and Parmar (2020), occurring in the upper part of the formation, are Middle Jurassic in age.

Kyrgyzstan: Averianov et al. (2005) recovered isolated theropod teeth by screen-washing sediment at microvertebrate sites in the Balabansai Formation near Tashkumyr, Jalal-Abad, Kyrgyzstan. The Balabansai Formation is a Bathonian–Callovian terrestrial to marginal marine deposit comprising sandstones, siltstones and clays unconformably overlain by red beds of the Lower Cretaceous Hodzhiabad Formation (Averianov et al., 2005; Jolivet et al., 2018; Nessov et al., 1994). As with the Kota Formation, the dating of

the Balabansai Formation relies on faunal comparisons with other terrestrial Jurassic deposits (Averianov et al., 2008; Evans et al., 1988; Martin & Averianov, 2004).

Madagascar: Isolated theropod teeth from the Mahajanga Basin of Madagascar (Maganuco et al., 2005) have been recovered from the "Isalo IIIb-Bathonien Facies Mixte Dinosauriens" a sequence of terrestrial sandstones, siltstones and claystones now included in the Sakaraha Formation and considered to be a Bathonian coastal plain environment (Geiger et al., 2004). The formation contains a rich, but poorly known, terrestrial vertebrate fauna including dinosaurs, mammals and crocodylomorphs (Burmeister et al., 1999; Flynn et al., 1998; Flynn et al., 1999; Maganuco et al., 2005; Young et al., 2015).

Materials and Methods

The material consists of isolated theropod teeth collected from the four Middle Jurassic (Bathonian, Great Oolite Group) localities in the UK: 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). I initially identified 164 isolated theropod teeth (Kirtlington 49, Hornsleasow 50, Watton Cliff four, Woodeaton 61) of which 149 were complete enough to warrant further investigation.

All teeth in the sample were processed as outlined in Chapter Two, with a combination of optical imaging using a Dino-Lite AM 7915 MZTL microscope and SEM on a LEO 1455VP microscope. I CT-scanned each (complete) tooth on a Nikon Metrology HMX ST 225 µCT and a Zeiss Versa µCT at a range of voxel resolutions from 4–30 µm and created 3D models from the CT volumes using Avizo (ThermoFisher 2014) and a range of Python, Matlab and Fiji scripts. See Chapter Two, for a full description of the process and Appendix Three for the processing scripts.

Five morphometric variables were collected from each tooth, which were measured directly from the images using Fiji v. 2.1 (Schindelin et al., 2012) and from the 3D models using Avizo v. 8.1 (ThermoFisher, 2014). The measurements are simple 2D linear distances (Fig. 5.1) 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). Where 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. Where required, the crown base ratio (CBR) is calculated as CBW/CBL and the denticle size density ratio (DSDI), a measure of the size difference between mesial and distal denticles (Rauhut & Werner, 1995), as MDM / DDM.

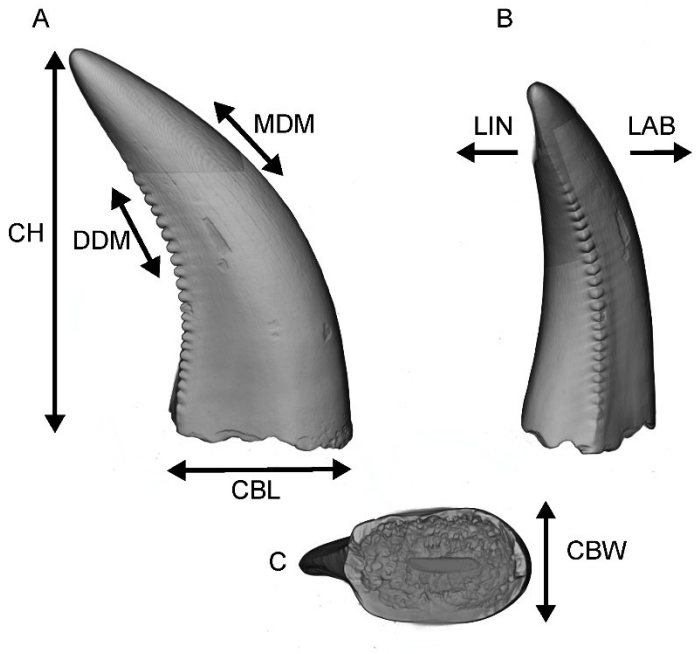


Fig. 5.1. Anatomical and morphometric terminology of theropod tooth crowns. Theropod tooth crown in labial (A), distal (B) and basal (C) views. Abbreviations: CH, crown height; CBL, crown base length; CBW, crown base width; MDM, mesial denticles per millimetre; DDM, distal denticles per millimetre; LIN, lingual; LAB, labial. After Hendrickx et al. (2019) and Wills et al. (2021).

Although other approaches, such as 3D data, are available (Hoyal Cuthill et al., 2019; MacLeod & Kolska Horwitz, 2020; Wills et al., 2021) I chose to use these 2D linear measurements as these variables are common to most published analyses of isolated theropod tooth datasets (e.g., Currie et al., 1990; Farlow et al., 1991; Hendrickx et al.,

2015b; Hendrickx et al., 2019; Hendrickx et al., 2020; Larson, 2008; Larson et al., 2016; Larson & Currie, 2013; Noto et al., 2022; Sankey et al., 2005; Williamson & Brusatte, 2014; Young et al., 2019), 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., Brusatte & Clark, 2015; Gates et al., 2015; Hendrickx et al., 2020; Larson & Currie, 2013; Williamson & Brusatte, 2014) and the machine learning analysis described in Chapter Three (Wills et al., 2021). Given this and the lack of comparative digital image based theropod tooth datasets I feel the approach I have taken is appropriate. This approach also allowed me to analyse the isolated teeth from microvertebrate sites in India, Madagascar and Kyrgyzstan using morphometric data in the published literature (Averianov et al., 2005; Maganuco et al., 2005; Prasad & Parmar, 2020). Morphometric measurements from a total of 63 isolated theropod teeth from these sites were included in the analysis: 35 from Madagascar, 21 from India and seven from Kyrgyzstan.

To determine the taxonomic identifications of the teeth I undertook a quantitative analysis of morphometric data using a mixture of machine learning models following the methodology outlined in Chapter Three (Wills et al., 2021). I employed 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. These techniques were chosen as they consistently gave the highest classification accuracies of all the methodologies tested in Chapter Three. The three models differ in their approach to learning, allowing me 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 sub-class distributions, RF is an ensemble comprised 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 (Breiman, 2001; Hastie & Tibshirani, 1996; 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 I combined several published datasets (Currie & Varrichio, 2004; Farlow et al., 1991; Gerke & Wings, 2016; Hendrickx et al., 2015a; Larson, 2008; Larson et al., 2016; Larson & Currie, 2013; Longrich, 2008; Rauhut et al., 2010; Sankey, 2008; Sankey et al., 2002; Smith, 2005; 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. 5.2). See Appendix Two for a summary of the data used, taxonomic groups chosen, and sample sizes employed in the analysis.

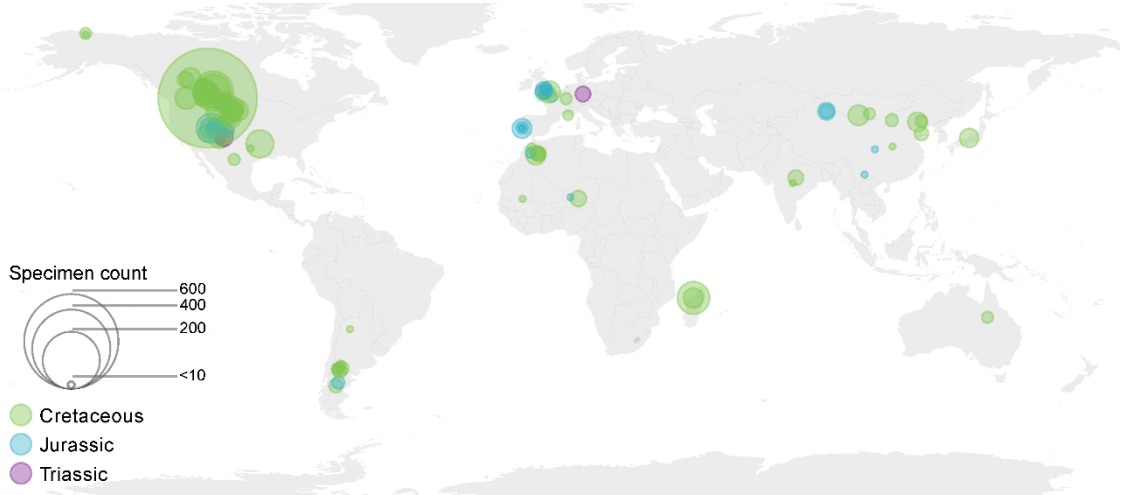


Fig. 5.2. Spatial and temporal distribution of training data samples (theropod tooth measurements) used to develop the machine learning models. Outliers and cases with missing data have been removed. Data sourced from: Currie and Varricchio (2004), Farlow et al. (1991), Gerke and Wings (2016), Hendrickx et al. (2015a), Larson (2008), Larson et al. (2016), Larson and Currie (2013), Longrich (2008), Rauhut et al. (2010), Sankey (2008), Sankey et al. (2002), Smith et al. (2005), Young et al. (2019).

Different definitions have been applied to these morphometric variables with Smith et al. (2005) and Hendrickx et al. (2015a) differing in their methods for measuring CBL and CH (Fig. 5.1) and I used the corrected data provided by Gerke and 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 outlined in Chapter Three and reported by Wills et al. (2021). Prior to training these models the data were cleaned to improve model performance. Firstly, I removed any outliers using a density-based spatial clustering algorithm (DBSCAN, Fig. 5.3) 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 impacts on the model accuracy and the resultant classification. Secondly, I removed any classes with fewer members than the number of predictive variables (five), and finally I removed cases with missing data as 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 one to allow the transformation of zero values), scaled and centred prior to analysis. I 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 as described in Chapter Three (Wills et al. 2021). From an initial dataset of 3,886 specimens, data cleaning resulted in a final set of 1,702 usable cases. I undertook an initial exploration of clade feature space on 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). I used both techniques as 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.

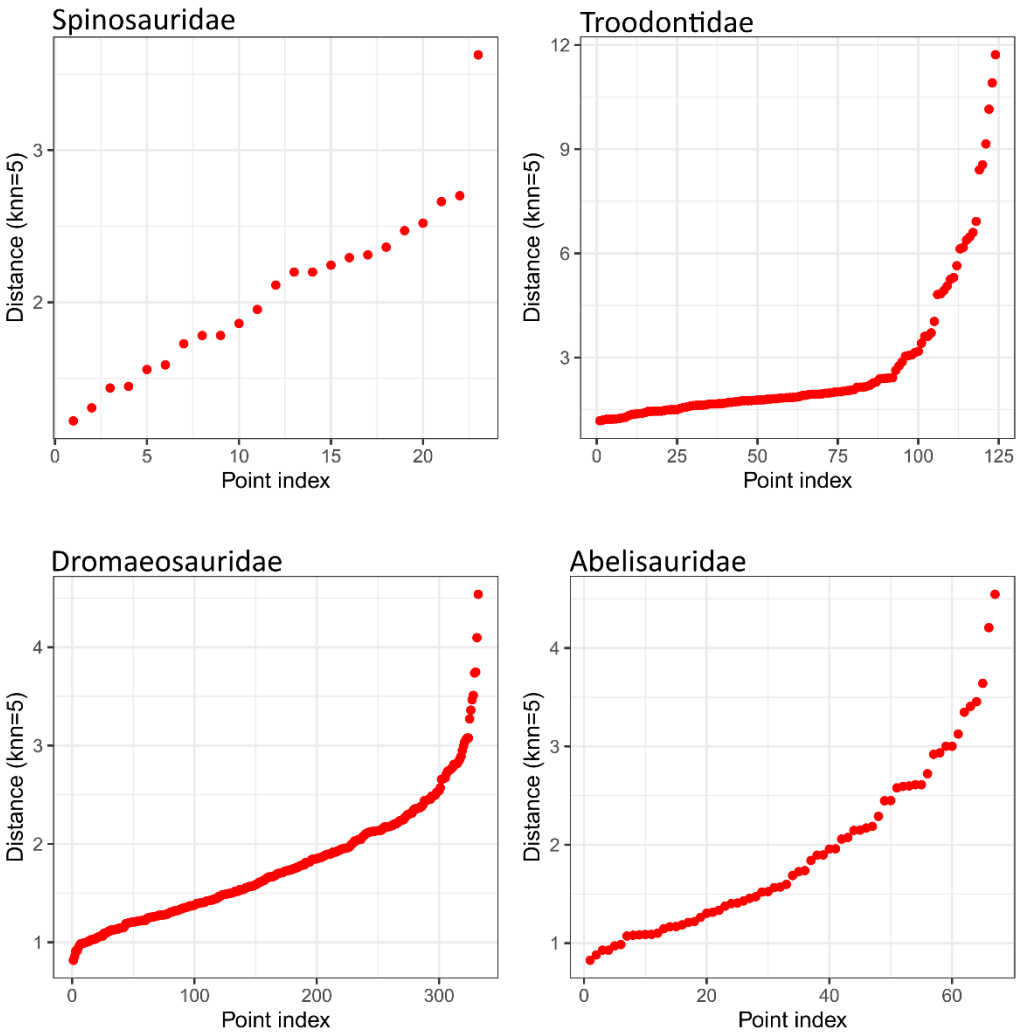


Fig. 5.3. An example of the DBSCAN, density-based spatial clustering algorithm, applied to four clades in the training dataset to detect outliers in the data. The algorithm orders the individual data points based on the distance to their five nearest neighbours. Data points plotting after the inflexion point in the graphs are treated as outliers and should be examined and possibly removed from models.

I undertook a series of non-parametric statistical analyses, permutational multivariate analysis of variance (PERMANOVA) using the Mahalanobis distance (Anderson, 2017; Anderson & Walsh, 2013), 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. As PERMANOVA only tests whether all the centroids in the data are equal I 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 (1,364 cases) and a testing dataset (338 cases). The models were developed on the training data and then assessed against the testing data. Testing data was 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. I used k-fold cross validation on the training set with 10 folds to give an overall model accuracy (Fig. 5.4). I also ran each model permutation using a range of tuning parameters to obtain the highest accuracy. For MDA I modelled the response using a range of sub-classes, from one to eight, for each taxonomic class; the random forest model was tuned by varying the random subset of predictors that the model uses at each split in the tree (mtry parameter) from two to five and I grew the forest to 2,000 trees to ensure stability; and for the C5.0 model I varied the number of model iterations from 1–100 and used both rule and tree based classifier models (Kuhn & Johnson, 2013b; Wills et al., 2021). In addition to the predicted class generated from the models I 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 randomisation input into the training. As a result, there may be slightly different results obtained from different training cycles of the models. A more detailed description of the techniques involved and descriptions of the differences between the machine learning algorithms can be found in Chapter Three, also see Wills et al. (2021).

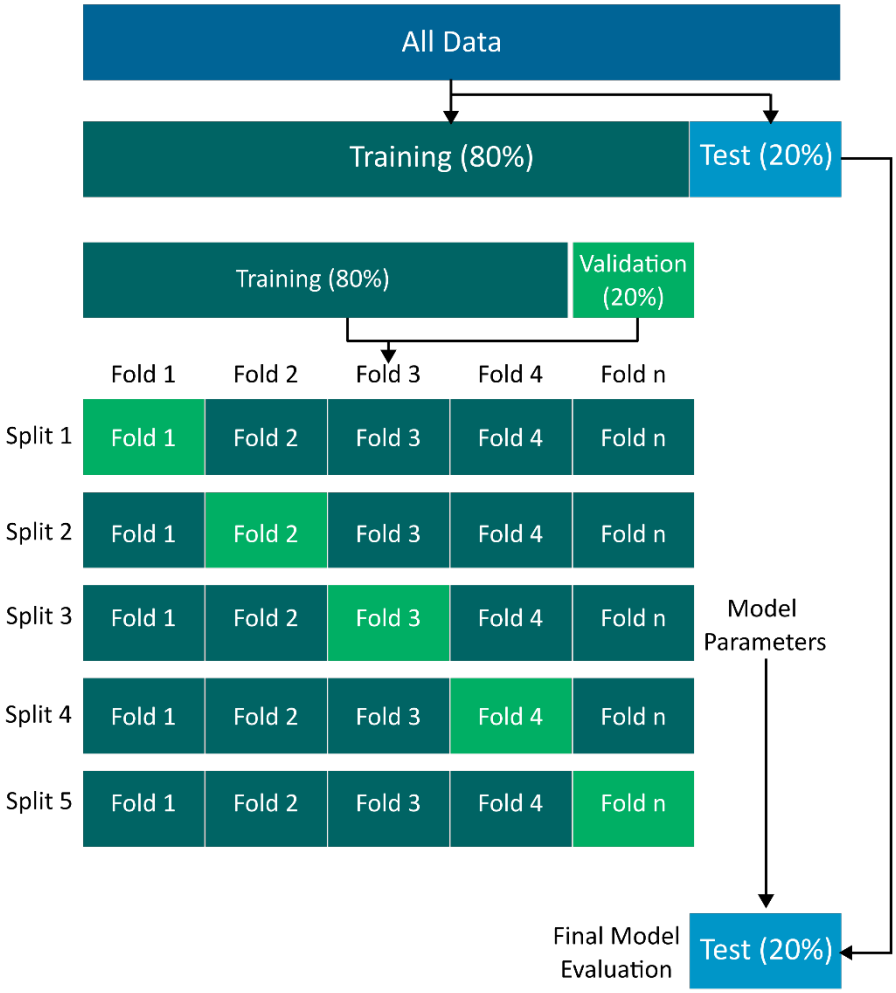


Fig. 5.4. Stratified k-fold cross validation as used in this study. The initial dataset is divided into training (80%) and test (20%) datasets. The test dataset is withheld from the model development to provide an unbiased assessment of model accuracy. During the model development and training, a further 20% of the training set is used for model validation. In this study 10 data folds were used. Modified after Dawson et al. (2023).

Dental terminology and nomenclature follows that outlined by Hendrickx et al. (2015b).

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), ranger (Wright & Ziegler, 2017) and caret (Kuhn, 2008) 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; chronosphere (Kocsis & Raja, 2019), divDyn (Kocsis et al., 2019) and rgplates (Kocsis &

Raja, 2021) for palaeogeographical reconstructions using the PALEOMAP plate model and data from Scotese (2016).

See Appendix Two for a complete description of the data used in the machine learning models, Appendix Four for a list of institutional abbreviations and Appendix Five for the full machine learning analysis results.

Results

Machine learning models

The difficulties in providing accurate quantitative assessments of theropod tooth morphological discrimination are highlighted in Figure 5.5. Here I 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 revealing between-group structures for some group-pairs as demonstrated by the pairwise PERMANOVA tests (Fig. 5.6). This is consistent with previous reports in the literature where attempts to distinguish theropod taxa using principal components analysis or linear discriminant analysis have reported high degrees of feature-space overlap between some taxonomic groups (e.g., Hendrickx et al., 2019; Noto et al., 2022; Young et al., 2019). This result is unsurprising as I am constrained in attempting to differentiate teeth with very similar gross morphology based on a small set of morphological measurements. As Macleod et al. (2021) point out, however, 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 5.1) suggest that these taxa are differentiable from most major theropod clades ($p < 0.01$).

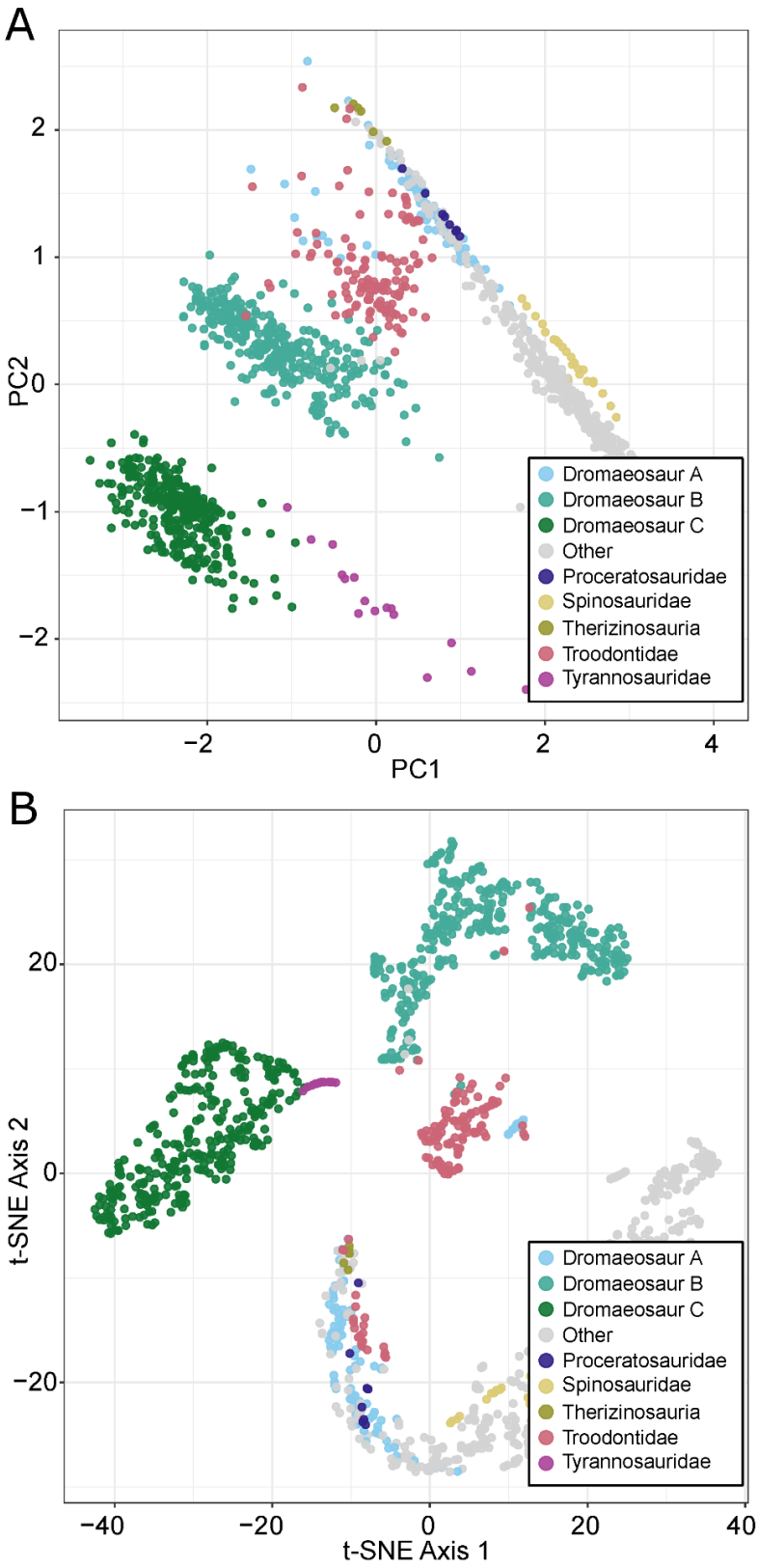


Fig. 5.5. Untrained ordinated feature-space occupation for teeth comprising the training data set. Formed by the first two principal component axes (A) and the first two t-SNE axes (B) showing the degree of morphospace overlap between different theropod clades irrespective of the ordination method used.

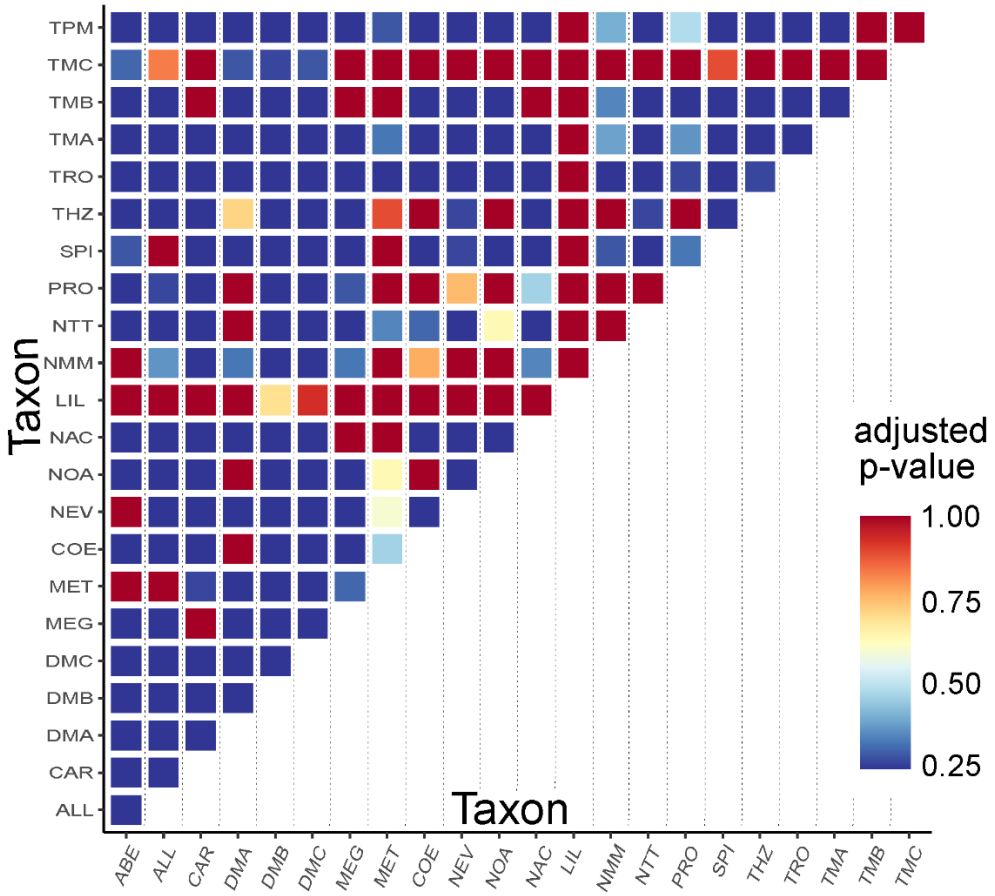


Fig. 5.6. Training data PERMANOVA Bonferroni adjusted p-values for pairwise clade groups using untrained t-SNE ordinated feature space based on three t-SNE dimensions. A p-value of < 0.01 for a taxon pair indicates a significant between group difference. ABE: Abelisauridae; ALL: Allosauridae; CAR: Carcharodontosauridae; DMA: Dromaeosaur morphotype A; DMB: Dromaeosaur morphotype B; DMC: Dromaeosaur morphotype C; MEG: Megalosauridae; MET: Metriacanthosauridae; COE: Coeolophysis; NEV: Neovenatoridae; NOA: Noasauridae; NAC: Other Ceratosauria; LIL: *Liliensternus*; NMM: other Megalosauroidea; NTT: other Tyrannosauroidea; PRO: Proceratosauridae; SPI: Spinosauridae; THZ: Therizinosauria; TRO: Troodontidae; TMA: Tyrannosauridae morphotype A; TMB: Tyrannosauridae morphotype B; TMC: Tyrannosauridae morphotype C; TPM: Tyrannosauridae premaxillary.

Table 5.1. By-group comparisons of maniraptoran clade-pairs. PERMANOVA Bonferroni adjusted p-values on untrained t-SNE ordinated feature space based on three t-SNE dimensions.

Taxon pairs	SumsOfSqs	F Model	R2	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

I also conducted PERMANOVA tests on the trained MDA feature-space scores generated from the training data (Fig. 5.7). 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.

Table 5.2. Machine learning model accuracies. [(1)] Random Forests: two randomly selected predictor variables at each tree node split and 2,000 trees; [(2)] Mixture discriminant analysis: eight sub-classes; [(3)] C5.0: tree-based model with 40 boosting iterations.

Machine learning model	Model accuracy	Testing data accuracy
RF[1]	85.6	88.4
MDA[2]	84.4	84.1
C5.0[3]	82.4	85.4

All three machine learning techniques show similar levels of accuracy (Table 5.2) with overall accuracies of each machine learning model 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 model at 84.1%. I 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 corresponding well to the accuracy returned from the test data. RF prediction errors decrease as the forest is grown to its full extent of 2,000 trees with the overall OOB error and most individual clade OOB errors settling after around 200 trees.

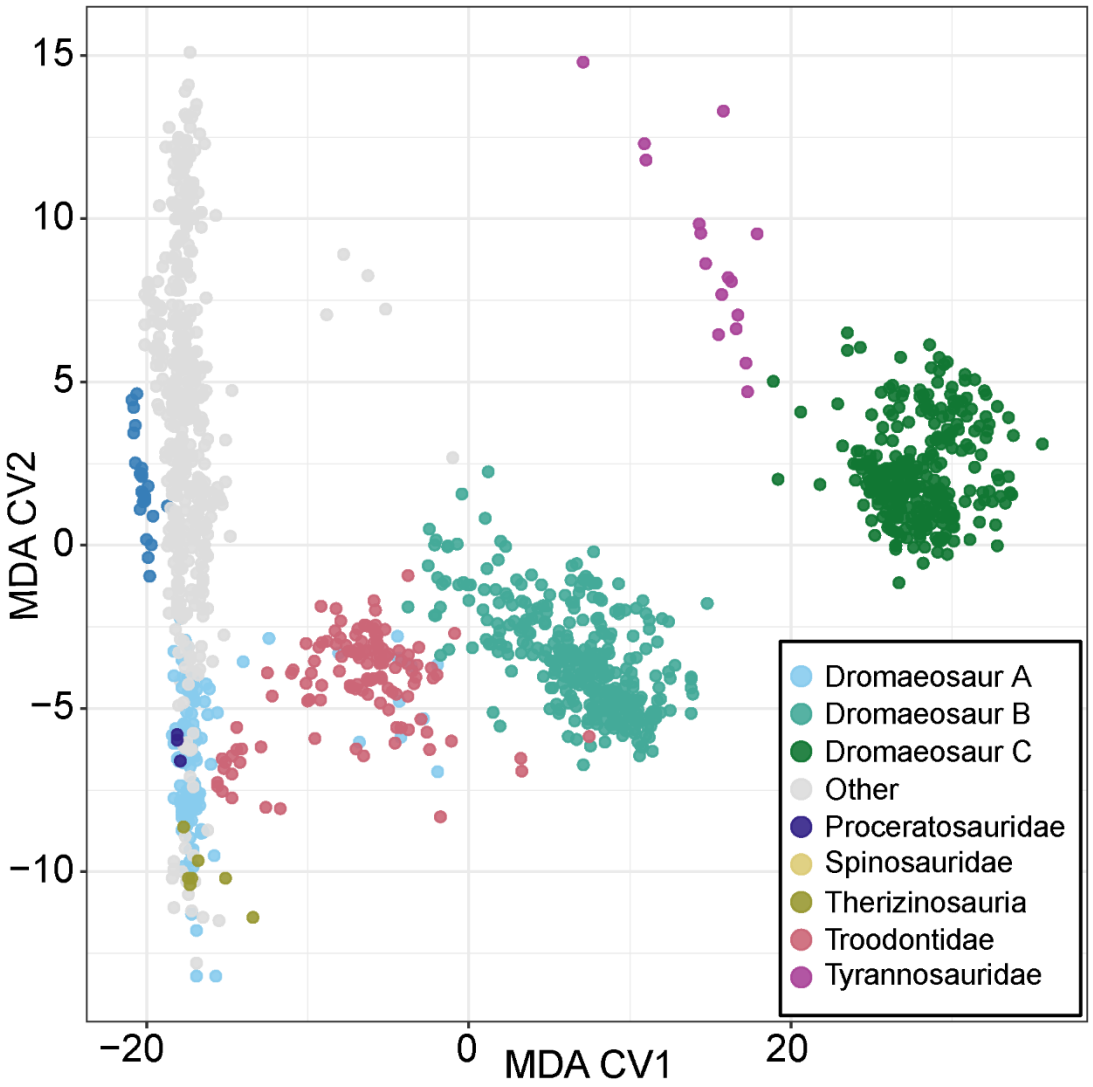


Fig. 5.7. Trained feature-space occupation of selected taxa from the training data based on two mixture discriminant analysis functions. Total between-group variance explained 97.5% (MDA CV1 = 90.4%, MDA CV2 = 7.1%).

The RF model responses achieved by varying the m_{try} tuning parameter range 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 MDA model responses achieved by varying the number of potential sub-classes in each taxonomic group range from an accuracy of 78.3% (one sub-class) to 84.4% (eight sub-classes) and the C5.0 model achieved the best response (82.4% accuracy) using a tree-based classifier with 40 boosting iterations.

Table 5.3 Machine learning classification accuracy by clade based on test data. (1) ensemble model accuracy; (2) individual model accuracies; (3) number of cases per clade in the training data.

Taxon	Accuracy (1)	RF (2)	C5.0 (2)	MDA (2)	Cases (3)
Neotheropoda: Coelophysis	91.7	100	99.8	100	9
Non averostran Neotheropoda: *Liliensternus*	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

At the individual clade level (Table 5.3; Fig. 5.8), the performance of both the ensemble model and the individual machine learning classifiers which make up this ensemble vary with classification accuracy ranging from 50–100% (Fig. 5.8). Maniraptoran clades show a high level of classification accuracy regardless of the machine learning model

employed, 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).

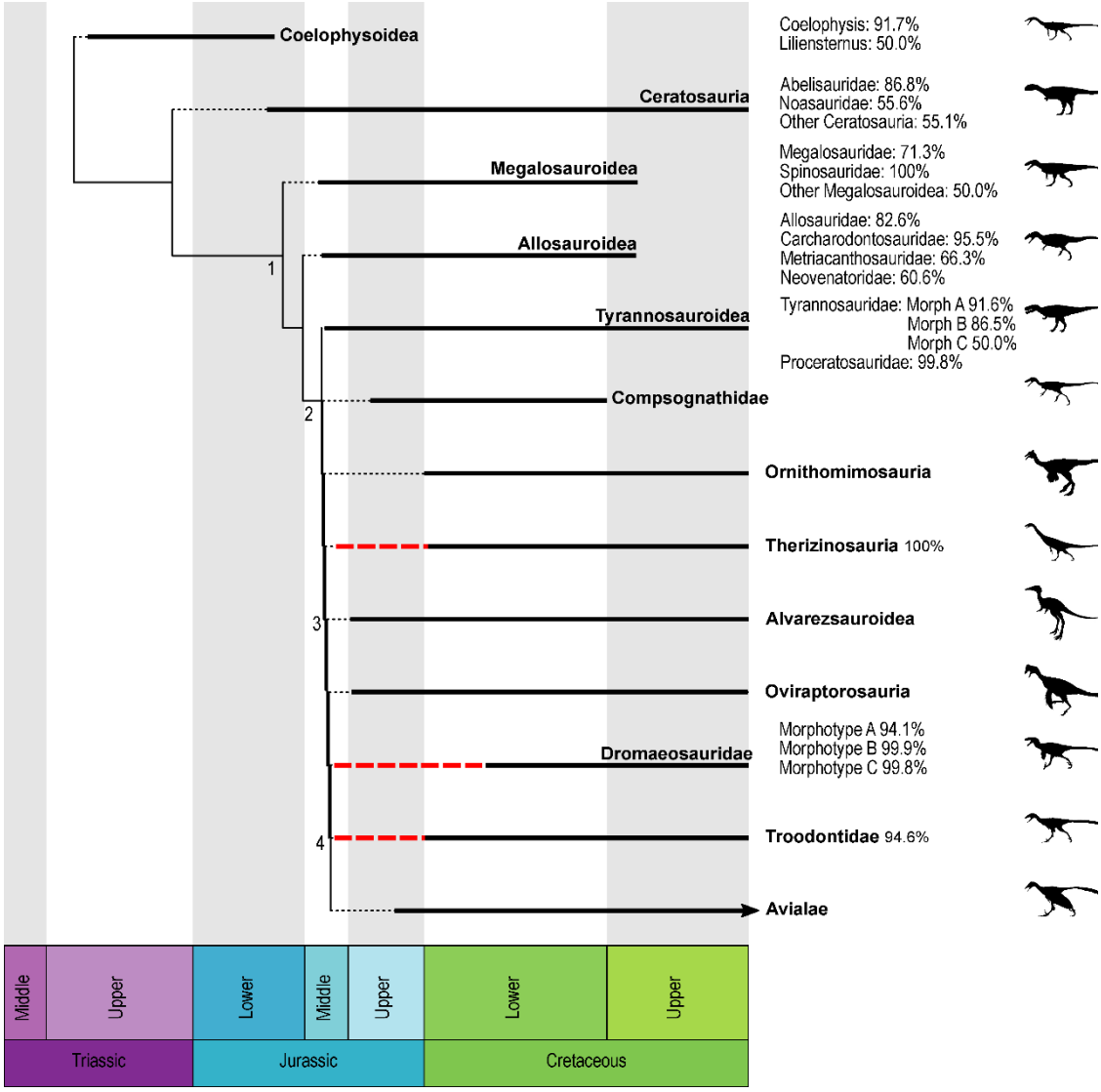


Fig. 5.8. Simplified time-calibrated theropod phylogeny showing the individual clade classification accuracies based on the machine learning ensemble and the range extensions (in red) implied by these results. (1) Tetanurae, (2) Coelurosauria, (3) Maniraptora, (4) Paraves. For Therizinosauria and Troodontidae I 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 and Foth (2020). All silhouettes taken from www.phylopic.org

The variation in clade accuracy is driven by several factors, including: the number of cases making up 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. (2021) pointed out, 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 5.4) 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 posterior probability mode. My confidence in the classifications is a combination of the machine-learning results from three independent classifiers and my post-hoc morphological analysis. In all machine learning systems there is likely to be a degree of misclassification and 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. 5.9) also add to my confidence in the machine learning prediction as 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. 5.10) 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.

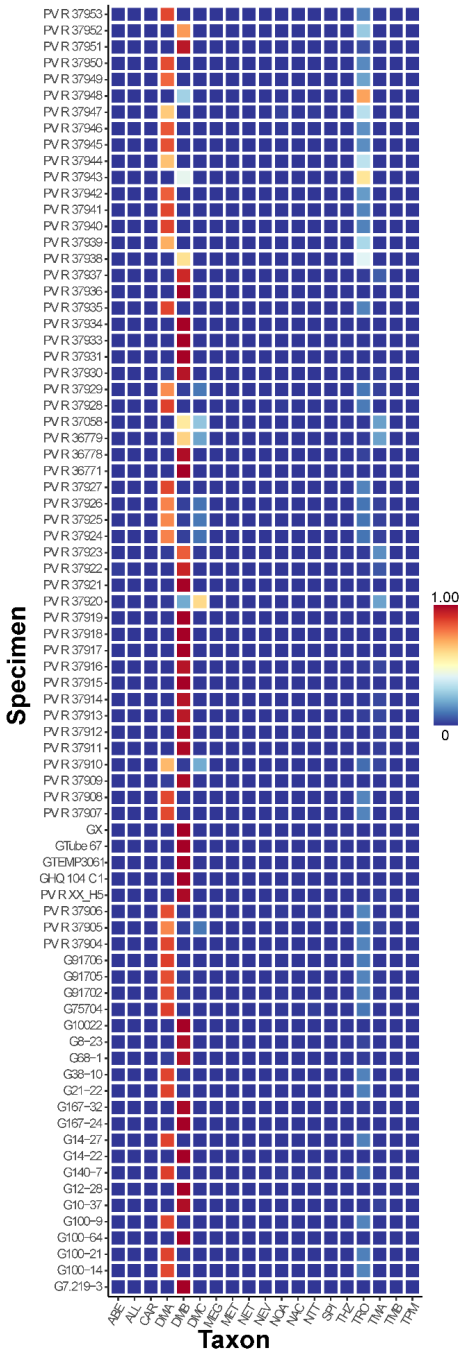


Fig. 5.9. Posterior probability of the final assigned taxon from the machine learning ensemble classifier for UK Bathonian teeth. Specimens with prefix PV R are in the Natural History Museum (NHMUK), those with prefix G are in the Museum of Gloucester (GLRCM). Colour scale represents the posterior probability ranging from zero to one. Taxon abbreviations are as follows. ABE: Abelisauridae; ALL: Allosauridae; CAR: Carcharodontosauridae; DMA: Dromaeosaur morphotype A; DMB: Dromaeosaur morphotype B; DMC: Dromaeosaur morphotype C; MEG: Megalosauridae; MET: Metriacanthosauridae; COE: *Coelophysis*; NEV: Neovenatoridae; NOA: Noasauridae; NAC: Other Ceratosauria; NTT: other Tyrannosauroidea; SPI: Spinosauridae; THZ: Therizinosauria; TRO: Troodontidae; TMA: Tyrannosauridae morphotype A; TMB: Tyrannosauridae morphotype B; TPM: Tyrannosauridae pre-maxillary.

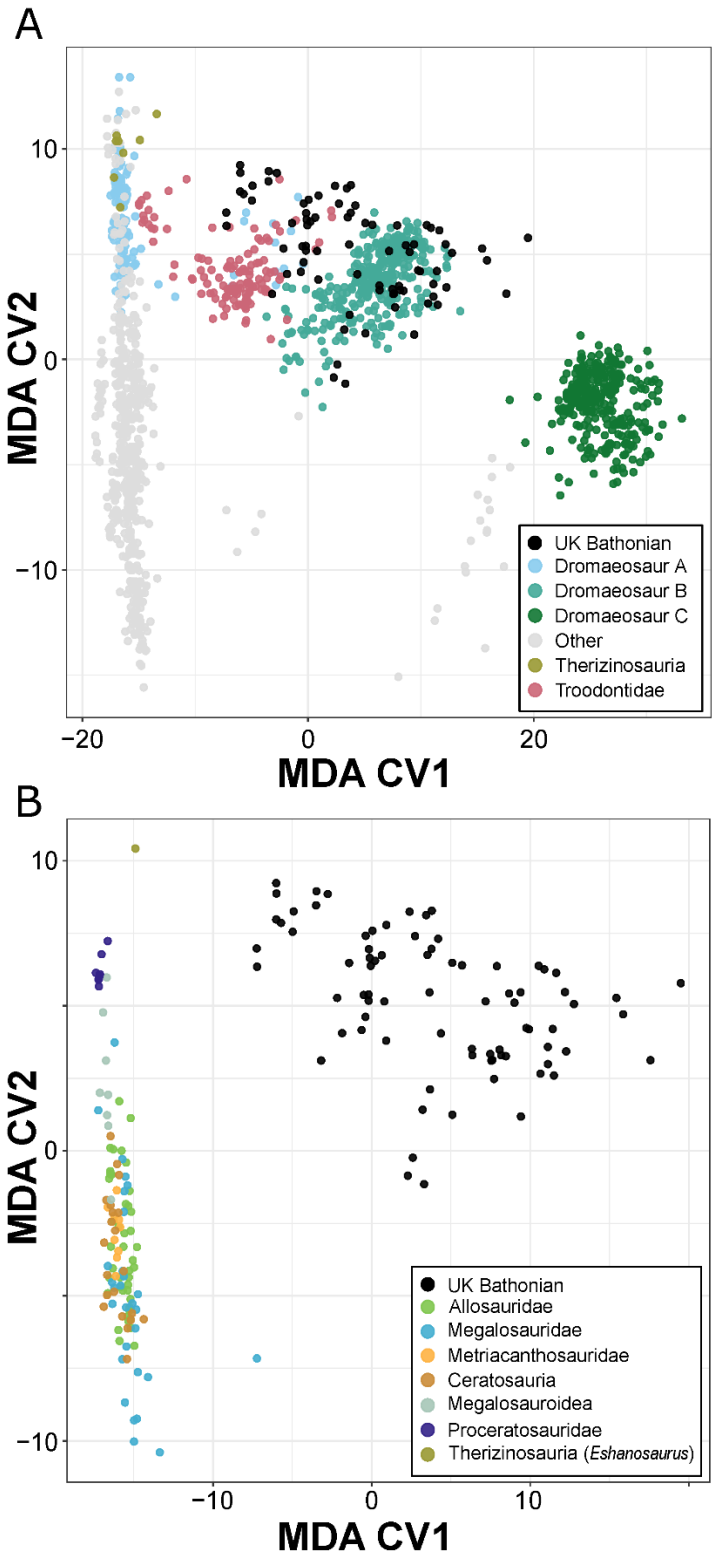


Fig. 5.10. Trained feature-space occupation of UK Bathonian teeth compared to training data based on two mixture discriminant analysis functions. (A) Compared to all taxa in the training data with Maniraptoran clades highlighted. (B) Compared to Jurassic taxa.

Table 5.4. Theropod tooth morphotypes, UK Bathonian sites. Morphotype: assigned tooth morphotype following machine learning and visual description. Majority vote: assigned tooth morphotype following simple majority vote of three machine learning models. Combined posterior probability: assigned tooth morphotype by combining posterior probabilities from three machine learning models. P: combined posterior probability value.

			Machine Learning		
Specimen	Locality	Morphotype	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
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

Systematic Palaeontology

Dromaeosauridae Morphotype A

THEROPODA Marsh, 1881

MANIRAPTORA Gauthier, 1986

PARAVES Sereno, 1997

DROMAEOSAURIDAE Matthew and Brown, 1922

Gen. et sp. indet. Morphotype A

Figure 5.11

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 1.45–7.79 mm (Fig. 5.12) 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

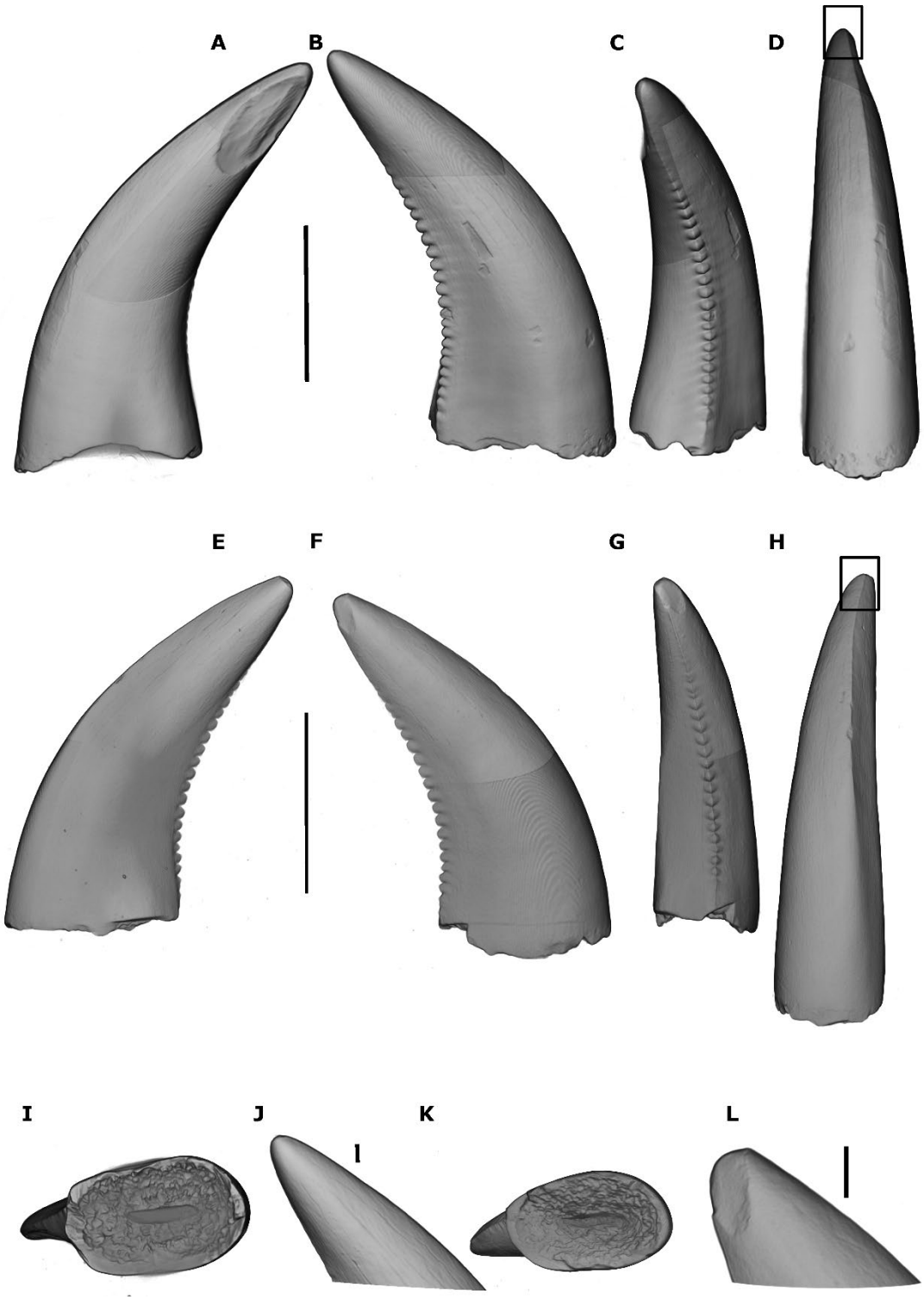


Fig. 5.11. Isolated crowns of indeterminate dromaeosaurs (Morphotype A) from Woodeaton Quarry (NHMUK PV R37059, A to D and I to J) and Kirtlington Quarry, (NHMUK PV R37925, E to H and K to L) in lingual (A, E), labial (B, F), distal (C, G), mesial (D, H), basal (I, K) and close-up of apical region (J, L) views. Scale bars = 1 mm for general views and 0.1 mm for apical close-up (boxed region).

