What are the best modern analogs for ancient South American mammal communities?
Evidence from ecological diversity analysis (EDA)

Angeline M. Catena and Darin A. Croft

ABSTRACT

Ecological diversity analysis (EDA) is a technique that uses ecological attributes of mammals to reconstruct the community structure and habitat of a fossil locality. EDAs of South American paleofaunas have generally relied on modern comparative datasets from that continent, but modern faunas from other continents may be more appropriate models considering the high-level taxonomic differences that exist between modern and fossil South American mammal communities. To test this hypothesis, we selected five, well-sampled fossil localities for which independent paleoenvironmental data (e.g., paleosols, ichnofossils) have been published: four from South America (La Venta, Colombia; Quebrada Honda, Bolivia; Santa Cruz, Argentina; Tinguiririca, Chile) and one from Europe (Rümikon, Switzerland). We coded the extinct species from these sites, as well as ca. 2,450 modern mammal species, for three ecological attributes: diet (eight categories), locomotor habit (six categories), and body mass (six categories). Percentages of species in each attribute category were used to compare the five paleofaunas to 179 modern faunas from six continents using correspondence analysis, hierarchical clustering, similarity percentage, and classification trees. The four South American paleofaunas were found to be most similar to Afrotropical, Indo-Malayan, and Palearctic modern faunas and, similarly, the European paleofauna most resembled faunas from a biome not currently present in the Palearctic. Our study highlights important differences in community structure between ancient and modern South American mammal faunas and suggests that modern mammalian communities from other continents are better analogues for ancient South American communities than modern South American faunas.

Angeline M. Catena. Diablo Valley College, 321 Golf Club Road, Pleasant Hill, California, 94523-1529, USA. Angeline.Catena@case.edu
Darin A. Croft. Department of Anatomy, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio, 44196-4930, USA. Darin.Croft@case.edu


Copyright: February 2020 Society of Vertebrate Paleontology. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. creativecommons.org/licenses/by/4.0/creativecommons.org/licenses/by-nc-sa/4.0/
INTRODUCTION

