Corresponding author:
Stephanie R. Killingsworth;
Email: skillingsworth@floridamuseum.ufl.edu

PALEOBIOLOGY | Paleontological
A PUBLICATION OF THE | SOCIETY

CAMBRIDGE
UNIVERSITY PRESS

Species occurrences of Mio-Pliocene horses (Equidae) from Florida: sampling, ecology, or both?

Stephanie R. Killingsworth[1] and Bruce J. MacFadden[2]

[1]Florida Museum of Natural History and Department of Geological Sciences, University of Florida, Gainesville, Florida 32611, U.S.A.
[2]Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, U.S.A.

Non-technical Summary

During the late Miocene and early Pliocene about 5.7 to 4.75 million years ago, a distinctive suite of four species of extinct horses (Family Equidae) were widespread in North America. This includes *Nannippus aztecus*, *Neohipparion eurystyle*, *Astrohippus stocki*, and *Dinohippus mexicanus*. In Florida, two additional horse species, *Pseudhipparion simpsoni* and *Cormohipparion emsliei*, are also typically found. Here we compare horses from four Florida fossil sites of this age, including three from the Bone Valley mines and a fourth from the recently discovered Montbrook site. Two of these sites have all six predicted species, one has five species, and one has only four species present. To explain these differences, we used species occurrences from research databases to better understand the relative abundances, species richness, and occurrences of these horses from these four sites. One site (Palmetto Mine Agrico), with five equid species, appears to lack the sixth species owing to ecological reasons. This is different from Montbrook, the site with only four of the six species. Results indicate that Montbrook is likely missing the two horse species for a couple of reasons: sampling bias and biological/ecological causes. Our results demonstrate that fossil sampling biases can account for observed horse species occurrences when the overall abundance of certain species is low. Nevertheless, other factors, including ecology and with sufficient resolution, perhaps also time, may also explain the distribution and occurrences of individual horse species at these and other fossil sites.

Abstract

During the late Miocene and early Pliocene (latest Hemphillian, Hh4 interval, 5.7 to 4.75 Ma) a distinctive suite of four species of extinct horses (Family Equidae) were widespread in North America. This includes *Nannippus aztecus*, *Neohipparion eurystyle*, *Astrohippus stocki*, and *Dinohippus mexicanus*. In Florida, two additional equid species, *Pseudhipparion simpsoni* and *Cormohipparion emsliei*, are also typically found at Hh4 localities. Here we compare horses from four Hh4 Florida fossil sites, including three from the Bone Valley mines and a fourth from the recently discovered Montbrook site. Two of these sites have all six expected species, one has five species, and one has only four species present. To explain these differences, we used species counts from research databases and rarefaction simulation to clarify the relative abundances, species richness, and occurrences of these horses from these four sites. The Palmetto Mine (Agrico) site, with five equid species, appears to lack the sixth species owing to ecological reasons. This is different from Montbrook, the site with only four of the six species. Results indicate that Montbrook is likely lacking two missing equid species for a couple of reasons: sampling bias and biological/ecological causes. Our results demonstrate that sampling biases can account for observed equid species richness when the overall abundance of certain equid species is low. Nevertheless, other factors, including ecology and with sufficient resolution, perhaps also time, may also explain the distribution and occurrences of individual species at these and other fossil sites. In a broader perspective, analyses such as this example provide an opportunity to address a persistent challenge in paleontology, that is, how do we explain absences of extinct taxa from the fossil record?

Introduction

Fossil horses (Family Equidae) underwent explosive evolution during the Neogene, resulting in diverse species that are widespread in North America (Simpson 1951; MacFadden 1992; Cirilli et al. 2022). In Florida, these are represented by three-toed (tridactyl) and one-toed (monodactyl) clades with high-crowned (hypsodont) teeth that are found at many localities. Here we investigate differences in equid species occurrences at four relatively well-sampled localities from the Mio-Pliocene of Florida in order to determine what factors are responsible for the observed patterns of diversity. These localities have been extensively collected and both

occur within the latest Hemphillian North American Land Mammal Age (NALMA; designated Hh4), an interval between about 5.7 and 4.75 Ma (Tedford et al. 2004; Hulbert 2016a). Two localities from the Bone Valley region, Kingsford Mine and Fort Green South Mine, have six equid species, whereas the Palmetto Mine (Agrico) site, also from the Bone Valley region, has five equid species, and another locality in north-central Florida, Montbrook, has four equid species (UF VP Collections Database 2023; see also either iDigBio [https://www.idigbio.org] or GBIF [https://www.gbif.org]).

This study examines why different numbers of equid species occur at these Hh4 localities in Florida. Specifically, as compared with the two Bone Valley localities with all six equid species, is the presence of only five species at Palmetto Mine (Agrico) and only four species at Montbrook an artifact of sampling bias or an indicator of the alpha-level diversity of the horses that existed during the Hh4 interval? In this study, we develop a predictive model to explain these different occurrences of Hh4 fossil horse species from key sites in this region. Occurrence data of extinct species in the fossil record are fundamental to interpreting foundational concepts in paleontology. In particular, knowledge of species' occurrences in space and time are important to reconstruct past biogeographic and macroevolutionary patterns. While the presence of a particular species at a given site or geological horizon is solid evidence of occurrence (unless taphonomically reworked), the absence of a species cannot be, at face value, taken as evidence that it did not actually exist there. Therefore, the absence of a species' occurrence is oftentimes equivocal as to its interpretation, but this represents potentially also important data to understand (NCSE 2007; Burnham 2008; Currie and Turner 2017). Realizing the importance of "absent occurrences" in the fossil record, many studies have tried to understand the probability that a particular fossil is absent from a locality, and whether it actually existed "on the paleo-landscape" (e.g., Nowak et al. 2000). Likewise, in our study presented here, we use rarefaction analysis (Oksanen et al. 2022; R Core Team 2022) to help us better understand absences of Hh4 equid taxa in the Mio-Pliocene of Florida. This model statistically predicts what equid species would ultimately occur at individual sites, particularly those that are less well sampled through collecting in the field.

Another persistent challenge in paleontology involves sampling bias, because portions of the past are either not preserved in the fossil record or are not sufficiently sampled to recover diversity patterns that actually existed at the time when fossils were deposited. Classic rarefaction techniques originally were used in the field of ecology (Heck et al. 1975), but Raup's (1975) early rarefaction studies transformed our perspective in paleobiology. In addition, great strides have recently been made in quantitative analyses that allow us to better understand potential sampling biases in the fossil record (Alroy et al. 2001; Bush et al. 2004; Alroy 2010; Chao and Jost 2012; Shimadzu 2018; Bokulich 2021) as well as in modern ecological studies (Chao et al. 2009, 2014). We assert that the results documented here also provide generalizable insights that may potentially inform similar paleoecological studies in the future.

Background and Conceptual Framework

Geological and Paleontological Setting

In this study, we focus on two of the largest (in terms of number of specimens recovered) and important Hh4 vertebrate fossil localities in eastern North America, that is, the Bone Valley region with more than 50,000 cataloged specimens (Hulbert 2015) and the Montbrook fossil locality with more than 73,000 cataloged specimens (Hulbert 2016b), both housed in the Florida Museum of Natural History (FLMNH) research collections (UF VP Collections Database 2023).

Bone Valley Region

The Hh4 mammals (sometimes also called the Palmetto Fauna [Hulbert 2015]) of the Bone Valley series of deposits are found within extensive open-pit mines of the Central Florida Phosphate District, which comprise a large area (~2000 km^2) spanning four Florida counties (Hillsborough, Polk, Manatee, and Hardee). East of Tampa and southwest of Orlando, this region is roughly centered at about latitude 27.5°N and longitude 82°W (Fig. 1). The Hh4 Bone Valley fossil vertebrates consist of about 85 sublocalities that are derived from the late Miocene and early Pliocene Upper Bone Valley Formation (also referred to as the Hawthorn Group, Peace River Formation, or Bone Valley Member [Scott et al. 2001]). This formation consists of primarily siliciclastic sediments with economically important minerals, including phosphates; the discovery of the fossils was a by-product of the mining operations. The depositional conditions have been interpreted to represent shallow-water marine environments, including low-energy bays and high-energy channels (Hulbert 2015).