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, though 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 0.05–0.27 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 exhibits a braided enamel texture comprising sinuous grooves and ridges that are orientated apicobasally (Hendrickx et al., 2015a; Hendrickx et al., 2019).

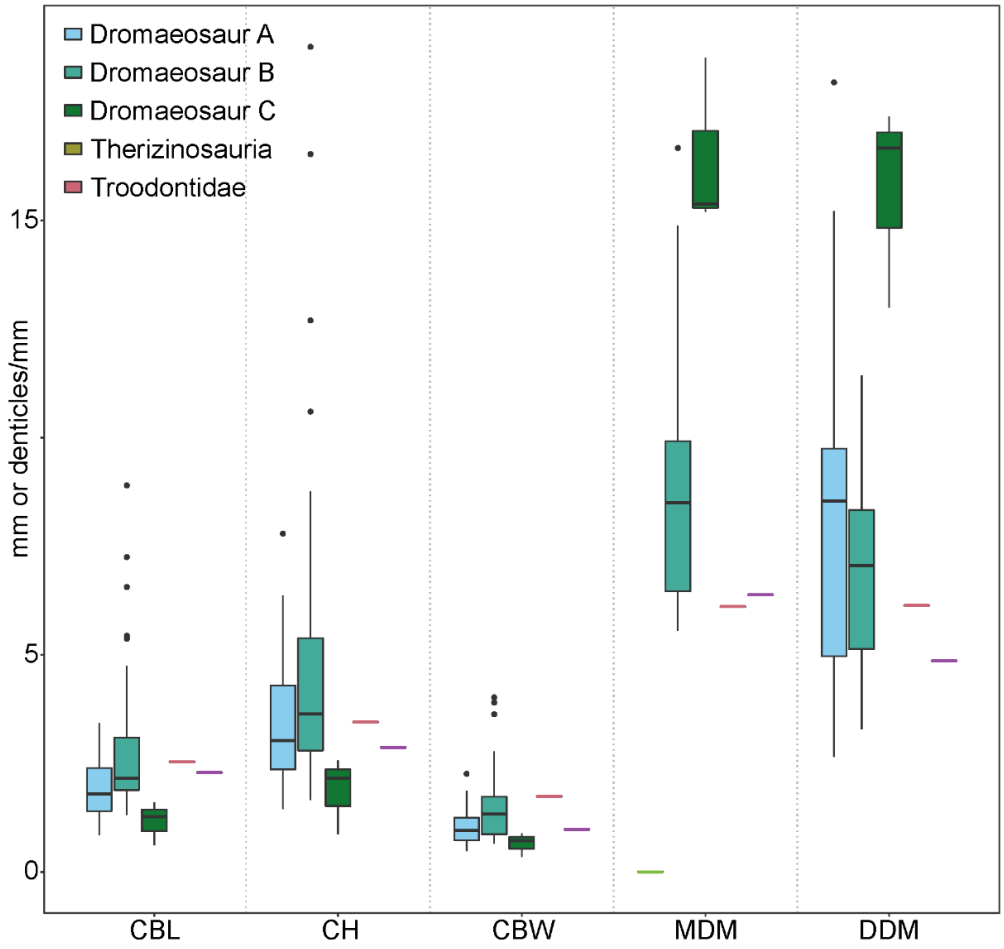


Fig. 5.12. Range of morphometric measurements across each maniraptoran morphotype. CBL, crown basal length; CH, crown height; CBW, crown basal width; MDM, mesial denticles per mm; DDM, distal denticles per mm.

Dromaeosauridae Morphotype B

Gen. et sp. indet. Morphotype B

Figure 5.13

Referred specimens. GCRLM GTUBE 67, G10022, G100-64, G10-37, G12-28, G14-22, G167-24, G68-1, G7.219-3, 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 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, so might encompass several different sub-groups with broadly similar morphology. Tooth crowns are ziphodont, slightly larger than Morphotype A (CH ranging from 1.66–19 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. 5.12) 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

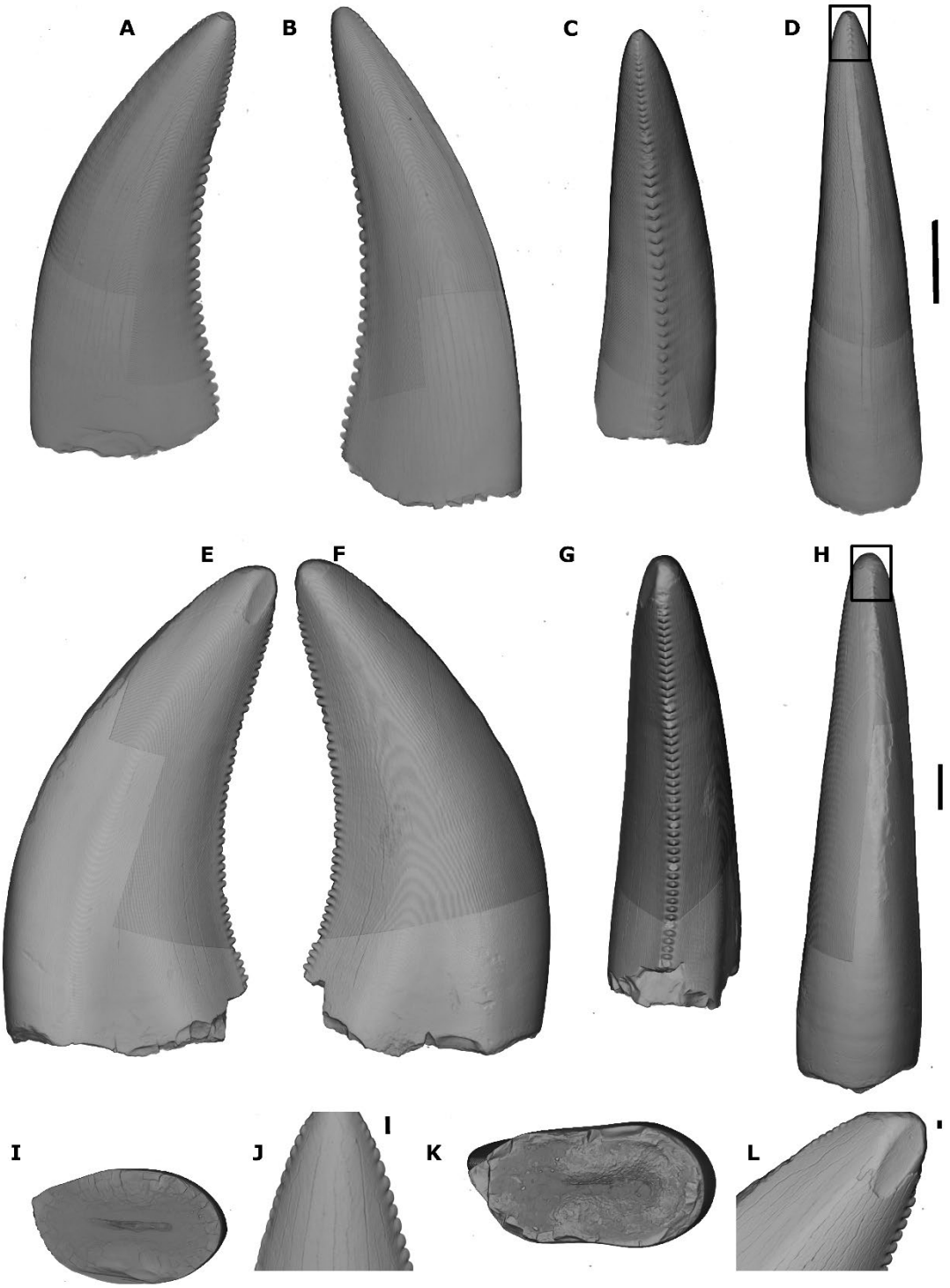


Fig. 5.13. Isolated crowns of indeterminate dromaeosaurs (Morphotype B) from Woodeaton Quarry (NHMUK PV R37916, A to D and I to J) and Hornsleasow Quarry, (GCLRM G167-24, E to H and K to L) in lingual (A, E), labial (B, F), distal (C, G), mesial (D, H), basal (I, K) and close up of apical region (J, L) views. Scale bars = 1 mm for general views and 0.1 mm for apical (boxed region).

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 a 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 a DSDI > 1.4 and it is possible they may represent either a variation within this morphotype or a separate morphotype. However, in the absence of any other morphological 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 sub-rectangular in shape with a convex external margin and are orientated perpendicular to the carina. The crown surface exhibits a braided enamel texture comprising sinuous grooves and ridges orientated apicobasally (Hendrickx et al., 2015a; Hendrickx et al., 2019).

Dromaeosauridae Morphotype C

Gen. et sp. indet. Morphotype C

Figure 5.14

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 0.61–1.6 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 8.7 per mm and distal denticles 7.0 per mm. Both mesial and distal denticles are rectangular to sub-rectangular 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 which is distinct from Morphotype B based on their smaller size, and greater serration density on both carinae.

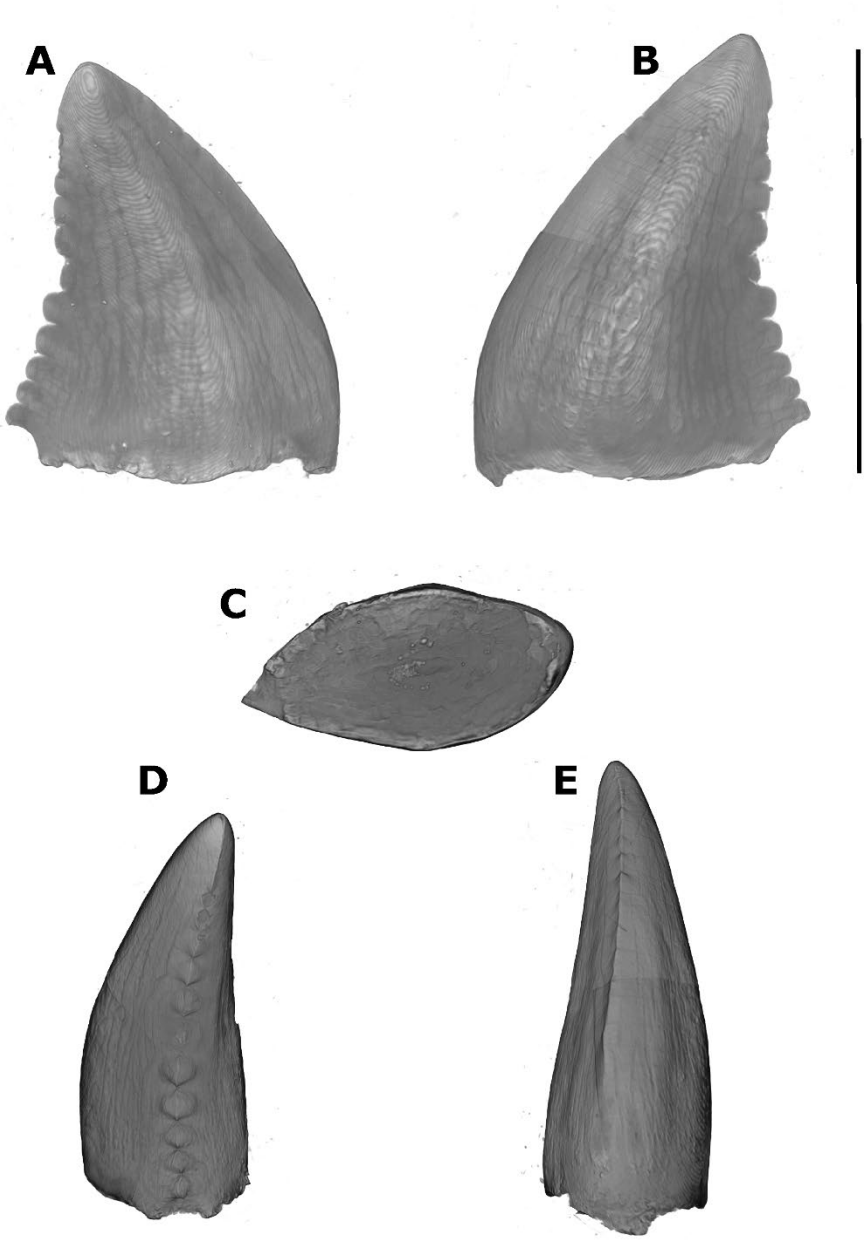


Fig. 5.14. Isolated crown of indeterminate dromaeosaur (Morphotype C) from Watton Cliff (NHMUK PV R37920), in lingual (A), labial (B), basal (C), distal (D) and mesial (E) views. Scale bar = 1 mm.

Troodontidae

TROODONTIDAE Gilmore, 1924

Gen. et sp. indet.

Figure 5.15

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 exhibits 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 mid-point whereas distal denticles increase in size slightly towards the apex. Distal denticles are sub-rectangular 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.

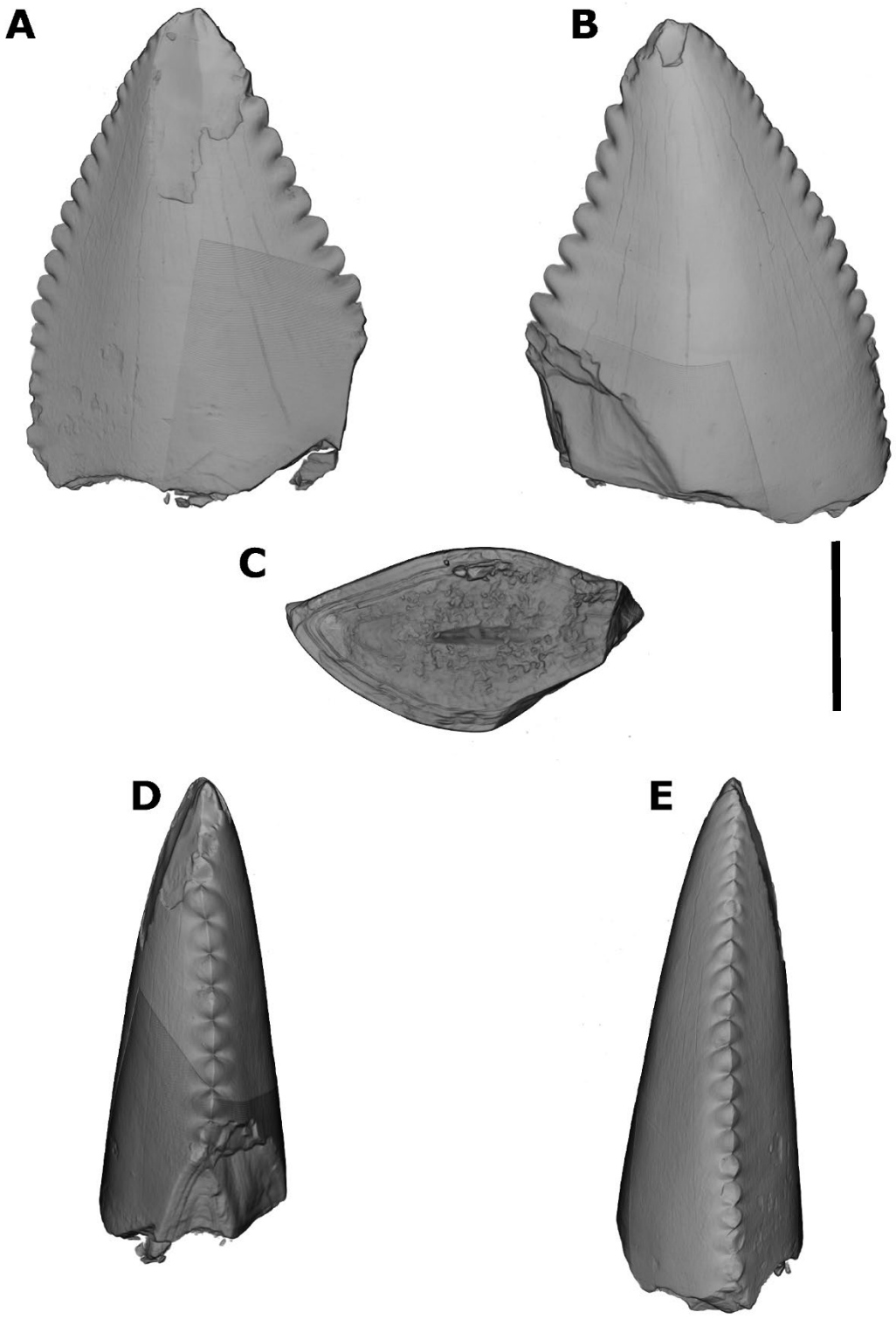


Fig. 5.15. Isolated crown of an indeterminate troodontid (GCLRM G8-23) from Hornsleasow Quarry, Gloucestershire in lingual (A), labial (B), basal (C), distal (D) and mesial (E) views. Scale bar = 1 mm.

Therizinosauroidea

THERIZINOSAUROIDEA Maleev, 1954

Gen. et sp. indet.

Figure 5.16

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 which is phylloform in shape, labiolingually compressed and sub-symmetrical 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 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 apicobasally, have a convex exterior margin and are slightly inclined apically.

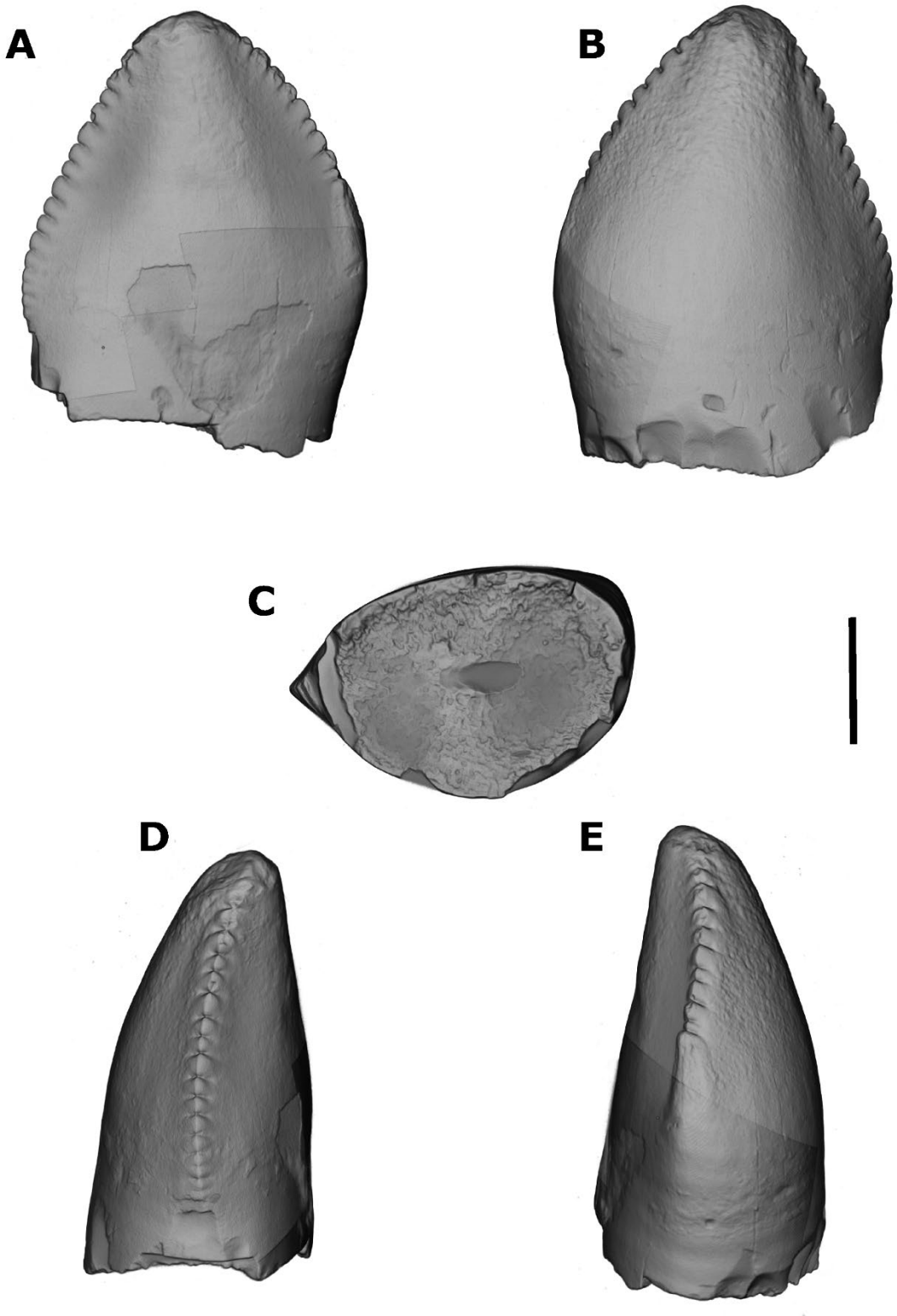


Fig. 5.16. Isolated crown of an indeterminate therizinosauroid (GCLRM G167-32) from Hornsleasow Quarry, Gloucestershire in lingual (A), labial (B), basal (C), distal (D) and mesial (E) views. Scale bar = 1 mm.

Morphological comparisons

Dromaeosaurid morphotypes. I interpret Morphotypes A–C as dromaeosaurids based on the machine learning classification and several morphological characters which, in conjunction with their small size, precludes 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 (Fig. 5.17). Dromaeosaurid teeth have the following combination of features (Fig. 5.18): relatively small size, with even larger-bodied taxa such as *Utahtraptor* 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 just the distal carina (Hendrickx et al., 2019; Prasad & Parmar, 2020; Sankey et al., 2002). 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; Larson et al., 2016; Williamson & Brusatte, 2014) and the Asian velociraptorine *Tsaagan* (Chiarenza et al., 2020; Norell et al., 2006). Larger distal denticles compared to mesial denticles has previously been used as a synapomorphy to identify velociraptorine teeth (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* (Chiarenza et al., 2020; Hendrickx et al., 2019; Larson, 2008; Larson et al., 2016; Larson & Currie, 2013). A lemniscate cross-section is a feature present in many deinonychosaurs, including most dromaeosaurids, but with the exception of some metriacanthosaurids, megaraptorans and tyrannosauroids is absent from non-maniraptoriform theropods (Hendrickx & Mateus, 2014; Hendrickx et al., 2019).

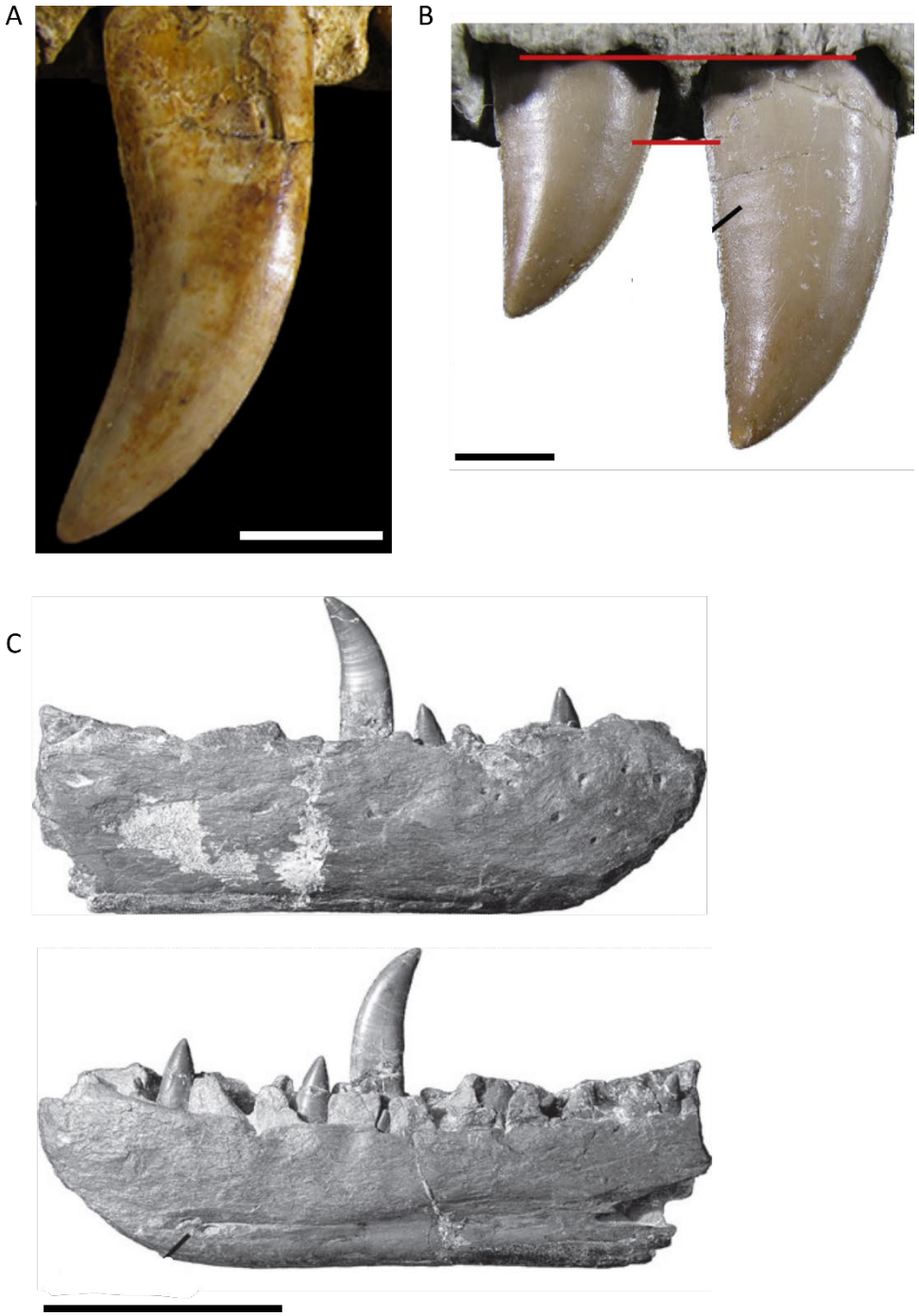


Fig. 5.17. Dental morphology of tyrannosaurids, abelisaurids and megalosaurids. A, maxillary teeth of the tyrannosaurid *Alioramus altai* (MPC-D 100-1844) in lingual view. (Hendrickx et al., 2019). B, left maxillary teeth of the abelisaurid *Majungasaurus crenatissimus* (FMNH PR 2100) in lingual view (Hendrickx et al., 2020). C, right dentary of *Megalosaurus bucklandii* (lectotype, OUMNH J.13505) from the Bathonian Taynton Limestone Formation, Oxfordshire (Benson et al., 2008). Scale bars: 1 cm (A and B), 10 cm (C).

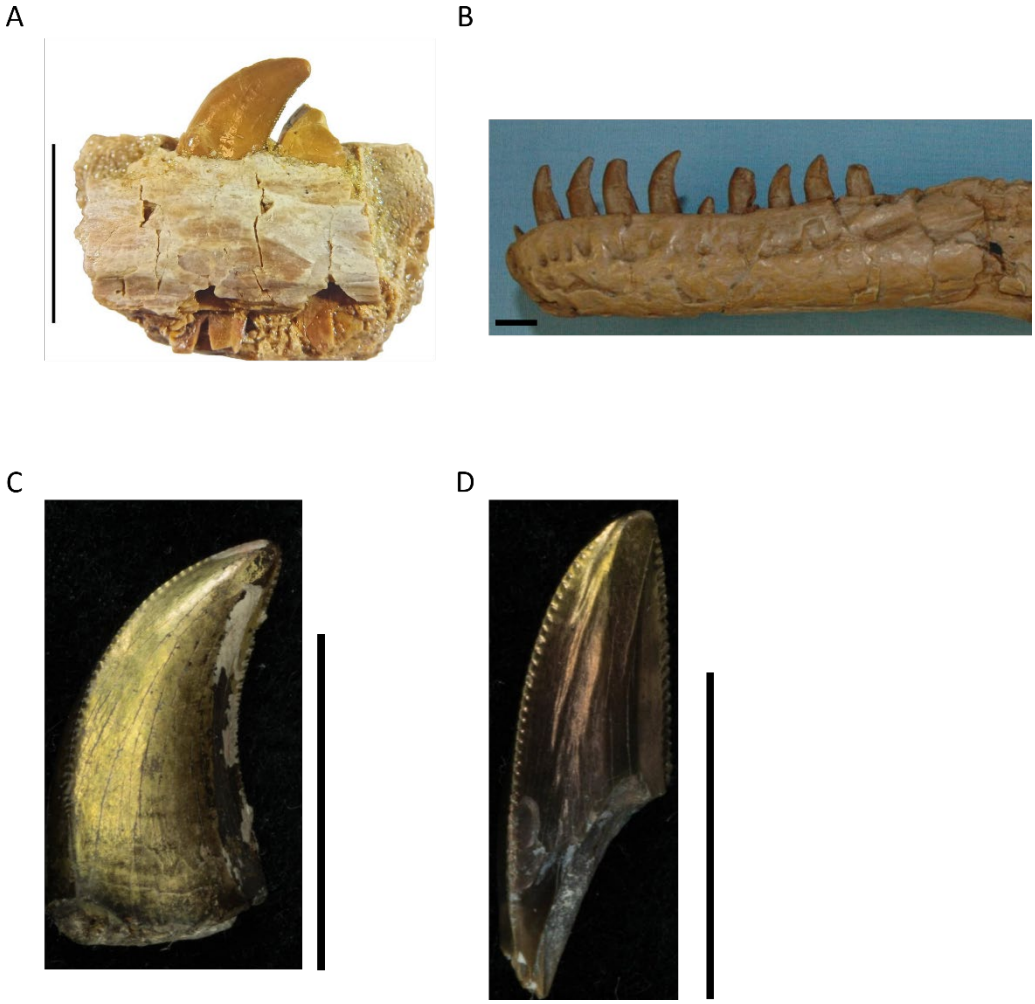


Fig. 5.18 Dromaeosaurid tooth morphology. A, indeterminate dromaeosaurid teeth (MCNA 14624) from the upper Campanian, Laño, Spain, in lingual view (Isasmendi et al., 2022). B, left mandible of *Dromaeosaurus albertensis* (TMP 1984.008.0001) from the Upper Cretaceous Judith River Formation, Dinosaur Provincial Park, Alberta, Canada. C, *Dromaeosaurus* tooth (TMP1981.026.0048), Dinosaur Park Formation, Campanian, Upper Cretaceous, Dinosaur Provincial Park, Alberta, Canada. D, *Dromaeosaurus* premaxillary tooth (TMP1973.013.0001), Horseshoe Canyon Formation, Maastrichtian, Upper Cretaceous, Munson, Canada. Images B, C and D Royal Tyrrell. Museum of Palaeontology. Scale bars 1 cm.

Troodontid morphotype. I refer the single tooth GCLRM G8-23 to Troodontidae on both morphological-based considerations and machine learning morphospace occupation. 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 (Fig. 5.19). The presence of denticles on both the mesial and distal carinae is seen in derived troodontids (Goswami et al., 2013; Hendrickx et al., 2019; Makovicky et al., 2003) with only *Troodon formosus* (Leidy, 1856)*, Zanabazar junior* (Norell et al., 2009), *Saurornithoides mongoliensis* (Hendrickx et al., 2019; Osborn, 1924) 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. I refer the single tooth GCLRM G167-32 to Therizinosauroidea on morphological-based considerations only, as this tooth was incorrectly classified as a dromaeosaurid in the machine learning analysis. The sub-symmetrical phylloform tooth with a basal constriction seen in GCLRM G167-32 are features shared with therizinosauroids (Fig. 5.20) such as *Falcarius* (Kirkland et al., 2005; Zanno, 2010a), *Erlikosaurus* (Barsbold & Perle, 1980) and *Eshanosaurus* (Barrett, 2009; Xu et al., 2001). 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 (Hendrickx et al., 2019; Zanno, 2010a). 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.

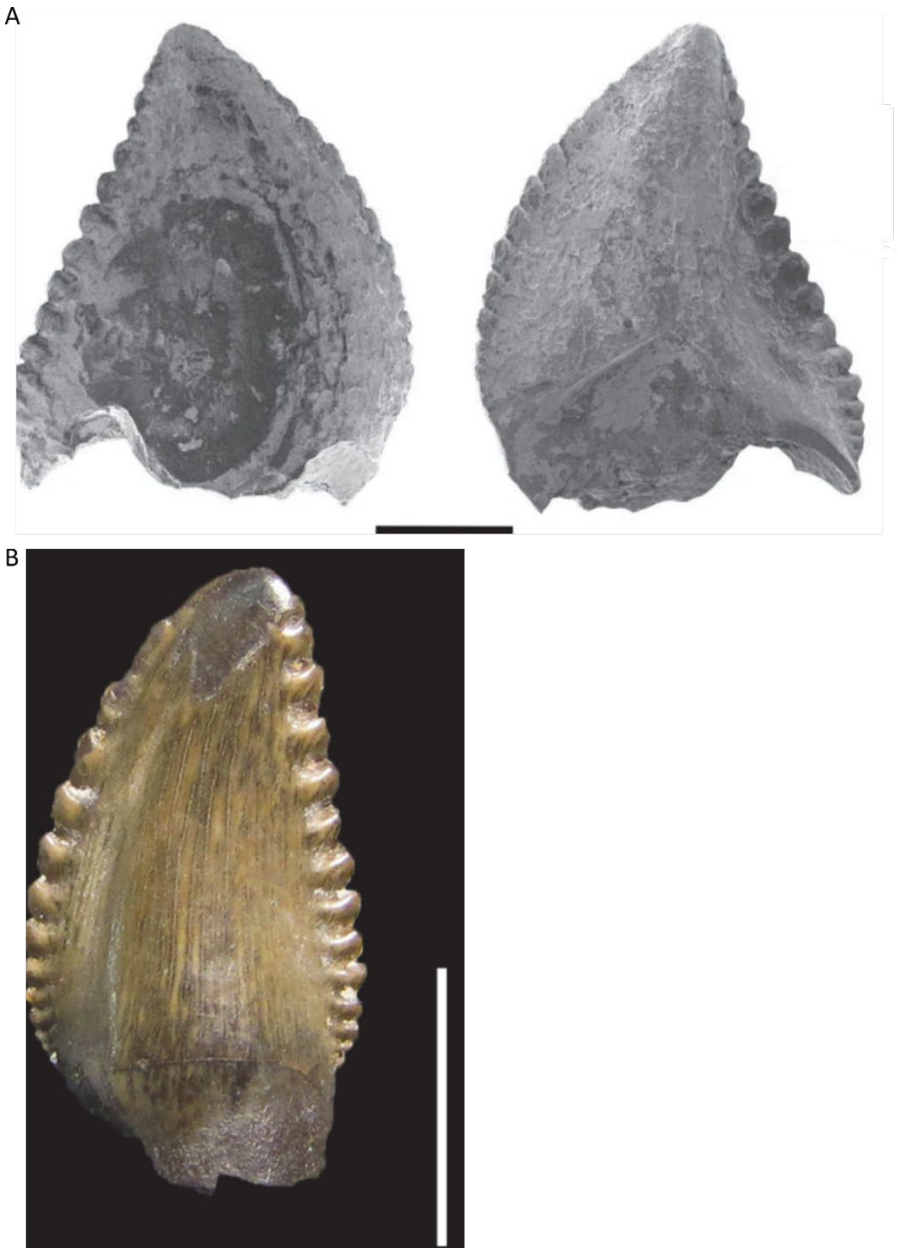


Fig. 5.19. Troodontid tooth morphology. A, troodontid tooth (DUGF/52) from the Maastrichtian, Late Cretaceous, Cauvery Basin of South India in labial and lingual views (Goswami et al., 2013). B, Isolated tooth of the troodontid *Troodon formosus* (DMNH 22837) in labial view (Hendrickx et al., 2019), locality unknown. Scale bars: 1 mm (A), 5 mm (B).

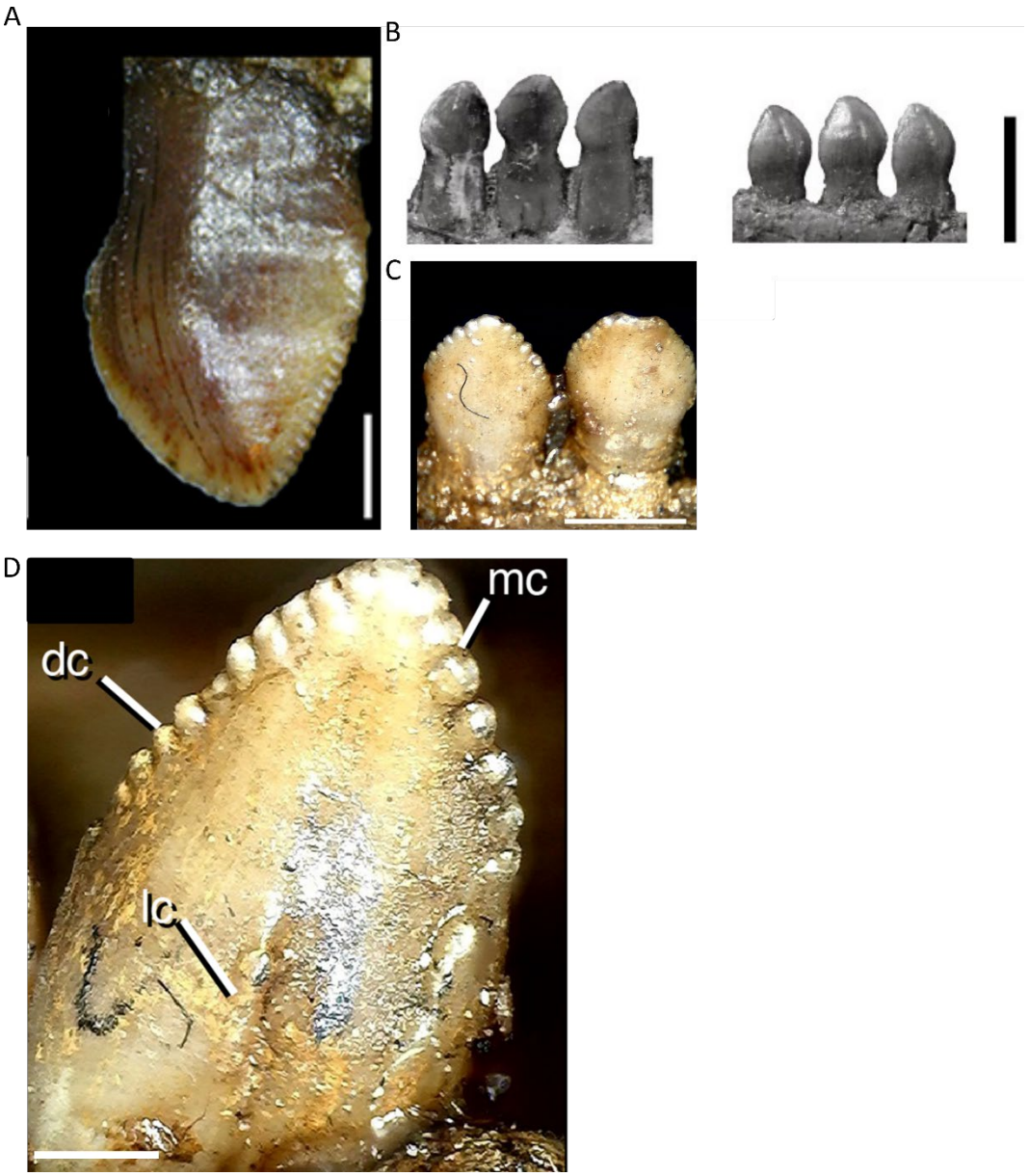


Fig. 5.20. Therizinosauroid tooth morphology. A, left maxillary tooth of *Falcarius utahensis* (UMNH VP 14545) in labial view (Hendrickx et al., 2019). B, left maxillary teeth of *Falcarius utahensis* (UMNH VP 14526) in lingual and labial views (Zanno, 2010a). C, dentary teeth of *Erlikosaurus andrewsi* (MPC-D 100/111) (Zanno et al., 2016). D, distalmost dentary teeth of *Segnosaurus galbinensis* (MPC-D 100/80) (Zanno et al., 2016). Scale bars: 1 mm (A, C and D), 0.5 mm (B). Annotation on D, dc, distal carina; mc, mesial carina; lc, lingual carina.

Comparisons to other UK Middle Jurassic theropod taxa

Among theropod dinosaurs known from the Middle Jurassic of the UK I can exclude larger non-maniraptoran taxa such as the megalosaurids *Megalosaurus* (Benson, 2010a; Buckland, 1824), *Magnosaurus* (Benson, 2010b; Huene, 1932), *Duriavenator* (Benson, 2008; Owen, 1883; Waldman, 1974), and *Eustreptospondylus* (Sadleir et al., 2008; Walker, 1964), and the basal tetanuran *Cruxicheiros* (Benson & Radley, 2010) on the basis of morphospace occupation (Fig. 5.10), size and overall morphology (Hendrickx et al., 2015a). 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 (Fig. 5.21) 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 (Rauhut et al., 2010; Woodward, 1910).

Middle Jurassic isolated theropod teeth from India, Madagascar, and Kyrgyzstan

In addition to analysing the isolated theropod teeth from the four British Bathonian microvertebrate sites, I used the same models to reassess the taxonomic affinities of putative Middle Jurassic maniraptorans described from the Kota Formation, India (Prasad & Parmar, 2020), the Balabansai Formation, Kyrgyzstan (Averianov et al., 2005) and the Sakaraha Formation, Madagascar (Maganuco et al., 2005). Isolated theropod teeth from these sites had been suggested to resemble maniraptorans based on gross morphology, but no quantitative assessment had been made. The results from the machine learning analysis confirm the presence of maniraptorans at all of these localities (Tables 4.5, 4.6 and 4.7).

All of the Kota Formation teeth, apart from DUGF/J30 and DUGF/J31, classify as dromaeosaurs irrespective of the machine learning model used. The classification of DUGF/J30 and DUGF/J31 as therizinosaurs is incorrect, being only weakly supported by the random forest and C5.0 models and not consistent with morphological evidence. In fact, the MDA classifier strongly supports a dromaeosaur classification of these teeth.

Fig. 5.21. Dentition of the Bathonian tyrannosauroid *Proceratosaurus bradleyi* (NHMUK PV R 4860). A, left maxilla and dentary in lateral view. B, Left 5th dentary tooth in labial view, C, Right 8th maxillary tooth in labial view.

The individual machine learning models and the ensemble classifier strongly support a dromaeosaur affinity for some of the Madagascan teeth with significant posterior probabilities (> 0.95) in favour of this classification (V5787, V5788, V5790, V5794, V5798, V5806, V5815, V5816, V5959). Other teeth are more problematic with very poor posterior probability support for the class assigned by the machine learning models and no support from gross morphology (e.g., V5779 has classified as a troodontid with a posterior probability of 0.35). Of the seven Balabansai Formation teeth from Kyrgyzstan that have complete morphometric measurements, five classify as dromaeosaurs (ZIN PH 14/42, 15/42, 18/42, 19/42 and 28/42) and are supported by all machine learning classifiers. The similarity of these to other small Middle Jurassic isolated teeth from the Bathonian sites of the UK, classified herein as maniraptoran theropods, was noted by (Averianov et al., 2005). The two remaining teeth (ZIN PH 7/42 and 9/42) classify either as tyrannosaurid or dromaeosaur teeth but are only weakly supported in either of these by the classifiers.

Table 5.5. Reclassification of the Kota Formation (Bathonian, Middle Jurassic, India) isolated theropod teeth (Prasad & Parmar, 2020). Majority vote: assigned tooth morphotype following simple majority vote of three machine learning models. Combined posterior probability: assigned tooth morphotype by combining posterior probabilities from three machine learning models. P: combined posterior probability value.

		Machine learning		
Specimen	Prasad and Parmar (2020) classification	Majority Vote	Combined posterior probability	P
DUGF/J19	Dromaeosauridae	Dromaeosaur C	Dromaeosaur B	0.58
DUGF/J20	Dromaeosauridae	Dromaeosaur C	Dromaeosaur B	0.53
DUGF/J22	Dromaeosauridae	Dromaeosaur A	Dromaeosaur B	0.79
DUGF/J23	Dromaeosauridae	Dromaeosaur C	Dromaeosaur B	0.51
DUGF/J27	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.79
DUGF/J28	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.79
DUGF/J30	Dromaeosauridae	Therizinosauria	Therizinosauria	0.51
DUGF/J31	Dromaeosauridae	Therizinosauria	Therizinosauria	0.51
DUGF/J32	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.74
DUGF/J36	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.74
DUGF/J37	Dromaeosauridae	Dromaeosaur B	Dromaeosaur B	0.99
DUGF/J45	Dromaeosauridae	Dromaeosaur B	Dromaeosaur B	0.89
DUGF/J46	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.81
DUGF/J52	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.74
DUGF/J53	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.63
DUGF/J56	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.78
DUGF/J60	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.63
DUGF/J61	Dromaeosauridae	Dromaeosaur B	Dromaeosaur B	0.75
DUGF/J62	Dromaeosauridae	Dromaeosaur A	Dromaeosaur A	0.68
DUGF/J64	Dromaeosauridae	Dromaeosaur C	Dromaeosaur C	0.93
DUGF/J65	Dromaeosauridae	Dromaeosaur C	Dromaeosaur B	0.41

Table 5.6. Reclassification of the Sakaraha Formation (Bathonian, Middle Jurassic, Madagascar) isolated theropod teeth. Majority vote: assigned tooth morphotype following simple majority vote of three machine learning models. Combined posterior probability: assigned tooth morphotype by combining posterior probabilities from three machine learning models. P: combined posterior probability value. Numbers in parentheses refer to the morphotypes (one to seven) erected by Maganuco et al. (2005).

		Machine Learning		
Specimen (MSNM)	Maganuco et al. (2005) classification	Majority Vote	Combined posterior probability	P
V5778	Neoceratosauria(1)	Abelisauridae	Tyrannosauridae	0.39
V5779	Neoceratosauria(3)	Troodontidae	Troodontidae	0.35
V5781	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.98
V5782	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.67
V5784	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.73
V5785	Theropoda(7)	Dromaeosaur B	Dromaeosaur B	0.94
V5787	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.95
V5788	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.99
V5789	Ceratosauria / Coelurosauria(4)	Dromaeosaur B	Dromaeosaur B	0.78
V5790	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	1.00
V5791	Coelurosauria (6)	Dromaeosaur A	Dromaeosaur A	0.67
V5792	Ceratosauria / Coelurosauria(4)	Dromaeosaur B	Dromaeosaur B	0.88
V5794	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.98
V5795	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.92
V5796	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.93
V5798	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.97
V5799	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.88
V5800	Ceratosauria / Coelurosauria(4)	Dromaeosaur B	Tyrannosauridae	0.42
V5806	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.96
V5807	Neoceratosauria(1)	Abelisauridae	Tyrannosauridae	0.37
V5808	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.75
V5809	Neoceratosauria(3)	Dromaeosaur B	Troodontidae	0.30
V5810	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.73
V5811	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.93
V5813	Ceratosauria / Coelurosauria(4)	Dromaeosaur B	Dromaeosaur B	0.86
V5814	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.81
V5815	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.95
V5816	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.95
V5817	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.84
V5818	Neoceratosauria(2)	Dromaeosaur B	Dromaeosaur B	0.87
V5820	Neoceratosauria(1)	Abelisauridae	Tyrannosauridae	0.34
V5822	Coelurosauria (6)	Dromaeosaur A	Dromaeosaur A	0.70
V5823	Coelurosauria (6)	Dromaeosaur A	Dromaeosaur A	0.71
V5824	Coelurosauria (6)	Dromaeosaur A	Dromaeosaur A	0.70
V5959	Coelurosauria (6)	Dromaeosaur B	Dromaeosaur B	0.98

Table 5.7. Reclassification of the Balabansai Formation (Bathonian, Middle Jurassic, Kyrgyzstan) isolated theropod teeth (Averianov et al., 2005). Majority vote: assigned tooth morphotype following simple majority vote of three machine learning models. Combined posterior probability: assigned tooth morphotype by combining posterior probabilities from three machine learning models. P: combined posterior probability value.