Paleontologists use a variety of approaches to reconstruct terrestrial habitats, climates, and ecological communities (Dark, 2008; Piperno, 2008; Shuman, 2009; Wright, 2017; Croft et al., 2018b). For example, sedimentological analyses can be used to determine the sources of sediments and characteristics of a depositional environment (Boggs, 2006; Nichols, 2009, Wright, 2017). Paleopedology can be used to deduce environmental factors such as mean annual precipitation and paleotopographic position (Caudill et al., 1996; Cerling et al., 1989; Retallack, 1994; Kraus, 1999; Retallack, 2001; Driese and Ober, 2005; Retallack, 2005; Kraus and Riggins, 2007), and ichnology can be used to determine soil moisture regimes, primary productivity, and shifts in ecosystems related to changes in environmental conditions (Rhoads, 1975; Retallack, 1984; Hasioitis, 2002, 2007; Kraus and Riggins, 2007; Smith et al., 2008; Hembree 2018). Plant fossils, including macrofossils (wood, leaves), pollen, and phytoliths, can provide taxonomic and ecomorphological information pertaining to floral communities, vegetative cover, and climate (Hooghiemstra, 1989; Wing and Greenwood, 1994; Gregory-Wodzicki et al., 1998; Dunn et al., 2015; Wright 2017; Kovarovic et al., 2018; Mander and Punyasena, 2018; Peppe et al., 2018; Strömberg et al., 2018). Ecological attributes and/or environmental constraints of vertebrates can be used to infer properties of the environments in which they lived (Fleming, 1973; Andrews et al., 1979; Eisenberg, 1981; Brown, 1995). Although the biological attributes of a fossil species cannot be observed directly, such attributes can be inferred with some confidence by comparing craniodental and postcranial anatomy to that of modern species (Kappelmann, 1988; Scott et al., 1990; DeGusta and Vrba, 2003, 2005; Kovarovic et al., 2018). By repeating this process for all species in a fossil assemblage and comparing the resulting community structure to that of modern faunas, the paleohabitat of a fossil site can be reconstructed (Andrews and Van Couvering, 1975; Nesbit Evans et al., 1979; Van Couvering 1980). This type of ecomorphological approach is considered to be taxon-independent because the proportions of species occupying broad ecological niches within a particular habitat type tend to be similar regardless of the taxonomic affinities of the species involved (Fleming, 1973; Andrews and Nesbet Evans, 1979; Andrews et al., 1979 Damuth, 1992). In theory, an EDA of any fauna for which the ecological characteristics of its species can be determined should faithfully reflect the habitat in which those species were living.
Ecological diversity analysis was originally developed and applied to late Cenozoic hominin-bearing sites in Africa (e.g., Andrews and Lord, 1979; Andrews et al., 1979; Van Couvering 1980), but the technique has also been used to reconstruct South American paleohabitats (e.g., Kay and Madden, 1997a, b; Flynn et al., 2003; Croft et al., 2008; Kay et al., 2012; Lintulaakso and Kovarovic, 2016). The study of these geologically older South American mammal faunas comes with a formidable set of challenges due to the continent’s biogeographic history. For the majority of the Cenozoic, South America was geographically and therefore biotically isolated from other landmasses and developed a highly endemic terrestrial mammal fauna (Patterson and Pascual, 1968; Foeden, 1972; Simpson, 1980; Marshall et al., 1983; Flynn and Wyss, 1998; Croft, 2016). During the Paleocene and most of the Eocene, South American mammal faunas consisted mainly of metatherians (opossums and many extinct relatives), several groups of endemic “ungulates” (notoungulates, astrapotheres, pyrotheres, and litopterns), and xenarthrans (relatives of sloths, armadillos, and anteaters) (Patterson and Pascual, 1968; Simpson, 1980; Marshall et al., 1983; Flynn and Wyss, 1998). Caviomorph rodents and platyrhine primates reached South America during the middle to late Eocene via waif dispersal from Africa (Antoine et al., 2012; Bond et al., 2015) and subsequently became major components of most ecosystems (Flynn and Wyss, 1998; Poux et al., 2006; Vuectich et al., 2010). Until the late Cenozoic, when North American mammals began to appear in South America (Webb, 1976; Woodburne, 2010; Cione et al., 2015), nearly all terrestrial, nonvolant mammals in South America pertain exclusively to these groups. During the Great American Biotic Interchange (GABI; Stehli and Webb, 1985) a protracted interval of faunal exchange between North and South America (Stehli and Webb, 1985; Webb, 2006; Woodburne et al., 2006; Woodburne, 2010), six orders and 18 families of North American mammals emigrated to South America including camelids (camels), cervids (deer), felids (cats), canids (dogs), and cricetid rodents (New World mice and rats) (for details, see Webb, 2006, Woodburne, 2010, Cione et al., 2015). A Pleistocene megafaunal extinction eliminated ~83% of South American megafaunal (> 44 kg) genera (Koch and Barnosky, 2006), including the last representatives of two orders of endemic ungulates (Notoungulata, Litopterna) (Cifelli, 1985; Marshall and Cifelli, 1990; Croft, 2000; Bond et al., 2001; Koch and Barnosky, 2006; Fiedel, 2009; Barnosky and Lindsey, 2010; Villavivencio et al., 2015). These two events substantially changed the high-level taxonomic composition of South American mammal communities relative to modern ones, with the result that modern South American mammal faunas bear very little resemblance to those that predominated for much of the Cenozoic (Croft, 2012).

Given the aforementioned differences between fossil and modern South American mammal faunas, one might reasonably question whether modern South American mammal faunas are the best ecological analogues for South American paleofaunas (cf. Croft and Townsend, 2005). Even though general resemblances in community structure may be sufficient for reconstructing paleohabitat (e.g., Kay and Madden, 1997a, b; Flynn et al., 2003; Croft et al., 2008), limiting the modern comparative dataset to the same biogeographic realm may preclude a more nuanced interpretation of community structure (Lintulaakso and Kovarovic, 2016).