For this study we selected three individual Hh4 localities from the late Miocene to early Pliocene Upper Bone Valley Formation. These include (Fig. 1, Supplementary Fig. 2):

1. The Kingsford Mine (FLMNH VP sites PO0018, PO034) occurs within T31S, R23E, Sec. 3, and consists of a 2.6 km2 (1 square mile) area with a total of 328 cataloged equid specimens identified to species.
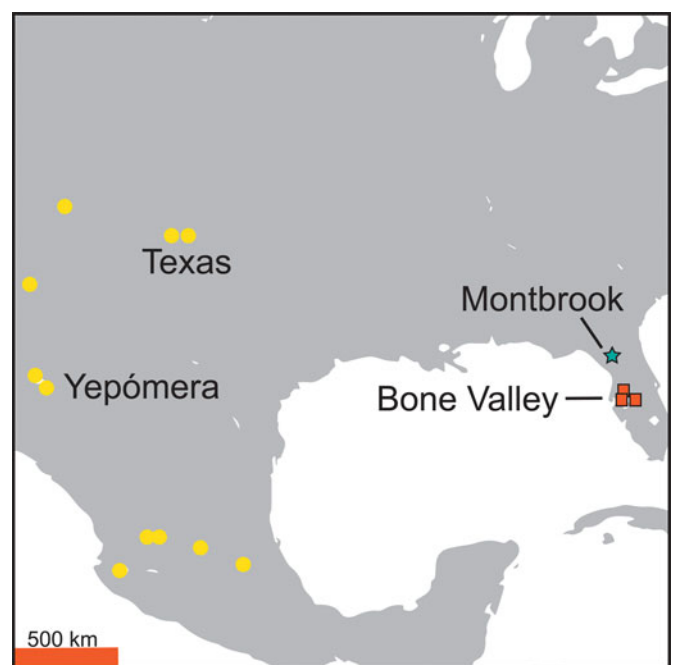

Figure 1. Map of a portion of southern North America showing representative fossil vertebrate localities that contain the four widespread species of late latest Hemphillian, Hh4 interval horses. For Florida, this includes the three Bone Valley sites (orange rectangles) and Montbrook (green star).

2. The Fort Green Mine South (FLMNH VP sites HR008, HR008b) occurs within T34S, R23E, Sec. 4–9, and consists of a 15.6 km2 (6 square mile) area with a total of 204 cataloged equid specimens identified to species.
3. The Palmetto Mine (Agrico) (FLMNH VP site PO001) occurs within T32S, R24E, Sec. 9, 10, 15, 16, 21, 22, and also consists of a 15.6 km2 (6 square mile) area with a total of 462 cataloged equid specimens identified to species.

These three Palmetto Fauna Bone Valley localities were selected because they have large numbers of taxonomically assignable (to genus and species) equid specimens and sample all, or the majority, of the equid species predicted to be present during Hh4 in this region.

During the second half of the twentieth century, field crews from the FLMNH, a myriad of amateur fossil collectors, fossil clubs, other museums, and former UF graduate student, John Waldrop, collected fossil vertebrate specimens from the Bone Valley region. Notably, Waldrop amassed a large private collection known as the Timberlane Research Organization (TRO) consisting of ∼20,000 specimens that he later donated to the FLMNH (Hulbert 2015).

While it is impossible to estimate the total collecting effort in the Bone Valley region during the second half of the twentieth century, it is safe to say that many tens of thousands, if not hundreds of thousands of hours, were expended. These efforts have resulted in more than four dozen species of terrestrial mammals identified from the Hh4 Palmetto Fauna (Hulbert 2015) of the Bone Valley region, mostly with an emphasis on larger taxa that are typically found through surface prospecting. In contrast, relatively little systematic screenwashing has been done, and therefore the microfauna is underrepresented.

Based on diagnostic mammals known to be from the latest Hemphillian (Hh4, NALMA) at localities throughout North America (Tedford et al. 2004), the Upper Bone Valley Formation falls within the Hh4 time interval via biochronology. Diagnostic Hh4 mammals include the gomphothere proboscidean *Rhynchotherium edense*, llamine camelid *Hemiauchania edensis*, peccary *Protherohyus* (sensu Prothero 2021) *brachydontus*, and as will also be described below, the horses *Dinohippus mexicanus* and *Pseudhipparion simpsoni* (Hulbert 2015).

Montbrook

The Montbrook locality (FLMNH VP site LV070) was discovered in late 2015 by the landowner, who brought petrified bones to the attention of the FLMNH. It is located on private property in a sand pit south of Williston, Florida, in rural Levy County, about 40 km southwest of Gainesville and about 200 km northwest of Bone Valley (Fig. 1, Supplementary Fig. 2). While the exact location is contained in the FLMNH collection files, all specimens have been curated into the research collections under a formal agreement with the landowner. Fossils from this site are recovered from a 3-m-thick section of siliciclastic fine sand, silts and clays interpreted to be reworked from the Hawthorn Formation. The environment of deposition is interpreted to represent a fluvial system near the ancient coastline (Hulbert 2016b). In contrast to the Upper Bone Valley faunas, which contain 16 taxa of marine mammals (whales, dolphins, and seals), no fully marine mammal species are yet known from Montbrook (Hulbert 2015, 2016b). According to Hulbert (2016b), the Montbrook vertebrate assemblage is latest Hemphillian in age, with Hh4 indicator taxa,

including the rhinoceros *Teleoceras*, gomphothere proboscidean *Rhyncotherium*, and canid *Borophagus*. It also preserves two other Hh4 taxa (Hulbert 2015), the llamine camelid *Hemiauchenia edensis* and the horse *P. simpsoni*.

With the exception of the COVID years 2020–2022, which saw only sporadic collecting, intensive excavations began in early 2016 and are ongoing. Fieldwork is generally done several days per week for 5 to 6 months during the fall, winter, and early spring. Excavations are done by FLMNH staff, students, and hundreds of volunteers annually. In addition, several of the devoted volunteers also assist in the vertebrate paleontology preparation laboratory cleaning fossils, performing general curation, and inputting data into the research collection digital database (UF VP Collections Database 2023). Up to now, we have concentrated our collecting efforts on a 400 m^2 area within the sand pit. Macrofossils are excavated from 1 m squares mapped into the locality, and these coordinates are recorded with the specimen. Although the exact number would be difficult to determine, many tens of thousands of person-hours have gone into the collection of the Montbrook fossils since 2016.

Of the ∼73,000 fossils cataloged from Montbrook, the overwhelming majority (43%) are turtles. Mammals are represented by over three dozen taxa, including families with multiple species. The database identifies >2800 mammal specimens to the individual species level (UF VP Collections Database 2023). Montbrook also contains a rich microfauna with numerous vertebrate taxa; the study of this fossil fraction offers huge potential to recover other smaller-sized species in the future. With regard to this study, horses (Family Equidae) are represented by >250 individual specimens, of which 107 are identifiable to species (Supplementary Figs. 1, 2).

Latest Hemphillian Horses (Family Equidae) of North America

During the latest Hemphillian, from about 5.7 to 4.75 Ma, a distinctive assemblage of horses inhabited much of southern North America, with large, representative collections known from Mexico, Texas, and Florida (Tedford et al. 2004; MacFadden 2008; Hulbert 2015; Fig. 2). Four of these equid species were widespread: the tridactyl *Nannippus aztecus* and *Neohipparion eurystyle* and the monodactyl *Astrohippus stocki* and *Dinohippus mexicanus*. Two other equid species in Hh4 faunas, the tridactyl *Pseudhipparion simpsoni* and *Cormohipparion emsliei*, occur in southeastern North America, and the latter species has also been reported from presumed Hh4 sites in Kansas (Webb and Hulbert 1986) and Oklahoma (Hulbert et al. 2005). In addition to their presence in these faunas, the relative abundances of the four widespread species vary geographically, with one-toed horse fossils relatively rare in Florida, yet numerically dominant in the faunas from Texas and Mexico (MacFadden 1986, 2008). We will also see differences in relative numerical abundances of these horses at the Florida Hh4 sites selected here and will discuss this pattern in more detail later.