		Machine learning		
Specimen	Averianov et al. (2005) classification	Majority Vote	Combined posterior probability	P
ZIN PH 7/42	Tetanurae indet.	Tyrannosauridae	Dromaeosaur B	0.18
ZIN PH 9/42	Tetanurae indet.	Tyrannosauridae	Dromaeosaur B	0.40
ZIN PH 14/42	Tetanurae indet.	Dromaeosaur B	Dromaeosaur B	0.80
ZIN PH 15/42	Tetanurae indet.	Dromaeosaur B	Dromaeosaur B	0.94
ZIN PH 18/42	Tetanurae indet.	Dromaeosaur B	Dromaeosaur B	0.96
ZIN PH 19/42	Tetanurae indet.	Dromaeosaur B	Dromaeosaur B	0.94
ZIN PH 28/42	Tetanurae indet.	Dromaeosaur B	Dromaeosaur B	0.87

Discussion

Applying 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 dromaeosaurid morphotypes (which might indicate multiple dromaeosaur taxa); a troodontid; and a therizinosauroid (Table 5.8). 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 to rigorously test these hypotheses. Buckley et al. (2010) and Buckley and 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 show 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 I distinguish the teeth only as morphotypes within a broader taxonomic framework.

Table 5.8. New theropod morphotypes from the Great Oolite Group, Bathonian, Middle Jurassic of Britain.

Watton Cliff, Dorset. Forest Marble Formation.	Dromaeosauridae indet. Morphotype B Dromaeosauridae indet. Morphotype C
Kirtlington Quarry, Oxfordshire. White Limestone Formation,	Dromaeosauridae indet. Morphotype A Dromaeosauridae indet. Morphotype B Dromaeosauridae indet. Morphotype C
Woodeaton Quarry, Oxfordshire. White Limestone Formation.	Dromaeosauridae indet. Morphotype A Dromaeosauridae indet. Morphotype B
Hornsleasow Quarry, Gloucestershire. Chipping Norton Limestone Formation.	Dromaeosauridae indet. Morphotype A Dromaeosauridae indet. Morphotype B Troodontidae indet. Therizinosauroidea indet.

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, increasing the known diversity of Middle Jurassic theropods from the UK and providing the oldest occurrences of troodontids and therizinosaurs worldwide (Fig. 5.8). These identifications provide the first definitive body-fossils consistent with predictions made by phylogenetic analyses, whose ghost lineages posited the likely presence of these clades at this time (Carrano et al., 2012; Holtz, 2000; Rauhut, 2003; Rauhut & Foth, 2020; Xu, QingYu, et al., 2010). Previous reports of Middle Jurassic maniraptoran occurrences have been disputed (Ding et al., 2020; Foth & Rauhut, 2017) 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 of stratigraphic uncertainties surrounding the placement of volcanic rocks within 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* (Huxley, 1868; Meyer, 1861; Owen, 1864) from the Late Jurassic of Germany, as well as the Yanliao Biota maniraptorans (Foth & Rauhut, 2017; Godefroit, Cau, et al., 2013; Godefroit, Demuynck, et al., 2013; Sullivan et al., 2014; Xu et al., 2015) 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). The confirmation of maniraptoran theropods from the Middle Jurassic of Gondwana as established by my re-evaluation of isolated teeth from India and Madagascar, notwithstanding the uncertainties around the absolute dating and stratigraphic position of those localities, adds further support to these results.

My results show that Maniraptora was not only established by the Bathonian but was already diverse at this time, in both Laurasia and Gondwana, and they also extend the known temporal ranges of all major maniraptoran clades significantly. Therizinosauroids, excluding the controversial occurrence of *Eshanosaurus* (Barrett, 2009; Xu et al., 2001), 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 (Joeckel et al., 2020; Kirkland et al., 2005; Senter et al., 2012; Zanno, 2010a) 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 at least 27 million years based on the dating of the Cedar Mountain Formation by Joeckel et al. (2020) (Fig. 5.8). However, more recent dates, based on chemostratigraphy and detrital zircons (Suarez et al., 2023), places the Yellow Cat Member of the Cedar Mountain Formation in the Aptian rather than the Berriasian which would increase this range extension to around 42 million years. Dromaeosaurs had an almost pan-global distribution during the Late Cretaceous, although 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; X. Zheng et al., 2009). Isolated teeth from the Middle and Late Jurassic of Laurasia and Gondwana have been assigned to the clade previously (Hendrickx & Mateus, 2014; Prasad & Parmar, 2020; Vullo et al., 2014; Zinke, 1998) but their identifications have not been widely accepted (Ding et al., 2020; Foth & Rauhut, 2017; Sellés et al., 2021). My 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 million years (Fig. 5.8). Based on comparisons with my data, it seems likely that some other published Jurassic records also represent this clade as has been shown by the re-analysis of records from India, Kyrgyzstan and Madagascar, although rigorous analysis will be needed to confirm this suggestion. Troodontids are known primarily from the Cretaceous of Asia, Europe and North America (Barsbold et al., 1987; Brown & Schlaikjer, 1943; Currie, 1987; Russell, 1946; Sellés et al., 2021) and possibly the Late Jurassic of China (Brusatte et al., 2014; Hu et al., 2009; Turner et al., 2012) 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; Goswami et al., 2013; Zinke, 1998) although many of these identifications have been questioned (Ding et al., 2020). Thus, this confirmed Middle Jurassic European troodontid pushes back the origin of this clade back by 27 million years (Fig. 5.8) from the Berriasian (*Geminiraptor*, Utah; Senter *et al.* 2010) to the Bathonian.

The presence of this diverse Middle Jurassic biota also suggests a need to re-visit the biogeographical scenarios that have been proposed to account for patterns in maniraptoran faunal distributions (Case et al., 2007; Ding et al., 2020; Rauhut et al., 2010; Zanno, 2010b). Two non-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 (Ding et al., 2020; Fastovsky & Weishampel, 1996; Upchurch et al., 2002; Zanno, 2010b), and faunal dispersal with dispersal routes shaped by the establishment of land bridges between continental masses (Ding et al., 2020; Dunhill et al., 2016; Upchurch et al.,

2002). It is also likely that regional extinction events played a part in shaping biogeographical distributions (Barrett et al., 2011; Benson et al., 2012; Sereno, 1997). The presence of Middle Jurassic Laurasian proceratosaurids and earliest Cretaceous Gondwanan ornithomimosaurs suggests that coelurosaurs were widespread before the breakup of Pangaea (Choiniere et al., 2012; Rauhut et al., 2010) 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 (Foth & Rauhut, 2017; Sellés et al., 2021; Zanno, 2010b) several different scenarios have been put forward to account for maniraptoran distributions, whilst accepting that more fossil evidence would be needed in order to test these. For example, Foth and 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 the 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 (Ding et al., 2020; Scotese, 2021; Zanno, 2010b). 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 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) suggest 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 (Ding et al., 2020; Goswami et al., 2013; Senter et al., 2010).

The confirmed presence of maniraptorans in the Middle Jurassic (Fig. 5.22) suggests a pan-Pangaean distribution was established before continental separation began at

~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.

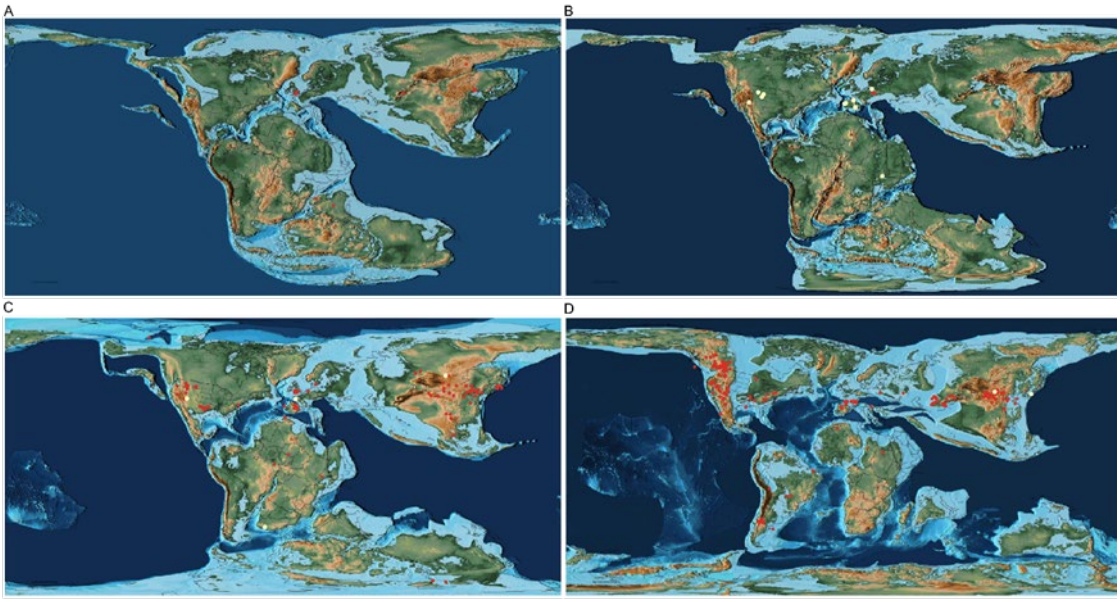


Fig. 5.22. Biogeographical history of Maniraptora from the Middle Jurassic to the Late Cretaceous. Middle Jurassic (A), Late Jurassic (B), Early Cretaceous (C), Late Cretaceous (D). Manirpatoran occurences shown in red circles and unconfirmed occurences in yellow circles. Middle Jurassic occurences include the UK results from this study and those from re-analysed data from India, Kyrgyzstan and Madagascar. Palaeogeographic maps from Scotese (2021).

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 (see Chapter Three; Wills et al. (2021). Using multiple machine learning algorithms. as applied here, allows the corroboration of results by checking predictions derived from another technique. It is also important to note the limitations of any technique and this 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, I 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, I

emphasise the importance of cross-checking results from machine learning analyses with more traditional morphological-based approaches.

Conclusions

Application of machine learning algorithms has allowed me to confirm, in a quantifiable framework, the presence of a diverse maniraptoran theropod fauna in the Middle Jurassic (Bathonian) of the UK. This greatly expands the known diversity of theropods from this time period in the UK. My sample includes the oldest-known occurrences of Troodontidae and Therizinosauroidea. This confirms a Middle Jurassic (or earlier) origin for Maniraptora. These results taken in conjunction with the confirmed presence, as a result of my reanalysis of published data, of maniraptoran theropods from India, Kyrgyzstan and Madagascar 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, providing a robust, testable framework for taxonomic identifications.

Chapter Six: Ornithischia

Abstract

Current research suggests ornithischians originated in the Middle to Late Triassic, achieving global dominance, with the Middle Jurassic a pivotal period in which the clade underwent rapid diversification and radiation. However, fossils of Middle Jurassic ornithischians are rare, with few named taxa and numerous occurrences of isolated teeth with disputed identifications. Here, I apply detailed morphological comparisons, to a suite of isolated ornithischian teeth from Bathonian microvertebrate sites in the UK, to assess the taxonomic affinities of the teeth. Undertaking these morphological comparisons has revealed a hitherto unknown, and highly diverse, ornithischian fauna from these sites which significantly increases the known diversity of ornithischian dinosaurs from this time period in the UK. The comparisons indicate the presence of six distinct ornithischian morphotypes: an indeterminate ornithischian, a heterodontosaurid, two indeterminate thyreophorans, a stegosaur and an ankylosaur. These results confirm the predictions made by phylogenetic studies that Ornithischia rapidly diversified in the Middle Jurassic, fill in temporal gaps in lineages and include one of the oldest global occurrences of stegosaurs. In addition, the mixture of non-eurypodan and eurypodan morphotypes identified here suggests that not only did non-eurypodans survive until at least the Middle Jurassic but they also co-existed in close temporal and spatial proximity with early eurypodans.

Introduction

Ornithischia, a speciose clade of mainly herbivorous dinosaurs, likely originated during the Middle to Late Triassic (Baron et al., 2017a; Boyd, 2015; Dieudonné et al., 2021), and later achieved high levels of abundance and a global distribution, with members of the clade discovered on every continent including Antarctica (Fig. 6.1). Nevertheless, representatives of this clade, although well-known from the Late Jurassic and Cretaceous (Weishampel et al., 2004), are poorly known from the Middle Jurassic (Fig. 6.2). This is in spite of numerous predictions made by multiple phylogenetic analyses that ornithischians underwent a major radiation at this time (Boyd, 2015; Butler et al., 2008; Dieudonné et al., 2021; Raven et al., 2023). Moreover, many of the earliest fossils of the major ornithischian clades (e.g., Ornithopoda, Ankylosauria, Stegosauria) are rare

specimens recovered from Middle Jurassic strata in Europe, east Asia and north Africa (e.g., Hui et al., 2022; Maidment et al., 2020; Maidment et al., 2021; Ruiz-Omeñaca et al., 2007). These lines of evidence suggest that the group achieved a global and diversified distribution by the Middle Jurassic, but with most of its history hidden by a poor fossil record (Boyd, 2015; Maidment et al., 2020; Maidment et al., 2021; Raven et al., 2023). Our understanding of the terrestrial vertebrate diversity of this period, including that of ornithischians, has been enhanced by material from microvertebrate sites (Evans et al., 2006; Evans & Milner, 1994; Wills et al., 2019; Wills, Underwood, et al., 2023). However, many of these specimens are excluded from broader scale analyses of distribution and diversity as they are difficult to incorporate into formal phylogenetic or biogeographic analyses, thereby excluding a rich source of data that can potentially inform us on taxon geographic ranges, palaeoecology and divergence dates.

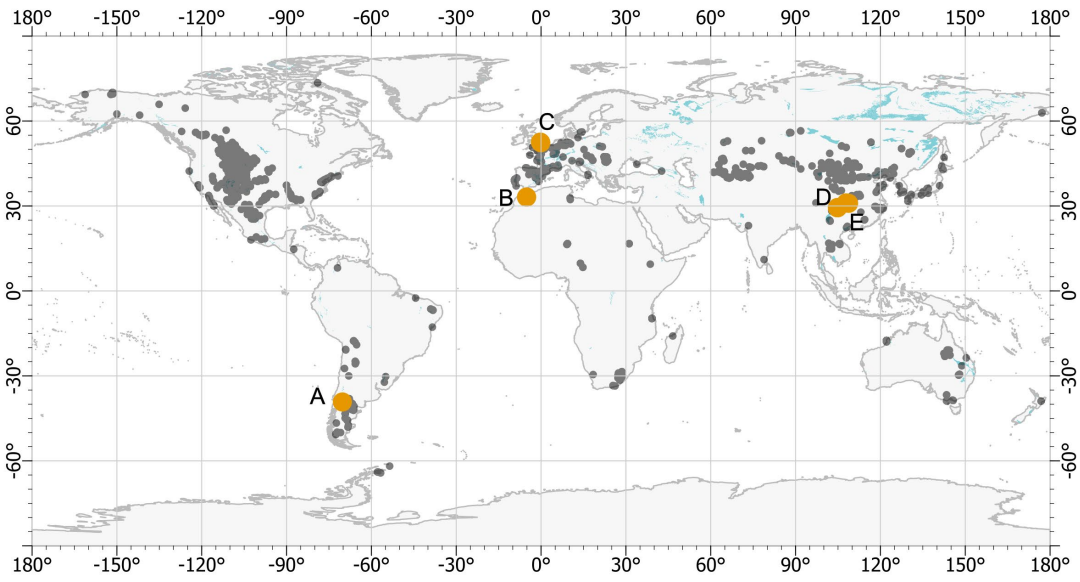


Fig. 6.1. Global distribution of ornithischian localities, with named Middle Jurassic taxa shown in brown: (A) early Bajocian, Argentina, *Isaberrysaura mollensis*; (B) late Bathonian / Callovian, Morocco, *Adratiklit boulahfa* and *Spicomellus afer*; (C) Callovian, UK, *Loricatosaurus priscus*, *Sarcolestes leedsi* and *Callovosaurus leedsi;* (D) Bajocian, China, *Yandusaurus hongheensis*, *Agilisaurus louderbacki*, *Hexinlusaurus multidens*, and *Xiaosaurus dashanpensis*. Bathonian / Callovian, China, *Huayangosaurus taibaii*; (E) Bajocian, China, *Bashanosaurus primitivus*. The occurrence data were downloaded from the Paleobiology Database on 20th January 2023.

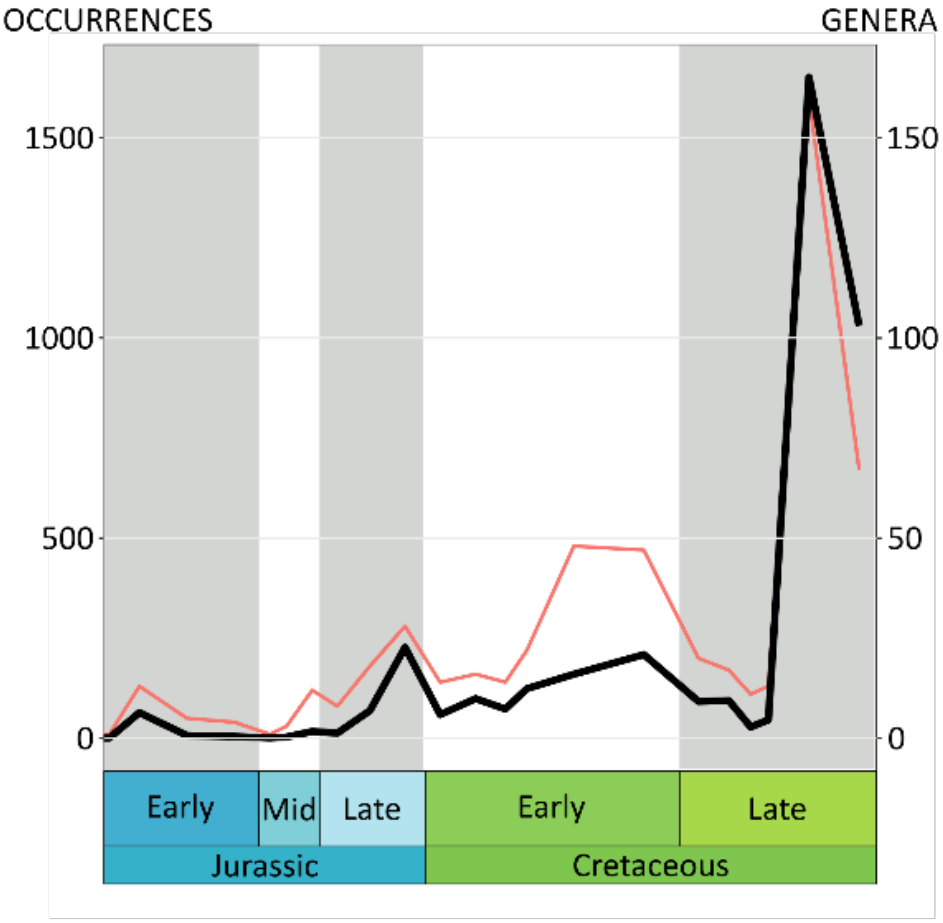


Fig. 6.2. Jurassic and Cretaceous ornithischian occurrences and generic diversity from the Paleobiology Database binned by stage. Occurrences are shown by the black line and generic diversity by the red line. The data were downloaded from the Paleobiology Database on 20th January 2023.

The ornithischian record of the UK Middle Jurassic is, in common with the rest of the world, sparse with the only named taxa occurring in the Callovian Oxford Clay Formation (Naish & Martill, 2008): the stegosaur *Loricatosaurus priscus* (Maidment et al., 2008; Nopcsa, 1911); the ankylosaur *Sarcolestes leedsi* (Galton, 1980a, 1983a; Lydekker, 1893); and the dryosaurid *Callovosaurus leedsi* (Galton, 1980b; Lydekker, 1889; Ruiz-Omeñaca et al., 2007). Indeterminate material is more common and includes: the fragmentary skeleton of an indeterminate stegosaur (Boneham & Forsey, 1992; Galton & Powell, 1983; Maidment et al., 2008) from near Chipping Norton, Oxfordshire (Sharps Hill Formation, Bathonian)**;** dermal armour referred to an indeterminate stegosaur (Galton & Powell, 1983; Maidment et al., 2008) from New Park Quarry, Gloucestershire

(Chipping Norton Formation, Bathonian); a stegosaurian fibula (Panciroli, Funston, et al., 2020) from the Isle of Eigg (Great Estuarine Group, Bathonian); an ulna and radius from the Bajocian of the Isle of Skye (Clark, 2001); indeterminate stegosaur, ornithopod and thyreophoran material (Galton, 1985; Maidment et al., 2008) from Fletton, Peterborough (Oxford Clay Formation, Callovian); a dermal spine (Barrett & Maidment, 2011) and a cervical centrum (Galton, 2017) referred to indeterminate thyreophorans and other indeterminate dinosaur material from the Inferior Oolite Formation (Aalenian to Bajocian) in Dorset; an isolated tooth referred to an indeterminate ornithopod from the Bathonian Great Oolite Group at Stonesfield, Oxfordshire (Galton, 1975); and isolated teeth from microvertebrate sites which have been referred tentatively to several ornithischian groups by previous authors, and which form the basis of this study (Benton & Spencer, 1995; Evans & Milner, 1994; Metcalf et al., 1992; Metcalf & Walker, 1994; Wills et al., 2019).

Here, I describe the ornithischian dinosaur remains from the four Middle Jurassic microvertebrate assemblages: Hornsleasow Quarry, Gloucestershire; Woodeaton and Kirtlington Quarries, Oxfordshire; and Watton Cliff, Dorset. This reveals the presence of heterodontosaurids, early-diverging thyreophorans, stegosaurs and ankylosaurs, as well as other indeterminate ornithischian taxa. These morphotypes fill important gaps in the known fossil records of these groups and extend the geographic range of the main thyreophoran clades, confirming that ornithischians had attained high diversity and a wide distribution by the Middle Jurassic.

See Appendix One for the Data Archiving Statement (for the micro-CT data) and Appendix Four for a list of institutional abbreviations.

Geological setting

The sites range in age from the basal Bathonian Chipping Norton Limestone Formation (Hornsleasow Quarry) through to the uppermost Bathonian Forest Marble Formation (Watton Cliff) and expose near-shore terrestrial sediments within an otherwise marine or marginal-marine succession. The oldest of the sites, Hornsleasow Quarry in the Chipping Norton Limestone Formation (*Z. zigzag* zone: (Cope et al., 1980), previously exposed a clay lens lying on a palaeokarst surface representing a coastal marsh pond (Metcalf et al., 1992; Metcalf & Walker, 1994; Vaughan, 1989). Woodeaton and

Kirtlington Quarries are found in the Bladon Member of the White Limestone Formation (*H. retrocostatum* zone) and are approximately coeval although it appears that the succession at Woodeaton is slightly lower in the stratigraphy (Wills et al., 2019). The microvertebrate horizon at Woodeaton represents a brackish-water lagoon that experienced seasonal aridity and fluctuating salinity with periodic influxes of seawater (Wills et al., 2019). The Kirtlington microvertebrate horizon, also known as the 'Mammal Bed' (Freeman, 1976a, 1979; McKerrow et al., 1969), formed in a shallow marginal marine environment during a period of marine regression along a coastal plain (Evans & Milner, 1994; Freeman, 1979; Palmer, 1979). Finally, the Watton Cliff microsite, in the Forest Marble Formation (*Discus* zone), is the only site where fully marine conditions prevailed, with moderate water depth and signs of weak storm influence on the sedimentation (such as rippled sand lenses). Here, the microvertebrates are commonly found in unconsolidated patches of bioclastic gravel within a strongly cemented unit which seem to be either channels or burrows representing deposition of an emergent shell bank, possibly during storm-related events (Holloway, 1983). For more detail on the geological setting of the sites see Chapter Four and Evans and Milner (1994), McKerrow et al. (1969), Metcalf et al. (1992), Wills et al. (2019) and Wills, Underwood, et al. (2023).

Materials and Methods

The sample consists of 578 isolated ornithischian teeth from four Middle Jurassic (Bathonian, Great Oolite Group) localities in the UK (Fig. 6.3): Hornsleasow Quarry, Gloucestershire (Chipping Norton Limestone Formation), 54 teeth; Woodeaton Quarry, Oxfordshire (White Limestone Formation), 29 teeth; Kirtlington Quarry, Oxfordshire (White Limestone Formation), 490 teeth; and Watton Cliff, Dorset (Forest Marble Formation), five teeth. It was collected over a period of several decades by teams from different institutions, including the Natural History Museum (Woodeaton, Watton Cliff), UCL (Kirtlington, Watton Cliff) and the University of Bristol/Museum of Gloucester (Hornsleasow). Except for the Hornsleasow material prefaced with GLRCM, which is housed in the Museum of Gloucester, all the specimens are in the collections of the Natural History Museum, London. Chapter Two details the methodology used to collect and process the specimens.

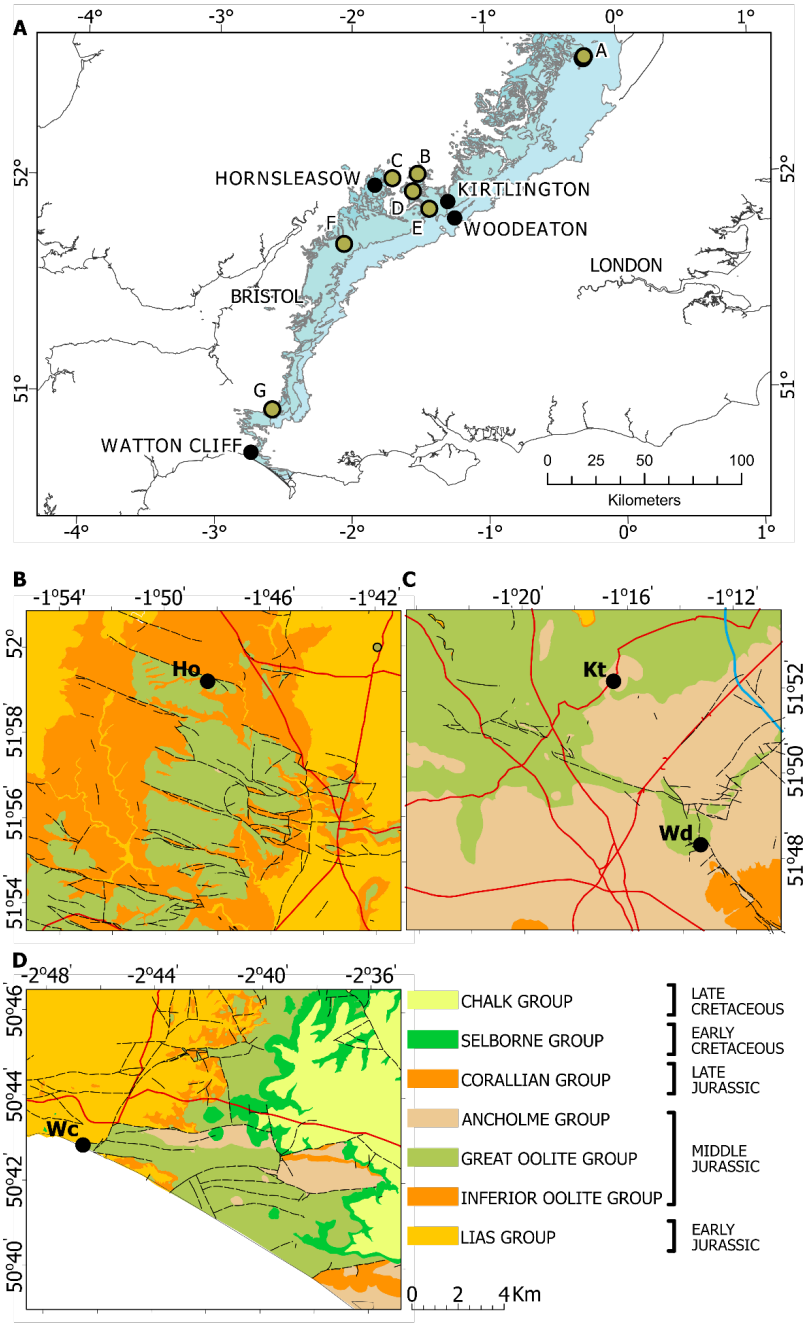


Fig. 6.3. Site localities, known Middle Jurassic ornithischian occurrences, and geological settings for specimens described in this study. A, site localities for this study and known Middle Jurassic ornithischian occurrences in central and southern England. (A), *Loricatosaurus priscus*, *Callovosaurus leedsi* and *Sarcolestes leedsi* Oxford Clay Formation, Callovian, Peterborough. (B) indeterminate ornithischian isolated teeth, Bathonian, Oxfordshire, (C) indeterminate Stegosauridae, Bathonian, New Park Quarry, Oxfordshire, (D) indeterminate Stegosauridae, Bathonian, Oxfordshire, (E) indeterminate ornithopod tooth, Bathonian, Oxfordshire, (F) isolated indeterminate ornithischian teeth, Bathonian, Oxfordshire, (G) Indeterminate dinosaur and thyreophoran remains, Aalenian to Bajocian, Dorset. B–D, geological setting of Hornsleasow quarry (Ho), Kirtlington (Kt) and Woodeaton (Wd) quarries, and Watton Cliff (Wc). Licence number 2017/024 ED British Geological Survey (c) NERC. All rights reserved.

I identified 509 teeth from the picked residue that were suitable for further investigation. All of the teeth in the sample were imaged with a combination of optical imaging using a Dino-Lite AM 7915 MZTL microscope and SEM on a LEO 1455VP microscope. I also CT-scanned each (complete) tooth on a Nikon Metrology HMX ST 225 µCT and a Zeiss Versa µCT at a range of voxel resolutions from 3–30 µm and created 3D models from the CT volumes using Avizo v. 8.1 (ThermoFisher, 2014). All measurements were taken either from the digital images or measured directly from the 3D models using Avizo.

As all of the tooth crowns are isolated, and it is unlikely for maxillary teeth to be preferentially represented over dentary teeth (or vice versa), it is not possible to distinguish between maxillary and dentary teeth, nor can I distinguish whether these pertain to the left or right sides of the jaws. Early-diverging ornithischians, such as thyreophorans and early neornithischians, show little morphological variation along the tooth rows, with most changes related to size, apart from those taxa which possess premaxillary or more specialised teeth, such as ornithopods and heterodontosaurids. To aid morphological comparisons I made the following assumptions: the convex tooth surface is the labial surface in all cases; and any offset of the crown apex is in a posterior or lingual direction. I refer all the teeth described below to various clades in Ornithischia based on a combination of features: the mesiodistal expansion of the crown relative to the root; the presence of a cingulum; apically inclined, relatively large denticles; a low triangular crown in lingual / labial views (for cheek teeth); and asymmetry of the crown in labial / lingual view (Butler et al., 2008; Irmis et al., 2007; Sereno, 1986). Although many of these features are shared with other Mesozoic archosaurs, such as the Late Triassic aetosaur *Revueltosaurus* (Boyd, 2015; Irmis et al., 2007; Nesbitt et al., 2007; Parker et al., 2005), and (with the exception of a basal cingulum) cannot be considered to be unique synapomorphies of ornithischians (Butler et al., 2008; Nesbitt et al., 2007), the Middle Jurassic age of these teeth implies that the most parsimonious assignment is to regard them as referrable to Ornithischia rather than to other archosaur clades.

Systematic palaeontology

Ornithischia indet.

DINOSAURIA Owen, 1842

ORNITHISCHIA Seeley, 1887

Gen. et sp. Indet.

Figure 6.4

Referred specimens: NHMUK PV R JR316, JR317, JR318, JR319, JR360, K1, K11

Localities and horizons: Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002).

Description: This morphotype is represented by seven small, isolated teeth ranging in maximum crown height from 2.4–2.7 mm (Fig. 6.4). The crowns are isosceles in outline, apicobasally elongate and mesiodistally asymmetrical with a distinct mesial shoulder in lateral view and the apex of the crown offset distally. In lateral view, the mesial margin is convex whereas the distal margin is straight. In mesial / distal view the apex of the crown is inclined lingually. The teeth are mesiodistally widest immediately above the base of the crown where mesiodistal / labiolingual expansion of the crown forms a cingulum. The cingulum is more prominent on the labial surface of the crown. Denticles are generally absent from both margins of the teeth, although a small apically pointing denticle is developed basally on the mesial margin of NHMUK PV R JR360. The crowns possess a series of prominent apicobasally extending ridges and deep grooves on both lingual and labial surfaces. The ridges pass across the cingulum becoming less prominent below this point to eventually merge with the preserved surface of the root. Apically, the ridges merge to form either a smooth surface around the crown apex (NHMUK PV R JR316) or continue to almost reach the apex (NHMUK PV R JR 319). The ridges on the lingual surface are more prominent than those on the labial surface, with most teeth possessing four or five lingual ridges and three labial ridges. The ridges converge towards the apex of the crown and diverge basally so that the corresponding groves widen towards the crown base. A distinct primary ridge is missing. The ridges are narrowest at the midpoint of the crown and increase slightly in width both apically and basally. In basal view the preserved part of the root is sub-circular to oval in outline.

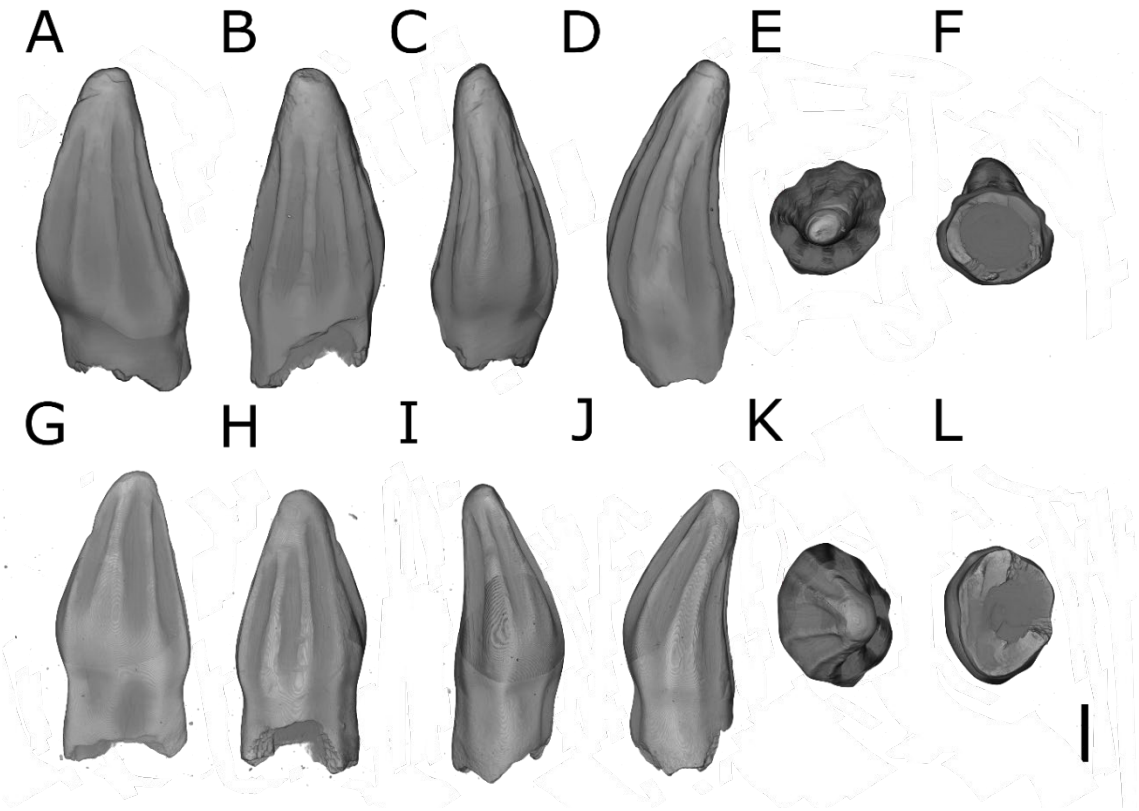


Fig. 6.4. Teeth of indeterminate ornithischians from Kirtlington Quarry, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic. NHMUK PV R JR316 (A to F), NHMUK PV R JR319 (G to L) in labial (A, G), lingual (B, H), mesial (C, I), distal (D, J), apical (E, K) and basal (F, L) views. Scale bar 0.5 mm.

Morphological comparisons: These distinctive highly ridged teeth occur only at Kirtlington Quarry where they were originally referred to either Ornithopoda or Crocodylia (Evans & Milner, 1994). Extensive sampling of the approximate coeval deposit at Woodeaton Quarry failed to find any further specimens, even though the microvertebrate faunas from both sites are broadly comparable (Wills et al., 2019). The partial preservation of a cylindrical root on some teeth, indicating thecodont implantation, allows referral to Archosauria. Triassic pseudosuchians and basal crocodylomorphs are not known to have survived into the Middle Jurassic (Nesbitt, 2011; Parker et al., 2005) limiting possible choices for these teeth to either terrestrial or semi-aquatic crocodyliforms or to Dinosauria.

During the Middle to Late Jurassic small-bodied non-pelagic sphenosuchian, protosuchian, goniopholidid and atoposaurid crocodyliforms were widely distributed, although within Europe the fauna was limited to goniopholidids, atoposaurids and

indeterminate neosuchians (Evans & Milner, 1994; Wills et al., 2014; Yi et al., 2017; Young et al., 2015). However, none of these clades have a tooth morphology resembling the specimens described here. The atoposaurid *Theriosuchus,* from the Middle Jurassic to Late Cretaceous of Europe (Schwarz & Salisbury, 2005), including the Bajoican–Bathonian Valtos Sandstone of the Isle of Skye (Young et al., 2015), possesses a strongly heterodont dentition with teeth that are clearly distinct from those described here (Schwarz & Salisbury, 2005). Middle to Late Jurassic goniopholidids such as *Sunosuchus* (Wings et al., 2010) have small ridged, slightly recurved teeth but lack the strongly developed ridges, mesial shoulder and cingulum present here. Within Dinosauria, sauropods display a range of tooth morphologies ranging from the cylindrical teeth of diplodocids and titanosaurs to the broader spatulate crowns of *Camarosaurus* (Upchurch et al., 2004), and can be excluded on the basis of gross morphology. The only two theropod clades with a basal constriction between crown and root (which could resemble a cingulum), Therizinosauroidea and Troodontidae, have a distinctly different overall crown morphology including prominent denticles on the mesial and distal margins (Hendrickx et al., 2019; Wills, Underwood, et al., 2023). The presence of a basal cingulum and morphological differences from other Middle Jurassic archosaur teeth suggests that the most parsimonious solution is to refer these teeth to Ornithischia. The lack of denticles and the high-crowned, slightly recurved morphology is reminiscent of ornithischian premaxillary teeth, although I am unable to identify a clade with identical tooth morphology. Premaxillary teeth occur in early-diverging ornithischians such as *Lesothosaurus* (Porro et al., 2015; Sereno, 1991), some early ornithopods (Barrett & Han, 2009; Brown & Druckenmiller, 2011; Galton, 1974), Early Jurassic thyreophorans such as *Scutellosaurus* (Breeden et al., 2021; Colbert, 1981), *Laquintasaura* (Barrett et al., 2014), heterodontosaurids (Butler et al., 2012; Sereno, 2012), ceratopsians and pachycephalosaurs (Sullivan, 2006), neornithischians such as *Thescelosaurus and Jeholosaurus* (Boyd, 2014), the ankylosaurs *Silvisaurus* (Eaton, 1960) and *Cedarpelta* (Carpenter, 2001), and the stegosaurs *Huayangosaurus* (Sereno & Dong, 1992) and *Isaberrysaura* (Salgado et al., 2017). Premaxillary crowns of *Lesothosaurus* are distally recurved with denticulate mesial and distal margins and lack a ridged surface ornamentation (Porro et al., 2015; Sereno, 1991). Some premaxillary crowns of basal ornithopods and neornithischians, such as *Hypsilophodon* and *Thescelosaurus* (Fig. 6.5),

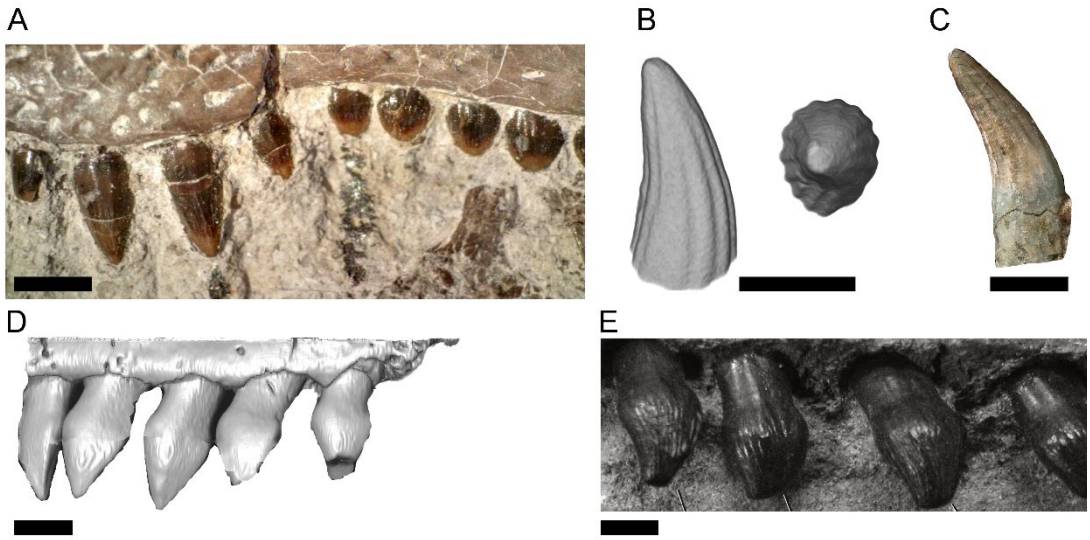


Fig. 6.5. Comparative crocodilian, basal thyreophoran and basal neornithischian teeth. Crocodilian (A and B) and premaxillary teeth of basal thyreophoran (C) and basal neornithischians (D and E). A, *Theriosuchus pusillus* (NHMUK PV OR 48228), right maxillary teeth in labial view. Purbeck Limestone Group, Berriasian, Early Cretaceous, Swanage, Dorset, UK (Owen, 1879). B, Goniopholidid crocodile (unregistered NHMUK), in labial and apical views. 'Mammal Bed', Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002), Kirtlington Quarry, Oxfordshire, UK. C, *Laquintasaura venezuelae* (MBLUZ P5050), premaxillary tooth in mesial/distal view. La Quinta Formation, Hettangian, Early Jurassic, Rio La Grita, Venezuela (Barrett et al., 2014). D, *Hypsilophodon foxii* (NHMUK PV R 197), left premaxillary teeth in labial view (posterior to the right). Vectis Formation, Barremian, Early Cretaceous, Isle of Wight, UK. E, *Thescelosaurus neglectus* (NCSM 15728), right premaxillary dentition in lateral view. Hell Creek Formation, Maastrichtian, Upper Cretaceous, South Dakota, USA. (Boyd, 2014).
Scale bars 2mm

also possess a superficially similar ridged ornamentation that extends from the crown apex to the base of the crown (Boyd, 2014; Brown & Druckenmiller, 2011; Galton, 1974); however, this ornamentation comprises much finer ridges than seen on the Kirtlington teeth. *Scutellosaurus* premaxillary teeth are finely denticulate, triangular with a sharp apex, and lack ridges or striations (Breeden et al., 2021; Colbert, 1981). In heterodontosaurids, the premaxillary teeth are either conical and lack a cingulum or, as in the case of *Frutiadens,* are subtriangular and expanded above the root (Butler et al., 2012). Those of early-diverging Late Jurassic ceratopsians, such as *Yinlong,* are semi-conical in shape, recurved distally with a sharp apex and lack ridges (Hu et al., 2022). The

premaxillary teeth of the ankylosaur *Silvisaurus* are low-crowned, denticulate and triangular in shape (Eaton, 1960). The stegosaur *Huayangosaurus* has phylliform, strongly denticulate, teeth with a distally offset apex (Sereno & Dong, 1992), and *Isaberrysaura* has conical, slightly asymmetric globose teeth that lack strong ornamentation (Salgado et al., 2017). Premaxillary teeth of the basal thyreophoran *Laquintasaura* have a very similar overall gross morphology with lingually inclined, isosceles shaped crowns that possess a distinctive highly ridged ornamentation on the labial and lingual surfaces, although the ridges are not as strongly developed (Barrett et al., 2014).

The teeth from Kirtlington Quarry cannot be referred with certainty to any known ornithischian and likely represent a new taxon. There are morphological similarities with the premaxillary teeth of *Laquintasaura* and these teeth may represent a related lineage, but in the absence of more diagnostic material I refrain from naming these specimens and refer them to Ornithischia indet.

Heterodontosauridae indet.

HETERODONTOSAURIDAE Kuhn, 1966

Gen. et sp. Indet.

Figure 6.6

Referred specimens: GLRCM GTUBE69, NHMUK PV R WD51.

Unregistered specimens: five

Localities and horizons: Hornsleasow Quarry, Chipping Norton Limestone Formation, Bathonian, Middle Jurassic (Metcalf et al., 1992), Kirtlington Quarry, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019), Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019).

Description: Two small (crown height [CH] 3.0–3.5 mm), sub-triangular, mesiodistally expanded crowns (Fig. 6.6) with coarse denticles on the mesial and distal margins (NHMUK PV R WD51 and GLRCM GTUBE69). Symmetrical in lateral view and asymmetrical in mesiodistal views with a cingulum developed above the crown-root

junction. In mesiodistal view the labial surface is slightly convex whereas the lingual surface is broadly concave with the crown base extending below the corresponding point on the labial surface. The maximum mesiodistal expansion of the crown occurs at a point approximately halfway up the lingual surface of the crown. The apicobasal length of the crown is greater than its mesiodistal width. Marginal denticles are restricted to the upper third of the crown, beginning at a point just above the maximum mesiodistal expansion of the crown and continuing to the crown apex, with equal numbers of denticles on both the mesial and distal margins. The apical denticle is connected to the crown base by a prominent broad apicobasal ridge which is present on both the lingual and labial surfaces of the teeth. The apically oriented marginal denticles are equal in size on both mesial and distal margins with the basal-most distal denticle shifted slightly towards the cingulum. GLRCM GTUBE69 has a preserved cylindrical root, with a circular cross-section that tapers basally.

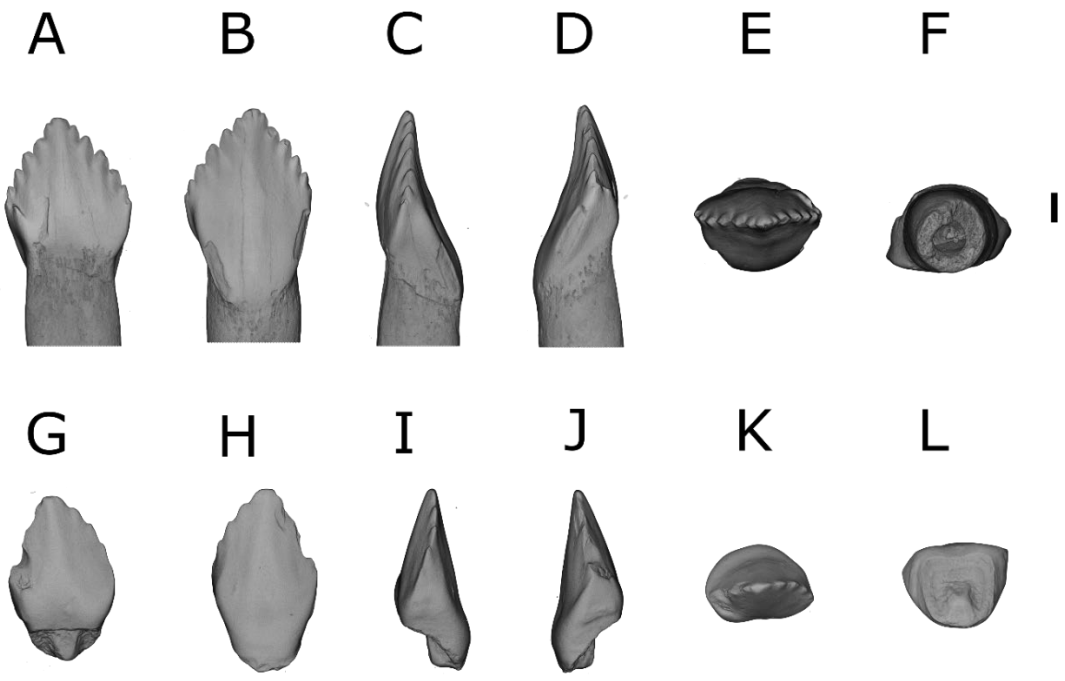


Fig. 6.6. Teeth of indeterminate heterodontosaurids from Hornsleasow Quarry (GLRCM GTUBE69), Chipping Norton Limestone Formation, Bathonian, Middle Jurassic and Woodeaton Quarry (NHMUK PV R WD51), Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic. GLRCM GTUBE69 (A to F), NHMUK PV R WD51 (G to L) in labial (A, G), lingual (B, H), mesial (C, I), distal (D, J), apical (E, K) and basal (F,L) views. Scale bar 0.5 mm.