The goal of this present study is to test if modern faunas from outside of South America are better analogues for South American fossil mammal communities than modern South American faunas. This proposition is evaluated via EDA using a global comparative dataset and five well-sampled paleofaunas, four from South America and one from Europe. These paleofaunas were chosen because they represent different geographic areas and depositional environments and have been the subject of paleoenvironmental reconstruction based on other types of evidence. We make three predictions: Prediction 1: our paleoenvironmental reconstructions based this global comparative dataset will be congruent with reconstructions based on other types of data; this is based on the previously-established efficacy of EDAs for reconstructing paleoenvironments. Prediction 2: the community structures of the South American paleofaunas will most closely resemble those of modern faunas from continents other than South America; this is based on observed taxonomic and body size differences between fossil and modern South American faunas. Prediction 3: any sampling biases (due to preservation or collection methods) shared among all five paleofaunas will not override underlying ecological signals (i.e., all five faunas will not plot together in multivariate analyses); this is based on the assumption that sampling biases should not be significant in these faunas, which were selected a priori to minimize such factors.
FOSSIL LOCALITIES

Localities Rationale

The paleofaunas chosen for this EDA (described below) derive from a small number of South American localities that have relatively precise stratigraphic control and an extensive (and up-to-date) faunal list. Within this limited subset of well-sampled localities, they were chosen because they represent distinct geographic areas (nearly the entire latitudinal range of South America), temporal intervals (early Oligocene to middle Miocene), paleohabitats (relatively arid grassland/shrubland to moist broadleaf forest), and depositional environments (floodplains, rivers, lakes, lahar deposits) (Croft, 2016). Thus, any similarities among these four sites can only reasonably be due to two factors: (1) taxonomic characteristics shared by South American paleofaunas; or (2) systematic sampling biases common to fossil mammal faunas in general. The European locality of Rümikon was included to distinguish between these two possibilities. Since the mammals of Rümikon are taxonomically distinct from those of the South American localities (no families and only two orders of mammals are shared between them), any features common to all five fossil localities must reflect some other sampling bias, such as time-averaging or under-sampling species with particular ecological characteristics.

La Venta, Colombia

La Venta (ca. 3° 20’ N, 75° 20’ W) is located in the Magdalena Valley of central Colombia (Figure 1) (Guerrero, 1997); it is composed of sedimentary units within the Honda Group, and the fossil bearing units date to the late middle Miocene (13.5–11.5 Ma, Serravallian Age; Flynn et al., 1997). The Monkey Beds is an approximately 15 m-thick, extremely fossiliferous interval within the Honda Group that has been constrained to an approximately 15,000-year interval within Polarity Chron C5AAn (Kay and Madden, 1997a), which spans 13.18–13.03 Ma based on Ogg (2012). La Venta is the richest pre-Pleistocene fossil locality within the tropical latitudes of South America (Croft, 2016); 66 species of non-volant mammals have been identified from La Venta in addition to a diversity of non-mammal vertebrates, and 50 mammal species are preserved within the Monkey Beds (Appendix 1). Based on paleobotany, paleopedology, and a previous EDA, the paleoenvironment of La Venta has been reconstructed as a river-associated, multi-storied tropical forest environment with a discontinuous canopy (Kay and Madden, 1997a). La Venta’s mean annual precipitation (MAP) is con-

FIGURE 1. Locations of the 179 modern ecoregions (colored areas) and the five fossil localities (stars) used in this study. Ecoregions are overlain on a grayscale global mean annual precipitation (MAP) map derived from Fick and Hijmans (2017), with lighter regions indicating areas of higher MAP and darker regions indicating areas of lower MAP. Abbreviations: LV, La Venta, Colombia; QH, Quebrada Honda, Bolivia; RU, Rümikon, Switzerland; SC, Santa Cruz, Argentina; TG, Tinguiririca, Chile.
servately estimated at 1,500–2,000 mm (Kay and Madden, 1997a, b; Spradley et al., 2019).