All of these Hh4 equid species were hypsodont (Fig. 3), that is, having high-crowned teeth with relatively flat occlusal surfaces adapted for crushing and grinding more abrasive vegetation such as grasses. These equids display hypsodonty indices > 1 (HI = ratio of unworn or little-worn crown height/anteroposterior width of occlusal surface). Based on stable isotope differences (δ^{13}C) of their teeth, the respective herbivore feeding niches of these horses is interpreted to represent variations in
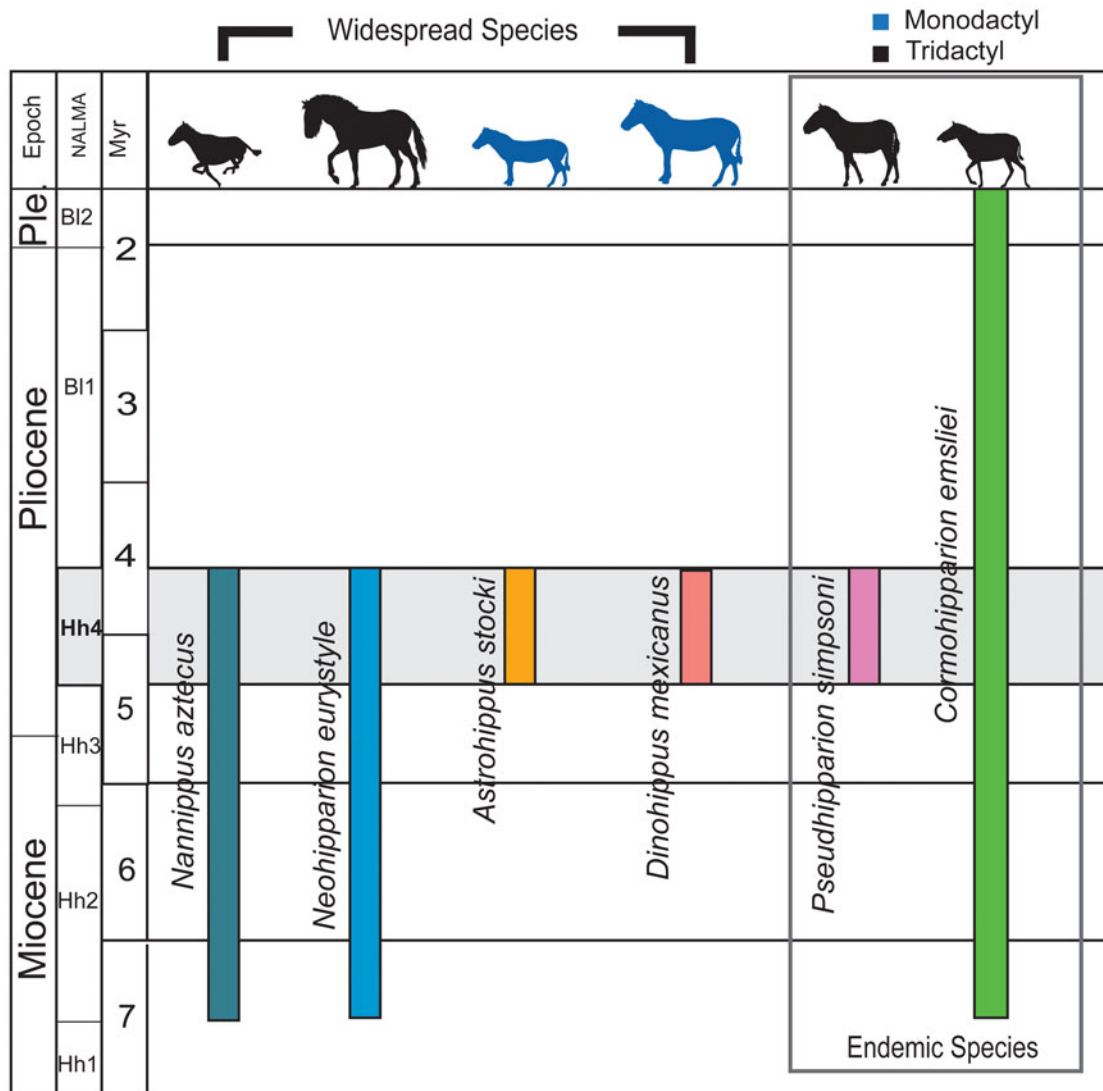
Figure 2. Geological range and North American Land Mammal Ages (NALMAs) of latest Hemphillian, Hh4 interval equid faunas in North America, including the four widespread and two endemic species.

the proportion of C_3 and C_4 plant foods that were consumed (MacFadden et al. 1999; MacFadden 2008). In addition, Webb and Hulbert (1986) also demonstrated that *P. simpsoni* developed a more advanced degree of hypsodonty, which they termed incipient hypselodonty, the latter referring to ever-growing teeth. Although hypselodonty is found in many clades of mammals (Damuth and Janis 2011), within the Equidae it is only found in *P. simpsoni*. MacFadden (2008) also showed that there was geographic variation in how the four widespread species varied their stable isotope niches in Florida, Texas, and northern Mexico.

Tooth enamel microwear is another tool used to analyze fossil herbivore diets, alone or along with other evidence, such as relative tooth hypsodonty and stable isotopes. MacFadden et al. (1999) described tooth microwear along with stable isotope analyses of Hh4 species from a sample of Bone Valley horses. More recently, Semprebon et al. (2016) provided a significantly more detailed and comprehensive study of the microwear of some of the Hh4 species studied here, including *N. aztecus*, *P. simpsoni*,

and *C. emsliei*. Their findings show that based on the microwear patterns, when compared with a large-scale microwear database of extant ungulates (Semprebon et al. 2016), these Mio-Pliocene horses were grazers or mixed feeders.

In the results presented, it will become important to understand the ancient ecologies of these coexisting equid species as a possible explanation for their relative abundance and occurrences in these four localities. Likewise, we will return to the earlier discussions about the interpreted paleoenvironments, as well as data from stable isotopes and microwear. In doing so, we will see how these additional data inform our analysis of the different species occurrences from the Bone Valley and Montbrook faunas from Florida.

Conceptual Framework: Sampling Bias and Florida Hh4 Horses

The issue of sampling bias has long been a challenge for realizing the true paleobiodiversity present in the fossil record (Sepkoski
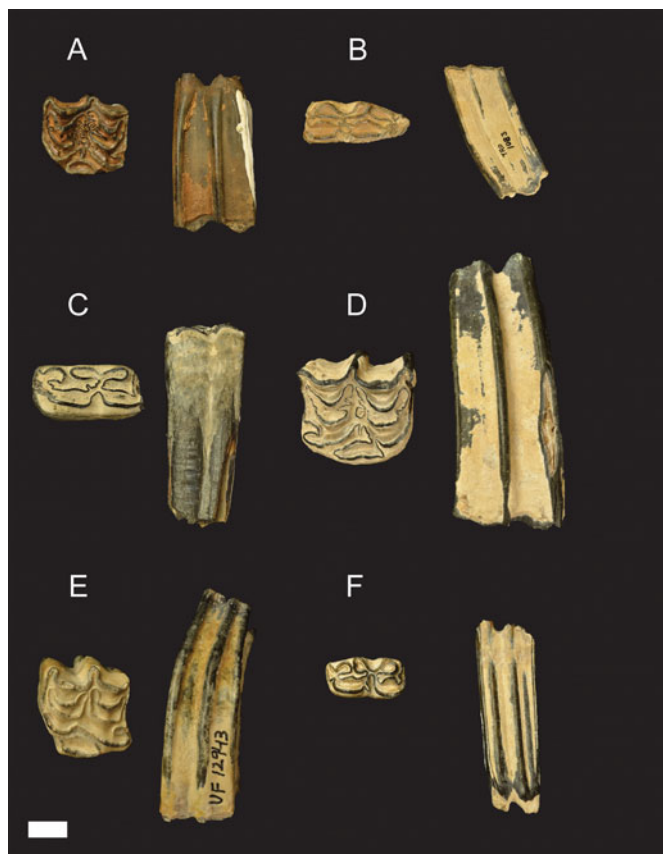
Figure 3. Representative specimens of the six species of latest Hemphillian, Hh4 interval horses studied in this report and retrieved from the UF VP Collections Database. (A) *Nannippus aztecus* (UF/TRO 50), LM1; (B) *Neohipparion eurystyle* (UF/TRO 1083), Lm3; (C) *Astrohippus stocki* (UF 220155), Rp4; (D) *Dinohippus mexicanus* (UF/TRO 1149), RM1; (E) *Pseudhipparion simpsoni* (UF 12943), LM1; (F) *Cormohipparion emsliei* (UF/TRO 861), Rm2. Abbreviations: R, right; L, left; P, upper premolar; M, upper molar; p, lower premolar; m, lower molar. Scale bar, 1 cm.

et al. 1981). Paleontologists seeking to understand this problem have consulted ecological sampling methods as analogs to account for species richness, or the number of taxa at a given site, as well as evenness, which is defined as the measure of relative abundance for each taxonomic group. In this study, we focus on quantitative assessment of species richness of horses, although we also will make general observations about evenness.