Morphological comparisons: These small teeth, with denticles restricted to the upper third of the crowns, are found at both Hornsleasow and Woodeaton Quarries. Similar teeth, although damaged and missing marginal denticles, have also been recovered from Kirtlington Quarry and I refer them to Heterodontosauridae based on gross morphology. Small triangular to sub-triangular denticulate cheek teeth are present in many ornithischian clades with the majority, such as early thyreophorans, ankylosaurs, stegosaurs and most neornithischians, having cheek teeth that are denticulate along their entire mesial and distal tooth margins. A denticle distribution limited to the apical third of the crown is restricted to heterodontosaurids (Fig. 6.7), apart from *Fruitadens*, and the Late Jurassic basal ceratopsian *Chaoyangsaurus* (Butler et al., 2012; Norman & Barrett, 2002; Zhao et al., 1999). Ceratopsians are currently unknown from the Middle Jurassic, with the earliest members of the clade such as *Yinlong* (Xu et al., 2006) and *Chaoyangsaurus* (Zhao et al., 1999) dating from the Late Jurassic. By contrast, heterodontosaurids have a wide stratigraphic distribution ranging from the upper Elliot Formation (Early Jurassic, Hettangian to Sinemurian) of southern Africa (Bordy et al., 2020; Porro et al., 2011; Viglietti et al., 2020) to the Purbeck Limestone Group (Early Cretaceous, Berriasian) of the UK (Galton, 1978; Norman & Barrett, 2002; Owen, 1879). The presence of heterodontosaurids in other Middle Jurassic deposits is disputed. *Manidens condorensis* (Pol et al., 2011) from the basal levels of the Cañadón Asfalto Formation (Argentina) is considered to be either Early (Becerra et al., 2018; Becerra et al., 2020) or Middle Jurassic in age (Pol et al., 2011; Sereno, 2012) with recent radiometric dating supporting a middle–late Toarcian age (Becerra et al., 2023; Pol et al., 2020). The age of *Tianyulong* from the Middle to Upper Jurassic Daohugou Beds (Yanliao Biota) in northeastern China is controversial because of stratigraphic uncertainties surrounding the placement of volcanic rocks within the sequence used to obtain radiometric dates (Sullivan et al., 2014; Xu et al., 2016) with recent work considering it to be either Early Cretaceous (Dieudonné et al., 2021) or Middle Jurassic (Callovian) (Becerra et al., 2023).

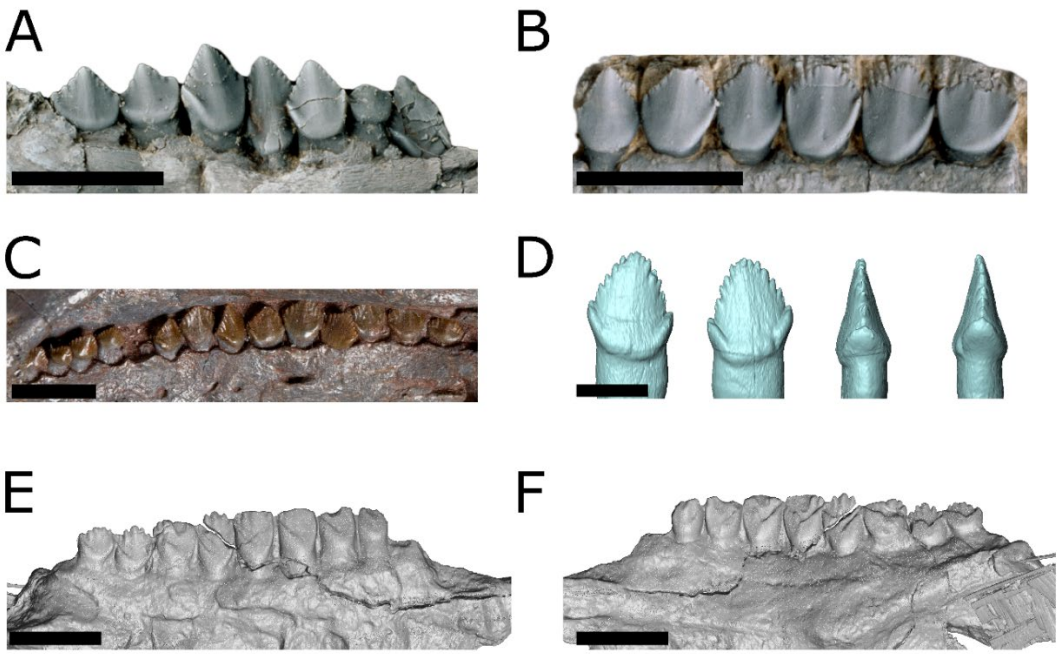


Fig. 6.7. Cheek teeth of various heterodontosaurid genera and the early ornithischian *Lesothosaurus.* (A) *Echinodon becklesii* (NHMUK PV R 48213), left dentary tooth crowns in lingual view. Purbeck Limestone Group, Berriasian, Early Cretaceous, Swanage, Dorset, UK. (B), *Echinodon becklesii* (NHMUK PV R 48211), right maxillary tooth crowns in labial view (inverted). Purbeck Limestone Group, Berriasian, Early Cretaceous, Swanage, Dorset, UK. (C) *Lesothosaurus diagnosticus* (NHMUK PV R 8501), maxillary tooth crowns in labial view (inverted). Upper Elliot Formation, Hettangian–Sinemurian, Early Jurassic, Lesotho. (D) *Fruitadens haagarorum* (LACM 115747), dentary tooth in labial, lingual, mesial and distal views (Butler et al., 2012). Brushy Basin Member, Morrison Formation, Tithonian, Upper Jurassic, Colorado, USA. (E) *Manidens condorensis* (MPEF PV 3809), maxillary tooth crowns in labial view (inverted). Cañadón Asfalto Formation, ?Toarcian, Upper Jurassic, Patagonia, Argentina. (F) *Manidens condorensis,* (MPEF PV 3809), maxillary tooth crowns in lingual view (inverted). Cañadón Asfalto Formation, ?Toarcian, Upper Jurassic, Patagonia, Argentina (Becerra et al., 2013). Scale bars 5mm.

Thyreophora indet. Morphotype A

THYREOPHORA Nopcsa, 1915 (sensu Butler, Uphurch & Norman, 2008)

Gen. et sp. indet.

Morphotype A

Figure 6.8

Referred specimens: NHMUK PV R WD90 and NHMUK PV R WD91

Unregistered specimens: nine

Localities and horizons: Hornsleasow Quarry, Chipping Norton Limestone Formation, Bathonian, Middle Jurassic (Metcalf et al., 1992), Kirtlington Quarry, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002), and Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019).

Description: Two largely complete sub-triangular to lanceolate, mesiodistally expanded, labiolingually compressed and coarsely denticulate crowns lacking the root (Fig. 6.8). The crowns are small with crown heights ranging from 2.32–4.15 mm and maximum mesiodistal widths from 1.69–2.52 mm. The crowns are slightly recurved distally and inclined lingually making the teeth asymmetrical in labial / lingual view. Both the lingual and labial crown surfaces are smooth and unornamented. In lingual / labial views the mesial margin is convex in outline whereas the distal margin is straight. The crowns have a basal swelling between crown and root which is more prominent on the lingual surface of the teeth but does not form a prominent cingulum. Both margins of the teeth have between five to seven prominent apically pointing denticles which are larger on the distal margin. There are no primary or secondary ridges on either the lingual or labial surfaces of the crowns. A broad central eminence is present on both lingual and labial surfaces, splitting the crown into two, extending from the base of the crown apically, narrowing as it approaches the apex and forming distinct depressions medially from both margins of the crown. Basally the eminence broadens and merges with the basal swelling. Narrow curved shelf-like ledges are present on the apical surface of the basal

swelling towards the mesial and distal margins of the crown. The roots are circular to elliptical in cross-section.

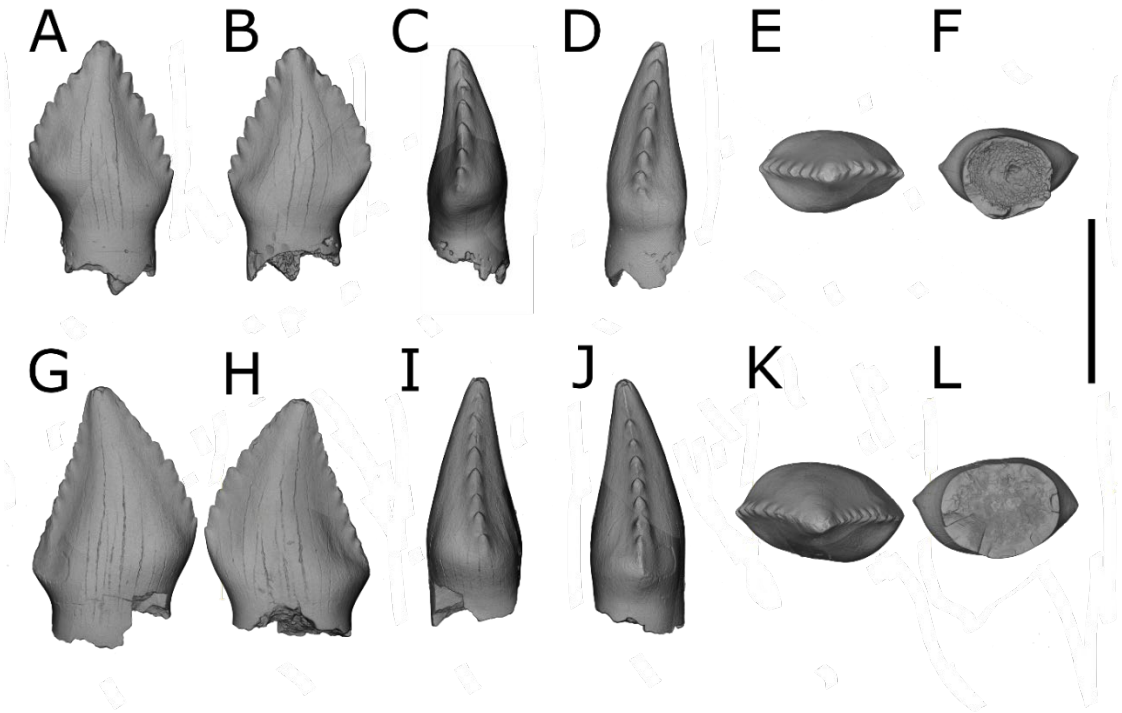


Fig. 6.8. Unworn teeth of indeterminate thyreophorans (Morphotype A) from Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019). NHMUK PV R WD90 (A to F), NHMUK PV R WD91 (G to L) in labial (A, G), lingual (B, H), mesial (C, I), distal (D, J), apical (E, K) and basal (F,L) views. Scale bar 2 mm.

Thyreophora indet. Morphotype B

Gen. et sp. Indet.

Morphotype B

Figure 6.9

Referred specimens: NHMUK PV R JR351 and NHMUK PV R JR352

Unregistered specimens: 132

Localities and horizons: Kirtlington Quarry, ‘Mammal Bed’, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002), Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle

Jurassic (Wills et al., 2019), Watton Cliff, Forest Marble Formation, Middle Jurassic (Hunter & Underwood, 2009).

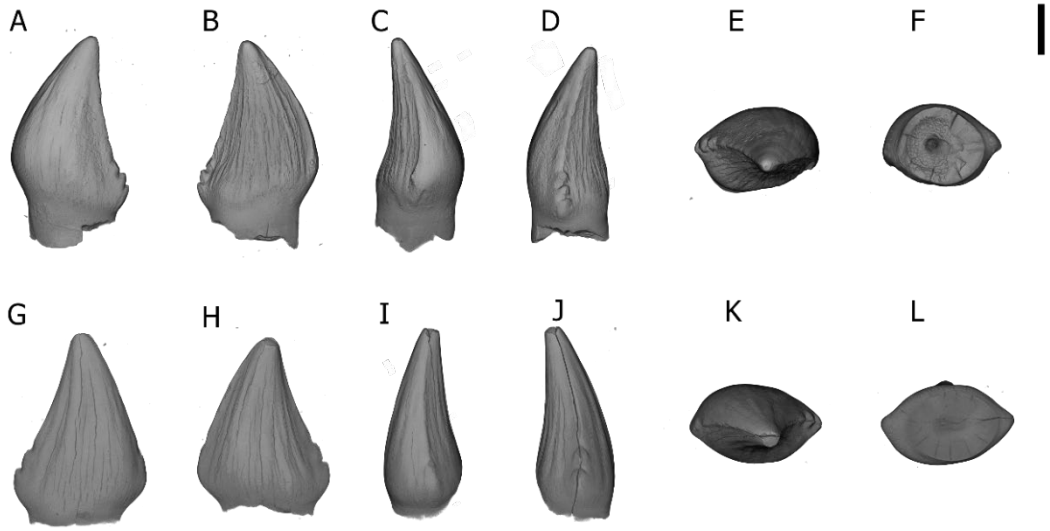


Fig. 6.9. Teeth of indeterminate thyreophorans (Morphotype B) from Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002). NHMUK PV R JR351 (A to F), NHMUK PV R JR352 (G to L), in labial (A, G), lingual (B, H), mesial (C, I), distal (D, J), apical (E, K) and basal (F, L) views. Scale bar 1 mm.

Description: This morphotype is represented by two isolated teeth with crown heights ranging from 2.52–3.12 mm and with maximum mesiodistal widths from 1.96–1.98 mm (Fig. 6.9). The crowns are isosceles in outline with a centrally located apex, apicobasally elongate, mesiodistally asymmetrical, often slightly recurved distally, and inclined lingually. The mesial margin is slightly convex whereas the distal margin is weakly concave. The teeth are widest mesiodistally immediately apical to the crown base forming a labiolingual swelling, or cingulum. In basal view, the roots are sub-circular to elliptical in outline. Coarse denticles, 2–4 in number, are present on the distal margin of the teeth starting immediately above the cingulum and extending approximately two-thirds of the way along the distal crown margin. The mesial margin is more variable with denticles either absent, a very poorly developed single denticle present immediately apically to the cingulum or two coarse denticles present immediately apically to the cingulum. Where mesial denticles are present they are smaller, both mesiodistally and

apicobasally, than the distal denticles. The denticles are not split into sub-denticles with the apical region of the denticles forming a single surface. The crowns possess apicobasally extending striations on both lingual and labial surfaces. The striations on the lingual surface form prominent fine ridges and extend across the cingulum whereas those on the labial surface are less pronounced and do not reach the cingulum.

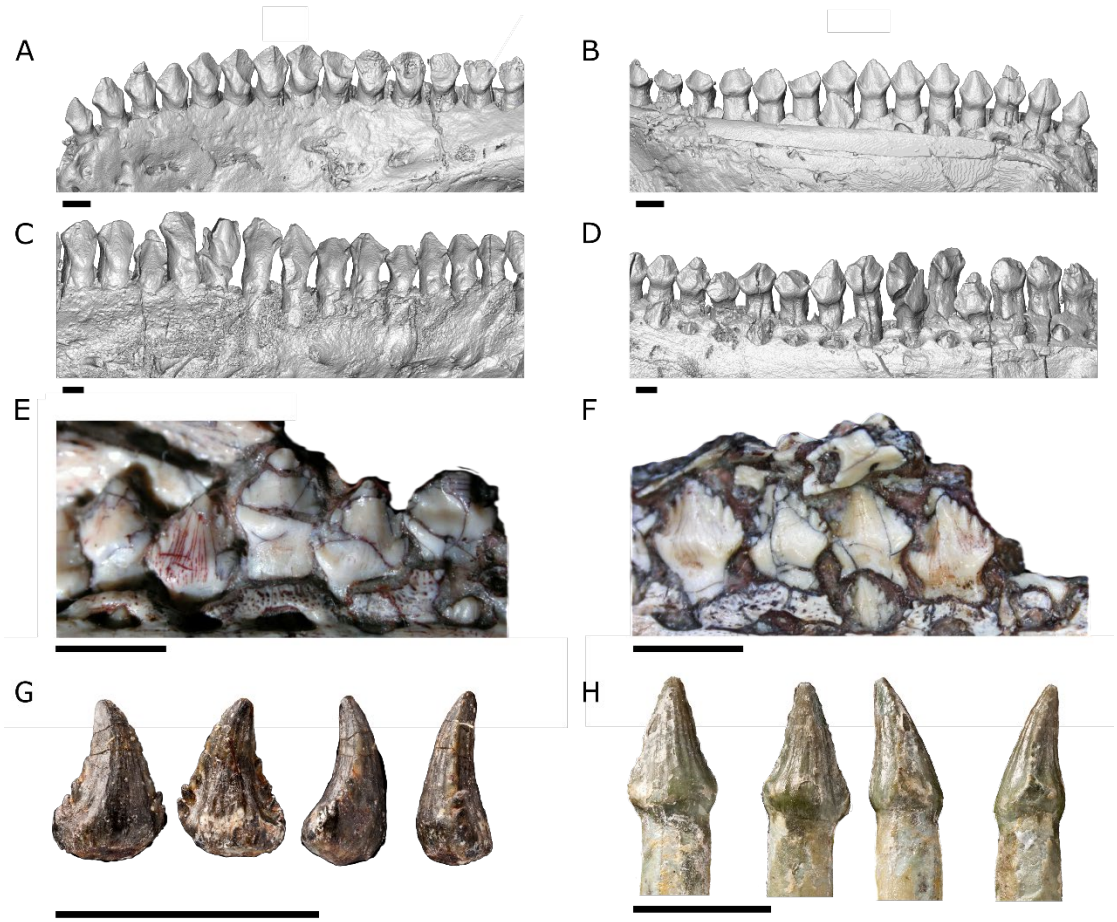


Fig. 6.10. Teeth of basal thyreophorans. *Scelidosaurus* (NHMUK PV R1111; A–D), Charmouth Mudstone Formation, Sinemurian, Early Jurassic, Dorset, UK; *Scutellosaurus* (MNA.V.175; E, F), Kayenta Formation, Sinemurian/Pliensbachian, Early Jurassic, Arizona, USA (Breeden et al., 2021); and the putative thyreophoran *Laquintasaura venezuelae* (MBLUZ P5063 and MBLUZ P5010; G and H, respectively), La Quinta Formation, Hettangian, Early Jurassic, Rio La Grita, Venezuela (Barrett et al., 2014). *Scelidosaurus* left dentary teeth in labial (A) and lingual (B) views, left maxillary teeth (inverted) in labial (C) and lingual views (D). *Scutellosaurus* maxillary (E) teeth (inverted) and dentary teeth (F) in lingual view. *Laquintasaura venezuelae* cheek teeth in lingual / labial and mesial / distal views. Scale bars 5mm.

Morphological comparisons:

The small sub-triangular crowns of morphotype A (Fig. 6.8) closely resemble the cheek teeth of basal thyreophorans such as *Scelidosaurus* and *Scutellosaurus* (Fig. 6.10) and basal ornithischians such as *Lesothosaurus* (Fig. 6.7) with unornamented labial and lingual surfaces, apart from a broad central eminence, and unequal mesial and distal margin lengths (Barrett, 2001; Breeden et al., 2021; Colbert, 1981; Sereno, 1991). Unlike more derived thyreophorans, and the basal ornithischian *Lesothosaurus*, the basal swelling has not developed into a distinct cingulum as is seen in ankylosaurs and stegosaurs (Coombs, 1990) but is similar to that present in *Scelidosaurus*, *Scutellosaurus* and *Emausaurus* (Barrett, 2001; Colbert, 1981; Haubold, 1990). In contrast to the symmetrical crowns of *Scutellosaurus* and *Lesothosaurus*, the crowns tips are offset distally, as seen in *Scelidosaurus* and *Emausaurus* but, as with *Scutellosaurus*, the crowns have equal numbers of denticles on each margin. Maxillary teeth in the early stegosaur *Isaberrysaura* have a similar overall morphology (Fig. 6.16) with distally offset lanceolate crowns bearing prominent denticles on both the mesial and distal margins and a smooth enamel surface (Salgado et al., 2017), but the broad central eminence as seen in Morphotype A is poorly developed. Neornithischians, such as *Agilisaurus*, *Hexinlusaurus, Yandusaurus, Hypsilophodon, Thescelosaurus, Jeholosaurus, Orodromeus,* and *Nanosaurus*, and pachycephalosaurs such as *Stegoceras* (Fig. 6.11) have a similar overall morphology to Morphotype A with triangular crowns bearing marginal denticles (Carpenter & Galton, 2018; Hudgins et al., 2022). Most have a distinct constriction at the base of the crown, with the exceptions of *Hypsilophodon* and *Jeholosaurus,* and a well-developed cingulum developed apically to this, for example in *Thescelosaurus* (Barrett et al., 2005; Boyd, 2014; Carpenter & Galton, 2018; Galton, 1974; Hudgins et al., 2022; Sullivan, 2006). Neornithischians, such as *Nanosaurus*, *Thescelosaurus*, *Hypsilophodon* and *Jeholosaurus*, also have an ornamented crown surface with a series of ridges extending apicobasally (Fig. 6.11), often supporting marginal denticles, which can be present on both lingual and labial surfaces or on some, but not all, teeth as in *Agilisaurus* (Barrett et al., 2005). This contrasts with the crowns of Morphotype A where a basal swelling rather than a distinct cingulum is present above the crown base constriction and, apart from the central eminence, the crown surface is devoid of ornamentation.

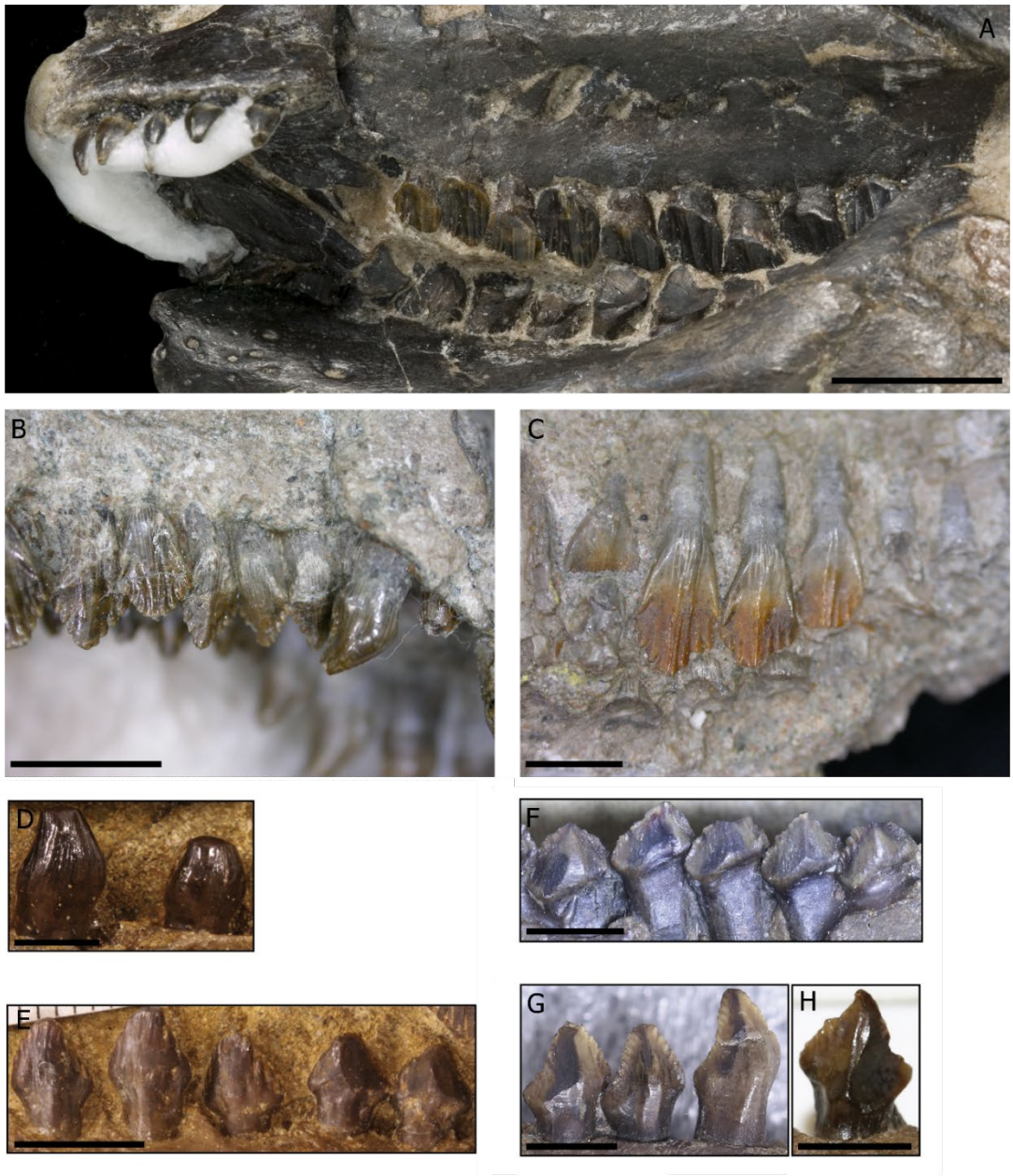


Fig. 6.11. Teeth of neornithischians (A to E) and pachycephalosaurids (G to H). A, left maxillary and dentary teeth (labial view) of *Hypsilophodon foxii* (NHMUK PV R197), Vectis Formation, Wealden Group, Barremian, Early Cretaceous, Isle of Wight, UK. B, right maxillary teeth (labial view) of *Agilisaurus louderbacki* (ZDM T6011), lower Shaximiao Formation, Bajocian, Middle Jurassic, Sichuan Province, China (Barrett et al., 2005). C, right maxillary teeth (labial view) of *Hexinlusaurus multidens* (ZDM T6001), lower Shaximiao Formation, Bajocian, Middle Jurassic, Sichuan Province, China (Barrett et al., 2005) D and E, premaxillary (labial view) and dentary (labial view) teeth of *Thescelosaurus neglectus* (NCSM 15728), Hell Creek Formation, Maastrichtian, Upper Cretaceous, Montana, USA (Hudgins et al., 2022). F and G, premaxillary (lingual view), maxillary (lingual view), and dentary (labial view) teeth of *Stegoceras validum* (UALVP 2), Dinosaur Park Formation, Campanian, Upper Cretaceous, Alberta, Canada (Hudgins et al., 2022). Scale bars: A = 1cm, B to H = 5mm. Images of *Agilisaurus* and *Hexinlusaurus* courtesy of Richard Butler.

The isosceles-shaped, apicobasally expanded, teeth with prominent striations on the lingual and labial surfaces of Morphotype B (Fig. 6.9) resemble teeth referred to the basal thyreophoran *Laquintasaura* (Fig. 6.10) (Baron et al., 2017b; Barrett et al., 2014). Basal ornithischians such as *Lesothosaurus* and other basal thyreophorans such as *Scelidosaurus*, *Scutellosaurus* and *Emausaurus* have low sub-triangular crowns that are symmetrical to slightly asymmetrical in lateral view with unornamented labial and lingual surfaces (Barrett, 2001; Colbert, 1981; Haubold, 1990). Heterodontosaurids have either low sub-equilateral triangular crowns such as with *Frutiadens* and *Tianyulong*, or more elongate crowns as seen in *Heterodontosaurus* and *Abrictosaurus,* and none possess the apicobasally extending striations seen here or in *Laquintasaura*, and in the latter two heterodontosaurid taxa the marginal denticles are restricted to the apical third of the crown (Barrett et al., 2014; Butler et al., 2012). The basal neornithischian *Hexinlusaurus* has high-crowned, rhomboidal teeth with apicobasally extending striations on both labial and lingual surfaces. However, unlike in both Morphotype B and *Laquintasaura,* these striations are not present in all teeth and are both less numerous and less prominent than seen here (Barrett et al., 2005). Other neornithischians, such as *Nanosaurus* have low, distally offset, sub-triangular crowns with a prominent primary ridge and secondary ridges that extend partway down the crown surface (Carpenter & Galton, 2018)**.**

Neither morphotype can be referred with any certainty to a known clade and therefore likely represent new taxa, with Morphotype A representing a basal thyreophoran similar to *Scelidosaurus*, *Scutellosaurus* and *Emausaurus,* and Morphotype B representing a *Laquintasaura*-type basal ornithischian (or thyreophoran) or possibly premaxillary teeth. I accept that in the case of Morphotype A it is not entirely possible to rule out that these teeth belong to a neornithischian or early eurypodan: however, given the morphological differences described above and the lack of Middle Jurassic basal neornithischians in the UK, I feel it is most conservative to refer them to Thyreophora at this time.

Ankylosauria

ANKYLOSAURIA Osborn, 1923

Gen. et sp. Indet.

Figure 6.12

Referred specimens: NHMUK PV R K228. NHMUK PV R WD2 and NHMUK PV R WD28

Unregistered specimens: 294

Localities and horizons: Kirtlington Quarry, 'Mammal Bed', Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002), Woodeaton Quarry, Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019), Watton Cliff, Forest Marble Formation, Middle Jurassic (Hunter & Underwood, 2009).

Description: Represented by three small, isolated teeth that, along with the unregistered material, make up the majority of the isolated teeth found at all of the localities studied (Fig. 6.12). The teeth are mesiodistally expanded and labiolingually compressed phylliform (leaf-shaped) crowns with prominent, apically inclined, denticles on both the mesial and distal margins and an apical cusp. Crown heights range from 1.26–4.08 mm with the maximum mesiodistal width, which occurs just apical to the crown-root junction, ranging from 1.55–3.96 mm. The crowns are asymmetrical in labial/lingual views with a slight distal offset to the crown apex and a pronounced cingulum that rings the tooth crown, which is slightly more prominent lingually. The cingulum is positioned more apically on the lingual surface of the crowns leading to strong asymmetry in mesial/distal views. The margin of the cingulum on the lingual surface of the crown can be denticulate, towards either the mesial or distal margin of the crown, with small, rounded, apically-pointed denticles decreasing in size towards the midline of the crown, although in some cases (NHMUK PV R K212) the cingulum lacks denticles. The crown narrows basally below the cingulum towards the root which, where preserved (PV R K212), then expands slightly before also narrowing basally. The preserved root is sub-circular in cross-section. The marginal denticles on both the mesial and distal margins are widely spaced, stand clear of the crown surface and are either aligned with the long axis of the crown or at angles of <45^{o} to the long axis.

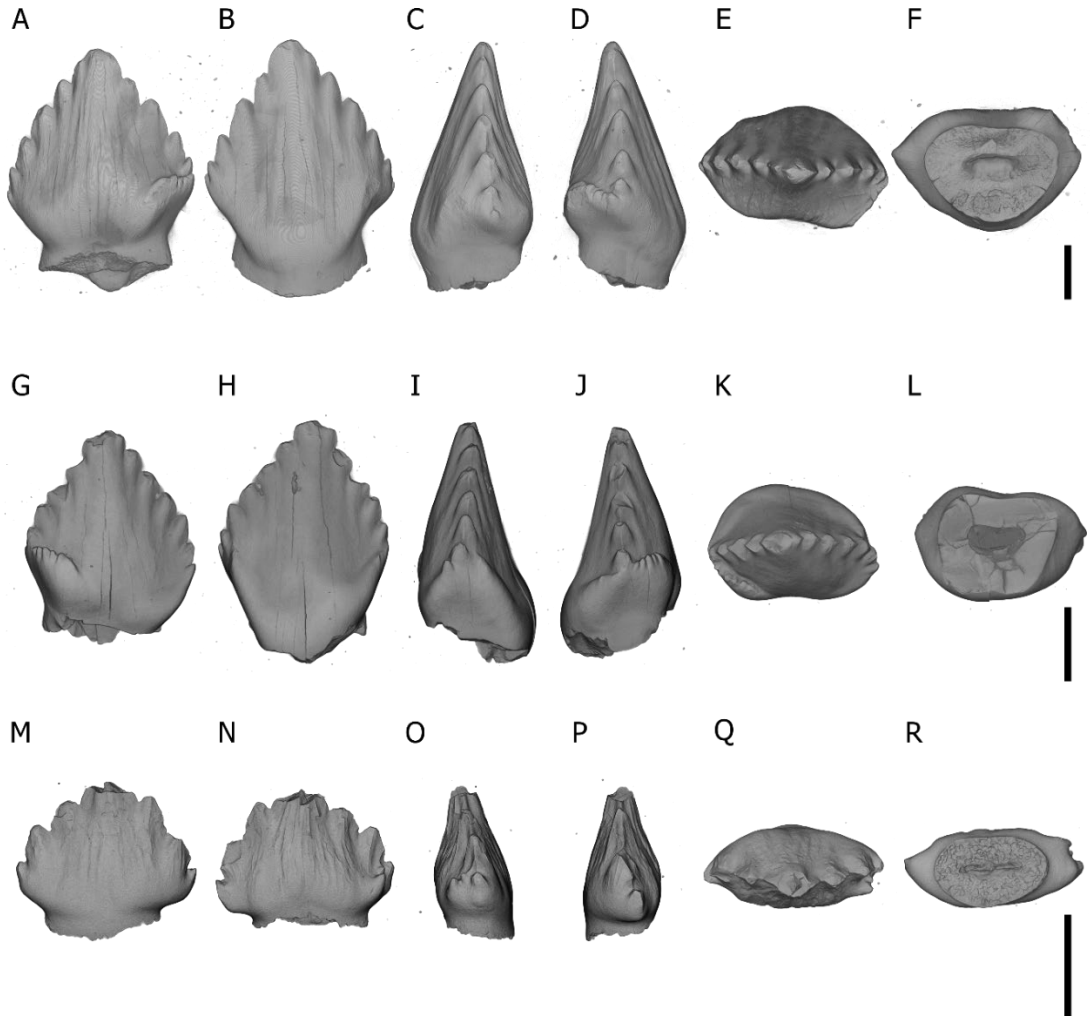


Fig. 6.12. Isolated ankylosaur teeth from Kirtlington Quarry (NHMUK PV R K228; A–F), 'Mammal Bed', Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Cox & Sumbler, 2002) and Woodeaton Quarry (NHMUK PV R WD 2 and NHMUK PV R WD28; G–L and M–R, respectively), Bed 23, Bladon Member, White Limestone Formation, Bathonian, Middle Jurassic (Wills et al., 2019) in labial (A, G, M,), lingual (B, H, N), mesial (C, I, O), distal (D, J, P), apical (E, K, Q), and basal (F, L, R) views. Scale bars 1 mm.

A minimum of three (PV R WD28) to a maximum of five (PV R WD2) marginal denticles are present on the mesial / distal margins adjacent to the primary ridge. The denticles taper apically to a sharp point with equal numbers of denticles on both the mesial and distal margins, apart from PV R K212 where there might be more denticles on the distal margin, although damage to its crown apically means this cannot be confirmed. The denticles are supported on the crown surface by a series of ridges that extend basally and are slightly more prominent on the lingual surface. The ridges on K228, WD2, WD28 are coincident with the notches between the marginal denticles whereas those on PV R

K212 are not. The apical denticle is either centrally positioned or slightly offset distally and is supported by a primary ridge that extends basally and merges with the cingulum. The primary ridge is poorly developed on PV R K212 compared to other teeth of this morphotype. Small wear facets are present on some crowns that truncate the tips of the marginal denticles and expose symmetrical enamel distribution around the tooth crown.

Morphological comparisons: Morphologically these teeth resemble those of ankylosaurs (Fig. 6.13) in having phylliform, labiolingually compressed crowns with a prominent apical cusp and marginal denticles along the mesial and distal margins, a distal offset to the apex, symmetrical distribution of enamel around the crown, ridges on the labial and lingual crown surfaces, and a basal cingulum (Barrett, 2001; Coombs, 1978, 1990; Mallon & Anderson, 2014; Vickaryous et al., 2004). Other thyreophorans such as the basal forms *Scelidosaurus* and *Scutellosaurus* have sub-triangular crowns with unornamented labial and lingual surfaces, apart from a broad primary ridge, and unequal mesial and distal margin lengths (Barrett, 2001; Colbert, 1981). Stegosaur teeth are broadly similar in morphology: however, stegosaur marginal denticles are usually rounded in lateral view compared to the more pointed form commonly seen in ankylosaurs (Barrett, 2001; Galton & Upchurch, 2004). A denticulate cingulum (as in PV R WD2 and PV R WD28) occurs in several ankylosaur taxa (Fig. 6.13), such as *Niobrarasaurus*, *Sauropelta*, the nodosaurid *Panoplosaurus,* and *'Priodontognathus'* (Arbour & Currie, 2015; Carpenter et al., 1995; Galton, 1980c; Mallon & Anderson, 2014), the stegosaur *Isaberrysaura* (Salgado et al., 2017), the Late Jurassic neornithischian *Nanosaurus* (Carpenter & Galton, 2018), Late Cretaceous ornithischians such as *Pachycephalosaurus* (Brown & Schlaikjer, 1943), some Late Triassic archosaurs (Parker et al., 2005)**,** and isolated ornithischian teeth from Portugal and Spain (Rauhut, 2001; Rauhut, 2002; Thulborn, 1973)**.**

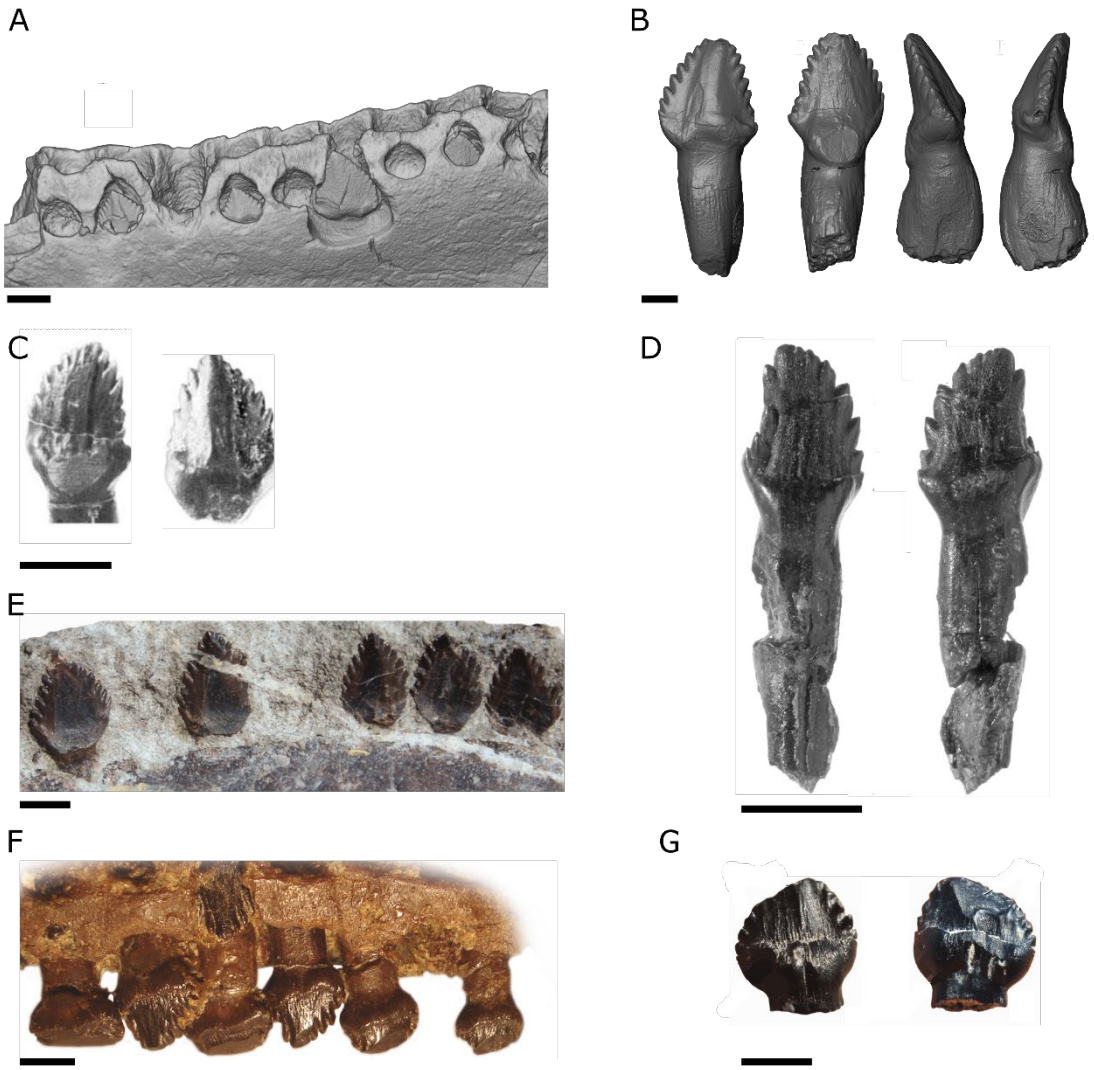


Fig. 6.13. Ankylosaurid and nodosaurid teeth. A, left mandible of *Sacrolestes leedsi* (NHMUK PV R 2682) in mediodorsal view (anterior to the left), Oxford Clay Formation, Callovian, Middle Jurassic, Peterborough, UK. B, indeterminate ankylosaur tooth (NHMUK UK PV R 2940), in lingual / labial and mesial / distal views, Purbeck Limestone Group, Berriasian, Early Cretaceous, Dorset, UK (Barrett & Maidment, 2011). C, isolated maxillary teeth of *Ankylosaurus magniventris* (AMNH 5895 and AMNH 5214), Hell Creek Formation, Maastrichtian, Upper Cretaceous, Montana, USA (Brown, 1908; Carpenter, 2004). D, isolated ankylosaurian teeth (NMV P186435) in lingual and labial views, Strzelecki Group, Albian, Early Cretaceous, Victoria, Australia (Barrett et al., 2010). E and G, right dentary in lateral view and isolated maxillary tooth (MTM 2007.25.2) in labial and lingual views of the nodosaurid ankylosaur, *Hungarosaurus tormai*, Csehbánya Formation, Santonian, Upper Cretaceous, Iharkút, Hungary (Ősi et al., 2014). F, partial left maxillary tooth row of the nodosaurid *Panoplosaurus mirus* (ROM 1215) in lingual view Dinosaur Park Formation, Campanian, Upper Cretaceous, Alberta, Canada (Mallon & Anderson, 2014). Scale bars 5 mm.

Stegosauria

STEGOSAURIA Marsh, 1877

Gen. et sp. Indet.

Figure 6.14

Referred specimens: GLRCM G1017, G1021, G75071, G75708, NHMUK PV R H1.

Unregistered specimens: ten

Localities and horizons: Hornsleasow Quarry, Chipping Norton Limestone Formation, Bathonian, Middle Jurassic (Metcalf et al., 1992).

Description: Five isolated stegosaurian teeth from Hornsleasow Quarry (Fig. 6.14) are characterised by labiolingually compressed spatulate crowns, which are symmetrical in lateral view, with a very pronounced cingulum, a centrally located apex and mesial and distal denticles that possess bluntly rounded tips (where preserved). Three of the teeth (GLRCM G1021, GLRCM G75071 and GLRCM G75708) are very worn with little surface ornamentation or traces of marginal denticles remaining. The other two teeth are represented by a complete crown missing the root (NHMUK PV R H1) and a crown (GLRCM G1017) with a partially preserved root. Wear facets are present apically on GLRCM G1017 and GLRCM G75708 and often truncate the individual denticles. The crowns range in height from 2.8–5.9 mm with maximum mesiodistal widths from 2.7–6.3 mm. The crowns are labiolingually compressed with the labiolingual width ranging from 1.8–3.4 mm. The crowns have pronounced marginal denticles on both the mesial and distal margins that are confluent with clearly defined ridges on both the lingual and labial crown surfaces. The unworn crown (NHMUK PV R H1) has seven marginal denticles on both its mesial and distal margins. The ridges terminate basally merging into a strongly developed cingulum that rings the crown and is slightly more pronounced on the labial side. The cingulum margin is devoid of denticles. The cingulum on GLRCM G1017 is strongly apically inclined towards both the mesial and distal margins whereas the cingulum on NHMUK PV R H1 is only weakly inclined apically at the mesial and distal margins. Both the cingulum and central primary ridge are more prominent on GLRCM G1017 than any of the other teeth. The denticles on the mesial margin of NHMUK PV R H1 are larger and fewer than those present on the distal margin. The second relatively

unworn tooth (GLRCM G1017) is damaged along its mesial and distal margins and the denticles cannot be observed clearly. The apical denticle on NHMUK PV R H1 and GLRCM G1017 is missing but was clearly present as it would have been confluent with an apicobasally extending primary ridge, which is more prominent on the labial surface of the crown, that merges with the cingulum basally. In (?)mesial view, the crown of NHMUK PV R H1 is inclined labially whereas GLRCM G1017 appears to have a slight lingual inclination. Where the root is preserved (GLRCM G1017) it occupies around 60% of the maximum crown mesiodistal width, is sub-oval to rounded in cross-section, and constricted at the crown root junction.

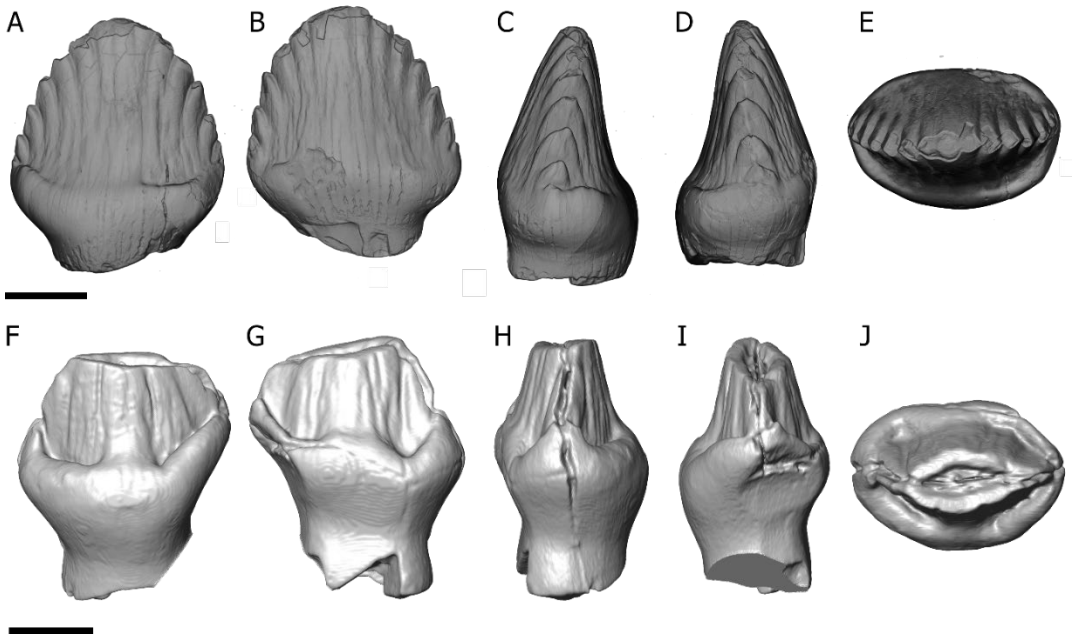


Fig. 6.14. Unworn teeth of indeterminate stegosaurs from Hornsleasow Quarry, Chipping Norton Limestone Formation, Bathonian, Middle Jurassic (Metcalf et al., 1992). NHMUK PV R H1 (A–E) and GLRCM G1017 (F–J) in labial (A, F), lingual (B, G), mesial (C, H), distal (D, I) and occlusal (E, J) views. Scale bar 2 mm.

Morphological comparisons: These isolated teeth, only found at Hornsleasow Quarry, resemble those of stegosaurs (Fig. 6.15) in having phylliform, almost symmetrical (in lingual / labial views), labiolingually compressed crowns with a prominent apical cusp and marginal denticles along the mesial and distal margins, a symmetrical distribution of enamel around the crown, ridges on the labial and lingual crown surfaces, and a strongly developed, ring-like basal cingulum that encircles the crown. Although this

morphology is somewhat similar to that seen in some ankylosaurs, there are several differences that support the referral of these teeth to Stegosauria. The well-developed cingulum is devoid of denticles, whereas in ankylosaurs the apical cingulum margin is often denticulate, as in *Niobrarasaurus* (Arbour & Currie, 2015), *Sauropelta* (Carpenter et al., 1995) and indeterminate teeth from Australia (Barrett et al., 2010)**.** The denticles are rounded at their tips (Barrett, 2001; Billon-Bruyat et al., 2010; Galton & Upchurch, 2004; Maidment et al., 2008; Skutschas et al., 2021), contrasting with the pointed morphology seen in many ankylosaurs (Barrett, 2001; Coombs, 1990; Vickaryous et al., 2004). The apical denticle (where preserved) is on the median vertical line of the crown whereas in ankylosaurs it is inclined posteriorly (Coombs, 1990). The primary ridge, which extends basally from the crown apex, is very strongly developed and divides the denticulate part of the crown into two symmetrical portions, which is more consistent with stegosaurs such as *Paranthodon*, *Huayangosaurus*, and *Stegosaurus* rather than the more muted expression of the ridge usually present in ankylosaurs (Coombs, 1990)**.** An exception to this generalised stegosaurian tooth morphology is found in *Isaberrysaura mollensis* (Toarcian– Bajocian, Middle Jurassic) where the maxillary teeth are more lanceolate in outline with a slight distal curvature giving an asymmetric appearance in lingual / labial views (Fig. 6.15). Unlike other stegosaurs the enamel surface of *Isaberrysaura* is smooth with no vertical ridges extending from the marginal denticles and the base of the crown lacks a well-developed cingulum with only a slight eminence present, being more reminiscent of basal thyreophorans such as *Scutellosaurus* and *Scelidosaurus* (Salgado et al., 2017).