Quebrada Honda, Bolivia

Quebrada Honda (ca. 21° 57’ S, 65° 25’ W) is located in southern Bolivia, approximately 65 km southwest of the city of Tarija (Figure 1), and its fossils derive from sediments of the Honda Group of Bolivia (unrelated to the Honda Group of Colombia) (MacFadden and Wolff, 1981). Quebrada Honda is coeval with La Venta; the entire sedimentary section of Quebrada Honda has been dated to the late middle Miocene (13.0–11.9 Ma; Serravallian Age), and the lower, most fossiliferous zones also fall within Polarity Chron C5AAa (MacFadden et al., 1990). At present, 33 species of non-volant mammals have been identified from Quebrada Honda (Appendix 1). Based on paleopedology and ichnology, the paleoenvironment of the most fossiliferous intervals of Quebrada Honda have recently been reconstructed as representing a seasonal, sub-humid to semi-arid wooded grassland or savanna-like environment with a MAP of ca. 1,000 mm (Catena et al., 2017).

Santa Cruz, Argentina

The localities of the coastal portion of the Santa Cruz Formation (ca. 50° 17’ W, 51° 36’ S) (Figure 1) are located in the southern Argentinian province of the same name and sample the late early Miocene (~17.9–16.2 Ma; Burdigalian Age; Perkins et al., 2012; Vizcaíno et al., 2012). Fossiliferous Levels 1–7 are fossil-rich stratigraphic intervals that crop out in four discrete but closely-spaced localities (Vizcaíno et al., 2012) and have been dated to a narrow time interval that spans 17.5–17.4 Ma (Fleagle et al., 2012; Perkins et al., 2012). Santa Cruz is one of the richest fossil localities in South America; over 100 mammal species have been reported (Croft, 2016; but see also Croft, 2013), and 48 species of non-volant mammals have been identified from Fossiliferous Levels 1–7 (Appendix 1). Based on sedimentology, ichnology, paleobotany, paleopedology, and an EDA, the paleoenvironment of Santa Cruz has been reconstructed as a mosaic of open temperate humid to semi-arid forest environments with pronounced seasonality and a MAP of approximately 1,000–1,500 mm (Kay et al., 2012; Raigeborna et al., 2019; Spradley et al., 2019).

Tinguiririca, Chile

Tinguiririca is located in central Chile near the town of Termas del Flaco (34° 59’ S, 70° 26’ W; Wyss et al., 1990; Flynn et al., 2003) (Figure 1). The fossiliferous horizons are composed of volcanioclastic sediments of the Abanico Formation (Charrier et al., 1996; Flynn et al., 2003) and have been dated to the early Oligocene (~31.5 Ma; Rupelian Age; Flynn et al., 2003, 2012). Twenty-six species of mammals have been identified from the site (Appendix 1). Faunal-based techniques, such as cenograms and EDAs, are presently the only options for paleohabitat reconstruction at Tinguiririca, as other biotic and sedimentary features, like phytoliths and paleosols, are not preserved (Flynn et al., 2003; Su and Croft, 2018). Tinguiririca has been interpreted as an open-habitat environment (Wyss et al., 1994; Flynn et al., 2003; Croft et al., 2008) with a MAP of approximately 600 mm (Flynn et al., 2003).

Rümikon, Switzerland

Rümikon is located in northern Switzerland near its border with Austria, approximately 25 km northeast of Zurich (47° 29’ N, 8° 47’ E; NOW Database) (Figure 1). The sediments of Rümikon are part of the Upper Freshwater Molasse (Kälin and Kempf, 2009), which is the uppermost formation of the Molasse Group sediments within the North Alpine Foreland Basin of Switzerland (Matter et al., 1980; Labhart, 2005). The sediments of Rümikon are placed in an early middle Miocene interval from 14.6–14.2 Ma (Langhian Age) based faunal correlations with Swiss and Bavarian localities within the Upper Freshwater Molasse that have been chronostratigraphically dated (Kälin and Engesser, 2001; Kälin and Kempf, 2009). Thirty-seven species of non-volant mammals have been identified from Rümikon (Appendix 1; Kälin and Kempf, 2009; NOW Database, 2017). Based on paleobotanical data and the paleobiology of fossil rodents, the habitats of MN 6 localities in the Upper Freshwater Molasse have been reconstructed as humid to sub-humid mixed and forested environments (Hantke, 1984; Kälin and Kempf, 2009). A general MAP range of approximately 1,300–1,500 mm has been estimated for the fossil localities of the Upper Freshwater Molasse (Hantke, 1984; Kälin and Kempf, 2009).