Consideration of differences in sampling effort among localities must be given before meaningful statistical tests on paleobiodiversity can be performed. Such differences may be present because of variables like the amount of time spent collecting and the overall size of the collection area, as well as geological influences on preservation, like weathering and erosion. Furthermore, comparisons among sites should consider analyzing similar taxa, time bins, and sites within proximity to one another. To correct for sampling variance among sites, a process that standardizes datasets to be compared, then replicates a subset of the data hundreds or even thousands of times via bootstrap simulation can be a useful tool. This approach, otherwise known as classic rarefaction (Raup 1975), is a robust option for addressing whether the presence or absence of taxa is an artifact of sampling or due to other factors such as ecological differences.

Some researchers suggest that rarefaction confounds issues of evenness and coverage, or the sum of the frequencies of the species sampled. To account for these aspects, previous studies (Alroy et al. 2001; Bush et al. 2004; Alroy 2010; Chao and Jost 2012; Bokulich 2021) introduced methods such as shareholder quorum sampling, arguing that results provide fairer sampling when different levels of evenness exist. In the case of rarefaction, Alroy (2010) questioned the method for its simplicity in treating samples fairly; however, Barnosky et al. (2005) emphasized that rarefaction only fails to recognize changes in species richness when localities being compared express varying diversities and total richness is not equal. Therefore, depending on the questions being asked, for example, the time intervals and amount of taxonomic diversity being studied, certain methods may be better than others. Shimadzu (2018) stresses that if the question being asked pertains to specific taxa within the community, then rarefaction will provide an accurate measure over coverage-based methods. Here we posit that in the case of comparing Hh4 horses from Florida, the rarefaction model is justified. Our justification is two-fold: (1) given the similarity in both age and representative taxa of the Bone Valley region sites to Montbrook, this outcome ensures a rational comparison between the larger and smaller sample sizes; and (2) the notion of a coverage deficit (Alroy 2010) is arguably addressed, as a rarefaction curve asymptotically plateaus, thereby conveying any coverage deficits from the sampling (Chao and Jost 2012).

Given the conceptual framework described, the question arises whether the horses that have been sampled from Kingsford Mine and Fort Green Mine South are good models from which to simulate sampling from the Palmetto Mine (Agrico) and Montbrook sites. While two fossil localities are almost never the same in all aspects, many important features of these Hh4 fossil sites justify comparison in our analyses. These include:

1. With >50,000 and >73,000 cataloged fossils, respectively, for the Bone Valley sites and Montbrook, it can be argued that these localities have been numerically well sampled after at least tens of thousands of person-hours of field collecting. These faunal assemblages are the largest of their kind in eastern North America and rival the largest Hh4 sites in western North America, that is, Yepómera from Chihuahua, Mexico (Lance 1950; Stirton 1954; Fig. 1).

2. The four sites studied here have diagnostic mammalian taxa that make them Hh4 in age (Tedford et al. 2004; Hulbert 2015, 2016a) and likely fall within an interval and duration of less than 1 Myr, between about 5.7 to 4.75 Ma (Fig. 2).

3. Preservation bias (Newell 1959) is not a factor in this analysis, because we are primarily comparing mineralized fossil horse teeth as evidence of species occurrences at the two localities. It is asserted that the relative preservability of the equid teeth is similar at both localities studied, because they fossilized in similar sedimentary environments.

4. Overall, the general number of terrestrial mammal species that occurs at both sites, at least four dozen from the Bone Valley sites and at least three dozen from Montbrook, suggest that medium- to large-bodied (i.e., non-microfaunal) components of each fauna, including some rare (e.g., carnivore) taxa are present at these localities.

5. These sites have yielded the widespread tridactyl horse species characteristically found at other Hh4 fossil localities in North America (Fig. 2), Neohipparion aztecus and Nannippus eurystyle. In addition, most of the Bone Valley sites also include two monodactyl species, Astrohippus stocki and Dinohippus mexicanus. They also record the presence of two additional species

from southeastern North America, *Pseudhipparion simpsoni* and *Cormohipparion emsliei*. Thus, despite some difference in the number of equid species at these sites that we analyze, the co-occurrence of the widespread species within the Bone Valley and Montbrook sites indicates the Hh4 biochron interval as defined by Tedford et al. (2004).

Materials and Methods

Database Queries

A comprehensive search of the UF VP Collections within the FLMNH database at the University of Florida (UF VP Collections Database 2023) was completed. The UF VP Collections Database is derived directly from the Specify Collections research database (Specify Collections Consortium 2023). The Specify-enabled UF VP Collections Database also forms the basis for biodiversity occurrence data within other well-known aggregators such as iDigBio (https://www.idigbio.org) and GBIF (https://www.gbif.org). While our research could have been conducted using either of the latter two aggregators, these are derivative and therefore provide no better accuracy in the search query results. As a research tool for this study, the UF VP Collections Database (2023) was the most accessible.

Concatenated searches for Family "Equidae" and faunal zone "Hh4" were conducted separately for Bone Valley Palmetto Fauna sites classified as Hh4, and then Montbrook (UF VP Collections Database 2023). For this study, we sought Bone Valley Palmetto Fauna sites with generally similar areas that also had large numbers of the Hh4 equid species. We therefore compared Bone Valley region sites that each reflected the presence of all six horse species (Kingsford Mine and Fort Green Mine South) to both Palmetto Mine (Agrico), a site with only five species, as well as Montbrook, located in northern Florida outside of the Bone Valley region (Fig. 1), with only four species represented.

Species Identification

The six equid species that can potentially occur at the sites studied here are confidently identified based on a suite of characters, including dental morphology and size. The dental morphology primarily includes differences in occlusal dental patterns, along with relative size (Lance 1950; Stirton 1955; MacFadden 1986). With regard to the latter, the relative size of some of the postcranials allowed unambiguous allocation of certain taxonomic identifications down to the species level.

Invariably, the level of taxonomic identification from the Family Equidae down to the species level could not be narrowed down for a number of the specimens that we reexamined due to their incomplete nature or lack of diagnostic characters. These were therefore not included in our analysis. All Bone Valley region samples are mostly represented by diagnostic teeth (95% of samples), in contrast to Montbrook with a total of 62 teeth and 45 assignable postcranial specimens for a total of 107 species-level confirmations from the original 252 specimens queried in the database search (Supplementary Fig. 3).

Quantitative and Model Analyses

As previously mentioned, we focus here on rarefaction, a resampling technique used to standardize all test sites being compared to examine relative abundance and, moreover, predictive occurrences of equid species. This is used to test our hypothesis that horse absences may be due to barriers that may be ecological or biological in nature. Previous studies indicate that rarefaction can be an ideal analytical tool for sites that share similar characteristics, for example, flora and fauna as well as chronology (Raup 1975; Bush et al. 2004). Placed in the current context of Hh4 horses from Florida, species occurrences from the better-sampled localities (Kingsford Mine and Fort Green Mine South) are used to simulate when rare species are predicted to occur from less-well sampled localities (Montbrook). Likewise, this method is used here to explain why species occurrences differ in well-sampled sites (Kingsford Mine and Fort Green Mine South) as compared with Palmetto Mine (Agrico). Data were analyzed using the well-established vegan package in R/RStudio via rarefaction by bootstrapping to simulate 1000 sampling efforts, ideal when computing to 95% confidence level or higher (Dixon 2003; R Studio Team 2020; Oksanen et al. 2022; R Core Team 2022).

The vegan package for R (Oksanen et al. 2022; R Core Team 2022) was used to read raw counts by species per site (Supplementary Figs. 1, 3) and assess the probability of collecting each of the six equid species from all of the Hh4 sites. Rarefaction simulations were run using the smaller footprint (2.6 km^2) Kingsford Mine site that accounts for all six equid species to subsample and compare with the two sites lacking all six equids (Palmetto Mine [Agrico] and Montbrook, 14.6 km^2 and 400 m^2, respectively). Another simulation for further comparison was made between Palmetto Mine (Agrico) and Montbrook to Fort Green Mine South, a second and slightly larger site (15.6 km^2 in size) possessing all six equids.