In addition to the specimens above there are a number of isolated teeth that, due to wear and damage, could not be attributed with any certainty to the morphotypes described: 18 indeterminate ornithischians and 20 indeterminate thyreophorans.

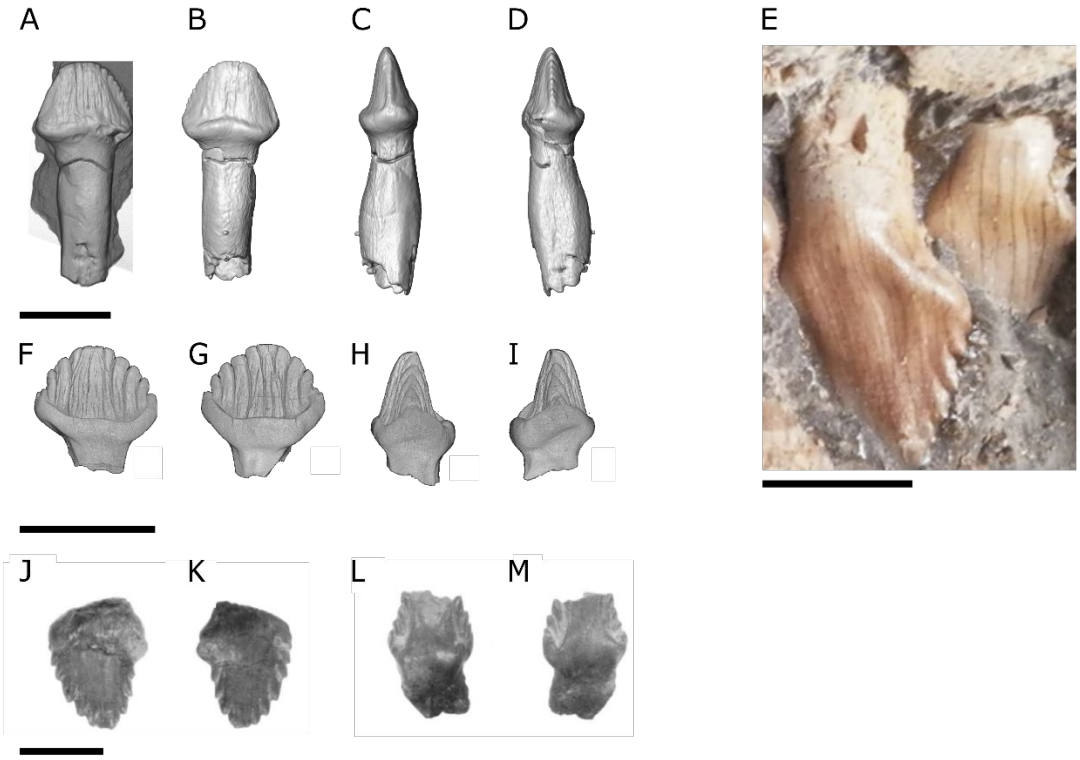


Fig. 6.15. Comparative stegosaur teeth. A–D, teeth of *Stegosaurus stenops* (NHMUK PV R 36730) in lingual / labial (A and B) and mesial / distal (C and D) views, Morrison Formation, Tithonian, Upper Jurassic, Wyoming, USA. E, maxillary teeth of the early stegosaur *Isaberrysaura mollensis* (MOZ-Pv 6459) in lateral view, Los Molles Formation, Toarcian–Bajocian, Middle Jurassic, Patagonia, Argentina (Salgado et al., 2017). F–I, indeterminate isolated stegosaur teeth (ZIN PH 65/246) in lingual / labial (F and G) and mesial / distal (H and I) views, Batylykh Formation, Early Cretaceous, Teete, Yakutia, Russia (Skutschas et al., 2021). J–M, maxillary (J and K) and dentary (L and M) teeth of *Huayangosaurus taibaii* (IVPP V6728) in labial (J and L) and lingual (K and M) views, lower Shaximiao Formation, Oxfordian, Upper Jurassic, Sichuan, China (Sereno & Dong, 1992). Scale bars 5 mm.

Discussion

There are relatively few named Middle Jurassic ornithischian taxa (Barrett et al., 2005; Galton, 1980b, 1983b; Hui et al., 2022; Maidment et al., 2020; Maidment et al., 2021; Nopcsa, 1911) and much of the record is made up of indeterminate material (e.g. Evans & Milner, 1994; Haddoumi et al., 2016; Metcalf & Walker, 1994; Prasad & Parmar, 2020; Wills et al., 2019). This morphological study of isolated ornithischian teeth confirms the presence of at least six distinct morphotypes in these Middle Jurassic microvertebrate assemblages, some of which probably represent new taxa. These include: a distinctive, rare, highly-ridged morphotype that, although sharing morphological similarities with the premaxillary teeth of the basal ornithischian / thyreophoran *Laquintasaura*, cannot

be referred with certainty to any known ornithischian taxon; small teeth with denticles restricted to the upper third of the crown that represent the hitherto unknown occurrence of heterodontosaurids in the Middle Jurassic of the UK; at least one morphotype of an early-diverging thyreophoran with similarities to the cheek teeth of *Scutellosaurus* and *Scelidosaurus*; an indeterminate thyreophoran from Kirtlington Quarry with morphological similarities to *Laquintasaura*; a stegosaur, which along with *Adratiklit boulahfa* from the Middle Jurassic of Morocco (Maidment et al., 2020) and *Loricatosaurus priscus* from the Callovian of the UK (Maidment et al., 2008; Nopcsa, 1911), represents one of the oldest stegosaurs worldwide; and an ankylosaur that appears to be the most abundant ornithischian these sites. Perhaps surprisingly, the sample includes no teeth that can clearly be assigned to Ornithopoda, a clade that might also be expected to be present given their occurrence elsewhere in Europe during the Middle Jurassic (Galton, 1980b; Ruiz-Omeñaca et al., 2007) and predictions made by ghost lineages present in ornithischian phylogenies (Boyd, 2015; Dieudonné et al., 2021; Poole, 2022).

These results (Table 6.1) increase the known diversity of Middle Jurassic ornithischians from the UK, lend support for the hypothesized timing of ornithischian diversification events, with the recovery of definitive ankylosaur and stegosaur remains, and have biogeographical distributions consistent with predictions made by phylogenetic analyses (Fig. 6.16).

When taken in conjunction with the theropod faunas from these localities described in Wills, Underwood, et al. (2023) and Chapter Five, several dromaeosaurid morphotypes, a troodontid and a therizinosauroid, these results highlight the global importance of these sites and the use of microvertebrate data in general.

I recognise that there is a degree of difficulty in accurately referring isolated ornithischian teeth to specific taxa due to the morphological similarities seen in different ornithischian clades and in non-dinosaurian Triassic archosaurs (Boyd, 2015; Irmis et al., 2007; Nesbitt et al., 2007). However, the Middle Jurassic age of these deposits renders non-dinosaurian identities unlikely. Moreover, I identify the teeth conservatively as morphotypes within a broader ornithischian taxonomic framework, in order to avoid definitive statements about absolute species-richness. For example, it is possible that

some of these morphotypes actually belonged to the same taxon, given known variation among the premaxillary / maxillary / dentary tooth rows of other ornithischians. In addition, it is possible that these teeth represent the same taxa as other isolated body fossils from the UK Middle Jurassic (Barrett & Maidment, 2011; Galton, 1980c, 1983a, 1983b, 1985; Galton & Powell, 1983; Maidment et al., 2008; Maidment, 2010; Nopcsa, 1911; Panciroli, Funston, et al., 2020) although associations between these materials are impossible in the absence of more complete specimens.

Table 6.1. New ornithischian morphotypes from the Great Oolite Group, Bathonian, Middle Jurassic of Britain.

Watton Cliff, Dorset. Forest Marble Formation.	Thyreophora indet. Morphotype B Ankylosauria indet.
Kirtlington Quarry, Oxfordshire. White Limestone Formation,	Ornithischia indet. Heterodontosauridae indet. Thyreophora indet. Morphotype A Thyreophora indet. Morphotype B Ankylosauria indet.
Woodeaton Quarry, Oxfordshire. White Limestone Formation.	Heterodontosauridae indet. Thyreophora indet. Morphotype A Thyreophora indet. Morphotype B Ankylosauria indet.
Hornsleasow Quarry, Gloucestershire. Chipping Norton Limestone Formation.	Heterodontosauridae indet. Thyreophora indet. Morphotype A Stegosauria indet.

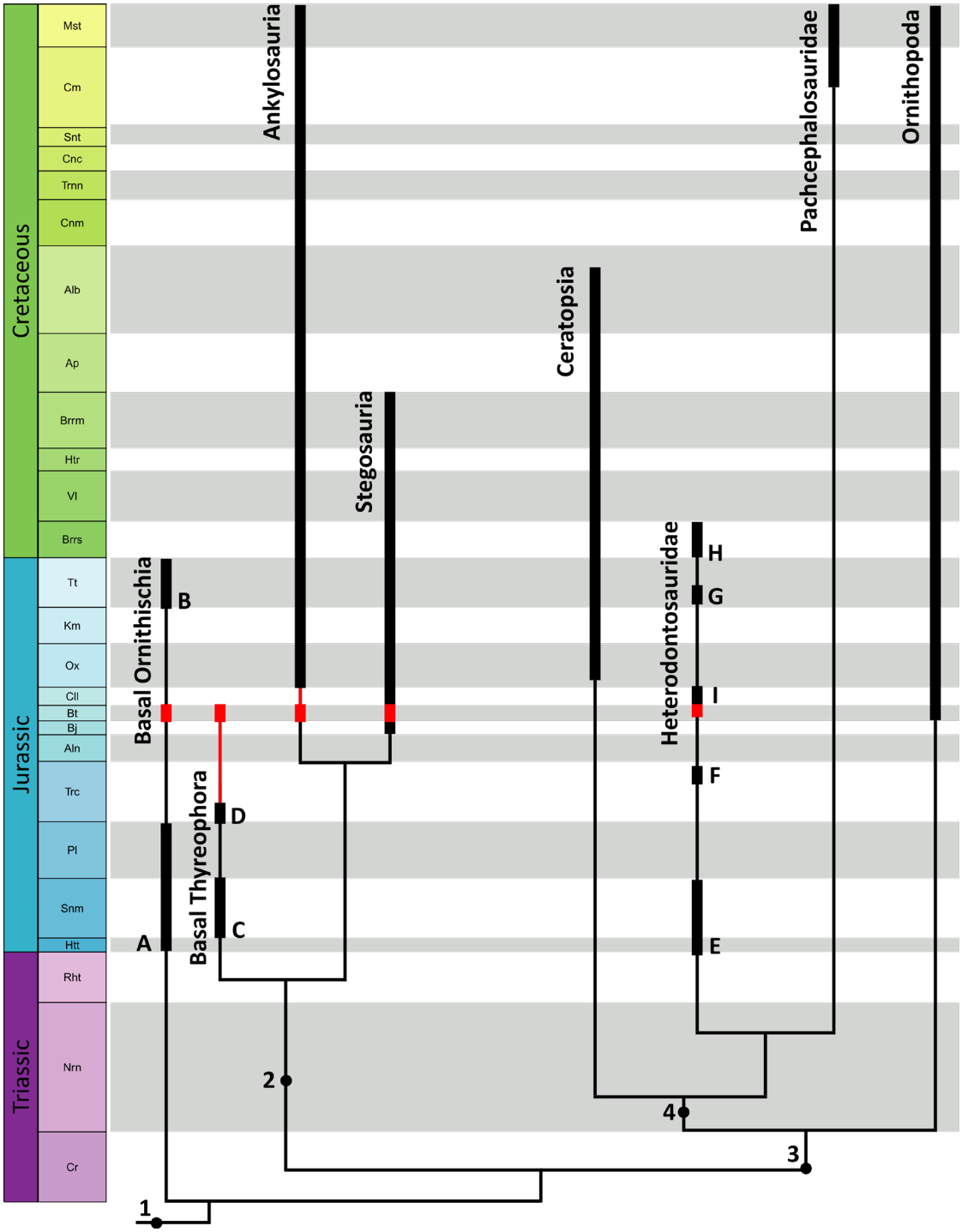


Fig. 6.16. Simplified time calibrated phylogeny of non-ornithopod Ornithischia based on Dieudonné et al. (2021). Occurrences and range extensions indicated by this study are shown in red. (1) Ornithischia, (2) Thyreophora, (3) Ceropoda, (4) Marginocephalia. (A) *Laquintasaura* and *Lesothosaurus*, (B) *Chilesaurus,* (C) *Scelidosaurus* and *Scutellosaurus*, (D) *Emausaurus*, (E) *Abrictosaurus, Lycorhinus* and *Heterodontosaurus*, (F) *Manidens*, (G) *Fruitadens*, (H) *Echinodon*, (I) *Tianyulong*. Dating of the occurrences of *Manidens* and *Tianyulong* from Becerra et al. (2023) being Toarcian and Callovian respectively.

Thyreophora

Thyreophorans have been an important part of terrestrial ecosystems from their initial appearance in the Early Jurassic, with a long ghost lineage stretching back to the Late Triassic (Dieudonné et al., 2021; Hui et al., 2022), to their eventual demise at the end of the Cretaceous (Fig. 6.16). Their early evolutionary history is poorly understood, with the Early Jurassic characterised by relatively few, non-eurypodan, members of the clade, such as *Scelidosaurus*, *Scutellosaurus* and *Yuxisaurus* (Breeden et al., 2021; Colbert, 1981; Owen, 1861, 1863; Raven et al., 2023), until the divergence and radiation of stegosaurs and ankylosaurs sometime in the Early to Middle Jurassic (Hui et al., 2022; Maidment et al., 2020; Maidment et al., 2021). The presence of other Early Jurassic forms such as *Laquintasaura* and *Lesothosaurus* is controversial as they may represent earlier-diverging ornithischians than thyreophorans (Baron et al., 2017b; Barrett et al., 2014; Dieudonné et al., 2021; Porro et al., 2015; Raven et al., 2023). Non-eurypodan thyreophorans are generally thought to have become extinct by the Middle Jurassic (Raven et al., 2023; Salgado et al., 2017); however, the mixture of non-eurypodan and eurypodan morphotypes identified here suggests that not only did non-eurypodans survive until at least the Middle Jurassic but they also co-existed in close temporal and spatial proximity with early eurypodans.

The poor record of Early to Middle Jurassic thyreophorans also opens the possibility that the teeth identified herein as basal thyreophorans might represent early eurypodans, as teeth from the early stegosaur *Isaberrysaura* have a similar morphology (Salgado et al., 2017, Fig 2), with the morphology described potentially being the plesiomorphic state for the clade. Isolated teeth from the Guimarota mine (Upper Jurassic, Kimmeridgian) of Portugal referred to '*Alocodon*' and regarded as ornithopod (Galton, 1980b; Rauhut, 2001; Thulborn, 1973) closely resemble some of the morphotypes described here. '*Alocodon*' maxillary / dentary teeth possess a denticulate cingulum, a feature found in ankylosaurs, *Isaberrysaura*, *Nanosaurus*, and some pachycephalosaurids (see above) and this, along with their overall morphology, suggests that '*Alocodon*' is likely to be an ankylosaur as previously suggested (Barrett et al. (2010). Other isolated teeth from Guimarota that Thulborn (1973) referred to as the premaxillary teeth of '*Alocodon*', with

no evidence for this apart from the association of these to the other teeth collected at the locality, bear a close resemblance to the 'Thyreophora Morphotype B' teeth described herein.

These results are congruent with recent analyses of thyreophoran biogeographic history which indicate a widespread distribution for early thyreophorans and a Laurasian distribution for Ankylosauria (Arbour & Currie, 2015; Raven, 2021; Raven et al., 2023). Under some of these analyses, Stegosauria are calculated to be ancestral to Asia and South America, which as Raven (2021) points out is likely to be a sampling effect; therefore, the addition of data points such as the morphotypes identified here, and other isolated tooth occurrences, should be included in these analyses going forward.

Heterodontosauridae

Heterodontosaurids are a rare component of Jurassic and Cretaceous terrestrial ecosystems with a contentious phylogenetic position and they have been regarded as either the earliest diverging ornithischians, basal ornithopods, early cerapodans, the sister group of Marginocephalia or members of Pachycephalosauria (Butler, 2005; Butler et al., 2007; Butler et al., 2008; Crompton & Charig, 1962; Dieudonné et al., 2021; Sereno, 1999; Xu et al., 2006). They range in age from the Early Jurassic to the Early Cretaceous and have a geographically widespread distribution (Becerra et al., 2023; Sereno, 2012); however, with the exception of the Early Jurassic deposits of southern Africa, in which their remains are abundant (Porro et al., 2011; Viglietti et al., 2020), most heterodontosaurid taxa are known from isolated occurrences (Fig. 6.17) and their temporal range is characterised by large gaps between known taxa.

The Early Jurassic (Hettangian–Sinemurian) of southern Africa yields both the earliest confirmed heterodontosaurids and the richest heterodontosaurid fauna: *Abrictosaurus*, *Heterodontosaurus*, *Lycorhinus and Pegomastax* (Crompton & Charig, 1962; Porro et al., 2011; Sereno, 2012). The next known occurrence of a heterodontosaurid is in the Toarcian, *Manidens condorensis* (Becerra et al., 2023; Pol et al., 2011), followed by the ?Middle Jurassic (Callovian) *Tianyulong confuciusi* (X.-T. Zheng et al., 2009). The teeth described herein fill in a gap in the temporal distribution of heterodontosaurids, and although the presence of a few isolated Middle Jurassic teeth does little to help untangle

the phylogenetic placement of the clade, it does confirm that heterodontosaurids were an ever-present, if rare, component of Jurassic and Early Cretaceous ecosystems with a global distribution across both Gondwana and Laurasia.

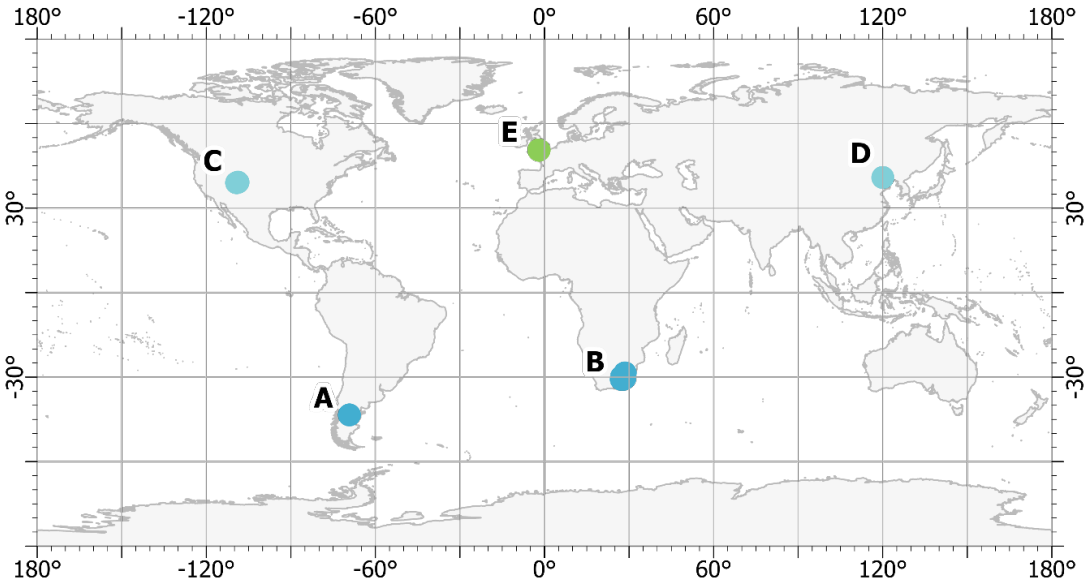


Fig. 6.17. Global distribution of named heterodontosaurid taxa. Blue circles, Jurassic. Green circles, Cretaceous. (A) *Manidens condorensis*, Cañadón Asfalto Formation, Toarcian, Lower Jurassic, Patagonia; (B) *Abrictosaurus consors*, *Heterodontosaurus tucki*, *Lycorhinus angustidens* and *Pegomastax africanus*, upper Elliot Formation, Hettangian – Sinemurian, Lower Jurassic, Lesotho and South Africa; (C) *Fruitadens haagarorum,* Morrison Formation, Tithonian, Upper Jurassic, USA; (D) *Tianyulong confuciusi* Tiaojishan Formation, Callovian, Middle Jurassic, China; (E) *Echinodon becklesii*, Purbeck Limestone Group, Berriasian, Lower Cretaceous, Swanage, Dorset, UK.

Conclusions

The morphotypes identified herein confirm the presence of a diverse ornithischian assemblage in the Middle Jurassic of the UK, increasing the known diversity from this important time period, with up to six distinctive morphotypes, some of which likely represent new taxa. Unfortunately, because of the morphological similarities seen in different ornithischian clades I can only place these teeth in a broad taxonomic framework and I refrain from naming any of these potential taxa. Nevertheless, this material is significant as it extends the geographical and temporal range of several ornithischian clades, fills temporal gaps, and lends additional support to the postulated

timings and origins of ornithischian diversification and dispersal events (Porro et al., 2011; Raven et al., 2023; Salgado et al., 2017). This material, when taken in conjunction with the recently described theropod taxa from the same sites (Wills, Underwood, et al., 2023), highlights the importance of the detailed description and identification of microvertebrate assemblages and the role these should play in our understanding of macroevolutionary and macroecological events.

Chapter Seven: Conclusions and future directions

Conclusions

The aim of this thesis has been to investigate the taxonomic affinities of isolated theropod and ornithischian teeth from four Bathonian (Middle Jurassic) microvertebrate sites. During the course of this work, I have been able to confirm the presence of several key dinosaur clades in these sites, confirming historic referrals, providing new, previously unknown records and confirming the predictions of phylogenetic analyses. I have also demonstrated that the application of new machine learning methodologies can greatly refine the morphological resolution possible from analysing microvertebrate samples. However, as noted in Chapters Three and Five, the currently available machine learning training data precludes making broader statements on actual diversity. The models at this point will not be able to distinguish between true diversity and morphological disparity as a result of ontogenetic, positional or taphonomic variation within the data.

The CT-scanning and machine learning methods outlined in Chapters Two and Three are the first quantitative assessment of the use of machine learning in solving a taxonomic problem, the classification of isolated dinosaur teeth. This is an entirely new approach to the problem, in which I used multiple machine learning algorithms to build and combine classification models in order to generate a robust and repeatable framework. Comparative assessment of the different machine learning techniques demonstrated that such models routinely outperform 'traditional' methods of classification, with decision tree-based approaches being the preferred approach. This assessment also highlights the negative impact that missing data, or the addition of imputed data, has on any classifier. The use of machine learning on isolated theropod teeth demonstrates that high levels of predictive accuracy are possible from simple morphometric data, as long as care is taken to understand the structure of the data in question and the assumptions that the various methods require. The models developed in this Chapter Three allowed their application to the isolated theropod teeth from the four Bathonian sites studied herein. As stated above, caution is still warranted in the use of machine learning at this point in time. Machine learning models are reliant on the training data used to build a model which is then itself used to make a prediction. In palaeontology

we have an incomplete fossil record, we have fragmentary specimens, collection bias, within taxon variation, and in many cases a relatively small number of specimens on which to base a model. Additionally, the currently available theropod training data is biased towards North American Cretaceous clades and reliant on previous identifications of isolated teeth. Therefore, we need to be cautious when applying this technique and validate any results against morphological comparisons until such time as we have a more robust training set.

Chapter Four provided new data on the geological setting of the four sites: Hornsleasow Quarry, Gloucestershire; Woodeaton and Kirtlington Quarries, Oxfordshire; and Watton Cliff, Dorset. These included a redescription of the Forest Marble Formation (Great Oolite Group, Bathonian, Middle Jurassic) section at Watton Cliff and a complete site description of the White Limestone Formation (Great Oolite Group, Bathonian, Middle Jurassic) at Woodeaton Quarry, including the new microvertebrate horizon discovered at the latter. This work has clarified the position of the White Limestone Formation / Forest Marble Formation boundary. It also allowed me to place the Kirtlington 'Mammal Bed' in the White Limestone Formation, rather than the Forest Marble Formation as previously suggested, at a slightly higher stratigraphic position than the approximately coeval microvertebrate horizon (Bed 23, Bladon Member, White Limestone Formation).

In Chapter Five I applied the machine learning methodologies developed in Chapter Three to a suite of isolated theropod teeth from the Bathonian microvertebrate sites described in Chapter Four. Additionally, I reassessed published tooth morphometric data for theropod specimens from Middle Jurassic sites in India, Kyrgyzstan, and Madagascar. This allowed me to confirm, in a quantifiable framework, the presence of a diverse maniraptoran theropod fauna in the Middle Jurassic (Bathonian) of the UK, including several dromaeosaur morphotypes, a troodontid and a therizinosaur. This expands the morphological disparity of theropods known from this time period in Europe and significantly extends the global ranges of Therizinosauroidea and Troodontidae by at least some 27 million years. The analysis of data from India, Kyrgyzstan and Madagascar also confirmed the presence of Middle Jurassic maniraptorans at these sites. My results indicate that not only were maniraptorans

present in the Middle Jurassic, as predicted by previous phylogenetic analyses, but had already radiated into a diverse global fauna that pre-dated the break-up of Pangaea.

Chapter Six described the ornithischian taxa present at the sites described in Chapter Four. Because of the lack of data to develop a suitable training set for machine learning analysis (see Future Directions, below), these descriptions were undertaken in conjunction with detailed morphological comparisons to known ornithischian taxa. These results revealed a hitherto unknown diversity of ornithischians from this time period in the UK. Six distinct ornithischian morphotypes are present: an indeterminate ornithischian, a heterodontosaurid, two indeterminate thyreophorans, a stegosaur and an ankylosaur. The stegosaur represents one of oldest global occurrences of the clade and the heterodontosaurid occurrence fills a major temporal gap in the lineage.

Future directions

Machine Learning

The methodology described in Chapter Three and applied in Chapter Five demonstrates the power of machine learning over other quantitative classification techniques. However, the fact that I could not apply this to ornithischian teeth, and the limitations outlined above, highlights a current disadvantage in applying machine learning more broadly to problems in vertebrate palaeontology, which is the lack of robust and independent training data on which to develop the models. I highlighted in Chapters Three and Five the restricted number of variables that these models were developed on due to the availability of, and measurements in, published datasets. I also discussed that using hand-measured morphometric data was sub-optimal compared to other data sources such as image data. The development of robust training datasets is crucial if machine learning is seen to be a viable option in vertebrate palaeontology. These training sets need to include teeth identified to clades based on features other than simply tooth morphology, such as those isolated from entire skulls or jaws, especially where those have complete (or nearly complete) tooth rows. This would allow the models to be used to analyse positional and ontogenetic variation and would begin to remove the reliance on class labels based themselves on previous identifications of isolated teeth. The addition of other clades, such as crocodile teeth, would potentially

allow a study of the degree in which root resorption impacts shape change and therefore the classification results. These techniques however are powerful, and when combined with advances in computer vision or established techniques such as geometric morphometrics and 3D landmarking, have the potential to revolutionise parts of our field.

I am currently working on several projects, in collaboration with the AI team at the Natural History Museum, which hope to address some of the issues around the lack of training data. This is by using either image data, 3D models or CT volumes to create multiple representations of one object, for example a tooth, with each representation essentially becoming one new training data case. We are exploring both a 2.5D method, where a virtual camera takes images around a 3D object at set rotations with each image then being fed into a classifier (Fig. 7.1), and a full 3D method where we extract image slices from a CT volume to build up the classifier. The CT slice method has previously been shown to be successful at detecting adrenal lesions from medical scans of patients (Sanson et al., 2023).

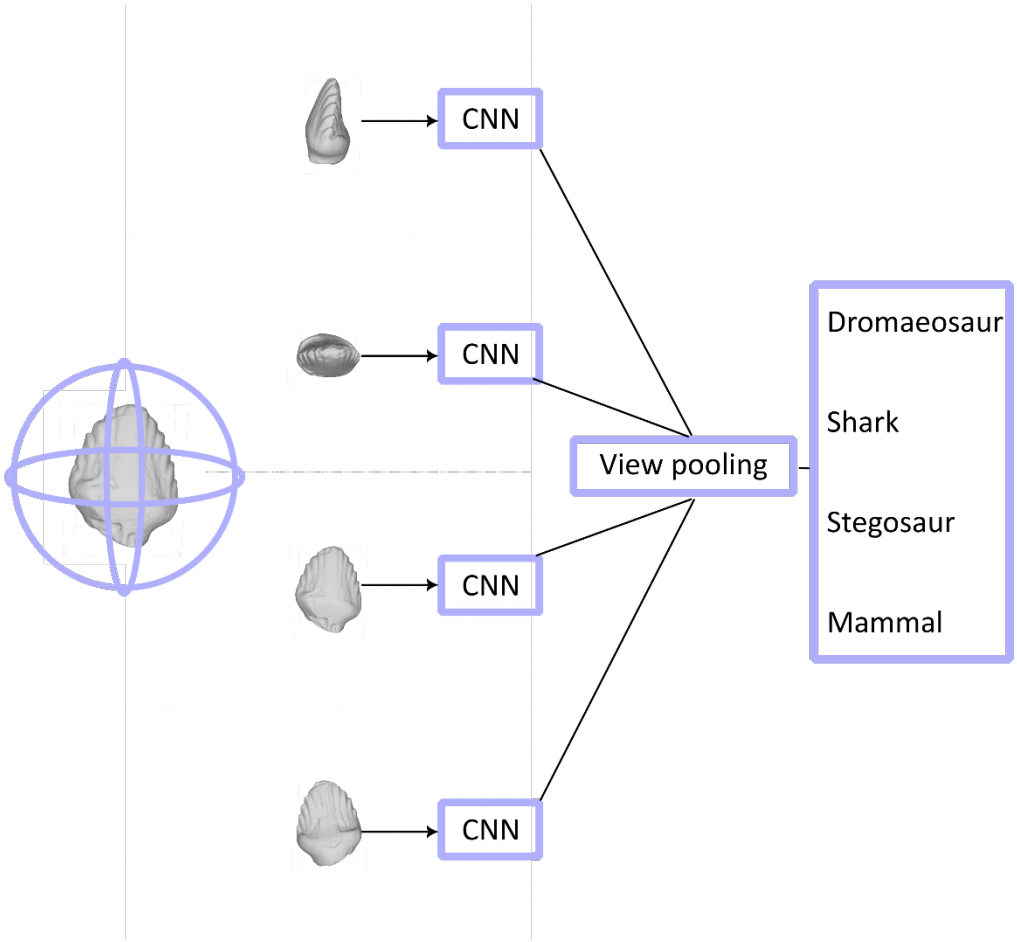


Fig. 7.1. Convolutional neural network (CNN) classifier generated from multiple views of the same object.

I also hope to expand the range of taxa, both dinosaur and non-dinosaur, that are included in the training set. I am working with colleagues at the Natural History Museum to use machine learning to classify images of shed shark teeth and with various colleagues in both the UK and USA to collect data from a Late Triassic site in New Mexico which will allow the inclusion of non-dinosaurian archosaurs in the data.

I am also involved with colleagues from the University of Oslo where we are hoping to apply machine learning to analyse images of fossils and identify evolutionary trends for tracking adaptation, speciation, and diversity events in the fossil record by incorporating machine learning classifications into phylogenetic analyses.

Woodeaton Quarry

During the picking of microvertebrate specimens from the sediment concentrate of Bed 23 at Woodeaton Quarry a substantial amount of fragmentary eggshell material was

discovered. This amounts to over 600 small (< 7 mm maximum dimension and < 0.5 mm in thickness) fragments of various taxa. The eggshell fragments (Fig. 7.2) have been imaged using a variety of methods including CT scanning and Electron backscatter diffraction (EBSD) and await description (Wills, Cavosie *et al*., In Prep). Although British Jurassic localities have yielded dinosaur and other reptilian remains for many years there is an absence of reptilian eggshell from this period. The paucity of eggshell from the British Jurassic is in common with the rest of the world where eggs and eggshell fragments in deposits older than the Cretaceous are relatively rare. Woodeaton has preserved a multi-taxic eggshell assemblage, probably the oldest multi-taxic assemblage currently known from the fossil record, where we can relate back to the probable egg-layers based on other lines of evidence.

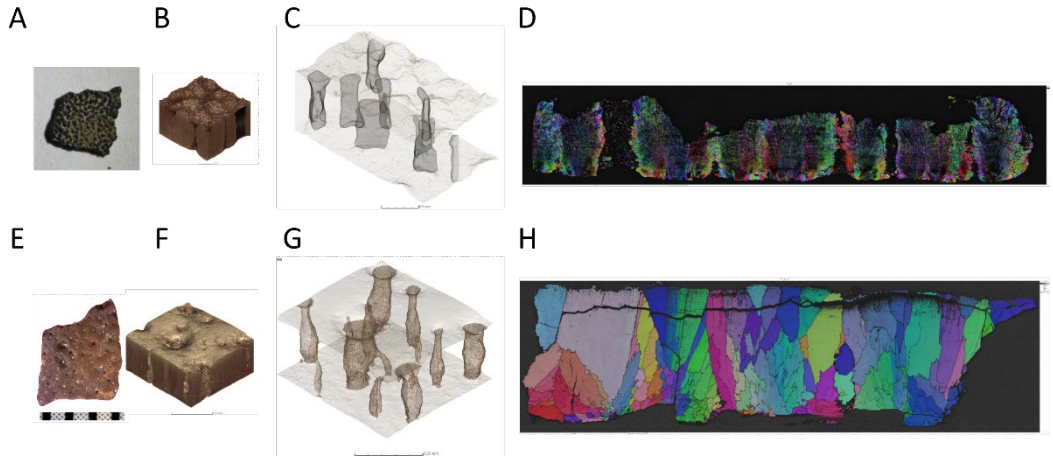


Fig. 7.2 Eggshell fragments recovered from Woodeaton Quarry, Oxfordshire. Surface images (A,E), 3D CT model (B,F), 3D internal pore model (C,G), EBSD scan (D,H). CT scans undertaken at the Natural History Museum, London and pore models developed courtesy of Vincent Fernandez (European Synchrotron Radiation Facility). EBSD scans courtesy of Aaron Cavosie (School of Earth and Planetary Science, Curtin University, Perth, WA).

Other sites

The discovery of maniraptoran theropods and a diverse ornithischian fauna from the four UK sites plus the confirmation of maniraptorans from India, Kyrgyzstan and Madagascar highlights the importance of reassessing isolated teeth from different sites and collections. The reassessment of published (and unpublished), but taxonomically disputed or tentative, data would be an interesting exercise into dinosaur diversity. The

Natural History Museum currently has access to most of the specimens from the Berriasian Sunnydown Farm site in Dorset. Much of this material is still undescribed and would lend itself to machine learning analysis. Another important site from the UK, where specimens remain undecided, would be the Tithonian microvertebrate deposit from Chicksgrove Quarry in Wiltshire although access to this data is problematic. In addition, I am working with colleagues from the Natural History Museum and University of Birmingham on a number of Middle Jurassic sites in Morocco which have yielded important stegosaur and ankylosaur material (Maidment et al., 2020; Maidment et al., 2021). We hope to apply machine learning analysis to the microfauna and isolated teeth recovered from these sites.

Overall summary

The traditional view of the dinosaur record from the British Middle Jurassic, sparse though it is, is the domination by larger-bodied taxa such as the tetanuran theropods *Magnosaurus nethercombensis* (Benson, 2010b), *Duriavenator hesperis* (Benson, 2008) and *Megalosaurus bucklandii* (Benson, 2010a; Buckland, 1824) with the early tyrannosauroid *Proceratosaurus bradleyi* (Rauhut et al., 2010) the only smaller bodied theropod taxon. A similar picture is seen within the ornithischians with large-bodied taxa such as the ankylosaur *Sarcolestes leedsi* (Galton, 1980a; R Lydekker, 1893) and the stegosaur *Loricatosaurus priscus* (Maidment et al., 2008; Nopcsa, 1911). The taxa recovered herein challenge this traditional view suggesting that Middle Jurassic ecosystems, at least in Britain, included a plethora of cryptic, small-bodied taxa. The question arises as to whether the taxa identified here are truly small-bodied taxa or juvenile examples of larger taxa. The presence of eggshell fragments at Woodeaton Quarry, from the same horizon as the isolated teeth, and tentatively assigned to maniraptoran and ornithischian ootaxa indicates that some of the smaller teeth might represent juveniles or hatchlings. The machine learning analysis is however clear in that the isolated theropod teeth are maniraptoran rather than small examples of known taxa. Other than stegosaurs, there are no other large-bodied ornithischians known in the UK Middle Jurassic. The addition of a heterodontosaurid represents a new small-bodied ornithischian taxa for this period in the UK and a Middle Jurassic ankylosaur is also new. When taken in conjunction with the other indeterminate ornithischian and

thyreophoran teeth I am confident that there are some genuinely small and otherwise unknown ornithischian taxa here.

We know that microvertebrate sites are an important source of faunal information, as they preferentially preserve small-bodied taxa that do not usually fossilise well. Consequently, information from sites like these should play a crucial role in advancing our understanding of ancient ecosystems whereas excluding them would leave our models incomplete and skewed. I hope that I have shown that by the careful application of new techniques this information can be obtained and used to help answer broader questions of dinosaur evolution. We know that the published literature only touches on the vast palaeontological holdings of museum collections and that there is a huge amount of 'dark data' that is still to be analysed (Marshall et al., 2018), one example would be the Tanzanian Tendaguru Formation holdings at the Natural History Museum. Machine learning is one of the tools that will help us unlock this 'dark data' resource. The work and methodologies presented here are simply a starting point for the application of machine learning within vertebrate palaeontology. Machine learning and artificial intelligence are fields that are undergoing massive leaps in technology and application and are impacting our everyday lives, with the rise of applications like ChatGPT as a current example. Vertebrate palaeontology needs to embrace this new technology and I hope the work presented in this thesis can be a small step in that direction.

References

Afanador, N. L., Smolinska, A., Tran, T. N., & Blanchet, L. (2016). Unsupervised random forest: a tutorial with case studies. *Journal of Chemometrics*, *30*(5), 232-241. https://doi.org/10.1002/cem.2790

Alifanov, V. R. (2014). The discovery of late Jurassic dinosaurs in Russia. *Doklady Earth Sciences*, *455*(2), 365-367. https://doi.org/10.1134/S1028334X14040011

Alifanov, V. R., Krasnolutksii, S. A., Markov, V. N., & Martynovich, N. V. (2001). About discovery of the Middle Jurassic dinosaurs in the Krasnoyarsk Territory. Scientific-Practical Conference "Problems of the struggle against illegal excavations and illegal turnover of the objects of archaeology, mineralogy and paleontology", Krasnoyarsk.

Amiot, R., Lécuyer, C., Buffetaut, E., Fluteau, F., Legendre, S., & Martineau, F. (2004). Latitudinal temperature gradient during the Cretaceous Upper Campanian–Middle Maastrichtian: δ18O record of continental vertebrates. *Earth and Planetary Science Letters*, *226*(1–2), 255-272. https://doi.org/http://dx.doi.org/10.1016/j.epsl.2004.07.015

Amiot, R., Wang, X., Zhou, Z., Wang, X., Lécuyer, C., Buffetaut, E., Fluteau, F., Ding, Z., Kusuhashi, N., Mo, J., Philippe, M., Suteethorn, V., Wang, Y., & Xu, X. (2015). Environment and ecology of East Asian dinosaurs during the Early Cretaceous inferred from stable oxygen and carbon isotopes in apatite. *Journal of Asian Earth Sciences*, *98*(0), 358-370. https://doi.org/http://dx.doi.org/10.1016/j.jseaes.2014.11.032

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA) [https://doi.org/10.1002/9781118445112.stat07841]. *Wiley StatsRef: Statistics Reference Online*, 1-15. https://doi.org/https://doi.org/10.1002/9781118445112.stat07841

Anderson, M. J., & 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*(4), 557-574. https://doi.org/https://doi.org/10.1890/12-2010.1

Anderson, P. E., Benton, M. J., Trueman, C. N., Paterson, B. A., & Cuny, G. (2007). Palaeoenvironments of vertebrates on the souther shore of Tethys: The nonmarine Early Cretaceous of Tunisia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *243*, 118-131.

Arbour, V. M., & Currie, P. J. (2015). Systematics, phylogeny and palaeobiogeography of the ankylosaurid dinosaurs. *Journal of Systematic Palaeontology*, 1-60. https://doi.org/10.1080/14772019.2015.1059985

Argast, S., Farlow, J. O., Gabet, R. M., & 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*(10), 927-930. https://doi.org/10.1130/0091-7613(1987)15<927:taofrt>2.0.co;2

Arkell, W. J. (1931). The Upper Great Oolite, Bradford Beds and Forest Marble of South Oxfordshire, and the Succession of Gastropod Faunas in the Great Oolite. *Quarterly Journal of the Geological Society*, *87*(1-4), 563-629. http://jgslegacy.lyellcollection.org/content/87/1-4/563.abstract

Arkell, W. J. (1933a). *The Jurassic System in Great Britain*. Clarendon Press.

Arkell, W. J. (1933b). New evidence on the Great Oolite succession at Bladon, near Woodstock, Oxfordshire. *Proceedings of the Geologists' Association*, *44*(2), 177-183. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(33)80018-6

Arkell, W. J. (1942). Stratigraphy and Structures East of Oxford. *Quarterly Journal of the Geological Society*, *98*(1-4), 187-204. http://jgslegacy.lyellcollection.org/content/98/1-4/187.abstract

Arkell, W. J. (1944). Stratigraphy and Structures East of Oxford. Part III: Islip. *Quarterly Journal of the Geological Society*, *100*(1-4), 61-73. http://jgslegacy.lyellcollection.org/content/100/1-4/61.abstract

Arkell, W. J. (1947). *The Geology of Oxford*. The Clarendon Press.

Arkell, W. J., Richardson, L., & Pringle, J. (1933). The Lower Oolites exposed in the Ardley and Fritwell Railway-cuttings, between Bicester and Banbury, Oxford. *Proceedings of the Geologists' Association*, *44*(3), 340-347. https://doi.org/https://doi.org/10.1016/S0016-7878(33)80009-5

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3.*

Averianov, A. O., Martin, T., & Bakirov, A. A. (2005). Pterosaur and dinosaur remains from the Middle Jurassic Balabansai Svita in the northern Fergana Depression, Kyrgyzstan (Central Asia). *Palaeontology*, *48*(1), 135-155. https://doi.org/10.1111/j.1475-4983.2004.00437.x

Averianov, A. O., Martin, T., Skutschas, P. P., Rezvyi, A. S., & Bakirov, A. A. (2008). Amphibians from the Middle Jurassic Balabansai Svita in the Fergana Depression, Kyrgyzstan (Central Asia). *Palaeontology*, *51*(2), 471-485. https://doi.org/10.1111/j.1475-4983.2007.00748.x

Avrahami, H. M., Gates, T. A., Heckert, A. B., Makovicky, P. J., & Zanno, L. E. (2018). A new microvertebrate assemblage from the Mussentuchit Member, Cedar Mountain Formation: insights into the paleobiodiversity and paleobiogeography of early Late Cretaceous ecosystems in western North America. *PeerJ*, *6*, e5883. https://doi.org/10.7717/peerj.5883

Baron, M. G., Norman, D. B., & Barrett, P. M. (2017a). A new hypothesis of dinosaur relationships and early dinosaur evolution [Article]. *Nature*, *543*(7646), 501-506. https://doi.org/10.1038/nature21700

http://www.nature.com/nature/journal/v543/n7646/abs/nature21700.html#supplementary-information

Baron, M. G., Norman, D. B., & Barrett, P. M. (2017b). Postcranial anatomy of *Lesothosaurus diagnosticus* (Dinosauria: Ornithischia) from the Lower Jurassic of southern Africa: implications for basal ornithischian taxonomy and systematics. *Zoological Journal of the Linnean Society*, *179*(1), 125-168. https://doi.org/10.1111/zoj.12434

Barrett, P. M. (2001). Tooth wear and possible jaw action of *Scelidosaurus harrisonii* Owen and a review of feeding mechanisms in other Thyreophoran dinosaurs. In K. Carpenter (Ed.), *The Armoured Dinosaurs*. Indiana University Press.

Barrett, P. M. (2006). A sauropod dinosaur tooth from the Middle Jurassic of Skye, Scotland. *Transactions of the Royal Society of Edinburgh: Earth Sciences*, *97*, 25-29.

Barrett, P. M. (2009). The affinities of the enigmatic dinosaur *Eshanosaurus deguchiianus* from the Early Jurassic of Yunnan Province, People's Republic of China [Article]. *Palaeontology*, *52*(4), 681-688. https://doi.org/10.1111/j.1475-4983.2009.00887.x

Barrett, P. M., Benson, R. B. J., Rich, T. H., & Vickers-Rich, P. (2011). First spinosaurid dinosaur from Australia and the cosmopolitanism of Cretaceous dinosaur faunas. *Biology Letters*, *7*(6), 933-936. https://doi.org/10.1098/rsbl.2011.0466

Barrett, P. M., Butler, R. J., & Knoll, F. (2005). Small-bodied ornithischian dinosaurs from the Middle Jurassic of Sichuan, China. *Journal of Vertebrate Paleontology*, *25*(4), 823-834. https://doi.org/10.1671/0272-4634(2005)025[0823:SODFTM]2.0.CO;2

Barrett, P. M., Butler, R. J., Mundil, R., Scheyer, T. M., Irmis, R. B., & Sánchez-Villagra, M. R. (2014). A palaeoequatorial ornithischian and new constraints on early dinosaur diversification. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1791), 20141147. https://doi.org/http://dx.doi.org/10.1098/rspb.2014.1147

Barrett, P. M., & Han, F.-L. (2009). Cranial anatomy of *Jeholosaurus shangyuanensis* (Dinosauria: Ornithischia) from the Early Cretaceous of China. *Zootaxa*, *2072*(1), 31-55. https://doi.org/10.11646/zootaxa.2072.1.2

Barrett, P. M., & Maidment, S. C. R. (2011). Dinosaurs of Dorset: Part III, the ornithischian dinosaurs (Saurischia, Ornithischia) with additional comments on the sauropods. *Proceedings of the Dorset Natural History and Archaeological Society*, *132*, 145-163.

Barrett, P. M., McGowan, A. J., & Page, V. (2009). Dinosaur diversity and the rock record. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1667), 2667-2674. https://doi.org/10.1098/rspb.2009.0352

Barrett, P. M., Pouech, J., Mazin, J.-M., & Jones, F. M. (2016). Teeth of embryonic or hatchling sauropods from the Berriasian (Early Cretaceous) of Cherves-de-Cognac, France. *Acta Palaeontologica Polonica*, *61*, 591-596.

Barrett, P. M., Rich, T. H., Vickers-Rich, P., Tumanova, T. Y. A., Inglis, M., Pickering, D., Kool, L., & Kear, B. P. (2010). Ankylosaurian dinosaur remains from the Lower Cretaceous of southeastern Australia. *Alcheringa: An Australasian Journal of Palaeontology*, *34*(3), 205-217. https://doi.org/10.1080/03115511003655430

Barrett, P. M., & Upchurch, P. (2005). Sauropodomorph diversity through time: macroevolutionary and palaeoecological implications. In K. A. Curry Rogers & J. A. Wilson (Eds.), *The Sauropods, Evolution and Paleobiology* (pp. 125-156). University of California Press.

Barron, A. J. M., Lott, G. K., & 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* (B. G. Survey, Ed.).

Barsbold, R., Osmólska, H., & Kurzanov, S. M. (1987). On a New Troodontid (Dinosauria, Theropoda) from the Early Cretaceous of Mongolia. *Acta Palaeontologica Polonica*, *32*(1-2), 121-132.

Barsbold, R., & Perle, A. (1980). Segnosauria, a new infraorder of carnivorous dinosaurs. *Acta Palaeontologica Polonica*, *25*(2), 187-195.

Barton, C. M., Woods, M. A., Bristow, C. R., Newell, A. J., Westhead, R. K., Evans, D. J., Kirby, G. A., Warrington, G., Riding, J. B., Freshney, E. C., Highley, D. E., Lott, G. K., Forster, A., & Gibson, A. (2011). Geology of south Dorset and south-east Devon and its World Heritage Coast. British Geological Survey, *Special Memoir for 1:50 000 geological sheets 328 Dorchester, 341/342 West Fleet and Weymouth and 342/343 Swanage, and partsof sheets 326/340 Sidmouth, 327 Bridport, 329 Bournemouth and 339 Newton Abbot, pp 171.*

Baszio, S. (1997). Systematic palaeontology of isolated dinosaur teeth from the latest Cretaceous of South Alberta, Canada. *Courier Forschungsinstitut Senckenberg*, *196*, 33-77.