METHODOLOGY

Faunal Lists

The faunal lists for La Venta, Quebrada Honda and Tinguiririca are from Croft (2016), updated based on Engelman et al. (2017, 2018), Brandoni et al. (2018), McGrath et al. (2018), and
Wyss et al. (2018). The faunal list for Santa Cruz is from Kay et al. (2012), and the faunal list for Rümkon is from the NOW database (http://www.helsinki.fi/science/now/), which was downloaded on November 01, 2017.

Modern faunal lists from 179 ecoregions are taken directly from the World Wild Life Fund’s WildFinder Database (WFDB) (World Wildlife Fund 2006; full database can be downloaded at http://www.worldwildlife.org/science/data.cfm). The WFDB contains presence/absence data for terrestrial vertebrates compiled from field guides and scientific works; the data are based on the ranges of extant species and do not include species that are present as human commensals or were introduced by humans (World Wildlife Fund, 2006). In order to reflect the historic, pre-disturbed, or potential vegetation of each ecoregion, the WFDB uses historic ranges of species with approximate distributions at 1,500 A.D. instead of current distributions, when possible (World Wildlife Fund, 2006).

Biogeographic realms are the largest biogeographic division of Earth; they share and are broadly characterized by the evolutionary history of the flora and fauna contained within them (Reid et al., 2005). The WFDB divides Earth into eight terrestrial biogeographic realms based on Olson et al. (2001), six of which were used in this study: Australasia (Appendix 2); Afrotopia (Appendix 3); Indo-Malaya (Appendix 4); Nearctica (Appendix 5); Neotropic (Appendix 6), which includes Central America and the tropical area of North America in addition to all of South America; and the Palearctic (Appendix 7, Figure 1). Oceana and Antarctic biogeographic realms were excluded due to their lack of terrestrial (non-marine) mammals. The WFDB (i.e., Olson et al., 2001) also recognizes 14 terrestrial biomes, which are floral and faunal communities characterized by adaptations to a particular environment (Campbell, 1996). Each biogeographic realm includes one or more of these biomes, and 11 of these biomes were included in this study. The polar boreal forest/taiga and tundra biomes (Biomes 6 and 11, respectively) were excluded, as the paleofaunas selected for this analysis are from tropical and temperate, rather than polar regions; the mangroves biome (Biome 14) was also excluded, as those ecoregions are poorly mapped and lack adequate species data (World Wildlife Fund, 2006). Each biome in a biogeographic realm is subdivided into one or more ecoregions, which are relatively large areas (mean = 150,000 km², median = 56,300 km²) that represent distinct biotas (Dasmann, 1973, 1974; Udvardy, 1975; Olson and Dinerman, 1998; Olson et al., 2001) and accurately describe terrestrial biodiversity patterns of vertebrates and plants (Smith et al., 2018). Each ecoregion is unique and identified by a six-digit alphanumeric code. As an example of this hierarchical structure, AA0105 represents the Central Range Montane Rain Forests ecoregion, which is a particular area of Biome 1 (tropical and subtropical moist broadleaf forests; middle pair of digits) within the Australasia biogeographic realm (prefix AA).

The ecoregions used in this study were selected to encompass maximum habitat variation while maintaining relatively equal representation among biomes and biogeographic realms. This was achieved by selecting up to five ecoregions for each biome within each biogeographic realm. For example, for the Indo-Malaya biogeographic realm, five ecoregions were selected within both Biome 1 and Biome 2. All ecoregions were included if five or fewer were represented in a particular biome in a particular biogeographic realm; if more than five ecoregions were available, five were chosen at random using a random number generator. Since precipitation is a major determinant of vegetational structure (Nicholson et al., 1990; Kay and Madden, 1997b), a dataset containing a 30-year average of mean annual precipitation (MAP) with spatial resolution of approximately 1 km² was obtained from WorldClim Version 2 for ArcGIS (Fick and Hijmans, 2017) for the 179 ecoregions used in this study (Table 1; Appendices 2–7). A linear regression with these data was used to estimate MAP for fossil localities (described below).