In the case of comparing Palmetto Mine (Agrico) with Kingsford Mine and Fort Green Mine South, we sampled 328 specimens (Palmetto Mine [Agrico]–Kingsford Mine) and 204 specimens (Palmetto Mine [Agrico]–Fort Green Mine South). Regarding the Montbrook simulations with Kingsford Mine and Fort Green Mine South, a sample total of 107 specimens was used in both cases, because Montbrook has the least number of specimens for all sites.

To simulate additional sampling efforts and minimize for statistical outliers, the abovementioned sampling regime was repeated 1000 times, each comparison using the *sapply* function to repeatedly call the *rarefy* function on the values in the sequence, producing counts with standard error (R Studio Team 2020; Oksanen et al. 2022; R Core Team 2022). The resultant rarefaction curves were plotted with lower and upper confidence intervals of 99% using the *ggplot* function in R (R Studio Team 2020; R Core Team 2022). For all R code and complete data output, see Supplementary Figures 4 and 5.

Results

Relative Abundances

Total raw specimen counts by fossil locality retrieved from the UF VP Collections Database indicate that six equid taxa were collected from the Kingsford Mine and Fort Green Mine South in the Bone Valley region. Raw specimen counts for Kingsford Mine (Fig. 4A) account for: *Nannippus aztecus* (*n* = 129), *Neohipparion eurystyle* (*n* = 103), *Astrohippus stocki* (*n* = 5), *Dinohippus mexicanus* (*n* = 14), *Pseudhipparion simpsoni* (*n* = 15), and *Cormohipparion emsliei* (*n* = 62); and for Fort Green Mine South (Fig. 4B): *N. aztecus* (*n* = 107), *N. eurystyle* (*n* = 50), *A. stocki* (*n* = 1), *D. mexicanus* (*n* = 9), *P. simpsoni* (*n* = 15), and
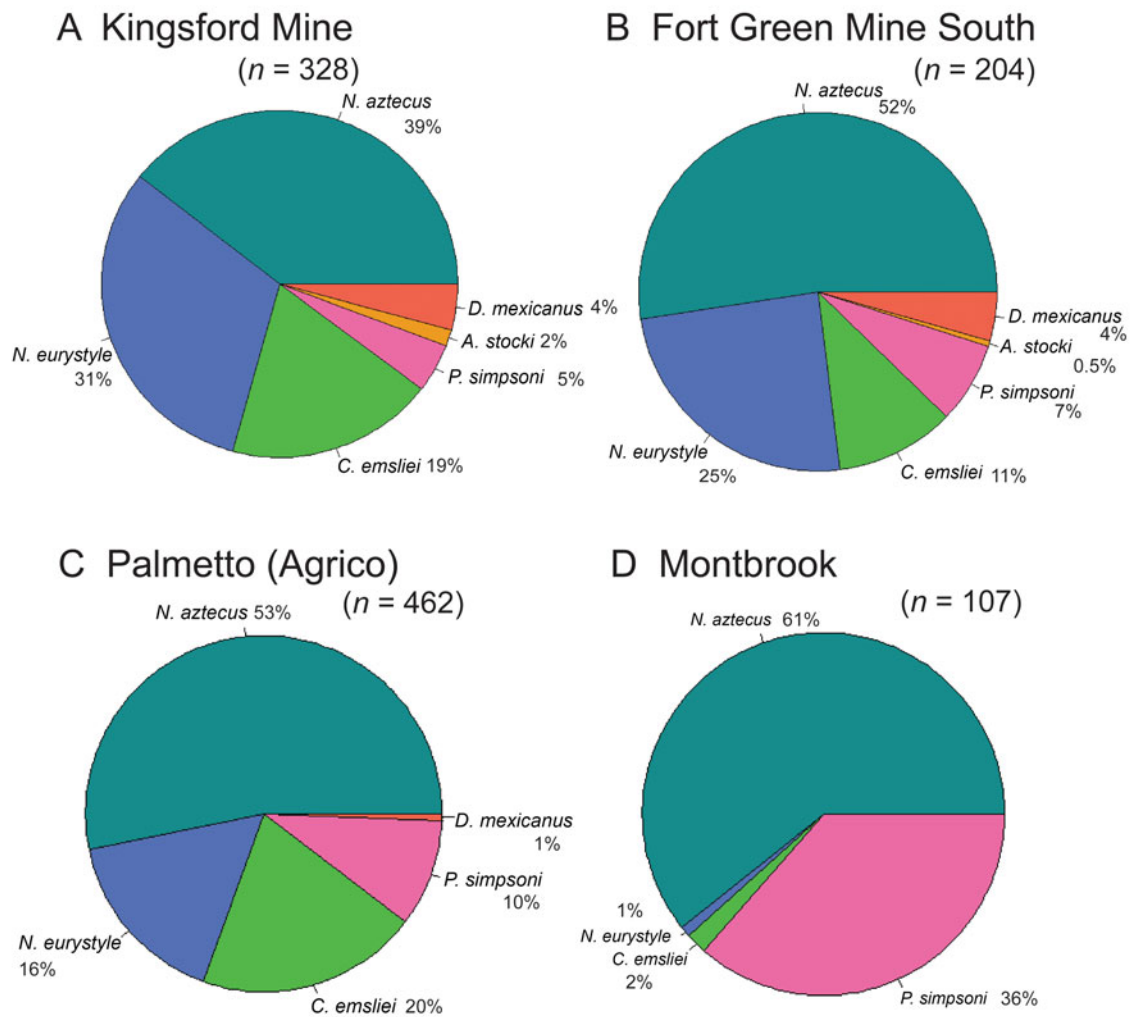
Figure 4. Pie charts showing the numerical abundances and percentages (cataloged raw specimen counts) of the different equid taxa (*Astrohippus stocki*, *Cormohipparion emsliei*, *Dinohippus mexicanus*, *Nannippus aztecus*, *Neohipparion eurystyle*, *Pseudhipparion simpsoni*) from the Kingsford Mine (A), Fort Green Mine South (B), Palmetto Mine (Agrico) (C), and Montbrook (D). These data were retrieved from the UF VP Collections Database (2023).

C. emsliei (*n* = 22). Whereas at Palmetto Mine (Agrico) (Fig. 4C), only five of the six equid species have been identified from the raw counts, including *N. aztecus* (*n* = 246), *N. eurystyle* (*n* = 75), *D. mexicanus* (*n* = 3), *P. simpsoni* (*n* = 45), and *C. emsliei* (*n* = 93). In further contrast, only four of the six taxa have been identified and collected from Montbrook (Fig. 4D), including *N. aztecus* (*n* = 65), *N. eurystyle* (*n* = 1), *P. simpsoni* (*n* = 39), and *C. emsliei* (*n* = 2). These results are also supplied in table format (Supplementary Fig. 3).

These results indicate that by 13, 34, and 309 sampling attempts, the fourth, fifth, and sixth horse species could be expected to occur in the Palmetto Mine (Agrico) using the Kingsford Mine for simulation (Fig. 5A, Table 1A). Whereas it would take 16, 42, and 378 sampling attempts when comparing Palmetto Mine (Agrico) using Fort Green Mine South (Fig. 5B, Table 1B). The Montbrook–Kingsford Mine (Fig. 5C, Table 1C) simulation indicates that by only 13 sampling events, a total of four taxa could be expected. With continued addition of sampling

attempts, the output suggests that by 34 attempts, the fifth species would be present, and by 309 attempts, we could expect the sixth species to be found. Regarding Montbrook, using the Fort Green Mine South simulation (Fig. 5D, Table 1D), sampling attempts produce the fourth, fifth, and sixth taxa at 14, 39, and 204 attempts, respectively. Rarefaction and sampling are calculated at the 99% confidence interval for these results (Fig. 5, Table 1), and the complete counts for sampling attempts for each comparison can be viewed in Supplementary Figure 5.

Sampling attempts were plotted using the *ggplot* function in R to produce a species richness curve for Hh4 horses of Florida (R Studio Team 2020; R Core Team 2022; Fig. 5). Figure 5 shows a mean asymptotic line with shading for the lower confidence limit (LCL) and upper confidence limit (UCL) calculated for 99% confidence.