Baszio, S. (2008). Information from microvertebrate localities: Potentials and limits. In J. T. Sankey & S. Baszio (Eds.), *Vertebrate microfossil assemblages: their role in paleoecology and paleobiogeography* (pp. 3-8). Indiana University Press.

Becerra, M. G., Pol, D., Marsicano, C. A., & Rauhut, O. W. M. (2013). The dentition of *Manidens condorensis* (Ornithischia; Heterodontosauridae) from the Jurassic Cañadón Asfalto Formation of Patagonia: morphology, heterodonty and the use of statistical methods for identifying isolated teeth. *Historical Biology*, *26*(4), 480-492. https://doi.org/10.1080/08912963.2013.794227

Becerra, M. G., Pol, D., Porro, L. B., Paulina-Carabajal, A., & Rauhut, O. W. M. (2023). Craniomandibular osteology of *Manidens condorensis* (Ornithischia: Heterodontosauridae) from the upper Lower Jurassic of Argentina. *Journal of Vertebrate Paleontology*, e2181087. https://doi.org/10.1080/02724634.2023.2181087

Becerra, M. G., Pol, D., Rössner, G. E., & Rauhut, O. W. M. (2018). Heterodonty and double occlusion in *Manidens condorensis*: a unique adaptation in an Early Jurassic ornithischian improving masticatory efficiency. *The Science of Nature*, *105*(7-8). https://doi.org/10.1007/s00114-018-1569-6

Becerra, M. G., Pol, D., Whitlock, J. A., & Porro, L. B. (2020). Tooth replacement in *Manidens condorensis* : baseline study to address the replacement pattern in dentitions of early ornithischians. *Papers in Palaeontology*. https://doi.org/10.1002/spp2.1337

Behrensmeyer, A. K. (1991). Terrestrial vertebrate accumulations. In P. A. Allison & D. E. G. Briggs (Eds.), *Taphonomy: Releasing the data locked in the fossil record* (pp. 560). Plenum Press.

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. (2010a). 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*(4), 882-935. https://doi.org/10.1111/j.1096-3642.2009.00569.x

Benson, R. B. J. (2010b). 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*(1), 131-146. https://doi.org/10.1080/14772011003603515

Benson, R. B. J., & Barrett, P. M. (2009). Dinosaurs of Dorset: Part I, the carniverous dinosaurs (Saurischia, Theropoda). *Proceedings of the Dorset Natural History and Archaeological Society*, *130*, 133-147.

Benson, R. B. J., Barrett, P. M., Powell, H. P., & Norman, D. B. (2008). The taxonomic status of *Megalosaurus bucklandii* (Dinosauria, Theropoda) from the Middle Jurassic of Oxfordshire, UK. *Palaeontology*, *51*(2), 419-424. https://doi.org/10.1111/j.1475-4983.2008.00751.x

Benson, R. B. J., & Radley, J. D. (2010). A new large-bodied theropod dinosaur from the Middle Jurassic of Warwickshire, United Kingdom. *Acta Palaeontologica Polonica*, *55*(1), 35-42. https://doi.org/10.4202/app.2009.0083

Benson, R. B. J., Rich, T. H., Vicker-Rich, P., & Hall, M. (2012). Theropod fauna from southern Australia indicates high polar diversity and climate-driven dinosaur provinciality. *PLoS ONE*, *7*(5), e37122. https://doi.org/10.1371/journal.pone.0037122

Benton, M. J., Cook, E., & Hooker, J. J. (2005). Mesozoic and Tertiary fossil mammals and birds of Great Britain. *Geological Conservation Review Series No. 32, Joint Nature Conservation Committee*.

Benton, M. J., & Spencer, P. S. (1995). British Mid Jurassic fossil reptile sites. In M. J. Benton & P. S. Spencer (Eds.), *Fossil Reptiles of Great Britain* (pp. 123-164). Springer Netherlands. https://doi.org/10.1007/978-94-011-0519-4_6

Bhattacharya, A., Ray, S., Datta, P. M., & Maulick, P. (1994). Fossil Charophyta from the Kota Formation of the Pranhita-Godavari valley, Andhra Pradesh, India. *Proceedings of Ninth International Gondwana Symposium 1*, 471-475.

Billon-Bruyat, J.-P., Mazin, J.-M., & Pouech, J. (2010). A stegosaur tooth (Dinosauria, Ornithischia) from the Early Cretaceous of southwestern France. *Swiss Journal of Geosciences*, *103*(2), 143-153. https://doi.org/10.1007/s00015-010-0028-y

Blagus, R., & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*, *14*(1), 1-16. https://doi.org/10.1186/1471-2105-14-106

Blob, R. W., & Fiorillo, A. R. (1996). The Significance of Vertebrate Microfossil Size and Shape Distributions for Faunal Abundance Reconstructions: A Late Cretaceous Example. *Paleobiology*, *22*(3), 422-435. https://doi.org/10.2307/2401098

Boneham, B., & Wyatt, R. J. (1993). The stratigraphical position of the Middle Jurassic (Bathonian) Stonesfield Slate of Stonesfield, Oxfordshire, UK. *Proceedings of the Geologists' Association*, *104*(2), 123-136.

Boneham, B. F. W., & Forsey, G. F. (1992). Earliest stegosaur dinosaur. *Terra Nova*, *4*, 628-632.

Bordy, E. M., Abrahams, M., Sharman, G. R., Viglietti, P. A., Benson, R. B. J., McPhee, B. W., Barrett, P. M., Sciscio, L., Condon, D., Mundil, R., Rademan, Z., Jinnah, Z., Clark, J. M., Suarez, C. A., Chapelle, K. E. J., & Choiniere, J. N. (2020). A chronostratigraphic framework for the upper Stormberg Group: Implications for the Triassic-Jurassic boundary in southern Africa. *Earth-Science Reviews*, *203*, 103120. https://doi.org/10.1016/j.earscirev.2020.103120

Boyd, C. A. (2014). The cranial anatomy of the neornithischian dinosaur *Thescelosaurus neglectus*. *PeerJ*, *2*, e669. https://doi.org/10.7717/peerj.669

Boyd, C. A. (2015). The systematic relationships and biogeographic history of ornithischian dinosaurs. *PeerJ*, *3*, e1523. https://doi.org/10.7717/peerj.1523

Bradshaw, M. J. (1978). A facies analysis of the Bathonian of eastern England, University of Oxford, PhD thesis.

Bradshaw, M. J., Cope, J. C. W., Cripps, D. W., Donovan, D. T., Howarth, M. K., Rawson, P. F., West, I. M., & Wimbledon, W. A. (1992). Jurassic. In J. C. W. Cope, J. K. Ingham, & P. F. Rawson (Eds.), *Atlas of palaeogeography and lithofacies. Geological Society, London, Memoirs, 13* (Vol. 13, pp. 110-120). http://mem.lyellcollection.org/content/13/1/107.abstract

Branco, P., Ribeiro, R. P., & Torgo, L. (2016). UBL: an R package for utility-based learning. *CoRR*, *abs/1604.08079*. http://arxiv.org/abs/1604.08079

Breeden, B. T., Raven, T. J., Butler, R. J., Rowe, T. B., & Maidment, S. C. R. (2021). The anatomy and palaeobiology of the early armoured dinosaur *Scutellosaurus lawleri* (Ornithischia: Thyreophora) from the Kayenta Formation (Lower Jurassic) of Arizona. *Royal Society Open Science*, *8*(7), 201676. https://doi.org/10.1098/rsos.201676

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Chapman and Hall / CRC Press.

Brett-Surman, M., Jabo, S., Kroehler, P., & Carrano, M. (2005). A new microvertebrate assemblage from the Upper Jurassic Morrison Formation, including mammals, theropods and sphenodontians. *Journal of Vertebrate Paleontology*, *25*(Supplement to Number 3), 39A.

Brinkman, D. B. (1990). Paleoecology of the Judith River Formation (Campanian) of Dinosaur Provincial Park, Alberta, Canada: Evidence from vertebrate microfossil localities. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *78*, 37-54.

Brinkman, D. B., Russell, A. P., & Peng, J.-H. (2005). Vertebrate microfossil sites and their contribution to studies of paleoecology. In P. J. Currie & E. B. Koppelhus (Eds.), *Dinosaur Provincial Park: A spectacular ecosystem revealed*. Indiana University Press.

Bristow, C. R., Barton, C. M., Freshney, E. C., Wood, C. J., Evans, D. J., Cox, B. M., Ivimey-Cook, H., & Taylor, R. T. (1995). Geology of the country around Shaftesbury, *British Geological Survey*, *Vol. Sheet 313 (England and Wales).*

Britzke, E. R., Duchamp, J. E., Murray, K. L., Swihart, R. K., & Robbins, L. W. (2011). Acoustic identification of bats in the eastern United States: A comparison of parametric and nonparametric methods. *The Journal of Wildlife Management*, *75*(3), 660-667. https://doi.org/10.1002/jwmg.68

Broderip, W. J. (1828). Observations on the jaw of a fossil mammiferous animal found in the Stonesfield Slate. *Zoological Journal*, *3*, 408-412.

Brown, B. (1908). The Ankylosauridae, a new family of armored dinosaurs from the Upper Cretaceous. *Bulletin of the American Museum of Natural History*, *12*, 187-201.

Brown, B., & 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*(5), 115-150.

Brown, C. M., & Druckenmiller, P. (2011). Basal ornithopod (Dinosauria: Ornithischia) teeth from the Prince Creek Formation (early Maastrichtian) of Alaska [Article]. *Canadian Journal of Earth Sciences*, *48*(9), 1342-1354. https://doi.org/10.1139/e11-017

Brunson, J. C. (2019). *ggalluvial: Alluvial Diagrams in 'ggplot2'. R package version 0.9.1.*

Brusatte, S. L., Butler, R. J., Barrett, P. M., Carrano, M. T., Evans, D. C., Lloyd, G. T., Mannion, P. D., Norell, M. A., Peppe, D. J., Upchurch, P., & Williamson, T. E. (2014). The extinction of the dinosaurs. *Biological Reviews*, 90: 628-642. https://doi.org/10.1111/brv.12128https://doi.org/doi: 10.1111/brv.12128

Brusatte, S. L., & Clark, N. D. L. (2015). Theropod dinosaurs from the Middle Jurassic (Bajocian–Bathonian) of Skye, Scotland. *Scottish Journal of Geology*. https://doi.org/10.1144/sjg2014-022

Buckland, W. (1824). Notice on *Megalosaurus*. *Transactions of the Geological Society of London*, *1*, 390-396.

Buckley, L. G., & 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., & 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*(9), 1227-1251. https://doi.org/10.1139/E10-029

Burmeister, K. C., Flynn, J. J., Parrish, J., & Wyss, A. (1999). New fossil vertebrates from the Northen Morondava Basin, Madagascar, and the recovery of microvertebrates from coprolites. *PaleoBios*, *19 Supplement to Number 1*, 3.

Butler, P. M., & Hooker, J. J. (2005). New teeth of allotherian mammals from the English Bathonian, including the earliest multituberculates. *Acta Palaeontologica Polonica*, *50*(2), 185-207.

Butler, R. J. (2005). The 'fabrosaurid' ornithischian dinosaurs of the Upper Elliot Formation (Lower Jurassic) of South Africa and Lesotho. *Zoological Journal of the Linnean Society*, *145*(2), 175-218. https://doi.org/10.1111/j.1096-3642.2005.00182.x

Butler, R. J., Porro, L. B., Galton, P. M., & Chiappe, L. M. (2012). Anatomy and Cranial Functional Morphology of the Small-Bodied Dinosaur *Fruitadens haagarorum* from the Upper Jurassic of the USA. *PLoS ONE*, *7*(4), e31556. https://doi.org/10.1371/journal.pone.0031556

Butler, R. J., Smith, R. M. H., & Norman, D. B. (2007). A primitive ornithischian dinosaur from the Late Triassic of South Africa, and the early evolution and diversification of Ornithischia. *Proceedings. Biological sciences*, *274*(1621), 2041-2046. https://doi.org/10.1098/rspb.2007.0367

Butler, R. J., Upchurch, P., & Norman, D. B. (2008). The phylogeny of the ornithischian dinosaurs [10.1017/S1477201907002271]. *Journal of Systematic Palaeontology*, *6*(01), 1-40.

Callomon, J. H., & Cope, J. C. W. (1995). The Jurassic Geology of Dorset. In P. D. Taylor (Ed.), *Field Geology of the British Jurassic* (pp. 51-103). The Geological Society.

Carpenter, K. (1982). Baby dinosaurs from the Late Cretaceous Lance and Hell Creek formations and a description of a new species of theropod. *Rocky Mountain Geology*, *20*(2), 123-134. http://rmg.geoscienceworld.org/content/20/2/123.short

Carpenter, K. (2001). Phylogenetic analysis of the Ankylosauria. In K. Carpenter (Ed.), *The Armoured Dinosaurs* (pp. 455-483). Indiana University Press.

Carpenter, K. (2004). Redescription of *Ankylosaurus magniventris* Brown 1908 (Ankylosauridae) from the Upper Cretaceous of the Western Interior of North America. *Canadian Journal of Earth Sciences*, *41*(8), 961-986. https://doi.org/10.1139/e04-043

Carpenter, K., Dilkes, D. W., & Weishampel, D. B. (1995). The dinosaurs of the Niobrara Chalk Formation (Upper Cretaceous, Kansas). *Journal of Vertebrate Paleontology*, *15*(2), 275-297.

Carpenter, K., & Galton, P. M. (2018). A photo documentation of bipedal ornithischian dinosaurs from the Upper Jurassic Morrison Formation, USA. *Geology of the Intermountain West*, *5*, 167 - 207.

Carrano, M. T., Benson, R. B. J., & Sampson, S. D. (2012). The phylogeny of Tetanurae (Dinosauria: Theropoda). *Journal of Systematic Palaeontology*, *10*(2), 211-300. https://doi.org/10.1080/14772019.2011.630927

Carrano, M. T., Oreska, M. P. J., & Lockwood, R. (2016). Vertebrate paleontology of the Cloverly Formation (Lower Cretaceous), II: Paleoecology. *Journal of Vertebrate Paleontology*, *36*(2), e1071265. https://doi.org/10.1080/02724634.2015.1071265

Carrano, M. T., & Velez-Juarbe, J. (2006). Paleoecology of the Quarry 9 vertebrate assemblage from Como Bluff, Wyoming (Morrison Formation, Late Jurassic). *Palaeogeography, Palaeoclimatology, Palaeoecology*, *237*(2–4), 147-159. https://doi.org/http://dx.doi.org/10.1016/j.palaeo.2005.11.018

Case, J. A., Martin, J. E., & Reguero, M. A. (2007). A dromaeosaur from the Maastrichtian of James Ross Island and the Late Cretaceous Antarctic dinosaur fauna. *In A. K. Cooper, C. R. Raymond, & the ISAES Editorial Team (eds.), Antarctica: A Keystone in a Changing World. Online Proceedings of the 10th ISAES. United States Geological Survey and the National Academies. USGS Open-File Report 2007-1047, Short Research Paper*, *083*, 1-4. https://doi.org/10.3133/0f2007-1047.srp083

Chandler, L. (1990). *Taxonomic and functional significance of serrated tooth morphology in theropod dinosaurs* Yale University, PhD thesis, New Haven, CT.

Channon, P. J. (1950). New and enlarged Jurassic sections in the Cotswolds. *Proceedings of the Geologists' Association*, *61*(4), 242-IN246. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(50)80032-9

Charig, A. J., & Crompton, A. W. (1974). The alleged synonymy of *Lycorhinus* and *Heterodontosaurus*. *Annals of the South African Museum*, *64*, 167-189.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321-357.

Chen, X., & Ishwaran, H. (2012). Random forests for genomic data analysis. *Genomics*, *99*(6), 323-329. https://doi.org/10.1016/j.ygeno.2012.04.003

Chiarenza, A. A., Fiorillo, A. R., Tykoski, R. S., McCarthy, P. J., Flaig, P. P., & Contreras, D. L. (2020). The first juvenile dromaeosaurid (Dinosauria: Theropoda) from Arctic Alaska. *PLoS ONE*, *15*(7), e0235078. https://doi.org/10.1371/journal.pone.0235078

Choiniere, J. N., Forster, C. A., & 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. https://doi.org/https://doi.org/10.1016/j.jafrearsci.2012.05.005

Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, *10*(10), 1632-1644. https://doi.org/10.1111/2041-210X.13256

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.

Clark, N. D. L. (2001). A thyreophoran dinosaur from the Early Bajocian (Middle Jurassic) of the Isle of Skye, Scotland. *Scottish Journal of Geology*, *37*(1), 19-26. https://doi.org/10.1144/sjg37010019

Clemens, W. A. (1960). *Fossil mammals of the type Lance Formation, Wyoming* University of California.

Clemens, W. A. (1966). Fossil mammals of the type Lance Formation, Wyoming. Part II. Marsupialia. *University of California Publications in Geological Sciences*, *62*, 1-105.

Clemens, W. A., Lillegraven, J. A., Lindsay, E. H., & Simpson, G. G. (1979). Where, when, and what: a survey of known Mesozoic mammal distribution. In J. A. Lillegraven, Z. Kielan-Jaworowska, & W. A. Clemens (Eds.), *Mesozoic mammals: the first two-thirds of mammalian history* (pp. 7-58). University of California Press.

Close, Roger A., Friedman, M., Lloyd, Graeme T., & Benson, Roger B. J. (2015). Evidence for a Mid-Jurassic adaptive radiation in Mammals. *Current Biology*, *25*(16), 2137-2142. https://doi.org/https://doi.org/10.1016/j.cub.2015.06.047

Coe, A. L., Argles, T. W., Rothery, D. A., & Spicer, R. A. (2010). *Geological Field Techniques*. Wiley-Blackwell.

Colbert, E. H. (1981). A primative ornithischian dinosaur from the Kayenta Formation of Arizona. *Museum of Northern Arizona Press*, *53*, 1-61.

Cook, E., & Ross, A. J. (1996). The stratigraphy, sedimentology and palaeontology of the Lower Weald Clay (Hauterivian) at Keymer Tileworks, West Sussex, southern England. *Proceedings of the Geologists' Association*, *107*(3), 231-239. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(96)80031-9

Coombs, W. P. (1978). The families of the ornithischain dinosaur order Ankylosauria. *Palaeontology*, *21*, 143-170.

Coombs, W. P. (1990). Teeth and taxonomy in ankylosaurs. In K. Carpenter & P. J. Currie (Eds.), *Dinosaur systematics: approaches and perspectives* (pp. 318). Cambridge University Press.

Cope, J. C. W., Duff, K. L., Parsons, C. F., Torrens, H., Wimbledon, W. A., & Wright, J. K. (1980). A correlation of Jurassic rocks in the British Isles Part Two: Middle and Upper Jurassic. *Geological Society of London, pp 1-109.*

Cope, J. C. W., Ingham, J. K., & Rawson, P. F. (1992). Altas of palaeogeography and lithofacies, *Geological Society, London, Memoir, 13, pp 1 – 153.*

Cope, J. W. (2012). Geology of the Dorset Coast (3rd ed.). *Geologists' Association London (2016). pp. 1–222.*

Corentin, B., & Salvador, B. (2018). A New Fossil Species of *Boa* Linnaeus, 1758 (Squamata, Boidae), from the Pleistocene of Marie-Galante Island (French West Indies). *Journal of Vertebrate Paleontology*, *38*(3). https://doi.org/10.1080/02724634.2018.1462829

Couronné, R., Probst, P., & Boulesteix, A.-L. (2018). Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics*, *19*, 1-14. https://doi.org/10.1186/s12859-018-2264-5

Cox, B. M., & Sumbler, M. G. (2002). British Middle Jurassic Stratigraphy. *Geological Conservation Review Series, No. 26* (Vol. 26). Joint Nature Conservation Committee.

Criminisi, A., Shotton, J., & Konukoglu, E. (2012). Decision forests: a unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and trends in computer graphics and vision*, *7*(2-3), 81-227. https://www.microsoft.com/en-us/research/publication/decision-forests-a-unified-framework-for-classification-regression-density-estimation-manifold-learning-and-semi-supervised-learning/

Cripps, D. W. (1986). A facies analysis of the upper great oolite group in central and eastern England. *Unpublished PhD thesis. Aston University. Birmingham.*

Crompton, A. W., & Charig, A. J. (1962). A new Ornithischian from the Upper Triassic of South Africa. *Nature*, *196*(4859), 1074-1077. https://doi.org/10.1038/1961074a0

Csiki, Z., Ionescu, A., & Grigorescu, D. (2008). The Budurone microvertebrate site from the Maastrichtian of the Hateg Basin – flora, fauna, taphonomy and paleoenvironment. *acta Palaeontologica Romaniae*, *6*, 49-66.

Cullen, T. M., & Evans, D. C. (2016). Palaeoenvironmental drivers of vertebrate community composition in the Belly River Group (Campanian) of Alberta, Canada, with implications for dinosaur biogeography. *BMC Ecology*, *16*(1), 52. https://doi.org/10.1186/s12898-016-0106-8

Cullen, T. M., Fanti, F., Capobianco, C., Ryan, M. J., & Evans, D. C. (2016). A vertebrate microsite from a marine-terrestrial transition in the Foremost Formation (Campanian) of Alberta, Canada, and the use of faunal assemblage data as a paleoenvironmental indicator. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *444*, 101-114. https://doi.org/http://dx.doi.org/10.1016/j.palaeo.2015.12.015

Currie, P. J. (1987). Bird-Like Characteristics of the Jaws and Teeth of Troodontid Theropods (Dinosauria, Saurischia). *Journal of Vertebrate Paleontology*, *7*(1), 72-81. http://www.jstor.org/stable/4523121

Currie, P. J., Rigby, K. J., & Sloan, R. E. (1990). Theropod teeth from the Judith River Formation of Southern Alberta, Canada. In K. Carpenter & P. J. Currie (Eds.), *Dinosaur Systematics Approaches and Perspectives* (pp. 107-125). Cambridge University Press.

Currie, P. J., & Varrichio, D. J. (2004). A new dromaeosaurid from the Horseshoe Canyon Formation (Upper Cretaceous) of Alberta, Canada. In P. J. Currie, E. B. Koppelhus, M. A. Shugar, & J. L. Wright (Eds.), *Feathered Dragons: Studies on the Transition from Dinosaurs to Birds* (pp. 112-132). Indiana University Press.

Curtis, K., & Padian, K. (1999). An Early Jurassic microvertebrate fauna from the Kayenta Formation of northeastern Arizona: Microfaunal change across the Triassic-Jurassic boundary. *PaleoBios*, *19*(2), 19-37.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783-2792. https://doi.org/10.1890/07-0539.1

D'Amore, D. C. (2009). A Functional Explanation for Denticulation in Theropod Dinosaur Teeth. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, *292*(9), 1297-1314. https://doi.org/10.1002/ar.20977

D'Emic, M. D., Whitlock, J. A., Smith, K. M., Fisher, D. C., & Wilson, J. A. (2013). Evolution of High Tooth Replacement Rates in Sauropod Dinosaurs. *PLoS ONE*, *8*(7), e69235. https://doi.org/10.1371/journal.pone.0069235

Darlington, J. (1988). Hornsleasow Quarry Archive Report. *Crickley Hill Archaeological Trust, pp 1-25.*

Dasgupta, S. (2021). A Review of Stratigraphy, Depositional Setting and Paleoclimate of the Mesozoic Basins of India. In S. Banerjee & S. Sarkar (Eds.), *Mesozoic Stratigraphy of India: A Multi-Proxy Approach* (pp. 1-37). Springer International Publishing. https://doi.org/10.1007/978-3-030-71370-6_1

Dawson, H. L., Dubrule, O., & John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. *Computers & Geosciences*, *171*, 105284. https://doi.org/https://doi.org/10.1016/j.cageo.2022.105284

De Spiegeleer, J., Madan, D. B., Reyners, S., & Schoutens, W. (2018). Machine learning for quantitative finance: fast derivative pricing, hedging and fitting. *Quantitative Finance*, *18*(10), 1635-1643. https://doi.org/10.1080/14697688.2018.1495335

Dietterich, T. G. (2001). Ensemble methods in machine learning. In J. Kittler & F. Roli (Eds.), *Multiple classier systems* (pp. 404). Springer-Verlag.

Dieudonné, P. E., Cruzado-Caballero, P., Godefroit, P., & Tortosa, T. (2021). A new phylogeny of cerapodan dinosaurs. *Historical Biology*, *33*(10), 2335-2355. https://doi.org/10.1080/08912963.2020.1793979

Dineley, D., & Metcalf, S. J. (1999). British Jurassic fossil fishes sites. In D. Dineley & S. J. Metcalf (Eds.), *Fossil Fishes of Great Britain, Geological Conservation Review Series No. 16* (Vol. 16, pp. 355 - 415). Joint Nature Conservation Committee.

Ding, A., Pittman, M., Upchurch, P., O'Connor, J., Field, D. J., & Xu, X. (2020). The biogeography of coelurosaurian theropods and its impact on their evolutionary history. In M. Pittman & X. Xu (Eds.), *Pennaraptoran theropod dinosaurs : past progress and new frontiers. (Bulletin of the American Museum of Natural History, no. 440)* (pp. 117-158). American Museum of Natural History.

Dodson, P. (1987). Microfaunal studies of dinosaur palaeoecology, Judith River Formation of southern Alberta. In P. J. Currie & E. H. Koster (Eds.), *4th symposium on Mesozoic terrestrial ecosystems* (Vol. 3, pp. 70-75). Tyrrell Museum of Palaeontology.

Dodson, P., & Dawson, S. (1991). Making the fossil record of dinosaurs. *Modern Geology*, *16*, 3-15.

Domingo, L., Barroso-Barcenilla, F., & Cambra-Moo, O. (2013). Paleoenvironmental reconstruction of the “Lo Hueco” fossil site (Upper Cretaceous, Cuenca, Spain): preliminary stable isotope analyses on crocodilians and dinosaurs. *PALAIOS*, *28*(3), 195-202. https://doi.org/10.2110/palo.2012.p12-097r

Domingo, L., Barroso-Barcenilla, F., & Cambra-Moo, O. (2015). Seasonality and Paleoecology of the Late Cretaceous Multi-Taxa Vertebrate Assemblage of “Lo Hueco” (Central Eastern Spain). *PLoS ONE*, *10*(3), e0119968. https://doi.org/10.1371/journal.pone.0119968

Dunhill, A. M., Bestwick, J., Narey, H., & Sciberras, J. (2016). Dinosaur biogeographical structure and Mesozoic continental fragmentation: a network-based approach. *Journal of Biogeography*, *43*, 1691-1704. https://doi.org/10.1111/jbi.12766

Džeroski, S. (2001). Applications of symbolic machine learning to ecological modelling. *Ecological Modelling*, *146*(1), 263-273. https://doi.org/https://doi.org/10.1016/S0304-3800(01)00312-X

Eagle, R. A., Tütken, T., Martin, T. S., Tripati, A. K., Fricke, H. C., Connely, M., Cifelli, R. L., & Eiler, J. M. (2011). Dinosaur Body Temperatures Determined from Isotopic (13C-18O) Ordering in Fossil Biominerals. *Science*, *333*(6041), 443-445. https://doi.org/10.1126/science.1206196

Eaton, T. H. (1960). A new armored dinosaur from the Cretaceous of Kansas. *The University of Kansas Paleontological Contributions: Vertebrata*, *8*, 1-24.

Eberth, D. A. (1990). Stratigraphy and sedimentology of vertebrate microfossil sites in the uppermost Judith River Formation (Campanian), Dinosaur Provincial Park, Alberta, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *78*, 1-36.

Eberth, D. A., Currie, P. J., Brinkman, D. B., Ryan, M. J., Braman, D. R., Gardner, J. D., Lam, V. D., Spivak, D. N., & Neuman, A. G. (2001). Alberta's dinosaurs and other fossil vertebrates: Judith River and Edmonton groups (Campanian-Maastrichtian) (Vol. 3). In C. L. Hill (ed), *Society of Vertebrate Paleontology, 61st Annual Meeting, Bozeman. Guidebook for the Field Trips: Mesozoic and Cenozoic Paleontology in the Western Plains and Rocky Mountains,* Museum of the Rockies Occasional Paper.

Eberth, D. A., Shannon, M., & Noland, B. (2007). A bonebeds database: classification, biases, and patterns of occurrence. In R. Rogers, D. A. Eberth, & A. Fiorillo (Eds.), *Bonebeds: genesis, analysis, and paleobiological significance*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226723730.003.0003

Ensom, P. C. (1977). A therapsid tooth from the Forest Marble (Middle Jurassic) of Dorset. *Proceedings of the Geologists' Association*, *88*(4), 201-205. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(77)80007-2

Ensom, P. C. (1987). A remarkable new vertebrate site in the Purbeck Limestone Formation on the Isle of Purbeck. *Proceedings of Dorset Natural History and Archaeological Society*, *108*, 205-206.

Ensom, P. C. (1988). Excavations at Sunnydown Farm, Langton Matravers, Dorset: Amphibians discovered in the Purbeck Limestone Formation. *Proceedings of Dorset Natural History and Archaeological Society*, *109*, 148-150.

Ensom, P. C., Evans, S. E., & Milner, A. R. (1991). Amphibians and reptiles from the Purbeck Formation (Upper Jurassic) of Dorset. *Fifth Symposium on Mesozoic Terrestiral Ecosystems and Biota, Short Papers. Contributions from the Palaeontological Museum, University of Oslo*, *364*, 19-20.

Ensom, P. C., Francis, J. F., Kielan-Jaworowska, Z., & Milner, A. R. (1994). The flora and fauna of the Sunnydown Farm footprint site and associated sites: Purbeck Limestone Formation, Dorset. *Proceedings of the Dorset Natural History and Archaeological Society*, *115*, 181-182.

Erickson, G. M. (1996). Incremental lines of von Ebner in dinosaurs and the assessment of tooth replacement rates using growth line counts. *Proceedings of the National Academy of Sciences*, *93*(25), 14623-14627. http://www.pnas.org/content/93/25/14623.abstract

Estes, R. (1964). Fossil vertebrates from the Late Cretaceous Lance Formation, eastern Wyoming. *University of California, Publications in Geological Sciences*, *49*, 1-180.

Evans, A. R., Wilson, G. P., Fortelius, M., & Jernvall, J. (2007). High-level similarity of dentitions in carnivorans and rodents [10.1038/nature05433]. *Nature*, *445*(7123), 78-81. https://doi.org/http://www.nature.com/nature/journal/v445/n7123/suppinfo/nature05433_S1.html

Evans, D. C., Barrett, P. M., & Seymour, K. L. (2012). Revised identification of a reported Iguanodon-grade ornithopod tooth from the Scollard Formation, Alberta, Canada. *Cretaceous Research*, *33*(1), 11-14. https://doi.org/http://dx.doi.org/10.1016/j.cretres.2011.07.002

Evans, S. E. (1990). The skull of Cteniogenys, a choristodere (Reptilia: Archosauromorpha) from the Middle Jurassic of Oxfordshire. *Zoological Journal of the Linnean Society*, *99*(3), 205-237. https://doi.org/10.1111/j.1096-3642.1990.tb00561.x

Evans, S. E. (1991). The postcranial skeleton of the choristodere *Cteniogenys* (Reptilia: Diapsida) from the Middle Jurassic of England. *Geobios*, *24*, 187-199.

Evans, S. E. (1992). Small reptiles and amphibians from the Forest Marble (Middle Jurassic) of Dorset. *Proceedings of Dorset Natural History and Archaeological Society*, *113*, 201-202.

Evans, S. E. (2003). At the feet of the dinosaurs: the early history and radiation of lizards. *Biological Reviews*, *78*(4), 513-551. https://doi.org/10.1017/S1464793103006134

Evans, S. E., Barrett, P. M., Hilton, J., Butler, R. J., Jones, M. E. H., Liang, M.-M., Parish, J. C., Rayfield, E. J., Sigogneau-Russell, D., & Underwood, C. J. (2006). The Middle Jurassic vertebrate assemblage of Skye, Scotland. In P. M. Barrett & S. E. Evans (Eds.), *Ninth international symposium on Mesozoic terrestrial ecosystems and biota, Manchester, UK* (pp. 36-39). Natural History Museum.

Evans, S. E., & Kermack, K. A. (1994). Assemblages of small tetrapods from the Early Jurassic of Britain. In N. C. Fraser & H.-D. Sues (Eds.), *In the shadow of the Dinosaurs - Early Mesozoic tetrapods* (pp. 435). Cambridge University Press.

Evans, S. E., & McGowan, G. J. (2002). Lissamphibian remians from the Purbeck Limestone Group, Southern England. In A. R. Milner & D. J. Batten (Eds.), *Life and environments in Purbeck times* (Vol. 68, pp. 103-119).

Evans, S. E., & Milner, A. R. (1994). Middle Jurassic microvertebrate assemblages from the British Isles. In N. C. Fraser & H.-D. Sues (Eds.), *In the shadow of the dinosaurs. Early Mesozoic tetrapods* (pp. 435). Cambridge University Press.

Evans, S. E., Milner, A. R., & Mussett, F. (1988). The earliest known salamanders (Amphibia: caudata) a record from the Middle Jurassic of England. *Geobios*, *21*, 539-552.

Evans, S. E., Prasad, G. V. R., & Manhas, B. K. (2002). Fossil lizards from the Jurassic Kota Formation of India. *Journal of Vertebrate Paleontology*, *22*(2), 299-312. https://doi.org/10.1671/0272-4634(2002)022[0299:FLFTJK]2.0.CO;2

Evans, S. E., & Searle, B. (2002). Lepidosaurian reptiles from the Purbeck Limestone Group of Dorset, Southern England. In A. R. Milner & D. J. Batten (Eds.), *Life and environments in Purbeck times* (Vol. 68, pp. 145-159).

Evans, S. E., & Waldman, M. (1996). Small reptiles and amphibians from the Middle Jurassic of Skye, Scotland. In M. Morales (Ed.), *The Continental Jurassic* (Vol. 60). Museum of Northern Arizona.

Fanti, F., & Miyashita, T. (2009). A high latitude vertebrate fossil assemblage from the Late Cretaceous of west-central Alberta, Canada: evidence for dinosaur nesting and vertebrate latitudinal gradient. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *275*(1–4), 37-53. https://doi.org/http://dx.doi.org/10.1016/j.palaeo.2009.02.007

Fanti, F., & Therrien, F. (2007). Theropod tooth assemblages from the Late Cretaceous Maevarano Formation and the possible presence of dromaeosaurids in Madagascar. *Acta Palaeontologica Polonica*, *52*(1), 155-166.

Farlow, J. O., Brinkman, D. B., Abler, W. L., & Currie, P. J. (1991). Size, shape and serration density of theropod dinosaur lateral teeth. *Modern Geology*, *16*, 161-198.

Fastovsky, D. E., & Weishampel, D. B. (1996). *The Evolution and Extinction of the Dinosaurs* (1st ed.). Cambridge University Press.

Feldesman, M. R. (2002). Classification trees as an alternative to linear discriminant analysis. *American Journal of Physical Anthropology*, *119*(3), 257-275. https://doi.org/10.1002/ajpa.10102

Finch, W. H., Bolin, J. H., & Kelley, K. (2014). Group membership prediction when known groups consist of unknown subgroups: a Monte Carlo comparison of methods [10.3389/fpsyg.2014.00337]. *Frontiers in Psychology*, *5*(Article 337), 1-12. https://www.frontiersin.org/article/10.3389/fpsyg.2014.00337

Fiorillo, A. R., & Currie, P. J. (1994). Theropod Teeth from the Judith River Formation (Upper Cretaceous) of South-Central Montana. *Journal of Vertebrate Paleontology*, *14*(1), 74-80. http://www.jstor.org/stable/4523546

Fiorillo, A. R., & Gangloff, R. A. (2001). Theropod teeth from the Prince Creek Formation (Cretaceous) of northern Alaska, with speculations on Arctic dinosaur paleoecology. *Journal of Vertebrate Paleontology*, *20*(4), 675-682.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*(2), 179-188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Flynn, J. J., Fox, S., Parrish, M., L, R., & A, W. (2006). Assessing diversity and palaeoecology of a Middle Jurassic microvertebrate assemblage from Madagascar. In J. D. Harris, S. G. Lucas, J. A. Spielmann, M. G. Lockley, A. R. Milner, & J. I. Kirkland (Eds.), *The Triassic-Jurassic Terrestrial Transition* (Vol. 37, pp. 476-489). Bulletin 37, New Mexico Museum of Hatural History and Science.

Flynn, J. J., Parish, J., M, Wyss, A., R, & Simpson, W., F. (1998). New Triassic and Jurassic vertebrates from Madagascar. *Journal of Vertebrate Paleontology*, *18*(3), 42A. https://doi.org/10.2307/4523942

Flynn, J. J., Parrish, J. M., Rakotosamimanana, B., Simpson, W. F., & Wyss, A. R. (1999). A Middle Jurassic mammal from Madagascar [10.1038/43420]. *Nature*, *401*(6748), 57-60. http://dx.doi.org/10.1038/43420

Foster, J. (2001). Taphonomy and palaeoecology of a microvertebrate assemblage from the Morrison Formation (Upper Jurassic) of Black Hills, Vrook COunty, Wyoming. *Geology Studies*, *46*, 13-33.

Foster, J. (2006). Microvertebrate sites in the Morrison Formation (Upper Jurassic) of the Western United States: Definition of taphonomic modes. *Journal of Vertebrate Paleontology*, *26*(3), 63A.

Foster, J. (2007). Jurassic West. *The dinosaurs of the Morrison Formation and their world*. Indiana University Press, pp 1-560.

Foster, J. R., & Heckert, A. B. (2011). Ichthyoliths and other microvertebrate remains from the Morrison Formation (Upper Jurassic) of northeastern Wyoming: A screen-washed sample indicates a significant aquatic component to the fauna. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *305*(1–4), 264-279. https://doi.org/http://dx.doi.org/10.1016/j.palaeo.2011.03.007

Foth, C., & Rauhut, O. W. M. (2017). Re-evaluation of the Haarlem Archaeopteryx and the radiation of maniraptoran theropod dinosaurs. *BMC Evolutionary Biology*, *17*(1), 236. https://doi.org/10.1186/s12862-017-1076-y

Fraser, D., & Theodor, J. M. (2011). Anterior dentary shape as an indicator of diet in ruminant artiodactyls. *Journal of Vertebrate Paleontology*, *31*(6), 1366-1375. https://doi.org/10.1080/02724634.2011.605404

Freeman, E. F. (1976a). Mammal teeth from the Forest Marble (Middle Jurassic) of Oxfordshire, England. *Science*, *194*(4269), 1053-1055.

Freeman, E. F. (1976b). 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.

Fricke, H. C., & Pearson, D. A. (2008). Stable isotope evidence for changes in dietary niche partitioning among hadrosaurian and ceratopsian dinosaurs of the Hell Creek Formation, North Dakota. *Paleobiology*, *34*(4), 534-552. https://doi.org/10.1666/08020.1

Fricke, H. C., & Rogers, R. R. (2000). Multiple taxon–multiple locality approach to providing oxygen isotope evidence for warm-blooded theropod dinosaurs. *Geology*, *28*(9), 799-802. http://geology.gsapubs.org/content/28/9/799.abstract

Fricke, H. C., Rogers, R. R., & Gates, T. A. (2009). Hadrosaurid migration: inferences based on stable isotope comparisons among Late Cretaceous dinosaur localities. *Paleobiology*, *35*(2), 270-288. https://doi.org/10.1666/08025.1

Galton, P. M. (1974). The ornithischian Dinosaur *Hypsilophodon* from the Wealden of the isle of Wight. *Bulletin of the British Museum (Natural History) Geology*, *25*(1), 1-152. https://doi.org/10.5962/p.313819

Galton, P. M. (1975). English hypsilophodontid dinosaurs (Reptilia: Ornithischia). *Palaeontology*, *18*, 741-752.

Galton, P. M. (1978). Fabrosauridae, the basal family of ornithischian dinosaurs (Reptilia: Ornithopoda). *Paläontologische Zeitschrift*, *52*(1), 138. https://doi.org/10.1007/BF03006735

Galton, P. M. (1980a). Armored dinosaurs (Ornithischia: Ankylosauria) from the Middle and Upper Jurassic of England. *Geobios*, *13*(6), 825-837. https://doi.org/https://doi.org/10.1016/S0016-6995(80)80038-6

Galton, P. M. (1980b). European Jurassic ornithopod dinosaurs of the families Hypsilophodontidae and Camptosauridae. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen*, *160*(1), 73-95.

Galton, P. M. (1980c). *Priodontognathus phillipsii* (Seeley), an ankylosaurian dinosaur from the upper Jurassic (or possibly lower Cretaceous) of England. *Neues Jahrbuch fur Geologie und Palaontologie, Monatschefte*, 477-489.

Galton, P. M. (1983a). Armoured dinosaurs (Ornithischia: Ankylosauria) from the Middle and Upper Jurassic of Europe. *Palaeontographica, Abteilung A*, *182*, 1-25.

Galton, P. M. (1983b). *Sarcolestes leedsi* Lydeeker, an ankylosaurian dinosaur from the Middle Jurassic of England. *Neues Jahrbuch fur Geologie und Palaontologie*, *1983*, 141-155.

Galton, P. M. (1985). British plated dinosaurs (Ornithischia, Stegosauridae). *Journal of Vertebrate Paleontology*, *5*, 211-254.

Galton, P. M. (2017). Purported earliest bones of a plated dinosaur (Ornithischia: Stegosauria): a "dermal tail spine" and a centrum from the Aalenian-Bajocian (Middle Jurassic) of England, with comments on other early thyreophorans. *Neues Jahrbuch fur Geologie und Palaontologie - Abhandlungen*, *285*(1), 1-10. https://doi.org/10.1127/njgpa/2017/0667

Galton, P. M., & Powell, H. P. (1983). Stegosaurian dinosaurs from the Bathonian (Middle Jurassic) of England, the earliest record of the family Stegosauridae. *Geobios*, *16*, 219-229.

Galton, P. M., & Upchurch, P. (2004). Stegosauria. In D. B. Weishampel, P. Dodson, & H. Osmolska (Eds.), *The Dinosauria* (2nd ed., pp. 343-362). University of California Press.

Gates, T. A., Zanno, L. E., & Makovicky, P. J. (2015). Theropod teeth from the upper Maastrichtian Hell Creek Formation “Sue” Quarry: New morphotypes and faunal comparisons. *Acta Palaeontologica Polonica*, *60*(1), 131-139. https://doi.org/http://dx.doi.org/10.4202/app.2012.0145

Gauthier, J. A. (1986). Saurischian monophyly and the origin of birds. In K. Padian (Ed.), *The Origin of Birds and the Evolution of Flight* (Vol. 8, pp. 1-55). Memoirs of the California Academy of Sciences.

Geiger, M., Clark, D. N., & Mette, W. (2004). Reappraisal of the timing of the breakup of Gondwana based on sedimentological and seismic evidence from the Morondava Basin, Madagascar. *Journal of African Earth Sciences*, *38*(4), 363-381. https://doi.org/https://doi.org/10.1016/j.jafrearsci.2004.02.003

Gerke, O., & 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*(7), e0158334. https://doi.org/10.1371/journal.pone.0158334

Gilbert, M. M., Bamforth, E. L., Buatois, L. A., & Renaut, R. W. (2018). Paleoecology and sedimentology of a vertebrate microfossil assemblage from the easternmost Dinosaur Park Formation (Late Cretaceous, Upper Campanian,) Saskatchewan, Canada: Reconstructing diversity in a coastal ecosystem. *Palaeogeography, Palaeoclimatology, Palaeoecology*. https://doi.org/https://doi.org/10.1016/j.palaeo.2018.01.016

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., Dong-Yu, H., Escuillie, F., Wenhao, W., & Dyke, G. (2013). A Jurassic avialan dinosaur from China resolves the early phylogenetic history of birds [Letter]. *Nature*, *498*, 359 – 362. https://doi.org/10.1038/nature12168

http://www.nature.com/nature/journal/vaop/ncurrent/abs/nature12168.html#supplementary-information

Godefroit, P., Demuynck, H., Dyke, G., Hu, D., Escuillié, F., & Claeys, P. (2013). Reduced plumage and flight ability of a new Jurassic paravian theropod from China [Article]. *Nature Communications*, *4*, 1394. https://doi.org/10.1038/ncomms2389

https://www.nature.com/articles/ncomms2389#supplementary-information

Goswami, A., Prasad, G. V. R., Verma, O., Flynn, J. J., & Benson, R. B. J. (2013). A troodontid dinosaur from the latest Cretaceous of India. *Nature Communications*, *4*, 1-5. https://doi.org/10.1038/ncomms2716

Govindan, A. (1975). Jurassic freshwater ostracods from the Kota limestone of India. *Palaeontology*, *18*(1), 207-216.

Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, *23*(1), 40-55. https://doi.org/10.1038/s41580-021-00407-0

Guthrie, R., Stukins, S., & Raub, T. (2014). The stratigraphy and palaeoenvironment of the Bathonian "Great Oolite Group" of Woodeaton Quarry, Oxfordshire. *Geophysical Research Abstracts*, *16, EGU2014*, 10476-10475.

Haddoumi, H., Allain, R., Meslouh, S., Metais, G., Monbaron, M., Pons, D., Rage, J.-C., Vullo, R., Zouhri, S., & Gheerbrant, E. (2016). Guelb el Ahmar (Bathonian, Anoual Syncline, eastern Morocco): First continental flora and fauna including mammals from the Middle Jurassic of Africa. *Gondwana Research*, *29*(1), 290-319. https://doi.org/http://dx.doi.org/10.1016/j.gr.2014.12.004

Halotel, J., Demyanov, V., & Gardiner, A. (2020). Value of Geologically Derived Features in Machine Learning Facies Classification. *Mathematical Geosciences*, *52*(1), 5-29. https://doi.org/10.1007/s11004-019-09838-0

Han, F., Clark, J. M., Xu, X., Sullivan, C., Choiniere, J., & Hone, D. W. E. (2011). Theropod Teeth from the Middle-Upper Jurassic Shishugou Formation of Northwest Xinjiang, China. *Journal of Vertebrate Paleontology*, *31*(1), 111-126. https://doi.org/10.1080/02724634.2011.546291

Hastie, T., & Tibshirani, R. (1996). Discriminant Analysis by Gaussian Mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 155-176. http://www.jstor.org/stable/2346171

Hastie, T., Tibshirani, R., & Friedman, J. (2009a). *The elements of statistical learning* (2 ed.). Springer-Verlag. https://doi.org/10.1007/978-0-387-84858-7

Hastie, T., Tibshirani, R., & Friedman, J. (2009b). Ensemble learning. In T. Hastie, R. Tibshirani, & J. Friedman (Eds.), *The elements of statistical learning* (2 ed., pp. 745). Springer-Verlag. https://doi.org/10.1007/978-0-387-84858-7

Hastie, T., Tibshirani, R., Leisch, F., Hornik, K., Ripley, B. D., & Narasimhan, B. (2020). mda: Mixture and Flexible Discriminant Analysis. R package version 0.5-2. https://CRAN.R-project.org/package=mda

Haubold, H. (1990). Ein neuer Dinosaurier (Ornithischia, Thyreophora) aus dem unteren Jura des nördlichen Mitteleuropa. *Revue de Paléobiologie*, *9*(1), 149-177.

Heckert, A. B. (2002). A revision of the Upper Triassic ornithischain dinosaur *Revueltosaurus,* with a description of a new species. In A. B. Heckert & S. G. Lucas (Eds.), *Upper Triassic Stratigraphy and Palaeontology. New Mexico Museum of Natural History and Science Bulletin* (Vol. 21, pp. 253-267).

Heckert, A. B. (2004). Late triassic microvertebrates from the Lower Chinle Group (Otischalkian-Adamanian:Carnian), southwestern USA. *New Mexico Museum of Natural History and Science*, *27*, 1-170.

Heckert, A. B., Crothers, J. P., Rose, L. J., Nesbitt, S. J., Pugh, I., Stocker, M., Barrett, P. M., Wills, S., Ward, D. J., Lauer, B., & Lauer, R. (2023). An exceptionally rich Upper Triassic microvertebrate site from the Early Revueltian (mid-Norian) Garita Creek Formation of east-central New Mexico, USA. In R. K. Hunt-Foster, J. I. Kirkland, & M. A. Loewen (Eds.), *The Anatomical Record. Special Issue: 14th Symposium on Mesozoic Terrestrial Ecosystems and Biota* (Vol. 306, pp. 129-131). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/ar.25219

Heckert, A. B., & Jenkins, H. S. (2005). The microvertebrate fauna of the Upper Triassic (Revueltian) Snyder Quarry, North-Central New Mexico. In *56th Field Conference Guidebook, Geology of the Chama Basin* (pp. 319-334). New Mexico Geological Society.

Heckert, A. B., Lucas, S. G., Sullivan, R. M., Hunt, A. P., & Spielmann, J. A. (2005). The vertebrate fauna of the Upper Triassic (Revueltian: Early-Mid Norian) Painted Desert Member (Petrified Forest Formation: Chinle Group) in the Chama Basin, northern New Mexico. In *56th Field Conference Guidebook, Geology of the Chama Basin* (pp. 302-318). New Mexico Geological Society.