To avoid using taxonomically under-sampled ecoregions with inaccurate faunal lists, only ecoregions with at least 20 species of non-volant mammals were included. This is an arbitrary threshold based on the observation that most terrestrial mammal faunas include at least 30 species (including a variable proportion of volant mammals; Ceballos and Ehrlich, 2006; see also Mares and Willig, 1994), and that most well-sampled South American fossil localities include at least 20 species (Croft, 2006). As with other ecological diversity analyses (e.g., Andrews et al., 1979; Reed, 1998; Kay and Madden 1997a, b; Kovarovic et al., 2002; Mendoza et al., 2004; Louys et al., 2011; Lintulaakso and Kovarovic, 2016), volant mammals (i.e., bats, order Chiroptera) and fully aquatic mammals (i.e., manatees/dugongs and whales, order Sirenia and infraorder Cetacea, respectively) were excluded from our analyses.
Ecological Variables

All extant and extinct mammal species were coded for three ecological attributes - body size, primary locomotor style, and diet - based primarily on the published literature (Kay and Madden, 1997a, b; Kay et al., 1997; Nowak, 1999; Flynn et al., 2003; MacDonald, 2006; Vizcaíno et al., 2006, 2012; Croft et al., 2008; Kay et al., 2012; Toledo et al., 2014; Bradham et al., 2015; Gainsbury et al., 2018; Appendices 1–8; Table 2). Data from congers were used in cases where such ecological data were unavailable.

Species were assigned to one of seven logarithmic body mass categories ranging from 0–10 g to 1,000–10,000+ kg (Table 2). The distribution of mammalian body sizes is right skewed, with smaller taxa (≤ 10 kg) naturally occurring at higher frequencies than larger taxa (Caughley 1987; Gardezi and da Silva, 1998; Kozlowski and Gawelczyk, 2002). A logarithmic scale was used to account for this greater frequency of small taxa and to more accurately characterize body mass ranges of the fauna.

The primary locomotor style ecological attribute included six categories: Arboreal, Scansorial, Terrestrial, Semi-fossorial, Fossorial, and Semi-aquatic (Table 2). These categories parallel those used by Kay and Madden (1997a, b) and Kay et al. (2012), except that we did not discriminate among terrestrial, cursorial, and ambulatory species (all were classified as Terrestrial).

The diet ecological attribute included eight categories (Table 2). Mammals display a wide range of dietary preferences that can vary seasonally and with resource availability (Jarman, 1971; Fedriani et al., 2001; Sinclair and Zeppelin, 2002). To more accurately characterize this dietary variation, we used a new method to code for diet in which each species was coded for two equally-weighted diet categories. Species with a relatively homogeneous diet (e.g., giant anteater, *Myrmecophaga tridactyla*), were coded twice for the same diet category (Diet 1 = In; Diet 2 = In), whereas species with a heterogeneous diet (e.g., monito del monte, *Dromiciops gliroides*), were coded for two different diet categories (Diet 1 = Fr; Diet 2 = In).