Figure 5 displays the respective rarefaction curves, a measure of species richness, and the predicted occurrences of the fourth, fifth, and sixth equid species for each test site. For Palmetto
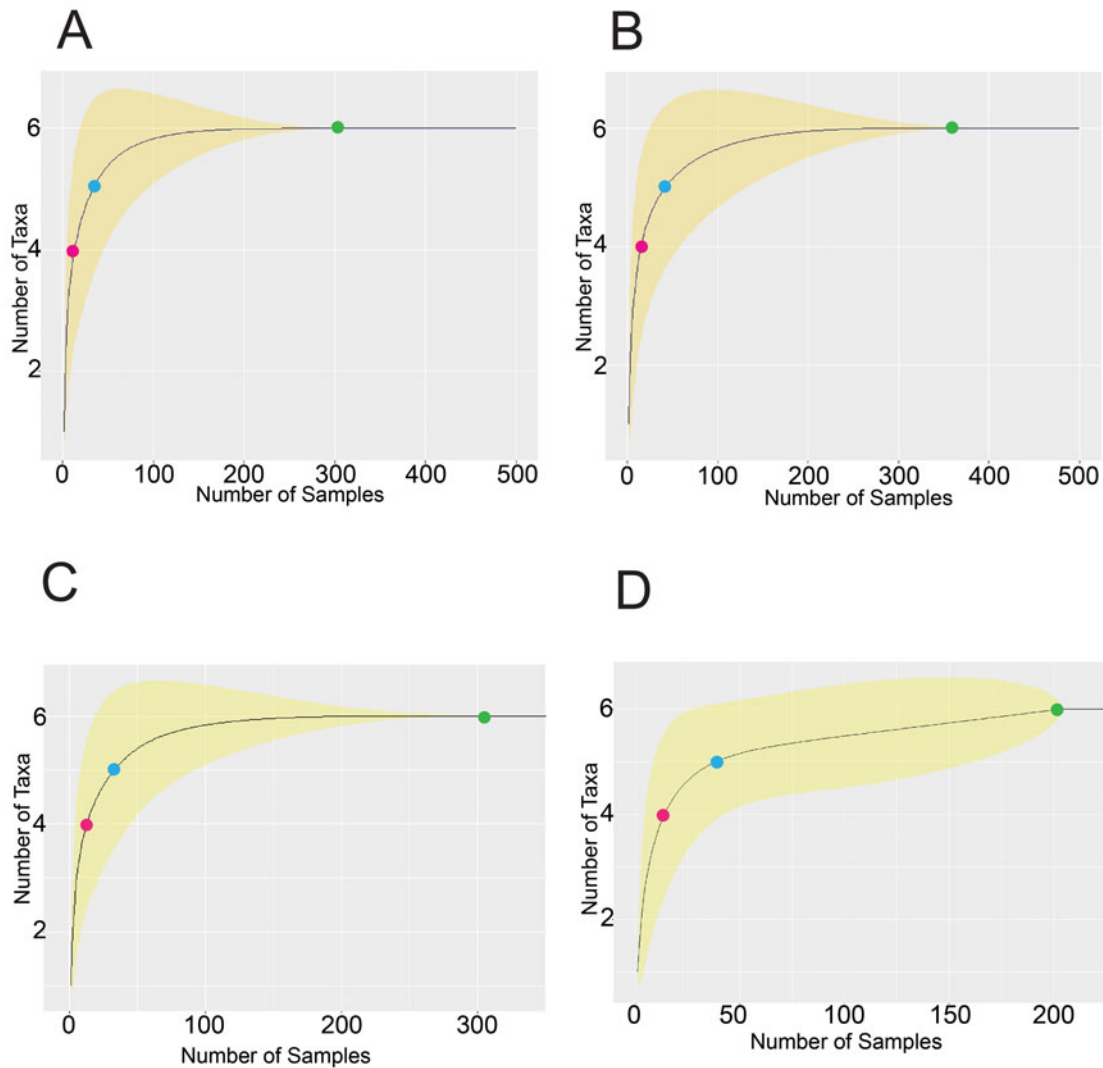
Figure 5. Rarefaction curves for species richness of equids from latest Hemphillian, Hh4 interval localities in Florida. Palmetto Mine (Agrico)–Kingsford Mine (A), Palmetto Mine (Agrico)–Fort Green Mine South (B), Montbrook–Kingsford Mine (C), and Montbrook–Fort Green Mine South (D). The x-axis specifies the level of sampling, and the y-axis indicates the predicted number of equid species. Colored sample points simulate the first occurrences of fourth species (red), fifth species (blue), and sixth species (green). Shading indicates 99% lower and upper confidence limits.

Mine (Agrico), using two comparative model simulation samples (Kingsford Mine and Fort Green Mine South; Fig. 5A,B, Table 1), a Wilcoxon test comparison shows no significant difference in predicted occurrences between the sites ($p = 1.0$). Similarly, Figure 5 illustrates the same prediction of occurrences for the fourth, fifth, and sixth equid species regarding Montbrook using the Kingsford Mine and Fort Green Mine South simulations (Fig. 5C,D, Table 1), with a Wilcoxon test producing a p-value of 0.3711 that also indicates no statistically significant difference in predicted occurrences in the comparative results.

Discussion

Relative Abundances

As depicted in Figure 5, the Hh4 equid species from Florida have different relative abundances in the four study localities. Although these numerical counts are not as precise a representation as the minimum number of individuals, for example, they nevertheless provide a general, or first-order, approximation of relative abundances and certainly presence/absence (occurrences, also species richness) worthy of discussion here.

Nannippus aztecus represents the largest numerical abundances and percentages at all four fossil localities at 39% (Kingsford Mine), 52% (Fort Green Mine South), 53% (Palmetto Mine [Agrico]), and 61% (Montbrook) (Fig. 4). *Neohipparion eurystyle* represents the second largest abundance and percentages for both Kingsford Mine (31%) and Fort Green Mine South (25%), while *Cormohipparion emsliei* ranks second for Palmetto Mine (Agrico) (20%) and *Pseudhipparion simpsoni* for Montbrook (36%). For all sites, *Dinohippus mexicanus* and *Astrohippus stocki* are least common, with percentages ranging from 0% to 4%.

Table 1. Model simulation of sampling events required to collect, respectively, four, five, and six species of equid taxa from the four Hh4 Hemphillian, Hh4 interval localities from Florida studied here. The Bone Valley sites (Kingsford Mine and Fort Green Mine South) with all six horse species are compared with Palmetto Mine (Agrico) (A, B) and Montbrook (C, D). All are calculated at 99% confidence levels. See Supplementary Fig. 5 for raw data.

Sampling attempts	Number of taxa	Standard deviation	Lower confidence limit (LCL)	Upper confidence limit (UCL)
A. Palmetto Mine (Agrico)–Kingsford Mine				
13	4.015315	0.804016	2.439473	5.591157
34	5.027459	0.737011	3.582943	6.471974
309	6	0.000616	5.998793	6.001206
B. Palmetto Mine (Agrico)–Fort Green Mine South				
16	4.058881	0.788337	2.513769	5.603992
42	5.000379	0.698823	3.630712	6.370046
378	6	0.000593	5.998838	6.001161
C. Montbrook–Kingsford Mine				
13	4.015315	0.804016	2.439473	5.591157
34	5.027459	0.737011	3.582943	6.471974
309	6	0.000616	5.998793	6.001206
D. Montbrook–Fort Green Mine South				
14	4.009486	0.795886	2.449579	5.569393
39	5.005773	0.558761	3.910621	6.100925
204	6	0	6	6

Assuming that these numerical counts accurately reflect relative abundances, as well as presence and/or absence of particular species, then our findings bring into question whether the differences in these occurrences result from ecogeography or sampling bias. Although sampling bias is modeled by the rarefaction analysis, the following comments about ecogeography are also relevant here.

Ecogeography

As noted earlier, although the fossil localities selected here have been intensively sampled, Montbrook is a single quarry that has yielded >73,000 cataloged specimens from a sand pit comprising an area of only about 400 m^2. In contrast, the Bone Valley localities comprising 2.6 km^2 (Kingsford Mine), 15.6 km^2 (Fort Green Mine South), and 15.6 km^2 (Palmetto Mine [Agrico]) have been intermittently collected (>50,000 cataloged specimens) (Supplementary Fig. 2). Species–area size and diversity theory (MacArthur and Wilson 1967; Barnosky et al. 2005; Sepkoski 2005) applied to paleontology might predict therefore that the greater areas of each of the Bone Valley localities relative to Montbrook may also have supported more habitats and species.