Hendrickx, C., & 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), 1-74. https://doi.org/http://dx.doi.org/10.11646/zootaxa.3759.1.1

Hendrickx, C., Mateus, O., & Araújo, R. (2015a). The Dentition of Megalosaurid Theropods. *Acta Palaeontologica Polonica*, *60*(3), 627-642. https://doi.org/10.4202/app.00056.2013

Hendrickx, C., Mateus, O., & Araújo, R. (2015b). A proposed terminology of theropod teeth (Dinosauria, Saurischia). *Journal of Vertebrate Paleontology*, e982797. https://doi.org/10.1080/02724634.2015.982797

Hendrickx, C., Mateus, O., Araújo, R., & 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*, 1-110 https://doi.org/https://doi.org/10.26879/820

Hendrickx, C., Tschopp, E., & 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. https://doi.org/https://doi.org/10.1016/j.cretres.2019.104312

Hervé, M. (2021). RVAideMemoire: Testing and Plotting Procedures for Biostatistics. *R package version 0.9-80*. https://CRAN.R-project.org/package=RVAideMemoire

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.

Hoffman, D. K., Edwards, H. R., Barrett, P. M., & Nesbitt, S. J. (2019). Reconstructing the archosaur radiation using a Middle Triassic archosauriform tooth assemblage from Tanzania. *PeerJ*, *7*, e7970. https://doi.org/10.7717/peerj.7970

Holden Bolin, J., & Finch, H. (2014). Supervised classification in the presence of misclassified training data: a Monte Carlo simulation study in the three group case [10.3389/fpsyg.2014.00118]. *Frontiers in Psychology*, *5*, 118. https://www.frontiersin.org/article/10.3389/fpsyg.2014.00118

Holden, J. E., Finch, W. H., & Kelley, K. (2011). A Comparison of Two-Group Classification Methods. *Educational and Psychological Measurement*, *71*(5), 870-901. https://doi.org/10.1177/0013164411398357

Holloway, S. (1983). The shell-detrital calcirudites of the Forest Marble Formation (Bathonian) of southwest England. *Proceedings of the Geologists' Association*, *94*(3), 259-266. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(83)80044-3

Holtz, T. R. (2000). A new phylogeny of the carnivorous dinosaurs. *Gaia*, *15*, 5-61.

Holtz, T. R., Brinkman, D. L., & Chandler, C. L. (1998). Denticle morphometrics and a possible omniverous feeding habit for the theropod dinosaur *Troodon*. *Gaia*, *15*, 159-166.

Hooker, J. J., Insole, A. N., Moody, R. T. J., Walker, C. A., & Ward, D. J. (1980). The distribution of cartilaginous fish, turtles, birds and mammals in the British Palaeogene. *Tertiary Research*, *3*, 1-45.

Horton, A., Sumbler, M. G., Cox, B. M., & Ambrose, K. (1995). Geology of the country around Thame, *Memoir of the British Geological Survey, Sheet 237 (England and Wales)*.

House, M. R. (1989). *Geology of the Dorset Coast*. Geologists' Association Guide, 1-162

Hoyal Cuthill, J. F., Guttenberg, N., Ledger, S., Crowther, R., & Huertas, B. (2019). Deep learning on butterfly phenotypes tests evolution's oldest mathematical model. *Science Advances*, *5*(8), eaaw4967. https://doi.org/10.1126/sciadv.aaw4967

Hu, D., Hou, L., Zhang, L., & Xu, X. (2009). A pre-Archaeopteryx troodontid theropod from China with long feathers on the metatarsus. *Nature*, *461*(7264), 640-643. https://doi.org/10.1038/nature08322

Hu, J., Forster, C. A., Xu, X., Zhao, Q., He, Y., & Han, F. (2022). Computed tomographic analysis of the dental system of three Jurassic ceratopsians and implications for the evolution of tooth replacement pattern and diet in early-diverging ceratopsians. *eLife*, *11*, e76676. https://doi.org/10.7554/eLife.76676

Hudgins, M. N., Currie, P. J., & Sullivan, C. (2022). Dental assessment of *Stegoceras validum* (Ornithischia: Pachycephalosauridae) and *Thescelosaurus neglectus* (Ornithischia: Thescelosauridae): paleoecological inferences. *Cretaceous Research*, *130*, 105058. https://doi.org/10.1016/j.cretres.2021.105058

Huene, F. v. (1932). Die fossile Reptil-Ordnung Saurischia, ihre Entwicklung und Geschichte. *Monographien zur Geologie und Palaontologie, Series 1*, *4*, 1-361.

Hui, D., Ning, L., Maidment, S. C. R., Guangbiao, W., Yuxuan, Z., Xufeng, H., Qingyu, M., Xunqian, W., Haiqian, H., & Guangzhao, P. (2022). New stegosaurs from the Middle Jurassic Lower Member of the Shaximiao Formation of Chongqing, China. *Journal of Vertebrate Paleontology*, *41*(5). https://doi.org/10.1080/02724634.2021.1995737

Hull, E. (1857). The geology of the country around Cheltenham*, Memoirs of the Geological Survey of Great Britain. Old series sheet 44*.

Hull, E. (1859). The geology of the country around Woodstock, Oxfordshire*. Memoirs of the Geological Survey of Great Britain. Old series sheet 45 SW.*

Hunt, A., & Lucas, S. (1994). Ornithischian dinosaurs from the Upper Triassic of the United States. In N. C. Fraser & H.-D. Sues (Eds.), *In the shadow of the dinosaurs. Early Mesozoic tetrapods* (pp. 435). Cambridge University Press.

Hunter, A. W., & Underwood, C. J. (2009). Palaeoenvironmental control on distribution of crinoids in the Bathonian (Middle Jurassic) of England and France. *Acta Palaeontologica Polonica*, *54*(1), 77-98.

Huxley, T. H. (1868). Remarks upon *Archaeopteryx lithographica*. *Proceedings of the Royal Society*, *98*, 1-5.

Irmis, R. B., Parker, W. G., Nesbitt, S. J., & Liu, J. (2007). Early ornithischian dinosaurs: the Triassic record. *Historical Biology*, *19*(1), 3-22. https://doi.org/10.1080/08912960600719988

Isasmendi, E., Torices, A., Canudo, J. I., Currie, P. J., & Pereda-Suberbiola, X. (2022). Upper Cretaceous European theropod palaeobiodiversity, palaeobiogeography and the intra-Maastrichtian faunal turnover: new contributions from the Iberian fossil site of Laño. *Papers in Palaeontology*. https://doi.org/10.1002/spp2.1419

Jackson, F. D., & Varricchio, D. J. (2016). Fossil egg and eggshells from the Upper Cretaceous Hell Creek Formation, Montana. *Journal of Vertebrate Paleontology*, e1185432. https://doi.org/10.1080/02724634.2016.1185432

Jain, S. L. (1980). The continental Lower Jurassic fauna from Kota Formation, India. In L. L. Jacobs (Ed.), *Aspects of vertebrate history* (pp. 99-123). Museum of Northern Arizona Press.

Joeckel, R. M., Ludvigson, G. A., Möller, A., Hotton, C. L., Suarez, M. B., Suarez, C. A., Sames, B., Kirkland, J. I., & Hendrix, B. (2020). Chronostratigraphy and terrestrial palaeoclimatology of Berriasian–Hauterivian strata of the Cedar Mountain Formation, Utah, USA. *Geological Society, London, Special Publications*, *498*(1), 75-100. https://doi.org/10.1144/sp498-2018-133

Jolivet, M., Boulvais, P., Barrier, L., Robin, C., Heilbronn, G., Ledoyen, J., Ventroux, Q., Jia, Y., Guo, Z., & Bataleva, E. A. (2018). Oxygen and Carbon Stable Isotope Composition of Cretaceous to Pliocene Calcareous Paleosols in the Tian Shan Region (Central Asia): Controlling Factors and Paleogeographic Implications. *Geosciences*, *8*(9).

Jolliffe, I. T. (2002). *Principle Component Analysis* (2nd ed.). Springer, pp 1-408.

Jones, M. E. H., Benson, R. B., Skutschas, P., Hill, L., Panciroli, E., Schmitt, A. D., Walsh, S., A, & Evans, S. E. (2022). Middle Jurassic fossils document an early stage in salamander evolution. *Proceedings of the National Academy of Sciences*, *119*(30), e2114100119. https://doi.org/10.1073/pnas.2114100119

Judd, J. W. (1875). The geology of Rutland and parts of Lincoln, Leicester, Northampton, Huntingdon and Cambridge*. Memoirs of the Geological Survey of Great Britain. Old series sheet 64*.

Keeping, H. (1910). On the discovery of Bembridge Limestone fossils on Creechbarrow Hill, Isle of Purbeck. *Geological Magazine, Decade V*, *7*, 436-439.

Kellaway, G. A., & Wilson, V. (1941). An outline of the geology of Yeovil, Sherborne and Sparkford Vale. *Proceedings of the Geologists' Association*, *52*, 131-174.

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, *6*, 9375-9389. https://doi.org/10.1109/ACCESS.2017.2788044

Kermack, D. M., Kermack, K. A., & Mussett, F. (1968). The Welsh Pantothere *Kuehneotherium praecursoris*. *Zoological Journal of the Linnean Society*, *53*, 87-175.

Kermack, K. A. (1988). British Mesozoic mammal sites. *Special Papers in Palaeontology*, *68*, 85-93.

Kermack, K. A., Kermack, D. M., Lees, P. M., & Mills, J. R. E. (1998). New multituberculate-like teeth from the Middle Jurassic of England. *Acta Palaeontologica Polonica*, *43*(4), 581-606.

Kermack, K. A., Lee, A. J., Lees, P. M., & Mussett, F. (1987). A new docodont from the Forest Marble. *Zoological Journal of the Linnean Society*, *89*, 1-39.

Kingsley, C. (1857, 26th December). Geological Discoveries at Swanage. *Illustrated London News*, (895).

Kirkland, J. I., & Wolfe, D. G. (2001). First definitive therizinosaurid (Dinosauria; Theropoda) from North America. *Journal of Vertebrate Paleontology*, *21*(3), 410-414. https://doi.org/10.1671/0272-4634(2001)021[0410:FDTDTF]2.0.CO;2

Kirkland, J. I., Zanno, L. E., Sampson, S. D., Clark, J. M., & DeBlieux, D. D. (2005). A primitive therizinosauroid dinsoaur from the Early Cretaceous of Utah. *Nature*, *435*(3468), 84-87.

Knoll, F., & Ruiz-Omeñaca, J. I. (2009). Theropod teeth from the basalmost Cretaceous of Anoual (Morocco) and their palaeobiogeographical significance. *Geological Magazine*, *146*(4), 602-616. https://doi.org/10.1017/s0016756809005950

Kocsis, Á. T., & Raja, N. B. (2019). chronosphere: Earth system history variables*. R package version 0.5.0.*

Kocsis, Á. T., & Raja, N. B. (2021). rgplates: R interface for the GPlates web service and desktop application. *R Package version 0.3.0.*

Kocsis, Á. T., Reddin, C. J., Alroy, J., & Kiessling, W. (2019). The R package divDyn for quantifying diversity dynamics using fossil sampling data. *Methods in Ecology and Evolution*, *10*(5), 735-743. https://doi.org/10.1111/2041-210x.13161

Kohn, M. J., Schoeninger, M. J., & Barker, W. W. (1999). Altered states: Effects of disgenesis on fossil tooth chemistry. *Geochimica et Cosmochimica Acta*, *63*(18), 2737-2747.

Kriwet, J., Rauhut, O. W. M., & Gloy, U. (1997). Microvertebrate remains (Pisces, Archosauria) from the Middle Jurassic (Bathonian) of southern France. *Neues Jahrbuch fur Geologie und Palaontologie - Abhandlungen*, *206*(1), 1-28.

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, *1*(5), 1-26. https://doi.org/10.18637/jss.v028.i05

Kuhn, M., & Johnson, K. (2013a). *Applied predictive modelling*. Springer-Verlag, pp 1-600, https://doi.org/10.1007/978-1-4614-6849-3

Kuhn, M., & Johnson, K. (2013b). Classification trees and rule-based models. In M. Kuhn & K. Johnson (Eds.), *Applied predictive modelling* (pp. 369-413). Springer-Verlag. https://doi.org/10.1007/978-1-4614-6849-3

Kuhn, M., & Johnson, K. (2013c). Measuring performance in classificaiton models. In M. Kuhn & K. Johnson (Eds.), *Applied predictive modelling* (pp. 247-273). Springer-Verlag. https://doi.org/10.1007/978-1-4614-6849-3

Kuhn, M., Weston, S., Culp, M., Coulter, N., & Quinlan, R. (2018). *C5.0 decision trees and rule-based model*. RuleQuest Research Pty Ltd.

Kuhn, O. (1966). *Die Reptilien, System und Stammesgeschichte*. Verlag Oeben, Krailling bei München.

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*(12), 1455-1468. http://www.ingentaconnect.com/content/nrc/cjes/2008/00000045/00000012/art00004

Larson, Derek W., Brown, Caleb M., & Evans, David C. (2016). Dental Disparity and Ecological Stability in Bird-like Dinosaurs prior to the End-Cretaceous Mass Extinction. *Current Biology*, *26*(10), 1325-1333. https://doi.org/http://dx.doi.org/10.1016/j.cub.2016.03.039

Larson, D. W., & Currie, P. J. (2013). Multivariate Analyses of Small Theropod Dinosaur Teeth and Implications for Paleoecological Turnover through Time. *PLoS ONE*, *8*(1), e54329. https://doi.org/10.1371/journal.pone.0054329

Lasseron, M. (2019). Enigmatic teeth from the Jurassic–Cretaceous transition of Morocco: The latest known non-mammaliaform cynodonts (Synapsida, Cynodontia) from Africa? *Comptes Rendus Palevol*, *18*(7), 897-907. https://doi.org/https://doi.org/10.1016/j.crpv.2019.05.002

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., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18-22.

Lindgren, J., Currie, P. J., Rees, J., Siverson, M., Lindström, S., & Alwmark, C. (2008). Theropod dinosaur teeth from the lowermost Cretaceous Rabekke Formation on Bornholm, Denmark. *Geobios*, *41*(2), 253-262. https://doi.org/http://dx.doi.org/10.1016/j.geobios.2007.05.001

Longrich, N. R. (2008). Small theropod teeth from the Lance Formation of Wyoming, USA. In J. T. Sankey & S. Baszio (Eds.), *Vertebrate Microfossil Assemblages: Their Role in Paleoecology and Paleobiogeography* (pp. 135-158). Indiana University Press.

Lydekker, R. (1889). On the remains and affinities of five genera of Mesozoic reptiles. *Quarterly Journal of the Geological Society of London*, *45*, 41-59.

Lydekker, R. (1893). On the jaw of a new carnivorous dinosaur from the Oxford Clay of Peterborough. *Quarterly Journal of the Geological Society of London*, *49*, 284-287.

Lydekker, R. (1893). On Two Dinosaurian Teeth from Aylesbury. *Quarterly Journal of the Geological Society*, *49*(1-4), 566-568. https://doi.org/10.1144/gsl.jgs.1893.049.01-04.64

Ma, X., & Lv, S. (2019). Financial credit risk prediction in internet finance driven by machine learning. *Neural Computing and Applications*, *31*(12), 8359-8367. https://doi.org/10.1007/s00521-018-3963-6

MacLeod, N. (2007). Automated taxon identification in systematics : theory, approaches and applications (Vol. 74). CRC Press, pp 1-368.

MacLeod, N. (2017). On the Use of Machine Learning Methods in Morphometric Analysis. Biological Shape Analysis, Proceedings of the 4th International Symposium on Biological Shape Analysis (ISBSA), School of Dentistry, UCLA, USA, 19 – 22 June 2015.

MacLeod, N. (2018). The quantitative assessment of archaeological artifact groups: Beyond geometric morphometrics. *Quaternary Science Reviews*, *201*, 319-348. https://doi.org/https://doi.org/10.1016/j.quascirev.2018.08.024

MacLeod, N. (2019). Artificial intelligence and machine learning in the earth sciences. *Acta Geologica Sinica (English Edition)*, *93*, 48-51.

Macleod, N., Canty, R. J., & Polaszek, A. (2021). Morphology-based identification of Bemisia tabaci cryptic species puparia via embedded group-contrast convolution neural network analysis. *Systematic Biology*. https://doi.org/10.1093/sysbio/syab098

MacLeod, N., & 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*(1), 113. https://doi.org/10.1186/s12915-020-00832-1

MacLeod, N., O'Neill, M. A., & Walsh, S. A. (2008). Automated tools for the identification of taxa from morphological data: Face recognition in wasps. In N. MacLeod (Ed.), *Automated taxon identification in systematics:Theory, approaches and applications* (pp. 339). CRC Press.

MacLeod, N., & Steart, D. (2015). Automated leaf physiognomic character identification from digital images. *Paleobiology*, *41*(4), 528-553. https://doi.org/10.1017/pab.2015.13

Maganuco, S. M., Cau, A., & Pasini, G. (2005). First description of theropod remains from the Middle Jurassic (Bathonian) of Madagascar. *Atti della Società Italiana di Scienze naturali e del Museo Civico di Storia naturale di Milano*, *146*(2), 165-202.

Maidment, S. C. R., Kirkpatrick, C., Craik-Smith, B., & Blythe, J. E. (2017). A new ornithischian dinosaur and the terrestrial vertebrate fauna from a bone bed in the Wealden of Ardingly, West Sussex. *Proceedings of the Geologists' Association*, *128*(3), 332-339. https://doi.org/https://doi.org/10.1016/j.pgeola.2017.03.006

Maidment, S. C. R., Norman, D. B., Barrett, P. M., & Upchurch, P. (2008). Systematics and phylogeny of Stegosauria (Dinosauria: Ornithischia). *Journal of Systematic Palaeontology*, *6*(4), 367-407. https://doi.org/10.1017/S1477201908002459

Maidment, S. C. R., Raven, T. J., Ouarhache, D., & Barrett, P. M. (2020). North Africa's first stegosaur: Implications for Gondwanan thyreophoran dinosaur diversity. *Gondwana Research*, *77*, 82-97. https://doi.org/https://doi.org/10.1016/j.gr.2019.07.007

Maidment, S. C. R., Strachan, S. J., Ouarhache, D., Scheyer, T. M., Brown, E. E., Fernandez, V., Johanson, Z., Raven, T. J., & Barrett, P. M. (2021). Bizarre dermal armour suggests the first African ankylosaur. *Nature Ecology & Evolution*, *5*(12), 1576-1581. https://doi.org/10.1038/s41559-021-01553-6

Maidment, S. R. (2010). Stegosauria: a historical review of the body fossil record and phylogenetic relationships. *Swiss Journal of Geosciences*, *103*(2), 199-210. https://doi.org/10.1007/s00015-010-0023-3

Maisch, M. W., & Matzke, A. T. (2003). Theropods (dinosauria, saurischia) from the middle Jurassic Toutunhe Formation of the Southern Junggar Basin, NW China. *Paläontologische Zeitschrift*, *77*(2), 281-292. https://doi.org/10.1007/BF03006942

Maisch, M. W., Matzke, A. T., Ye, J., Prfretzschner, H.-U., Sun, G., Stöhr, H., & Grossmann, F. (2003). Fossil vertebrates from the Middle and Upper Jurassic of the Southern Junggar

basin (NW China), results of the Sino-german expeditions 1999-2000. *Neues Jahrbuch fur Geologie und Palaontologie*, *Monatshefte*, 297-313.

Makovicky, P. J., Norell, M. A., Clark, J. M., & Rowe, T. (2003). Osteology and relationships of *Byronosaurus jaffei* (Theropoda: Troodontidae). *American Museum Novitates*, *3402*, 1.

Maleev, E. A. (1954). Noviy cherepachoobrazhniy yashcher v Mongolii [New tortoise-like saurian from Mongolia]. *Priroda*, *1954*(3), 106-108.

Mallon, J. C., & Anderson, J. S. (2014). The Functional and Palaeoecological Implications of Tooth Morphology and Wear for the Megaherbivorous Dinosaurs from the Dinosaur Park Formation (Upper Campanian) of Alberta, Canada. *PLoS ONE*, *9*(6), e98605. https://doi.org/10.1371/journal.pone.0098605

Marramà, G., & Kriwet, J. (2017). Principal component and discriminant analyses as powerful tools to support taxonomic identification and their use for functional and phylogenetic signal detection of isolated fossil shark teeth. *PLoS ONE*, *12*(11), e0188806. https://doi.org/10.1371/journal.pone.0188806

Marsh, O. C. (1877). Notice of new dinosaurian reptiles from the Jurassic formation. *American Journal of Science and Arts*, *14*, 514-516.

Marsh, O. C. (1881). Principle characters of American Jurassic dinosaurs. Part V. *American Journal of Science, Series 3*, *21*, 417-423.

Marshall, C. R., Finnegan, S., Clites, E. C., Holroyd, P. A., Bonuso, N., Cortez, C., Davis, E., Dietl, G. P., Druckenmiller, P. S., Eng, R. C., Garcia, C., Estes-Smargiassi, K., Hendy, A., Hollis, K. A., Little, H., Nesbitt, E. A., Roopnarine, P., Skibinski, L., Vendetti, J., & White, L. D. (2018). Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital revolution. *Biology Letters*, *14*(9), 20180431. https://doi.org/10.1098/rsbl.2018.0431

Marsland, S. (2015). *Machine learning: An algorithmic perspective*. CRC Press, pp 1-452.

Martin, T., & Averianov, A. O. (2004). A new Docodont (Mammalia) from the Middle Jurassic of Kyrgyzstan, Central Asia. *Journal of Vertebrate Paleontology*, *24*(1), 195-201. https://doi.org/10.1671/15

Matthew, W. D., & 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*(6), 367-385.

Maugis, C., Celeux, G., & Martin-Magniette, M. L. (2011). Variable selection in model-based discriminant analysis. *Journal of Multivariate Analysis*, *102*(10), 1374-1387. https://doi.org/https://doi.org/10.1016/j.jmva.2011.05.004

Mazin, J. M., Billon-Bruyat, J.-P., Pouech, J., & Hantzpergue, P. (2006). The Purbeckian site of Cherves-De-Cognac (Berrisian, Early Cretaceous, Southwest France): A continental ecosystem accumulated in an evaporitic littoral depositional environment. In P. Barrett & S. Evans (Eds.), *Ninth international symposium on Mesozoic terrestrial ecosystems and biota, Manchester, UK* (pp. 187). Natural Histroy Museum.

McKerrow, W. S., Johnson, R. T., & Jakobson, M. E. (1969). Palaeoecological studies in the Great Oolite at Kirtlington, Oxfordshire. *Palaeontology*, *12*, 56-83.

Mears, E. M., Rossi, V., MacDonald, E., Coleman, G., Davies, T. G., Arias-Riesgo, C., Hildebrandt, C., Thiel, H., Duffin, C. J., Whiteside, D. I., & Benton, M. J. (2016). The Rhaetian (Late Triassic) vertebrates of Hampstead Farm Quarry, Gloucestershire, UK. *Proceedings of the Geologists' Association*, *127*(4), 478-505. https://doi.org/http://dx.doi.org/10.1016/j.pgeola.2016.05.003

Mehta, S., & Shukla, D. (2015, 10-12 Sept. 2015). Optimization of C5.0 classifier using Bayesian theory. 2015 International Conference on Computer, Communication and Control (IC4),

Melstrom, K. M., & Irmis, R. B. (2019). Repeated Evolution of Herbivorous Crocodyliforms during the Age of Dinosaurs. *Current Biology*, *29*(14), 2389-2395.e2383. https://doi.org/https://doi.org/10.1016/j.cub.2019.05.076

Melville, R. V., & Freshney, E. C. (1982). *The Hampshire Basin and Adjoining Areas*. HMSO.

Metcalf, S. J. (1995). The palaeoenvironment and palaeoecology of a middle Jurassic vertebrate-bearing fen-type paleosol in a coastal carbonate regime. *PhD thesis, University of Bristol, Bristol.*

Metcalf, S. J., Vaughan, R. F., Benton, M. J., Cole, J., Simms, M. J., & 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.

Metcalf, S. J., & Walker, R. J. (1994). A new bathonian microvertebrate locality in the English Midlands. In N. C. Fraser & H.-D. Sues (Eds.), *In the shadow of the dinosaurs. Early Mesozoic Tetrapods* (pp. 322-331). Cambridge University Press.

Meyer, H. v. (1861). [Die Feder von Solenhofen] [The feather of Solnhofen]. *Neues Jahrbuch für Mineralogie, Geognosie, Geologie und Petrefakten-Kunde*, *1861*, 678-679.

Miall, A. D. (1997). *The Geology of Stratigraphic Sequences*. Springer-Verlag Berlin. https://doi.org/https://doi.org/10.1007/978-3-662-03380-7

Milla Carmona, P. S., Lazo, D. G., & Soto, I. M. (2016). Giving taxonomic significance to morphological variability in the bivalve Ptychomya Agassiz. *Palaeontology*, *59*(1), 139-154. https://doi.org/10.1111/pala.12215

Mitchell, T. M. (2006). *The discipline of machine learning*. http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf

Monson, T. A., Armitage, D. W., & Hlusko, L. J. (2018). Using machine learning to classify extant apes and interpret the dental morphology of the chimpanzee-human last common ancestor. *PaleoBios*, *35*, 1-20.

Moore, C. (1867). On Abnormal Conditions of Secondary Deposits when connected with the Somersetshire and South Wales Coal-Basin; and on the age of the Sutton and Southerndown Series. *Quarterly Journal of the Geological Society*, *23*(1-2), 449-568. https://doi.org/10.1144/gsl.jgs.1867.023.01-02.63

Mousavi, S. M., Zhu, W., Sheng, Y., & Beroza, G. C. (2019). CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection. *Scientific Reports*, *9*(1), 10267. https://doi.org/10.1038/s41598-019-45748-1

Murchison, R. I. (1834). *Outline of the geology of the neighbourhood of Cheltenham*. Davies, pp 1-40.

Naish, D., & Martill, D. M. (2008). Dinosaurs of Great Britain and the role of the Geological Society of London in their discovery: Ornithischia. *Journal of the Geological Society*, *165*(3), 613. https://doi.org/10.1144/0016-76492007-154

Nesbitt, S. J. (2011). The early evolution of archosaurs : relationships and the origin of major clades. *Bulletin of the American Museum of Natural History*, *352*, 1-292.

Nesbitt, S. J., Irmis, R. B., & Parker, W. G. (2007). A critical re-evaluation of the Late Triassic dinosaur taxa of North America. *Journal of Systematic Palaeontology*, *5*(2), 209-243. https://doi.org/10.1017/S1477201907002040

Nessov, L. A., Kielan-Jaworowska, Z., Hurun, J. H., Averianov, A. O., Fedorov, P. V., Potapov, D. O., & Froyland, M. (1994). First Jurassic mammals from Kyrgyzstan. *Acta Palaeontologica Polonica*, *39*(3), 315-326.

Nopcsa, F. (1911). Notes on British dinosaurs IV: *Stegosaurus priscus* sp. nov. *Geological Magazine*, *9*, 481-484.

Nopcsa, F. (1915). Die dinosaurier der Siebenbürgischen landesteile Ungarns. *Mitteilungen aus dem Jahrbuche der Königlich Ungarischen Geologischen Anstalt*, *23*, 1-24.

Norell, M., Clark, J. M., Turner, A. H., Makovicky, P. J., Barsbold, R., & 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., & Rowe, T. (2009). A review of the Mongolian Cretaceous dinosaur *Saurornithoides* (Troodontidae: Theropoda). *American Museum Novitates*, *3654*, 1-63.

Norman, D. B., & Barrett, P. M. (2002). Ornithischian dinosaurs from the Lower Cretaceous (Berriasian) of England. *Special Papers in Palaeontology*, *68*, 161-189.

Noto, C. R., D'Amore, D. C., Drumheller, S. K., & 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*. https://doi.org/https://doi.org/10.7717/peerj.12782

Odling, M. (1913). The Bathonian rocks of the Oxford district. *Quarterly Journal of the Geological Society of London*, *69*, 484-513.

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., & Wagner, H. (2020). vegan: Community Ecology Package. *R package version 2.5-7*. https://CRAN.R-project.org/package=vegan

Onojeghuo, A. O., Blackburn, G. A., Wang, Q., Atkinson, P. M., Kindred, D., & Miao, Y. (2018). Mapping paddy rice fields by applying machine learning algorithms to multi-temporal Sentinel-1A and Landsat data. *International Journal of Remote Sensing*, *39*(4), 1042-1067. https://doi.org/10.1080/01431161.2017.1395969

Oreska, M. P. J., Carrano, M. T., & Dzikiewicz, K. M. (2013). Vertebrate paleontology of the Cloverly Formation (Lower Cretaceous), I: faunal composition, biogeographic relationships, and sampling. *Journal of Vertebrate Paleontology*, *33*(2), 264-292. https://doi.org/10.1080/02724634.2012.717567

Ortega, F., Bardet, N., Barroso-Barcenilla, F., Callapez, P. M., Cambra-Moo, O., Daviero- Gómez, V., Díez Díaz, V., Domingo, L., Elvira, A., Escaso, F., García-Oliva, M., Gómez, B., Houssaye, A., Knoll, F., Marcos-Fernández, F., Martín, M., Mocho, P., Narváez, I., Pérez-García, A., . . . Sanz, J. L. (2015). *The biota of the Upper Cretaceous site of "Lo Hueco" (Cuenca, Spain)* (Vol. 41) [fossil vertebrates; biodiversity; Campanian-Maastrichtian, Cretaceous; Cuenca; Spain]. http://revistas.ucm.es/index.php/JIGE/article/view/48657

Osborn, H. F. (1923). Two Lower Cretaceous dinosaurs from Mongolia. *American Museum Notivates*, *95*, 1-10.

Osborn, H. F. (1924). Three new Theropoda, Protoceratops zone, central Mongolia. *American Museum Novitates*, *144*, 1-12.

Ősi, A., Barrett, P. M., Földes, T., & Tokai, R. (2014). Wear Pattern, Dental Function, and Jaw Mechanism in the Late Cretaceous Ankylosaur *Hungarosaurus*. *The Anatomical Record*, *297*(7), 1165-1180. https://doi.org/10.1002/ar.22910

Ősi, A., Prondvai, E., Mallon, J., & Bodor, E. R. (2016). Diversity and convergences in the evolution of feeding adaptations in ankylosaurs (Dinosauria: Ornithischia). *Historical Biology*, *29*(4), 539-570. https://doi.org/10.1080/08912963.2016.1208194

Ostrom, J. H., & Wellnhofer, P. (1990). *Tricerotops*: an example of flawed systematics. In K. Carpenter & P. J. Currie (Eds.), *Dinosaur Systematics Approaches and Perspectives* (pp. 245-254). Cambridge University Press.

Owen, R. (1838). A paper on the "Phascolotherium", being the second part of the "Description of the Remains of Marsupial Mammalia from the Stonesfield Slate. *Proceedings of the Geological Society of London*, *3*, 17-21.

Owen, R. (1840). Odontography; or, A treatise on the comparative anatomy of the teeth, their physiological relations, mode of development, and microscopic structure, in the vertebrate animals. *Hippolyte Bailliere, London,* pp 1-760.

Owen, R. (1842). Report on British fossil reptiles, part II. *Report of the British Association for the Advancement of Science*, *11*, 60-204. http://archive.org/stream/cbarchive_108440_1842reportonbritishfossilrepti9999/1842reportonbritishfossilrepti9999#page/n43/mode/2up

Owen, R. (1854). On some fossil reptilian and mammalian remains from the Purbecks. *Quarterly Journal of the Geological Society*, *10*, 420-423.

Owen, R. (1855). Monograph on the fossil reptilia of the Wealden and Purbeck Formations, Part II, Dinosauria. *Monograph of the Palaeontographical Society, London*, *8*(27), 1-54.

Owen, R. (1861). Monograph on the fossil reptilia of the Wealden and Purbeck Formations, Part V, Lacertilia. *Monograph of the Palaeontographical Society, London*, *12*(53), 31-39.

Owen, R. (1863). Monograph on the British fossil Reptilia from the Oolitic Formations. Part Two. *Sceildosaurus harrisonii*. *Palaeontographical Society Monographs*, *14*(Number 60), 1-26.

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 of London*, *153*, 33-47.

Owen, R. (1879). Monograph on the fossil Reptilia of the Wealden and Purbeck formations: supplement IX. *Palaeontographical Society Monographs*. https://www.biodiversitylibrary.org/item/222785

Owen, R. (1883). On the skull of *Megalosaurus*. *Quarterly Journal of the Geological Society of London*, *39*, 334-347.

Padian, K. (1990). The Ornithischian Form Genus *Revueltosaurus* from the Petrified Forest of Arizona (Late Triassic: Norian; Chinle Formation). *Journal of Vertebrate Paleontology*, *10*(2), 268-269. https://doi.org/10.2307/4523322

Palmer, T. J. (1973). Field meeting in the Great Oolite of Oxfordshire: 5–7 May 1972. *Proceedings of the Geologists' Association*, *84*(1), 53-64. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(73)80006-9

Palmer, T. J. (1974). Some palaeoecological studies of the Middle and Upper Bathonian of Central England and Northern France. *Unpublished PhD thesis* 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., & Jenkyns, H. C. (1975). A carbonate island barrier in the Great Oolite (Middle Jurassic) of Central England. *Sedimentology*, *22*, 125-135.

Panciroli, E., Benson, R. B. J., Walsh, S., Butler, R. J., Castro, T. A., Jones, M. E. H., & Evans, S. E. (2020). Diverse vertebrate assemblage of the Kilmaluag Formation (Bathonian, Middle Jurassic) of Skye, Scotland. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 1-22. https://doi.org/10.1017/S1755691020000055

Panciroli, E., Funston, G. F., Holwerda, F., Maidment, S. C. R., Foffa, D., Larkin, N., Challands, T., Depolo, P. E., Goldberg, D., Humpage, M., Ross, D., Wilkinson, M., & Brusatte, S. L. (2020). First dinosaur from the Isle of Eigg (Valtos Sandstone Formation, Middle Jurassic), Scotland. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *111*(3), 157-172. https://doi.org/10.1017/s1755691020000080

Parker, W. G., Irmis, R. B., Nesbitt, S. J., Martz, J. W., & Browne, L. S. (2005). The Late Triassic pseudosuchian *Revueltosaurus callenderi* and its implications for the diversity of early ornithischian dinosaurs. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1566), 963-969. https://doi.org/10.1098/rspb.2004.3047

Pei, R., Li, Q., Meng, Q., Norell, M. A., & 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*, 66.

Peng, J., Russell, A. P., & Brinkman, D. B. (2001). Vertebrate microsite assemblages (exclusive of mammals) from the Foremost and Oldman Formations of the Judith River Group

(Campanian) of southeastern Alberta: an illustrated guide. *Provincial Museum of Alberta, Natural History Occasional Paper*, *25*.

Penn, I. E. (1982). Middle Jurassic stratigraphy and correlation of the Winterborne Kingston Borehole, Dorset. In G. H. Rhys, G. K. Lott, & M. A. Calver (Eds.), *The Winterborne Kingston Borehole, Dorset, England. Report of the Institute of Geological Sciences, No.81/3.* (pp. 53-76). Institue of Geological Sciences.

Peterson, J. E., Coenen, J. J., & Noto, C. R. (2014). Fluvial transport potential of shed and root-bearing dinosaur teeth from the late Jurassic Morrison Formation. *PeerJ*, *2*, e347. https://doi.org/10.7717/peerj.347

Phillips, J. (1871). Geology of Oxford and the valley of the Thames. *Clarendon Press,* pp 1-598.

Pol, D., Ramezani, J., Gomez, K., Carballido, J. L., Carabajal, A. P., Rauhut, O. W. M., Escapa, I. H., & Cúneo, N. R. (2020). Extinction of herbivorous dinosaurs linked to Early Jurassic global warming event. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1939), 20202310. https://doi.org/10.1098/rspb.2020.2310

Pol, D., Rauhut, O. W. M., & Becerra, M. G. (2011). A Middle Jurassic heterodontosaurid dinosaur from Patagonia and the evolution of the heterodontosaurids. *Naturwissenschaften.*, *98*(5), 369-379.

Poole, K. (2022). Phylogeny of iguanodontian dinosaurs and the evolution of quadrupedality. *Palaeontologia Electronica*. https://doi.org/10.26879/702

Porro, L. B., Butler, R. J., Barrett, P. M., Moore-Fay, S., & Abel, R. L. (2011). New heterodontosaurid specimens from the Lower Jurassic of southern Africa and the early ornithischian dinosaur radiation. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *101*(3-4), 351-366. https://doi.org/10.1017/S175569101102010X

Porro, L. B., Witmer, L. M., & Barrett, P. M. (2015). Digital preparation and osteology of the skull of *Lesothosaurus diagnosticus* (Ornithischia: Dinosauria). *PeerJ*, *3*, e1494. https://doi.org/10.7717/peerj.1494

Pouech, J., Mazin, J. M., & Billon-Bruyat, J.-P. (2006). Microvertebrate diversity from Cherves-de-Cognac (Lower Cretaceous, Berriasian: Charente, France). In P. Barrett & S. Evans (Eds.), *Ninth international symposium on Mesozoic terrestrial ecosystems and biota, Manchester, UK* (pp. 187). Natural History Museum.

Prabhakar, M. (1989). Palynological evidence and its significance for the Kota Formation in the Pranhita-Godavari basin. *Proceedings of*

12th Indian Colloquium on Micropalaeontology and Stratigraphy, 59-65.

Prasad, G. V. R., & Parmar, V. (2020). First Ornithischian and Theropod Dinosaur Teeth from the Middle Jurassic Kota Formation of India: Paleobiogeographic Relationships. In G. V. R. Prasad & R. Patnaik (Eds.), *Biological Consequences of Plate Tectonics: New Perspectives on Post-Gondwana Break-up–A Tribute to Ashok Sahni* (pp. 1-30). Springer International Publishing. https://doi.org/10.1007/978-3-030-49753-8_1

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann Publishers Inc.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Rauhut, O. W. M. (2000). The dinosaur fauna from the Guimarota mine. In T. Martin & B. Krebs (Eds.), *Guimarota a Jurassic Ecosystem* (pp. 75-82). Verlag Dr. Friedrich Pfeil.

Rauhut, O. W. M. (2001). Herbivorous dinosaurs from the Late Jurassic (Kimmeridgian) of Guimarota, Portugal. *Proceedings of the Geologists' Association*, *112*(3), 275-283. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(01)80007-9

Rauhut, O. W. M. (2002). Dinosaur teeth from the Barremian of Uña, Province of Cuenca, Spain. *Cretaceous Research*, *23*(2), 255-263. https://doi.org/http://dx.doi.org/10.1006/cres.2002.1003

Rauhut, O. W. M. (2003). A tyrannosauroid dinosaur from the Upper Jurassic of Portugal. *Palaeontology*, *46*(5), 903-910. https://doi.org/10.1111/1475-4983.00325

Rauhut, O. W. M., & Foth, C. (2020). The Origin of Birds: Current Consensus, Controversy, and the Occurrence of Feathers. In O. W. M. Rauhut & C. Foth (Eds.), *The Evolution of Feathers. Fascinating Life Sciences*. Springer. https://doi.org/https://doi.org/10.1007/978-3-030-27223-4_3

Rauhut, O. W. M., Milner, A. C., & Moore-Fay, S. (2010). Cranial osteology and phylogenetic position of the theropod dinosaur *Proceratosaurus bradleyi* (Woodward, 1910) from the Middle Jurassic of England [Article]. *Zoological Journal of the Linnean Society*, *158*(1), 155-195. https://doi.org/10.1111/j.1096-3642.2009.00591.x

Rauhut, O. W. M., & Werner, C. (1995). First record of the family Dromaeosauridae (Dinosauria: Theropoda) in the Cretaceous of Gondwana (Wadi Milk Formation, northern Sudan). *Paläontologische Zeitschrift*, *69*(3/4), 475-489.

Rausch, J. R., & Kelley, K. (2009). A comparison of linear and mixture models for discriminant analysis under nonnormality. *Behavior Research Methods*, *41*(1), 85-98. https://doi.org/10.3758/BRM.41.1.85

Raven, T. J. (2021). The taxonomic, phylogenetic, biogeographic and macroevolutionary history of the armoured dinosaurs (Ornithischia: Thyreophora). PhD thesis, *University of Brighton.*

Raven, T. J., Barrett, P. M., Joyce, C. B., & Maidment, S. C. R. (2023). The phylogenetic relationships and evolutionary history of the armoured dinosaurs (Ornithischia: Thyreophora). *Journal of Systematic Palaeontology*, *21*(1), 2205433. https://doi.org/10.1080/14772019.2023.2205433

Richardson, L. (1929). The country around Moreton-in-the-Marsh. *Memoir of the Geological Survey of Great Britain. Sheet 217 (England and Wales)*.

Richardson, L. (1933). The country around Cirencester. *Memoir of the Geological Survey of Great Britain. Sheet 235 (England and Wales)*.

Richardson, L., Arkell, W. J., & Dines, H. G. (1946). The country around Whitney. *Memoir of the Geological Survey of Great Britain. Sheet 236 (England and Wales)*.

Richter, U., Mudroch, A., & Buckley, L. (2012). Isolated theropod teeth from the Kem Kem Beds (Early Cenomanian) near Taouz, Morocco. *Paläontologische Zeitschrift*, 1-19. https://doi.org/10.1007/s12542-012-0153-1

Riffenburgh, R. H. (2012). Chapter 19 - Modeling Concepts and Methods. In R. H. Riffenburgh (Ed.), *Statistics in Medicine (Third Edition)* (pp. 393-414). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-384864-2.00019-6

Robinson, P. L. (1957). The Mesozoic fissures of the Bristol Channel area and their vertebrate faunas. *Zoological Journal of the Linnean Society*, *43*, 260-282.

Rogers, R. R., & Brady, M. E. (2010). Origins of microfossil bonebeds: insights from the Upper Cretaceous Judith River Formation of north-central Montana. *Paleobiology*, *36*(1), 80-112. https://doi.org/10.1666/0094-8373-36.1.80

Roli, F., Giacinto, G., & Vernazza, G. (2001). *Methods for Designing Multiple Classifier Systems*. Springer.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. Boston, MA.

Ruiz-Omeñaca, J. I., Pereda Suberbiola, X., & Galton, P. M. (2007). *Callovosaurus leedsi*, the earliest dryosaurid dinosaur (Ornithischia: Euornithopoda) from the Middle Jurassic of England. In K. Carpenter (Ed.), *Horns and Beaks: Ceratopsian and Ornithopod Dinosaurs.* (pp. 3-16). Indiana University Press.

Russell, L. S. (1946). The lower jaw of the theropod dinosaur *Troodon*. *Transactions of the Royal Society of Canada, series 3*, *60*, 171.

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach*. Prentice Hall Press.

Sadleir, R., Barrett, P. M., & Powell, H. P. (2008). The anatomy and systematics of *Eustreptospondylus oxoniensis*, a theropod dinosaur from the Middle Jurassic of Oxfordshire, England. *Monograph of the Palaeontographical Society London*, *160*(627), 1-82.

Salgado, L., Canudo, J. I., Garrido, A. C., Moreno-Azanza, M., Martínez, L. C. A., Coria, R. A., & Gasca, J. M. (2017). A new primitive Neornithischian dinosaur from the Jurassic of Patagonia with gut contents. *Scientific Reports*, *7*(1), 42778. https://doi.org/10.1038/srep42778

Salzberg, S. L. (1994). C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine Learning*, *16*(3), 235-240. https://doi.org/10.1007/BF00993309

Samman, T., Powell, G. L., Currie, P. J., & Hills, L. V. (2005). Morphometry of the teeth of western North American tyrannosaurids and its applicability to quantitative classification. *Acta Palaeontol. Pol.*, *50*(4), 757-776.

Sankey, J. T. (2008). Diversity of Latest Cretaceous (Late Maastrichtian) small Theropods and Birds: Teeth from the Lance and Hell Creek Foramtiosn. In J. T. Sankey & S. Baszio (Eds.), *Vertebrate Microfossil Assemblages: Their Role in Paleoecology and Paleobiogeography*. Indiana university Press.

Sankey, J. T., Brinkman, D. B., Guenther, M., & Currie, P. J. (2002). Small Theropod and Bird Teeth from the Late Cretaceous (Late Campanian) Judith River Group, Alberta. *Journal of Paleontology*, *76*(4), 751-763. https://doi.org/10.2307/1307107

Sankey, J. T., Standhardt, B. R., & Schiebout, J. A. (2005). Theropod teeth from the Upper Cretaceous (Campanian–Maastrichtian), Big Bend National Park, Texas. In K. Carpenter (Ed.), *The Carnivorous Dinosaurs* (pp. 127-152). Indiana University Press.

Sanson, T. S. P., Fahmy, W. F. H., François, L., Cherian, G., Alexander, C., Simon, L., Charlie, C., & Giuseppe, S. (2023). Detecting adrenal lesions on 3D CT scans using a 2.5D deep learning model. *medRxiv*, 2023.2002.2022.23286184. https://doi.org/10.1101/2023.02.22.23286184

Savage, R. (1993). Vertebrate fissure faunas with special reference to Bristol Channel Mesozoic faunas. *Journal of the Geological Society*, *150*(6), 1025-1034. https://doi.org/10.1144/gsjgs.150.6.1025

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.

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., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676-682. https://doi.org/10.1038/nmeth.2019

Schuettpelz, E., Frandsen, P. B., Dikow, R. B., Brown, A., Orli, S., Peters, M., Metallo, A., Funk, V. A., & Dorr, L. J. (2017). Applications of deep convolutional neural networks to digitized natural history collections [10.3897/BDJ.5.e21139]. *Biodiversity Data Journal*, *5*, e21139. https://doi.org/10.3897/BDJ.5.e21139

Schwarz, D., & Salisbury, S. W. (2005). A new species of *Theriosuchus* (Atoposauridae, Crocodylomorpha) from the Late Jurassic (Kimmeridgian) of Guimarota, Portugal. *Geobios*, *38*(6), 779-802. https://doi.org/http://dx.doi.org/10.1016/j.geobios.2004.04.005

Scotese, C. R. (2016). Tutorial: PALEOMAP PaleoAtlas for GPlates and the PaleoData plotter program.

Scotese, C. R. (2021). An Atlas of Phanerozoic Paleogeographic Maps: The Seas Come In and the Seas Go Out. *Annual Review of Earth and Planetary Sciences*, *49*(1), 679-728. https://doi.org/10.1146/annurev-earth-081320-064052

Seeley, H. G. (1887). On the classification of the fossil animals commonly named Dinosauria. *Proceedings of the Royal Society of London*, *43*, 165-171.

Sellés, A. G., Vila, B., Brusatte, S. L., Currie, P. J., & Galobart, À. (2021). A fast-growing basal troodontid (Dinosauria: Theropoda) from the latest Cretaceous of Europe. *Scientific Reports*, *11*(1), 4855. https://doi.org/10.1038/s41598-021-83745-5

Sellwood, B. W., & 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*(2), 189-IN188. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(74)80023-4

Senter, P. (2007). A new look at the phylogeny of coelurosauria (Dlnosauria: Theropoda). *Journal of Systematic Palaeontology*, *5*(4), 429-463. https://doi.org/10.1017/S1477201907002143

Senter, P., Kirkland, J. I., Bird, J., & Bartlett, J. A. (2010). A new troodontid theropod dinosaur from the Lower Cretaceous of Utah. *PLoS ONE*, *5*(12), e14329. https://doi.org/10.1371/journal.pone.0014329

Senter, P., Kirkland, J. I., & Deblieux, D. D. (2012). *Martharaptor greenriverensis*, a New Theropod Dinosaur from the Lower Cretaceous of Utah. *PLoS ONE*, *7*(8), e43911. https://doi.org/10.1371/journal.pone.0043911

Sereno, P. (2012). Taxonomy, morphology, masticatory function and phylogeny of heterodontosaurid dinosaurs. *ZooKeys*, *226*(0), 1-225. https://doi.org/10.3897/zookeys.226.2840

Sereno, P. C. (1986). Phylogeny of the bird-hipped dinosaurs (Order Ornithischia). *National Geographic Research*, *2*, 234-256.

Sereno, P. C. (1991). *Lesothosaurus*, "Fabrosaurids," and the early evolution of Ornithischia. *Journal of Vertebrate Paleontology*, *11*(2), 168-197. https://doi.org/10.1080/02724634.1991.10011386

Sereno, P. C. (1997). The origin and evolution of dinosaurs. *Annual Review of Earth and Planetary Science*, *25*, 435-489.