### TABLE 1. Data for the 11 biomes used in this study. MAP, mean annual precipitation.

<table>
<thead>
<tr>
<th>Biome Number</th>
<th>Biome Name</th>
<th>N</th>
<th>MAP Range</th>
<th>MAP Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tropical and subtropical moist broadleaf forests</td>
<td>22</td>
<td>997 - 3,672 mm</td>
<td>2,193 mm</td>
</tr>
<tr>
<td>2</td>
<td>Tropical and subtropical dry broadleaf forests</td>
<td>10</td>
<td>243 - 1,902 mm</td>
<td>1,047 mm</td>
</tr>
<tr>
<td>3</td>
<td>Tropical and subtropical coniferous forests</td>
<td>10</td>
<td>631 - 2,321 mm</td>
<td>1,406 mm</td>
</tr>
<tr>
<td>4</td>
<td>Temperate broadleaf and mixed forests</td>
<td>20</td>
<td>642 - 2,528 mm</td>
<td>1,239 mm</td>
</tr>
<tr>
<td>5</td>
<td>Temperate coniferous forests</td>
<td>12</td>
<td>238 - 1,244 mm</td>
<td>914 mm</td>
</tr>
<tr>
<td>6</td>
<td>Tropical and subtropical grasslands, savannas, and</td>
<td>16</td>
<td>481 - 1,894 mm</td>
<td>1,070 mm</td>
</tr>
<tr>
<td>7</td>
<td>shrublands</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Temperate grasslands, savannas, and shrublands</td>
<td>15</td>
<td>267 - 964 mm</td>
<td>512 mm</td>
</tr>
<tr>
<td>9</td>
<td>Flooded grasslands and savannas</td>
<td>16</td>
<td>182 - 1,559 mm</td>
<td>742 mm</td>
</tr>
<tr>
<td>10</td>
<td>Montane grasslands and shrublands</td>
<td>17</td>
<td>165 - 3,387 mm</td>
<td>921 mm</td>
</tr>
<tr>
<td>11</td>
<td>Mediterranean forests, woodlands and sclerophyll</td>
<td>16</td>
<td>275 - 868 mm</td>
<td>541 mm</td>
</tr>
<tr>
<td>12</td>
<td>forests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Deserts and xeric shrublands</td>
<td>24</td>
<td>54 - 609 mm</td>
<td>229 mm</td>
</tr>
</tbody>
</table>

### TABLE 2. Categories for the three ecological attributes used in this analysis. Abbreviations for each attribute are in parentheses.

<table>
<thead>
<tr>
<th>Diet</th>
<th>Body Mass</th>
<th>Primary Locomotor Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carnivorous (Ca):</td>
<td>Size 0 (S0): 0-10 g</td>
<td>Arboreal (Ar)</td>
</tr>
<tr>
<td>Frugivorous (Fr):</td>
<td>Size 1 (S1): 10-100 g</td>
<td>Fossorial (Fo)</td>
</tr>
<tr>
<td>Graminivorous (Gr):</td>
<td>Size 2 (S2): 100 g-1 kg</td>
<td>Scansorial/semi-arboreal (Sc)</td>
</tr>
<tr>
<td>Granivorous (Se):</td>
<td>Size 3 (S3): 1-10 kg</td>
<td>Semi-aquatic (Sa)</td>
</tr>
<tr>
<td>Insectivorous (In):</td>
<td>Size 4 (S4): 10-100 kg</td>
<td>Semi-fossorial (Sf)</td>
</tr>
<tr>
<td>Lignivorous (Li):</td>
<td>Size 5 (S5): 100-1,000 kg</td>
<td>Terrestrial/cursorial/ambulatory (Te)</td>
</tr>
<tr>
<td>Omnivorous (Om):</td>
<td>Size 6 (S6): 1,000-10,000+ kg</td>
<td></td>
</tr>
<tr>
<td>Ramivorous (Ra):</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The primary locomotor style ecological attribute included six categories: Arboreal, Scansorial, Terrestrial, Semi-fossorial, Fossorial, and Semi-aquatic (Table 2). These categories parallel those used by Kay and Madden (1997a, b) and Kay et al. (2012), except that we did not discriminate among terrestrial, cursorial, and ambulatory species (all were classified as Terrestrial).

The diet ecological attribute included eight categories (Table 2). Mammals display a wide range of dietary preferences that can vary seasonally and with resource availability (Jarman, 1971; Fedriani et al., 2001; Sinclair and Zeppelin, 2002). To more accurately characterize this dietary variation, we used a new method to code for diet in which each species was coded for two equally-weighted diet categories. Species with a relatively homogeneous diet (e.g., giant anteater, *Myrmecophaga tridactyla*), were coded twice for the same diet category (Diet 1 = In; Diet 2 = In), whereas species with a heterogeneous diet (e.g., monito del monte, *Dromiciops gliroides*), were coded for two different diet categories (Diet 1 = Fr; Diet 2 = In).
Species with a diet spanning three or more diet categories were double-coded as omnivorous (e.g., common opossum, *Didelphis marsupialis*, Diet 1 = Om; Diet 2 = Om), and species with a single primary diet but a secondary diet spanning three or more categories (e.g., nine-banded armadillo, *Dasypus novemcinctus*), were coded as omnivorous for one of the diet categories (Diet 1 = In; Diet 2 = Om). Because these categories were weighted equally, designating the primary diet as “Diet 1” or “Diet 2” had no effect on analyses. Since each species was coded twice, the denominator for computing the percentage of species in each dietary category was twice the number of species in each ecoregion (2N). The resulting percentages of each diet category were used in the subsequent analyses.