Taking another line of evidence, based on stable isotope ecology, MacFadden et al. (1999; also MacFadden 2008) demonstrated that the six equid species from Bone Valley region have significantly different carbon isotope ($\delta^{13}C$) values, indicating habitat and dietary differentiation for the coexisting horses. Likewise, using enamel microwear, Semprebon et al. (2016) found dietary differences in the species *N. aztecus*, *P. simpsoni*, and *C. emsliei* that indicate different proportions of mixed feeding and grazing in these species.

These different lines of ecogeographic evidence, collecting area, and dietary niches would predict that taken together, the Bone Valley localities would have a greater species richness than Montbrook (see also Barnosky et al. 2005). Our results, however, only partially support this inference. Whether or not ecogeography played a major part in the different numbers of species that occur at the four Florida Hh4 sites is not necessarily informative, unless paired with results from rarefaction analysis.

Rarefaction Analysis

The rarefaction results produced from this study highlight that an either-or scenario of sampling bias/ecogeography may be too simplistic an explanation for the absence of *A. stocki* at Palmetto Mine (Agrico) and *D. mexicanus* and *A. stocki* at Montbrook. Raw counts from the simulation further demonstrate the complex story of horses in Florida during the latest Hemphillian. The simulations indicate that for each of the comparisons (Table 1), the fifth equid taxon would be present by the 34th and 42nd sampling for Montbrook and 34th and 39th sampling for Palmetto Mine (Agrico) collection attempts. This is informative, because Montbrook, with the number of horse specimens collected at 107 species-identified samples, suggests the probability of collecting the fifth equid species should have occurred by this level of sampling. It therefore implies an actual absence of the fifth horse species at Montbrook when this fauna existed during the Hh4. While the simulation supports the absence of the fifth equid species at Montbrook, it conveys different results for the sixth equid species. Rarefaction simulations do not predict the sixth species occurrence until 204 and 309 sampling attempts (Table 1). This simulated number of sampling events far exceeds the smaller sample size of 107 specimens currently recorded at Montbrook, thereby suggesting that the sixth equid species could be absent due to either sampling bias or ecological reasons.

We also note that the Palmetto Mine (Agrico), our other test site for simulation and overall largest in sample size of the four localities studied here, contains three *D. mexicanus* specimens from the 462 equid specimens collected, accounting for <1%

of the total. Using the rarefaction simulations, the expected fifth equid species for this site would occur by 34 and 42 sampling attempts, and furthermore, we would predict that multiple *D. mexicanus* specimens would be present based on this probability. However, this lower number of *D. mexicanus* suggests a paleobiological influence driven, for example, by habitat changes.

Regarding the absence of the sixth species, that is, the rarest species, *A. stocki*, at Palmetto Mine (Agrico), the model suggests that it would occur by 204 (Fort Green South Mine) and 309 (Kingsford Mine) sampling attempts. Thus, regarding Palmetto Mine (Agrico), which has 462 specimens collected, the sixth species should statistically be present based on model predictions, but it is not. The absence of *A. stocki* is therefore best explained by the actual lack of this species in this region during the late Hemphillian. Collectively, the influence of biological and ecological drivers for these differences between the Hh4 sites is supported by the rarefaction simulations, thus providing an explanation for the equid fossil record for all sites.

Given the size of each of the sites and their paleogeography, as well as the ecological analyses of these horses, it is natural to ask whether we can distinguish from our predictive models which of the two missing taxa at Montbrook is *A. stocki* and which is *D. mexicanus*. It seems reasonable to infer, given our current knowledge base, that *D. mexicanus* would be the fifth species to occur at Montbrook if a fifth taxa were to be found, because of its presence at all three Bone Valley sites. In contrast, *A. stocki* is only known from two of the Bone Valley sites, and its relative abundance is extremely low for both sites. This suggests that it would be the sixth equid species, if it were to occur at Montbrook.

Significance and Conclusions

The results of our study emphasize the complexity of Hh4 equid diversity in Florida. Rarefaction provides a mechanism for understanding the relative abundances and presence and/or absence of equid species in Florida during the latest Hemphillian. These outcomes corroborate previous studies described earlier that analyze stable isotopes and tooth microwear of the horses in this area. The paucity of the monodactyl horses *Astrohippus stocki* and *Dinohippus mexicanus* in the Hh4 faunas from Florida is potentially interesting for understanding the differences in paleoenvironments and equid species richness of Florida relative to other North American Hh4 sites (MacFadden 2008). While high abundances are well documented at other North American sites for both *D. mexicanus* and *A. stocki*, most notably from Yepómera, Mexico (Lance 1950), Florida's environment at this same time seems to have provided a less hospitable habitat for these monodactyl horses.

The results of this study open up new possibilities for future research. Given the large number (>85) of Hh4 sites from the Bone Valley region, in addition to the three studied here (Kingsford Mine, Fort Green Mine South, and Palmetto Mine [Agrico]), detailed studies and subsampling of the other sites, combined with ecological data such as stable isotopes and microwear, could potentially elucidate ancient ecosystem dynamics, such as the continuity or patchiness of habitats.

Other future extensions of this study might include a more comprehensive simulation that compares the broader key Hh4 sites across North America. Analyzing the horse biodiversity at other North American Hh4 sites and their respective abundance and species richness data might add to our understanding of equid ecogeography across the continent. Taking into account the Yepómera site in Chihuahua, Mexico, for example, might provide greater context for understanding the differences in horse species abundance across North America during this pivotal time.

After the latest Hemphillian, the equid genera *Neohipparion*, *Pseudhipparion*, *Astrohippus*, and *Dinohippus* became extinct, although the latter is the closest sister taxon of the genus *Equus* (Bennett 1980; Cirilli et al. 2021). Likewise, the species *Nannippus aztecus* became extinct, but the genus is represented in the Blancan by *Nannippus peninsulatus*. The sixth equid taxon discussed, *Cormohipparion emsliei*, has a relictual biochron and endemic range that extend into the late Blancan, that is, the early Pleistocene until about 1.8 Ma (Fig. 2) in southern North America. Future studies that explore our understanding of the adaptations and ecology of this species and why it persisted for almost 3 Myr longer than the other five species studied here are of potential interest in further refining our understanding of the rich record of fossil horses in North America.

Continued sampling at Montbrook as well as potential new Hh4 site discoveries will advance our understanding of horse species richness and abundance in Florida. As with any fossil locality, greater sampling efforts can build upon what we know. For example, it will be interesting to confirm whether the occurrence predictions for the two species from Montbrook validate the rarefaction simulations presented.

Our study also has broad significance as an example of a model for predicting occurrences of extinct taxa with continued sampling at a particular locality. Like some previous studies (e.g., Nowak et al. 2000; Burnham 2008), this method addresses the challenge of how we interpret absences of taxa in the fossil record. Because this is fundamental to the conceptual framework driving our interpretation of paleobiogeography and macroevolution, similar studies of sampling (bias) and occurrence data (presence–absence) will remain important considerations in future studies in paleobiology.

Likewise in paleontology, we oftentimes try to find a simple solution for complex systems. One of the points made by this paper is there is not always a simple explanation for whether sampling or ecology results in the pattern of equid species occurrences observed from the Mio-Pliocene of Florida. Instead, our study predicts that both of these factors come into play to explain the presence or absence of these horses at particular localities.

Fossil horses (Family Equidae) are a classic textbook example of macroevolution interpreted from the fossil record. This study uses modern analytical techniques to better understand sampling, which is a persistent challenge in the fossil record, and the nuances of positing simple explanations for complex systems. The results presented here also serve as a model for similar studies of other groups in the future.

Literature Cited

Alroy J. 2010. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. *Palaeontology* **53**:1211–1235.

Alroy J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fursich, T. A. Hansen, *et al.* 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences USA* **98**:6261–6266.

Barnosky, A. D., M. A. Carrasco, and E. B. Davis. 2005. The impact of the species-area relationship on estimates of paleodiversity. *PLoS Biology* **3**: e266.

Bennett, D. 1980. Stripes do not a zebra make, Part I: A Cladistic analysis of *Equus* [*sic*]. *Systematic Zoology* **29**:282–287.