Sereno, P. C. (1999). The Evolution of Dinosaurs. *Science*, *284*(5423), 2137-2147. https://doi.org/10.1126/science.284.5423.2137

Sereno, P. C., & Dong, Z. (1992). The Skull of the Basal Stegosaur *Huayangosaurus taibaii* and a Cladistic Diagnosis of Stegosauria. *Journal of Vertebrate Paleontology*, *12*(3), 318-343. http://www.jstor.org.ezproxy.nhm.ac.uk/stable/4523456

Shi, T., & Horvath, S. (2006). Unsupervised learning with random forest predictors. *Journal of Computational and Graphical Statistics*, *15*(1), 118-138. https://doi.org/10.1198/106186006X94072

Sigogneau-Russell, D. (2001). Docodont nature of *Cyrtlatherium*, an upper Bathonian mammal from England. *Acta Palaeontologica Polonica*, *46*(3), 427-430.

Sigogneau-Russell, D. (2003a). Docodonts from the British Mesozoic. *Acta Palaeontologica Polonica*, *48*(3), 357-374.

Sigogneau-Russell, D. (2003b). Holotherian mammals from the Forest Marble (Middle Jurassic of England). *Geodiversitas*, *25*(3), 501-537.

Sigogneau-Russell, D., Evans, S. E., Levine, J., & Russell, D. (1998). The Early cretaceous microvertebrate locality of Anoual, Morocco: A glimpse at the small vertebrate assemblages of Africa. In S. Lucas, J. Kirkland, & J. Estep (Eds.), *Lower and Middle Cretaceous Terrestrial Ecosystems* (Vol. 14, pp. 177-181). New mexico Museum of natural History and Science.

Skutschas, P. P., Gvozdkova, V. A., Averianov, A. O., Lopatin, A. V., Martin, T., Schellhorn, R., Kolosov, P. N., Markova, V. D., Kolchanov, V. V., Grigoriev, D. V., Kuzmin, I. T., & Vitenko, D. D. (2021). Wear patterns and dental functioning in an Early Cretaceous stegosaur from Yakutia, Eastern Russia. *PLoS ONE*, *16*(3), e0248163. https://doi.org/10.1371/journal.pone.0248163

Slater, T. S., Duffin, C. J., Hildebrandt, C., Davies, T. G., & Benton, M. J. (2016). Microvertebrates from multiple bone beds in the Rhaetian of the M4–M5 motorway junction, South Gloucestershire, U.K. *Proceedings of the Geologists' Association*, *127*(4), 464-477. https://doi.org/http://dx.doi.org/10.1016/j.pgeola.2016.07.001

Smith, J. B. (2002). An examination of dental morphology and variation in theropod dinosaurs: Implications for the identification of shed teeth. *PhD thesis, University of Pennsylvania, Pennsylvania.*

Smith, J. B. (2005). Heterodonty in Tyrannosaurus rex: Implications for the taxanomic and systematic utility of Theropod dentitions. *Journal of Vertebrate Paleontology*, *25*(4), 865-887.

Smith, J. B., Vann, D. R., & Dodson, P. (2005). Dental morphology and variation in Theropod Dinosaurs: Implications for the taxanomic identification of isolated teeth. *The Anatomical Record Part A*, *285A*, 699-736.

Snyder, K., McLain, M., Wood, J., & Chadwick, A. (2020). Over 13,000 elements from a single bonebed help elucidate disarticulation and transport of an Edmontosaurus thanatocoenosis. *PLoS ONE*, *15*(5), e0233182. https://doi.org/10.1371/journal.pone.0233182

Son, N.-T., Chen, C.-F., Chen, C.-R., & Minh, V.-Q. (2018). Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. *Geocarto International*, *33*(6), 587-601. https://doi.org/10.1080/10106049.2017.1289555

Stekhoven, D. J. (2013). *missForest: Nonparametric Missing Value Imputation using Random Forest*.

Stekhoven, D. J., & Buehlmann, P. (2012). MissForest - non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112-118.

Stockmann, U., Minasny, B., & McBratney, A. B. (2014). How fast does soil grow? *Geoderma*, *216*, 48-61. https://doi.org/https://doi.org/10.1016/j.geoderma.2013.10.007

Strahan, A. (1898). The Geology of the Isle of Purbeck and Weymouth. *Memoirs of the Geological Survey, England and Wales*, 278.

Strickson, E., Prieto-Márquez, A., Benton, M. J., & Stubbs, T. L. (2016). Dynamics of dental evolution in ornithopod dinosaurs [Article]. *Scientific Reports*, *6*(28904), (28904). https://doi.org/10.1038/srep28904

http://www.nature.com/articles/srep28904#supplementary-information

Suarez, C., Forster, C., Sharman, G., Fekete, J., Suarez, M., Zanno, L., & Tucker, R. (2023). Updates on the chronostratigraphic framework of the Cedar Mountain Formation of Utah and implications for faunal comparisons and the cretaceous terrestrial revolution. In R. K. Hunt-Foster, J. I. Kirkland, & M. A. Loewen (Eds.), *The Anatomical Record. Special Issue: 14th Symposium on Mesozoic Terrestrial Ecosystems and Biota* (Vol. 306, pp. 238-241). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/ar.25219

Sues, H.-D., & Averianov, A. (2013). Enigmatic teeth of small theropod dinosaurs from the Upper Cretaceous (Cenomanian–Turonian) of Uzbekistan. *Canadian Journal of Earth Sciences*, *50*(3), 306-314. https://doi.org/10.1139/e2012-033

Sullivan, C., Wang, Y., Hone, D. W. E., Wang, Y., Xu, X., & Zhang, F. (2014). The vertebrates of the Jurassic Daohugou Biota of northeastern China. *Journal of Vertebrate Paleontology*, *34*(2), 243-280. https://doi.org/10.1080/02724634.2013.787316

Sullivan, R. M. (2006). A taxanomic reivew of the Pachycephalosauridae (Dinosauria: Ornithischia). In S. G. Lucas & R. M. Sullivan (Eds.), *Late Cretaceous vertebrates from the Western Interior. New Mexico Museum of Natural History and Science Bulletin 35*.

Sumbler, M. G. (1984). The stratigraphy of the Bathonian White Limestone and Forest Marble formations of Oxfordshire. *Proceedings of the Geologists' Association*, *95*, 51-64.

Sumbler, M. G., Barron, A. J. M., Morigi, A. N., Cox, B. M., Ivimey-Cook, H. C., Horton, A., Lott, G. K., Warrington, G., Williamson, I. T., Pharaoh., T. C., Forster, A., Buckley, D. K., & Robinson, V. K. (2000). Geology of the Cirencester district*, British Geological Survey, memoir for 1:50000 geological sheet 235 (England & Wales)*.

Sumbler, M. G., & Wyatt, R. J. (1999). Correlation of the Bathonian (Middle Jurassic) succession in the Minchinhampton-Burford district: a critique of Wyatt (1996). *Proceedings of the Geologists' Association*, *110*(1), 53-64. https://doi.org/http://dx.doi.org/10.1016/S0016-7878(99)80006-6

Sweetman, S. (2007). Aspects of the microvertebrate fauna of the Early Cretaceous (Barremian) Wessex Formation of the Isle of Wight, Southern England. *PhD thesis, University of Portsmouth, Portsmouth.*

Sweetman, S. C. (2004). The first record of velociraptorine dinosaurs (Saurischia, Theropoda) from the Wealden (Early Cretaceous, Barremian) of southern England. *Cretaceous Research*, *25*(3), 353-364. https://doi.org/http://dx.doi.org/10.1016/j.cretres.2004.01.004

Sweetman, S. C. (2006). The tertapod microbiota of theWessex Formation (Lower Cretaceous, Barremian) of the Isle of Wight, UK. In P. Barrett & S. Evans (Eds.), *Ninth international symposium on Mesozoic terrestrial ecosystems and biota, Manchester, UK* (pp. 187). Natural History Museum.

Tałanda, M., Fernandez, V., Panciroli, E., Evans, S. E., & Benson, R. J. (2022). Synchrotron tomography of a stem lizard elucidates early squamate anatomy. *Nature*, *611*(7934), 99-104. https://doi.org/10.1038/s41586-022-05332-6

The MathWorks Inc. (2023). *MATLAB version: 9.14.0 (R2023a)*. The MathWorks Inc.

ThermoFisher. (2014). *Avizo v 8.1.1*. ThermoFisher Scientific.

Thulborn, R. A. (1973). Teeth of ornithischian dinosaurs from the Upper Jurassic of Portugal, with description of a hypsilophodontid (*Phyllodon henkeli* gen. et sp. nov.) from the Guimarota lignite. In *Contribuiçao para o conhecimento da Fauna do Kimeridgiano da Mina de Lignito Guimarota (Leiria, Portugal) III Parte, Serviços Geológicos de Portugal, Memória 22 (Nova Série)* (pp. 469).

Thulborn, R. A. (1992). Taxonomic characters of *Fabrosaurus australis,* an ornithischian dinosaur from the Lower Jurassic of Southern Africa. *Geobios*, *25*(2), 283-292.

Torices, A., Reichel, M., & Currie, P. J. (2014). Multivariate analysis of isolated tyrannosaurid teeth from the Danek Bonebed, Horseshoe Canyon Formation, Alberta, Canada. *Canadian Journal of Earth Sciences*, *51*(11), 1045-1051. https://doi.org/10.1139/cjes-2014-0072

Torrens, H. (1969a). *Excursion Number 1. Guide for Dorset and South Somerset*, International field symposium on the British Jurassic.

Torrens, H. (1969b). *Excursion to the Costwolds and Oxfordshire*, International field symposium on the British Jurassic.

Turner, A. H., Makovicky, P. J., & Norell, M. A. (2012). A Review of Dromaeosaurid Systematics and Paravian Phylogeny. *Bulletin of the American Museum of Natural History*, 1-206. https://doi.org/10.1206/748.1

Underwood, C. J. (2004). Environmental controls on the distribution of neoselachian sharks and rays within the British Bathonian (Middle Jurassic). *Palaeogeography, Palaeoclimatology, Palaeoecology*, *203*(1-2), 107-126.

Underwood, C. J., & Ward, D. J. (2004). Neoselachian sharks and rays from the British Bathonian (Middle Jurassic) [Article]. *Palaeontology*, *47*(3), 447-501. https://doi.org/10.1111/j.0031-0239.2004.00386.x

Upchurch, P., Andres, B., Butler, R. J., & Barrett, P. M. (2014). An analysis of pterosaurian biogeography: implications for the evolutionary history and fossil record quality of the

first flying vertebrates. *Historical Biology*. https://doi.org/DOI: 10.1080/08912963.2014.939077

Upchurch, P., Barrett, P. M., & Dodson, P. (2004). Sauropoda. In D. B. Weishampel, P. Dodson, & H. Osmolska (Eds.), *The Dinosauria* (2nd ed., pp. 259 - 322). University of California Press.

Upchurch, P., Hunn, C. A., & 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 of London. Series B: Biological Sciences*, *269*(1491), 613-621. https://doi.org/10.1098/rspb.2001.1921

Valiant, L. (1984). A theory of the learnable. *Communications of the ACM*, *27*, 1134-1142.

Valkenburgh, B. V., & Molnar, R. E. (2002). Dinosaurian and Mammalian Predators Compared. *Paleobiology*, *28*(4), 527-543. https://doi.org/10.2307/3595499

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1-67.

van den Berg, T., Whiteside, D. I., Viegas, P., Schouten, R., & Benton, M. J. (2012). The Late Triassic microvertebrate fauna of Tytherington, UK. *Proceedings of the Geologists' Association*, *123*(4), 638-648. https://doi.org/http://dx.doi.org/10.1016/j.pgeola.2012.05.003

Van Der Lubbe, T., Richter, U., & Knotschke, N. (2009). Velociraptorine dromaeosaurid teeth from the Kimmeridgian (Late Jurassic) of Germany. *Acta Palaeontologica Polonica*, *54*(3), 401-408.

Vaughan, R. F. (1989). The excavation at Hornsleasow Quarry. *Interim report for the Nature Conservance Council*. Gloucester County Museums.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer.

Vickaryous, M. K., Maryanska, T., & Weishampel, D. B. (2004). Ankylosauria. In D. B. Weishampel, P. Dodson, & H. Osmolska (Eds.), *The Dinosauria* (2nd ed., pp. 363-392). University of California Press.

Viglietti, P., McPhee, B. W., Bordy, E. M., Sciscio, L., Barrett, P. M., Benson, R. B. J., Wills, S., Chapelle, K. E. J., Dollman, K. N., Mdezaki, C., & Choiniere, J. N. (2020). Biostratigraphy of the *Massospondylus* Assemblage Zone (Stormberg Group, Karoo Supergroup), South Africa. *Geological Society of South Africa*, *123*, 249-262. https://doi.org/10.25131/sajg.123.0018

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., & Tong, H. (2014). Palaeontology of the Purbeck-type (Tithonian, Late Jurassic) bonebeds of Chassiron (Oléron Island, western France). *Comptes Rendus Palevol*, *13*(5), 421-441. https://doi.org/10.1016/j.crpv.2014.03.003

Vullo, R., & Neraudeau, D. (2010). Additional dinosaur teeth from the Cemonian (Late Cretaceous) of Charentes, southwestern France. *Comptes Rendus Palevol*, *9*, 121-126.

Vullo, R., Neraudeau, D., & Lenglet, T. (2007). Dinosaur Teeth from the Cenomanian of Charentes, Western France: Evidence for a mixed Laurasian-Gondwanan assemblage. *Journal of Vertebrate Paleontology*, *27*(4), 931-943.

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

Waldman, M., & Evans, S. E. (1994). Lepidosauromorph reptiles from the Middle Jurassic of Skye. *Zoological Journal of the Linnean Society*, *112*(1–2), 135-150. https://doi.org/http://dx.doi.org/10.1006/zjls.1994.1037

Waldman, M., & Savage, R. J. G. (1972). The first Jurassic mammal from Scotland. *Journal of the Geological Society*, *128*(2), 119-125. https://doi.org/10.1144/gsjgs.128.2.0119

Walker, A. D. (1964). Triassic reptiles from the Elgin area: *Ornithosuchus* and the origin of carnosaurs. *Philosophical Transactions of the Royal Society of London Series B*, *248*, 53-134.

Ward, D. J. (1981). A simple machine for bulk processing of clays and silts. *Tertiary Research*, *3*, 121-124.

Ward, D. J. (1984). Collecting islolated microvertebrate fossils. *Zoological Journal of the Linnean Society*, *82*, 245-259.

Weishampel, D. B., Barrett, P. M., Coria, R. A., Le Loeuff, J., Xing, X., Xijin, Z., Sahni, A., Gomani, E. M. P., & Noto, C. R. (2004). Dinosaur Distribution. In D. B. Weishampel, P. Dodson, & H. Osmolska (Eds.), *The Dinosauria* (2nd ed., pp. 861). University of California Press.

Welch, B. L. (1939). Note on discriminant functions. *Biometrika*, *31*, 218-220.

West, I. M. (2012). *Bridport, West Cliff to Eype Mouth - Geological Field Guide*. http://www.southampton.ac.uk/~imw/Bridport-West.htm

Whitenack, L. B., & Gottfried, M. D. (2010). A Morphometric Approach for Addressing Tooth-Based Species Delimitation in Fossil Mako Sharks, Isurus (Elasmobranchii: Lamniformes). *Journal of Vertebrate Paleontology*, *30*(1), 17-25. https://doi.org/10.1080/02724630903409055

Whiteside, D. I., Duffin, C. J., Gill, P. G., Marshall, J. E. A., & Benton, M. J. (2016). The Late Triassic and Early Jurassic fissure faunas from Bristol and South Wales: Stratigraphy and setting. *Palaeontologia Polonica*, *67*, 257-287.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Wilke, C. O. (2019). *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 0.9.4.*

Williams, H., Duffin, C. J., Hildebrandt, C., Parker, A., Hutchinson, D., & Benton, M. J. (2022). Microvertebrates from the Rhaetian bone beds at Westbury Garden Cliff, near Gloucester, UK. *Proceedings of the Geologists' Association*. https://doi.org/10.1016/j.pgeola.2022.01.002

Williamson, T. E., & 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*(4), e93190. https://doi.org/10.1371/journal.pone.0093190

Wills, S., Barrett, P. M., & Walker, A. (2014). New dinosaur and crocodilian material from the Middle Jurassic (Bathonian) Kilmaluag Formation, Skye, Scotland. *Scottish Journal of Geology*, *50*, 183-190. https://doi.org/http://dx.doi.org/10.1144/sjg2014-005

Wills, S., Bernard, E. L., Brewer, P., Underwood, C. J., & 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*(2), 170-186. https://doi.org/https://doi.org/10.1016/j.pgeola.2019.02.003

Wills, S., Cavosie, A. J., Fernandez, V., Underwood, C. J., Ward, D. J., Bernard, E. L., & Barrett, P. M. (2023). High diversity Middle Jurassic dinosaur faunas revealed through microvertebrate analysis. In R. K. Hunt-Foster, J. I. Kirkland, & M. A. Loewen (Eds.), *The Anatomical Record. Special Issue: 14th Symposium on Mesozoic Terrestrial Ecosystems and Biota* (Vol. 306, pp. 258-261). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/ar.25219

Wills, S., Underwood, C. J., & Barrett, P. M. (2021). Learning to see the wood for the trees: machine learning, decision trees, and the classification of isolated theropod teeth [https://doi.org/10.1111/pala.12512]. *Palaeontology*, *64*(1), 75-99. https://doi.org/https://doi.org/10.1111/pala.12512

Wills, S., Underwood, C. J., & Barrett, P. M. (2023). Machine learning confirms new records of maniraptoran theropods in Middle Jurassic UK microvertebrate faunas

[https://doi.org/10.1002/spp2.1487]. *Papers in Palaeontology*, *9*(2), e1487. https://doi.org/https://doi.org/10.1002/spp2.1487

Wilson, G. P., Evans, A. R., Corfe, I. J., Smits, P. D., Fortelius, M., & Jernvall, J. (2012). Adaptive radiation of multituberculate mammals before the extinction of dinosaurs [10.1038/nature10880]. *Nature*, *483*(7390), 457-460. https://doi.org/http://www.nature.com/nature/journal/v483/n7390/abs/nature10880.html#supplementary-information

Wilson, J. A., & Upchurch, P. (2009). Redescription and reassessment of the phylogenetic affinities of *Euhelopus zdanskyi* (Dinosauria: Sauropoda) from the early cretaceous of China. *Journal of Systematic Palaeontology*, *7*(2), 199-239. https://doi.org/10.1017/S1477201908002691

Wilson, L. E. (2008). Comparative taphonomy and paleoecological reconstruction of two microvertebrate accumulations from the Late Cretaceous Hell Creek Formation (Maastrichtian), eastern Montana. *PALAIOS*, *23*, 289-297. https://doi.org/10.2110/palo.2007.p07-006r

Wilson, V., Welch, F. B. A., Robbie, J. A., & Green, G. W. (1958). Geology of the country around Bridport and Yeovil. *Memoirs of the Geological Survey of Great Britain*.

Wings, O., Schwarz-Wings, D., Pfretzschner, H.-U., & Martin, T. (2010). Overview of Mesozoic crocodylomorphs from the Junggar Basin, Xinjiang, Northwest China, and description of isolated crocodyliform teeth from the Late Jurassic Liuhuanggou locality. *Palaeobiodiversity and Palaeoenvironments*, *90*(3), 283-294. https://doi.org/10.1007/s12549-010-0033-1

Wings, O., Tütken, T., Fowler, D. W., Martin, T., Pfretzschner, H.-U., & Sun, G. (2015). Dinosaur teeth from the Jurassic Qigu and Shishugou Formations of the Junggar Basin (Xinjiang/China) and their paleoecologic implications. *Paläontologische Zeitschrift*, *89*(3), 485-502. https://doi.org/10.1007/s12542-014-0227-3

Wood, S. V. (1839). Announcing the discovery of fossil quadrumanous remains, near Woodbridge, Suffolk. *The Magazine of Natural History*, *New Series 3*, 444-445.

Wood, S. V. (1844). Record of the discovery of an alligator with several new Mammalia in the freshwater strata at Hordwell. *Annals and Magazine of Natural History*, *14*, 319-351.

Woodward, A. S. (1910). On a Skull of *Megalosaurus* from the Great Oolite of Minchinhampton (Gloucestershire). *Quarterly Journal of the Geological Society*, *66*(1-4), 111-115. http://jgslegacy.lyellcollection.org/content/66/1-4/111.abstract

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*, 1-628.

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*(1), 1 - 17. https://doi.org/10.18637/jss.v077.i01

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. In B. M. Cox & M. G. Sumbler (Eds.), *British Middle Jurassic stratigraphy, Geological Conservation Review Series, No. 26* (Vol. 26, pp. 234-239). Joint Nature Conservation Committee.

Wyatt, R. J. (2011). A gamma-ray correlation of boreholes and oil wells in the Bathonian stage succession (Middle Jurassic) of the Wealden shelf subcrop*. British Geological Survey open report, OR/11/048*.

Xu, X., Choiniere, J. N., Pittman, M., Tan, Q., Xiao, D., Li, Z., Clark, J. M., Norell, M. A., Hone, D. W. E., & Sullivan, C. (2010). A new dromaeosaurid (Dinosauria: Theropoda) from the Upper Cretaceous Wulansuhai Formation of Inner Mongolia, China. *Zootaxa*, *2403*, 1-9.

Xu, X., Forster, C. A., Clark, J. M., & Mo, J. (2006). A basal ceratopsian with transitional features from the Late Jurassic of northwestern China. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1598), 2135-2140. https://doi.org/10.1098/rspb.2006.3566

Xu, X., QingYu, M., & DongYu, H. (2010). Pre-*Archaeopteryx* coelurosaurian dinosaurs and their implications for understanding avian origins. *Chinese Science Bulletin*, *55*(35), 3971-3977. https://doi.org/10.1007/s11434-010-4150-z

Xu, X., Upchurch, P., Mannion, P. D., Barrett, P. M., Regalado-Fernandez, O. R., Mo, J., Ma, J., & Liu, H. (2018). A new Middle Jurassic diplodocoid suggests an earlier dispersal and diversification of sauropod dinosaurs. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05128-1

Xu, X., Zhao, X., & Clark, J. M. (2001). A New Therizinosaur from the Lower Jurassic Lower Lufeng Formation of Yunnan, China. *Journal of Vertebrate Paleontology*, *21*(3), 477-483. https://doi.org/10.2307/20061976

Xu, X., Zheng, X., Sullivan, C., Wang, X., Xing, L., Wang, Y., Zhang, X., O'Connor, J. K., Zhang, F., & Pan, Y. (2015). A bizarre Jurassic maniraptoran theropod with preserved evidence of membranous wings. *Nature*, *521*, 70. https://doi.org/10.1038/nature14423

https://www.nature.com/articles/nature14423#supplementary-information

Xu, X., Zhou, Z., Sullivan, C., Wang, Y., & 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.

Xu, X., Zhou, Z., & Wang, X. (2000). The smallest known non-avian theropod dinosaur. *Nature*, *408*, 705-708.

Yi, H., Tennant, J. P., Young, M. T., Challands, T. J., Foffa, D., Hudson, J. D., Ross, D. A., & Brusatte, S. L. (2017). An unusual small-bodied crocodyliform from the Middle Jurassic of Scotland, UK, and potential evidence for an early diversification of advanced neosuchians. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *107*(1), 1-12. https://doi.org/10.1017/S1755691017000032

Young, C. M. E., Hendrickx, C., Challands, T. J., Foffa, D., Ross, D. A., Butler, I. B., & 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. https://doi.org/10.1144/sjg2018-020

Young, M. T., Tennant, J. P., Brusatte, S. L., Challands, T. J., Fraser, N. C., Clark, N. D. L., & Ross, D. A. (2015). The first definitive Middle Jurassic atoposaurid (Crocodylomorpha, Neosuchia), and a discussion on the genus Theriosuchus. *Zoological Journal of the Linnean Society*, n/a-n/a. https://doi.org/10.1111/zoj.12315

Zanno, L. E. (2010a). Osteology of *Falcarius utahensis* (Dinosauria: Theropoda): characterizing the anatomy of basal therizinosaurs. *Zoological Journal of the Linnean Society*, *158*(1), 196-230. https://doi.org/10.1111/j.1096-3642.2009.00464.x

Zanno, L. E. (2010b). A taxonomic and phylogenetic re-evaluation of Therizinosauria (Dinosauria: Maniraptora). *Journal of Systematic Palaeontology*, *8*(4), 503-543. https://doi.org/10.1080/14772019.2010.488045

Zanno, L. E., Tsogtbaatar, K., Chinzorig, T., & Gates, T. A. (2016). Specializations of the mandibular anatomy and dentition of *Segnosaurus galbinensis* (Theropoda: Therizinosauria). *PeerJ*, *4*, e1885. https://doi.org/10.7717/peerj.1885

Zavorka, S., & Perrett, J. J. (2014). Minimum sample size considerations for two-group linear and quadratic discriminant analysis with rare populations. *Communications in Statistics - Simulation and Computation*, *43*(7), 1726-1739. https://doi.org/10.1080/03610918.2012.744041

Zhao, X., Cheng, Z., & Xu, X. (1999). The Earliest Ceratopsian from the Tuchengzi Formation of Liaoning, China. *Journal of Vertebrate Paleontology*, *19*(4), 681-691. http://www.jstor.org.ezproxy.nhm.ac.uk/stable/4524038

Zheng, X.-T., You, H.-L., Xu, X., & Dong, Z.-M. (2009). An Early Cretaceous heterodontosaurid dinosaur with filamentous integumentary structures. In (Vol. 458, pp. 333-336): Nature Publishing Group.

Zheng, X., Xu, X., You, H., Zhao, Q., & Song, Z. (2009). A short-armed dromaeosaurid from the Jehol Group of China with implications for early dromaeosaurid evolution. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2009.1178

Zinke, J. (1998). Small theropod teeth from the Upper Jurassic coal mine of Guimarota (Portugal). *Paläontologische Zeitschrift*, *72*(1-2), 179-189. https://doi.org/10.1007/BF02987825

Appendix One: Data archiving statement

All scripts and machine learning results are available on GitHub at:

https://github.com/simonwills/Bathonian_Dinosaurs

Chapter Three, machine learning methods

Data for this study are available in the Dryad Digital Repository:

https://doi.org/10.5061/dryad.1zcrjdfq9

Chapter Five, Theropoda

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

Appendix Two: Summary of data sources

Machine learning

All theropod tooth morphometric measurements for this thesis are available on GitHub at: https://github.com/simonwills/Bathonian_Dinosaurs

The repository contains three data files: trainingData.csv, UKBathonianTeeth.csv, UK Bathonian Theropod Teeth Full Data.csv

trainingData.csv

1702 individual theropod tooth linear measurements used to train the machine learning models, sourced from published datasets (Currie & Varrichio, 2004; Farlow et al., 1991; Gerke & Wings, 2016; Hendrickx et al., 2015a; Larson, 2008; Larson et al., 2016; Larson & Currie, 2013; Longrich, 2008; Rauhut et al., 2010; Sankey, 2008; Sankey et al., 2002; Smith, 2005; Young et al., 2019). Any specimens with missing data in the morphometric variables have been removed.

File description

Dimensions: 29 columns x 1703 rows

Variables:

* ID: numeric unique ID for each specimen

* SPECIMENID: original specimen ID as per museum registers. Character data.

* CBL, CH, CBW, ADM, PDM: morphometric variables (Fig A2.1). Continuous data. CBL = Crown Base Length, CH = Crown Height, CBW = Crown Base Width, ADM = Anterior Denticles / mm, PDM = Posterior Denticles / mm. A value of 0 is a true value, it does not indicate missing data.

* Original_Taxon: taxon assigned by source data

* Clade, Clade1, Clade2, Combined_clade, Higher_level_taxonomy: differing levels of taxonomy for the specimen

* Data Source: Reference to the source dataset

* Reference: Reference to the original publication (which may be different to where the data was sourced from)

* Period, Lower_Stage, Upper_Stage, Group, Formation, Member: Stratigraphic position of specimen

* Max_ma, Min_ma, mid_ma: chronostratigraphy based on stage boundaries

* Lat, Long: latitude and longitude of locality in decimal degrees

* P_lat, P_lon: palaeolatitude and longitude of locality in decimal degrees

* Locality: locality name if known

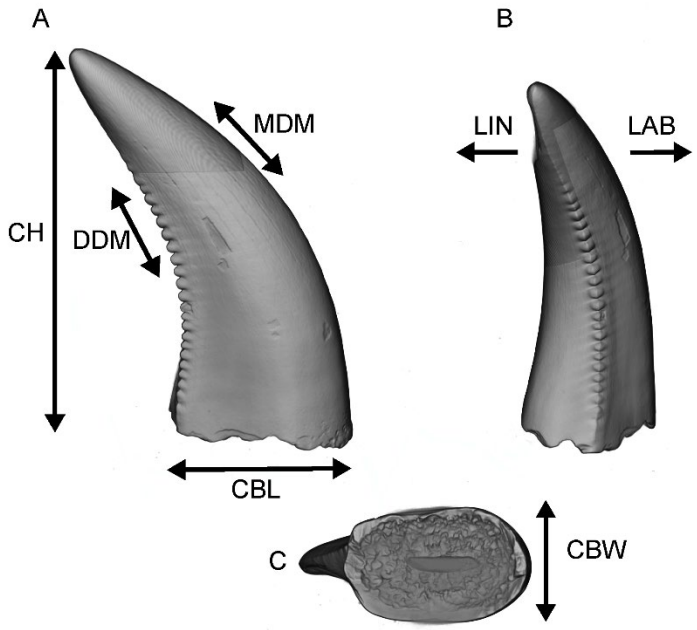


Fig. A2.1. Anatomical and morphometric terminology of theropod tooth crowns. Theropod tooth crown in labial (A), distal (B) and basal (C) views. Abbreviations: CH, crown height; CBL, crown base length; CBW, crown base width; MDM, mesial denticles per millimetre; DDM, distal denticles per millimetre; LIN, lingual; LAB, labial. After Hendrickx et al. (2019) and Wills et al. (2021).

UKBathonianTeeth.csv and UK Bathonian Theropod Teeth Full Data.csv

Individual theropod tooth linear measurements from Hornsleasow Quarry, Woodeaton Quarry, Kirtlington Quarry and Watton Cliff with 93 and 149 specimens respectively.

The file structure mirrors that of **trainingData.csv** with **UKBathonianTeeth.csv** containing only those specimens with complete morphometric data, and UK Bathonian **Theropod Teeth Full Data.csv** the full dataset including any missing data.

Comparative data

Theropoda

The specimens listed below were scanned and measured as part of this thesis.

Institution / Source	Specimen numbers	Description
Wealden Group, Barremian, Early Cretaceous		
Natural History Museum, London	NHMUK PV R 5226 NHMUK PV R 5227 NHMUK PV R 9551	Isolated theropod teeth, Spinosauridae, *Baryonyx walkeri* Smokejacks Brickworks, Surrey
Purbeck Group, Berriasian, Early Cretaceous		
Natural History Museum, London	NHMUK PV OR 48208 NHMUK PV R 15870 NHMUK PV R 15871 NHMUK PV R 15872 NHMUK PV R 15873 NHMUK PV R 15874 NHMUK PV R 15875 NHMUK PV R 15876 NHMUK PV R 15877 NHMUK PV R 15878 NHMUK PV R 15889	Isolated theropod teeth, Dromaeosauridae, *Nuthetes destructor* Durlston Bay, Dorset
Forest Marble Formation, Bathonian, Middle Jurassic		
Natural History Museum, London	NHMUK PV OR 39476	Isolated theropod tooth, Tetanurae indet. Stanton, Wiltshire
White Limestone Formation, Bathonian, Middle Jurassic		
Natural History Museum, London	NHMUK PV R 5797	Isolated theropod tooth, Tetanurae indet. Kirtlington Quarry, Oxfordshire
Taynton Limestone Formation, Bathonian, Middle Jurassic		
Oxford University Museum of Natural History	OXUMNH J 29762 OXUMNH J 29776 OXUMNH J 29809 OXUMNH J 29810 OXUMNH J 29855 OXUMNH J 29856 OXUMNH J 29863 OXUMNH J 48171	Isolated theropod teeth. *Megalosaurus bucklandii* apart from J 29776 Theropoda indet. Stonesfield, Oxfordshire

	OUMNH J 13505	Right dentary of *Megalosaurus bucklandii*
Natural History Museum, London	NHMUK PV OR 31834 NHMUK PV OR 42024 NHMUK PV OR 47963 NHMUK PV R 234 NHMUK PV OR 28608	Isolated theropod teeth. *Megalosaurus bucklandii*
Charlbury Formation, Bathonian, Middle Jurassic		
Natural History Museum, London	NHMUK PV OR 2635	Isolated theropod tooth, Tetanurae indet. Huntsman's Quarry, Gloucestershire
Chipping Norton Limestone Formation, Bathonian, Middle Jurassic.		
Oxford University Museum of Natural History	OXUMNH J 23014 OXUMNH J 23050	Isolated theropod teeth, Tetanurae indet. Sarsden, Oxfordshire
	OXUMNH J 29855 OXUMNH J 29856	Isolated theropod teeth, Tetanurae indet. Sarsgrove, Oxfordshire

Table A2.1. Theropod specimens imaged and measured as part of this thesis.

Ornithischia

The specimens listed below were scanned and measured as part of this thesis.

Institution / Source	Specimen numbers	Description
Wessex Formation, Wealden Group, Barremian, Early Cretaceous		
Natural History Museum, London	NHMUK PV R 193 NHMUK PV R 197 NHMUK PV R 2472 NHMUK PV R 5830	Skull and jaws. Isolated teeth Neornithischian, *Hypsilophodon foxii* Isle of Wight
Purbeck Group, Berriasian, Early Cretaceous		
Natural History Museum, London	NHMUK PV OR 48211 NHMUK PV OR 48212 NHMUK PV OR 48213 NHMUK PV OR 48215	Fragments of jaws and isolated teeth. Heterodontosauridae, *Echinodon becklesii* Durlston Bay, Dorset
Cañadón Asfalto Formation, ?Toarcian, Late Jurassic		
Museo Egidio Feruglio, Trelew, Argentina	MPEF PV 3809	Maxilla Heterodontosauridae, *Manidens condorensis* Patagonia
Brushy Basin Member, Morrison Formation, Tithonian, Late Jurassic		

Natural History Museum of Los Angeles County, Los Angeles, USA	LACM 115747 LACM 128258	Maxilla and dentary. Heterodontosauridae, *Fruitadens haagarorum* Colorado, USA
Morrison Formation, Late Jurassic		
Natural History Museum, London	NHMUK PV R 36730	Isolated teeth. Stegosauridae, *Stegosaurus stenops* Wyoming, USA
Oxford Clay Formation, Callovian, Middle Jurassic		
Natural History Museum, London	NHMUK PV R 2682	Mandible. Ankylosauridae, *Sarcolestes leedsi* Peterborough
Charmouth Mudstone Formation, Sinemurian, Early Jurassic		
Natural History Museum, London	NHMUK P R 1111	Jaws with teeth. Thyreophora, *Scelidosaurus harrisonii* Charmouth, Dorset
La Quinta Formation, Hettangian, Early Jurassic		
Museo de Biología de la Universidad del Zulia, Zulia, Venezuela	MBLUZ 008 MBLUZ 011 MBLUZ 012 MBLUZ 1396 MBLUZ 93.002 MBLUZ P-1395 MBLUZ P-1397 MBLUZ P-1398 MBLUZ P-1400	Isolated ornithischian teeth. Ornithischia indet. Or Thyreophora indet. *Laquintasaura venezuelae* Rio La Grita, Venezuela

Table A2.2. Ornithischian specimens imaged and measured as part of this thesis.

Appendix Three: R, Python and Matlab scripts

All scripts for this thesis are available on GitHub at:

https://github.com/simonwills/Bathonian_Dinosaurs

A summary and brief description of each script is given below.

R scripts

The R scripts were initially created on R version 3.6.0 using R Studio 1.3.1056. They have also been tested on R version 4.0.2

R Script name	Usage
Combined_Results.R	Alluvial plots (See Chapter Five)
DataDistribution.R	Produces a map showing the distribution of data points
DBSCAN Outliers.Rmd	Density-based spatial clustering of applications for outlier identification
DeepTimePlots.Rmd	Plots taxonomic ranges of taxa from the Paleobiology Database
GetPBDBClean.Rmd	Downloads and cleans occurrence data from the Paleobiology Database
PBDB_Microvertebrates.Rmd	Downloads and cleans microvertebrate occurrence data from the Paleobiology Database
Permanova.R	Permutational multivariate analysis of variance
PredictiveModels.R	Generation of machine learning models

Table A3.1. R scripts and usage

Bycladeaccuracy.R

Combines the results from machine learning analyses into one data frame and calculates the majority vote across the different models. The results are then presented as an alluvial plot showing the different 'flows' of data from the models to the final classification.

DataDistribution.R

Produces a world map showing the distribution of data (e.g., training data) by geological period. The input to the script is a dataframe containing clade, period, lower stage, upper stage, latitude, longitude.

Latitude and longitude must be in decimal degrees with degrees west and south expressed as negative numbers.

DBSCAN Outliers.Rmd

Runs the Density-based spatial clustering of applications (DBSCAN) data clustering algorithm on an input dataframe containing either raw morphometric data or ordinated morphometric data. The dataframe must include specimen ID, taxon and the row / ordinated data as columns with each specimen represented by a row of data. The output is presented as a series of plots per taxon.

DeepTimePlots.Rmd

Plots taxonomic ranges of taxa from the Paleobiology Database binned by stage, period or epoch. To use run the data through the GetPBDBClean.Rmd script first.

GetPBDBClean.Rmd

Downloads occurrence data from the Paleobiology Database and undertakes some data cleaning prior to use in other scripts. The taxa to be downloaded are controlled in the block of code between lines 23 and 31 (as below) and can be modified.

Code block

```
23 ```{r}

#url <-
"https://paleobiodb.org/data1.2/occs/list.csv?&interval_id=2&base_name=ornithischia^Krzyzanowskisaurus#^Gregaripus^Eoanomoepus^Shenmuichnus^%pus&show=class,coords,paleoloc,strat"

# interval_id 2 = Mesozoic

url <-
"https://paleobiodb.org/data1.2/occs/list.csv?&interval_id=2&base_name=abrictosaurus,lycorhinus,lanasaurus,heterodontosaurus,echinodon,fruitadens,tianyulong,manidens,pegomastax,geranosaurus&show=class,coords,paleoloc,strat"

taxa_df <- read.csv(url)

31 ```
```

End code block

PBDB_Microvertebrates.Rmd

Script to download dinosaur occurrences from the PBDB, extract those with references to microfossils and then plot binned occurrence data.

Permanova.R

Permutational multivariate analysis of variance (PERMANOVA) tests on ordinated morphometric data. The script undertakes both an overall PERMANOVA test and a pairwise (by clade) test. The results are plotted as a heatmap.

PredictiveModels.R

The main R script used to generate machine learning models. This was used for both Chapter Three and Chapter Five. The script generates three different models: Random Forest and C50 decision tree models and a mixture discriminant analysis model.

The script reads morphometric measurements from .csv files and runs various predictive models i.e., classifications on the data. As a minimum the input file must contain the following columns: Specimen ID, grouping variable, morphometric variables.

The first row must be the column names. Subsequent rows are individual specimens or cases.

The grouping variable can be anything you want e.g., genus, species, family, geological period. You can have multiple columns containing different grouping variables and the script will ask you to choose the variable you want to group on for the classification.

There is no limit on the number of morphometric variables and the script will ask you to choose the columns that represent the variables you wish to use in the analysis.

The following packages are required: caret, MASS, mda, randomForest, C50 and the script will attempt to download them and install if they are not already on the system.

Step by step instructions are as below

1. Choose the input .csv file

2. Choose the column you wish to group the results by

3. Choose which groups you want to analyse – this can be all the groups in the data set or you can chose a subset.

4. Choose the columns which contain the morphometric measurement data.

5. Choose the column which contains the specimen ID.

6. The script then does some housework on the data:

    a. Gives some consistency to the specimen ID and group column names

    b. Refactors the data ensuring all the names are acceptable to R

    c. Checks for missing data and ONLY retains those cases where data is complete (lines 54 to 58). If you do not want to drop missing data then simply comment out these lines.

    d. Drops any groups where there are less group members then there are morphometric variables (lines 61 to 66).

    e. Offers to log + 1 transform the data

    f. Offers to scale and centre the data

7. The script then splits the data set into two data partitions, one for training the model and one for testing the model. The split is 80:20

8. Drop any extra columns in the data – only retain Specimen ID, group and morphometric variables

9. Pull out the specimen ID to two files (training and testing) and then drop this column. This data can be added back later.

10. Set up the control for the model (line 107). This is where the script specifies cross validation.

11. Define prior probabilities (line 108) – in this case equal priors based on the numbers of groups in the training data.

12. Run the models: C50, MDA, RF

    a. Each run creates a model (e.g. C5Fit1), training and testing results and training and testing confusion matrices.

    b. A series of .csv files with the results

Please note that the C5 and Random Forest models can take a long time to run as the script tries many permutations of model parameters.

Predictions on new data

To create predictions based on new data call the R predict function on the dataframe that contains the new data.

If you are going to scale and centre the variables this should be done on a combined training / unknown dataset.

The dataframe must be in the same format as the training / testing data that was used to create the model in the first place i.e. contains a column called GroupID and the morphometric variables ONLY.

The specimen ID's are not required to run the predictions and can be added back later when the prediction has run.

Make sure that the morphometric variables columns are in the same order as the model training data and are called by the same names.

Python scripts

All Python scripts have been developed using the Anaconda (version 2.4.0) distribution of Python (version 3.10.9), IPython (version 8.10) and Spyder IDE (version 5.4.1).

Python Script name	Usage
Contactsheet2.py	Pdf contact sheet of images
Ply2png.py	Convert .ply file to png image
Remesh_meshlab.py	Remesh point clouds
Snapshots.py	Takes snapshots of meshes
StackClipv1_3.py	Clips CT volume to sub-volumes

Table A3.2. Python scripts and usage

Contactsheet2.py

Reads in a directory containing PNG images and creates a pdf contact sheet with multiple scaled images per page.

The contact sheet size can be controlled by the variables sheet_width and sheet_height.

The individual images sizes on the contact sheet are controlled by the variables image_width and image_height.

Ply2png.py

Renders a directory of ply meshes to png images.

Remesh_meshlab.py

Uses the pymeshlab library to remesh a directory of point clouds (in ply format) to meshes. The script first computes all the normals for the point cloud, uses a Poisson filter to generate the surface and then force flips the resultant face orientations.

Snapshots.py

Reads a directory of meshes in ply format and creates a series of snapshots around each mesh at a series of pre-defined angles. The snapshots are saved as png images. The script scales each mesh to unit size before taking the images, firstly in xy, xz, yx, yz, zx and zy views and then every 15 degrees rotated around the x, y and z axes in turn. The camera positions and focal points can be adjusted if needed. Before running this

script it is advised to run the Matlab script AlignPlyDirectory.m to align the meshes to their long axis.

StackClipv1_3.py

Script designed to read an entire volume of CT data (as a directory of tiff files) containing multiple specimens and clip out each specimen into its own sub-volume of tiff files. The script is controlled by a csv file containing the specimen label and the min x, max y, min y, max y, min z, max z coordinates of the specimen within the original data volume.

Matlab scripts

Tested on Matlab R2023a

Matlab Script name	Usage
AlignPlyDirectory.m	**Aligns mesh to major axis**

Table A3.3. Matlab scripts and usage

AlignPlyDirectory.m

Used to align a series of meshes, in ply format, to their major axis. Script is used prior to generating snapshots of meshes using the Snapshots.py script.

Appendix Four: Institutional abbreviations

AMNH, American Museum of Natural History, New York, USA

DUGF, University of Delhi, Geology Department, Delhi, India.

FMNH, Field Museum of Natural History, Chicago, USA

GLRCM, Museum of Gloucester, Gloucester, UK

IVPP, Institute of Vertebrate Paleontology and Paleoanthropology, Beijing, People's Republic of China

LACM, Natural History Museum of Los Angeles County, Los Angeles, USA

MBLUZ, Museo de Biología de la Universidad del Zulia, Zulia, Venezuela

MNA, Museum of Northern Arizona, Flagstaff, USA

MOZ, Museo Olsacher Zapala, Zapala, Argentina

MPCD, Institute of Paleontology and Geology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia

MPEF, Museo Egidio Feruglio, Trelew, Argentina

MSNM, Museo di Storia Naturale di Milano, Milan, Italy

MTM, Magyar Természettudományi Múzeum, Budapest, Hungary

NCSM, North Carolina Museum of Natural Sciences, Raleigh, USA

NHMUK, Natural History Museum, London, UK

OUMNH, Oxford University Museum of Natural History, Oxford, UK

ROM, Royal Ontario Museum, Toronto, Canada

TMP, Royal Tyrrell Museum of Palaeontology, Drumheller, Alberta, Canada

UALVP, Laboratory for Vertebrate Paleontology, University of Alberta, Edmonton, Canada

UMNH, Natural History Museum of Utah, University of Utah, Salt Lake City, USA

ZDM, Zigong Dinosaur Museum, Dashanpu, People's Republic of China

ZIN PH, Zoological Institute of the Russian Academy of Sciences, St. Petersburg, Russia

Appendix Five: Machine learning results

All results for this thesis are available on GitHub at:

https://github.com/simonwills/Bathonian_Dinosaurs

A summary and brief description of each table is given below.

UK Bathonian data

Isolated teeth from Hornsleasow, Woodeaton, Kirtlington and Watton Cliff

Dataset	**Description**
UK Bathonian Random Forest	Random Forest classification results. File contains the specimen numbers (column B), The machine learning model (column C), the class assigned by the model (column D), the posterior probabilities for each class (columns E to V)
UK Bathonian MDA	Mixture Discriminant Analysis classification results. File contains the specimen numbers (column B), The machine learning model (column C), the class assigned by the model (column D), the posterior probabilities for each class (columns E to V)
UK Bathonian C50	C50 classification results. File contains the specimen numbers (column B), The machine learning model (column C), the class assigned by the model (column D), the posterior probabilities for each class (columns E to V)
UK Bathonian Post Probabilities	Combined classification results. File contains the combined machine learning results from all three models. Specimen number (column A), locality (column B), assigned class (column C), majority vote class (column D), combined posterior probability class (column E), maximum combined posterior probability (column F)
UK Bathonian MDA functions	First six MDA functions derived from the MDA classification. Specimen number (column B), machine learning model (column

	C), assigned class (column D), MDA functions (columns E to J)

Kyrgyzstan data

Analysis of isolated theropod teeth from Kyrgyzstan.

Balabansai Formation (Bathonian, Middle Jurassic), Averianov et al. (2005).

Dataset	**Description**
KYG_C5	C5.0 classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the C5.0 classification result (column I), the posterior probabilities for each class (columns J to AF)
KYG_MDA	Mixture discriminant analysis classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the MDA classification result (column I), the posterior probabilities for each class (columns J to AF)
KYG_RF	Random Forest classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the random forest classification result (column I), the posterior probabilities for each class (columns J to AF)
KYG_Comb_PP	Combined classification results. Specimen number (column B), MDA, RF and C5.0 classifications (columns C to E), majority vote classification (column F), combined posterior probabilities of all three classifiers (columns H to AD), combined posterior classification (column AE), maximum posterior probability (column AF)

Indian Data

Analysis of isolated theropod teeth from India

Kota Formation (Bathonian, Middle Jurassic, India) isolated theropod teeth (Prasad & Parmar, 2020)

Dataset	Description
India_C5	C5.0 classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the C5.0 classification result (column I), the posterior probabilities for each class (columns J to AF)
India_MDA	Mixture discriminant analysis classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the MDA classification result (column I), the posterior probabilities for each class (columns J to AF)
India_RF	Random Forest classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the random forest classification result (column I), the posterior probabilities for each class (columns J to AF)
Indian_results	Combined classification results. Specimen number (column B), MDA, RF and C5.0 classifications (columns C to E), majority vote classification (column F), combined posterior probabilities of all three classifiers (columns H to AD), combined posterior classification (column AE), maximum posterior probability (column AF)

Madagascar Data

Analysis of isolated theropod teeth from Madagascar

Sakaraha Formation (Bathonian, Middle Jurassic, Madagascar) isolated theropod teeth, Maganuco et al. (2005)

Dataset	**Description**
Madagascar_C5	C5.0 classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the C5.0 classification result (column I), the posterior probabilities for each class (columns J to AF)
Madagascar_MDA	Mixture discriminant analysis classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the MDA classification result (column I), the posterior probabilities for each class (columns J to AF)
Madagascar_RF	Random Forest classification results. File contains the specimen numbers (column C), original morphometric data (Columns D to H), the random forest classification result (column I), the posterior probabilities for each class (columns J to AF)
Madagascar_results	Combined classification results. Specimen number (column B), MDA, RF and C5.0 classifications (columns C to E), majority vote classification (column F), combined posterior probabilities of all three classifiers (columns H to AD), combined posterior classification (column AE), maximum posterior probability (column AF)