Bokulich, A. 2021. Using models to correct data: paleodiversity and the fossil record. *Synthese* **198** (Suppl. 24):5919–5940.

Burnham, R. J. 2008. Hide and go seek: what does presence mean in the fossil record. *Annals of the Missouri Botanical Garden* **95**:51–71.

Bush, A. M., M. J. Markey, and C. R. Marshall. 2004. Removing bias from diversity curves: the effects of spatially organized biodiversity on sampling-standardization. *Paleobiology* **30**:666–686.

Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**:2533–2547.

Chao, A., R. K. Colwell, C. W. Lin, and N. J. Gotelli. 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* **90**:1125–1133.

Chao, A., C. H. Chiu, T. C. Hsieh, T. Davis, D. A. Nipperess, and D. P. Faith. 2014. Rarefaction and extrapolation of phylogenetic diversity. *Methods in Ecology and Evolution* **6**:380–388.

Cirilli, O., L. Pandolfi, L. Rook, and R. L. Bernor. 2021. Evolution of Old World *Equus* and origin of the zebra-ass clade. *Scientific Reports* **11**:10156.

Cirilli O., H. Machado, J. Arroyo-Cabrales, C. I. Barrón-Ortiz, E. Davis, C. N. Jass, A. M. Jukar, Z. Landry, A. H. Marín-Leyva, and L. Pandolfi. 2022. Evolution of the Family Equidae, Subfamily Equinae, in North, Central and South America, Eurasia and Africa during the Plio-Pleistocene. *Biology* **11**(9):1258.

Currie, A., and D. Turner. 2017. The missing fossils matter as much as the ones we have found. *AEON*. https://aeon.co/ideas/the-missing-fossils-matter-as-much-as-the-ones-we-have-found, accessed 9 September 2023.

Damuth J., and C. M. Janis. 2011. On the relationship between hypsodonty and feeding ecology in ungulate mammals, and its utility in palaeoecology. *Biological Reviews of the Cambridge Philosophical Society* **86**:733–758.

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14**:927–930.

Heck, K. L., Jr., G. Van Belle, and D. Simberloff. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology* **56**:1459–1461.

Hulbert, R. C., Jr. 2015. Palmetto Fauna. Florida Museum (of Natural History). https://www.floridamuseum.ufl.edu/florida-vertebrate-fossils/sites/palmetto-fauna, accessed 13 March 2023.

Hulbert, R. C., Jr. 2016a. Hemphillian North American Land Mammal Age. Florida Museum (of Natural History). https://www.floridamuseum.ufl.edu/florida-vertebrate-fossils/land-mammal-ages/hemphillian, accessed 26 April 2023.

Hulbert, R. C., Jr. 2016b. Montbrook site. Florida Museum (of Natural History). https://www.floridamuseum.ufl.edu/florida-vertebrate-fossils/sites/montbrook, accessed 14 March 2023.

Hulbert, R. C., Jr., N. J. Czaplewski, and S. D. Webb. 2005. New records of *Pseudhipparion simpsoni* (Mammalia, Equidae) from the late Hemphillian of Oklahoma and Florida. *Journal of Vertebrate Paleontology* **25**:737–740.

Lance, J. F. 1950. Palaeontología y estratigrafía del Plioceno de Yepómera, estado de Chihuahua 1a Parte: equidos, excepto *Neohipparion*. *Instituto de Geología, Universidad Nacional Autónoma de México* **54**:1–81.

MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Rev. ed. Princeton University Press, Princeton, N.J. http://www.jstor.org/stable/j.ctt19cc1t2.

MacFadden, B. J. 1986. Late Hemphillian monodactyl horses (Mammalia, Equidae) from the Bone Valley formation of central Florida. *Journal of Paleontology* **60**:466–475.

MacFadden, B. J. 1992. *Fossil horses: systematics, paleobiology, and evolution of the family Equidae*. Cambridge University Press, Cambridge. https://doi.org/10.2307/1382454.

MacFadden, B. J. 2008. Geographic variation in diets of ancient populations of 5-million-year-old (early Pliocene horses) from southern North America. *Palaeogeography, Palaeoclimatology, Palaeoecology* **266**:83–94.

MacFadden, B. J., N. Solounias, and T. E. Cerling. 1999. Ancient diets, ecology, and extinction of 5-million-year-old horses from Florida. *Science* **283**:824–827.

[NCSE] National Center for Science Education. 2007. Absence of fossil evidence. https://ncse.ngo/absence-fossil-evidence, accessed 9 September 2023.

Newell, N. D. 1959. The nature of the fossil record. *Proceedings of the American Philosophical Society* **103**:264–285.

Nowak, R. S, C. L. Nowak, and R. J. Tausch. 2000. Probability that a fossil absent from a sample is also absent from the paleolandscape. *Quaternary Research* **54**:144–154.

Oksanen, J., G. Simpson, F. Blanchet, R. Kindt, P. Legendre, P. Minchin, R. O'Hara, *et al.* 2022. vegan: community ecology package, R package version 2.6-4. https://CRAN.R-project.org/package=vegan.

Prothero, D. R. 2021. The systematics of North American peccaries (Mammalia: Artiodactyla: Tayassuidae). *New Mexico Museum of Natural History and Science Bulletin* **85**.

Raup, D. M. 1975. Taxonomic diversity estimation using rarefaction. *Paleobiology* **1**:333–342.

R Core Team. 2022. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Studio Team. 2020. *RStudio: integrated development for R*. R Studio, Boston.

Scott, T. M., K. M. Campbell, F. R. Rupert, J. D. Arthur, T. M. Missimer, J. M. Lloyd, J. W. Yon, and J. G. Duncan. 2001. Geological map of the State of Florida. Scale 1:750,000. https://www.nbbd.com/godo/ef/geology/USGS-Florida-lg.jpg, accessed 1 May 2023.

Semprebon, G. M., F. Rivals, N. Solounias, and R. C. Hulbert Jr. 2016. Paleodietary reconstruction of fossil horses from the Eocene through Pleistocene of North America. *Palaeogeography, Palaeoclimatology, Palaeoecology* **442**:110–127.

Sepkoski, J. 2005. Stephen Jay Gould, Jack Sepkoski, and the "Quantitative Revolution" in American Paleobiology. *Journal of the History of Biology* **38**:209–237.

Sepkoski, J., R. Bambach., D. Raup, and J. W. Valentine. 1981. Phanerozoic marine diversity and the fossil record. *Nature* **293**:435–437.

Shimadzu, H. 2018. On species richness and rarefaction: size- and coverage-based techniques quantify different characteristics of richness change in biodiversity. *Journal of Mathematical Biology* **77**:1363–1381.

Simpson, G. G. 1951. *Horses: the story of the horse family in the modern world and through sixty million years of history*. Oxford University Press, New York.

Specify Collections Consortium. 2023. Home page. https://www.specifysoftware.org, accessed 25 April 2023.

Stirton, R. A. 1954. Late Miocene mammals from Oaxaca, Mexico. *American Journal of Science* **252**:634–638.

Stirton, R. A. 1955. Two new species of the equid genus *Neohipparion* form the middle Pliocene, Chihuahua, Mexico. *Journal of Paleontology* **29**:88–902.

Tedford, R. H., L. B. Albright III, A. D. Barnosky, I. Ferrusquia-Villafranca, R. M. Hunt Jr., J. E. Storer, C. C. Swisher III, M. R. Voorhies, S. D. Webb, and D. P. Whistler. 2004. Mammalian biochronology of the Arikareean through Hemphillian interval (late Oligocene through early Pliocene epochs). Pp. 169–231 *in* M. O. Woodburne, ed. *Late Cretaceous and Cenozoic mammals of North*

America. Columbia University Press, New York. https://doi.org/doi:10.7312/wood13040-008.

[UF VP] **University of Florida Vertebrate Paleontology Collections Database**. 2023. Home page. https://www.floridamuseum.ufl.edu/vertpaleo-search, accessed 28 February 2023.

Webb, S. D., and R. C. Hulbert Jr. 1986. Systematics and evolution of *Pseudhipparion* (Mammalia, Equidae) from the Late Neogene of the Gulf Coastal Plain and the Great Plains. Pp. 237–272 *in* K. M. Flanagan and J. A. Lillegraven, eds. *Vertebrates, phylogeny, and philosophy. University of Wyoming Contributions to Geology Special Paper* **3**.