**Data Analyses**

The variables used in the analyses were percentages of species in each ecological attribute category within each fauna (Appendices 8, 9). These 21 variables (one for each ecological attribute category) were analyzed in PAST v. 3.1 (Hammer et al., 2001) using correspondence analysis (CA), linear regression, hierarchical cluster analysis (HCA), and similarity percentage (SIMPER) analyses. These variables were also analyzed via classification trees (CTs) in JMP Pro 13 (SAS Institute, Cary, NC). All analyses were performed with and without the Size 0 variable to determine whether a lack of very small species in a fossil fauna (as commonly occurs in the absence of screen-washing) might bias EDA-based paleohabitat interpretations. The analyses used in this study do not assume multivariate normal distributions of the data (Hammer and Harper, 2006). Nonetheless, all analyses were also performed with and without an arcsine data transformation. There were no discernable differences in the results between the transformed and non-transformed data; the results of the analyses using the non-transformed data are presented here.

Correspondence analysis is an exploratory technique used to visualize trends and groupings in a multivariate dataset (Hammer and Harper, 2006; LeGendre and Legendre, 2012). In our study, the CA plotted faunas and ecological variables in the same space and maintained correspondence between the two, permitting ecological variables most highly associated with particular faunas to be determined. The CA results were also used with linear regression analysis to infer paleoprecipitation; MAP data for modern faunas was regressed on first axis score (CA1 scores), and paleoprecipitation was estimated based on the CA1 values of the fossil faunas. We used these results to test all three of our predictions. The CA and the associated linear regression were used to broadly compare the results of this study to those of the previous studies and were also used to compare the paleofaunas to the modern faunas grouped by biome and biogeographic realm. The positions of the paleofaunas on the CA axes relative to each other were used to test Prediction 3.

Hierarchical cluster analysis is an exploratory technique used to identify trends and groupings in a multivariate data set based on a similarity index (Hammer and Harper, 2006; Legendre and Legendre, 2012). HCA determines which faunas are most similar to one another in terms of their ecological variables (Hammer and Harper, 2006). Unweighted pair group method with arithmetic mean (UPGMA) clustering was used in the HCAs. This is an agglomerative clustering method in which clusters are based on the average distance between all members of the group; it is commonly used in data sets with ecological variables (Hammer and Harper, 2006; Legendre and Legendre, 2012). The chord similarity measure was used in the HCA; it is used to compare faunal assemblages by species abundances (Legendre and Legendre, 2012) and has greater accuracy in high-dimensional datasets than other similarity indexes, such as Euclidian distance and Average distance (Shirkhorshidi et al., 2015). Six HCAs, one including all five paleofaunas and five with one paleofauna each, were performed in order to visualize the relationships of the fossil and modern faunas as a group and individually. The HCAs with individual paleofaunas were used to test which biomes and specific ecoregions the fossil faunas most closely resembled; they were also used to further assess Prediction 2.

A SIMPER analysis quantifies and ranks the average percentage contribution of different variables to the dissimilarity between samples (Clarke, 1993). SIMPER analyses were used to assess which ecological variables were primarily responsible for similarities and differences between the individual paleofaunas and the groups of modern faunas within each of the 11 biomes. The chord similarity measure was also used in the SIMPER analyses. Dissimilarities between the paleofaunas and the modern faunas (grouped by biomes) were used to test Prediction 3.

Classification trees belong to the more general category of statistical analyses known as clas-
sification and regression trees (Breiman et al., 1984), which are used to construct predictive models for complex multivariate systems (De'ath and Fabricus, 2000). We used three CTs to create models and further explore the relationships between the ecological variables and the habitats represented by the modern faunas. The first CT model (CT1) was built to classify modern faunas by vegetative cover; ecoregions of Biomes 1, 2, 3, 4, and 5 were coded as “closed”, those of Biomes 7, 8, 9, 10, and 12 were coded as “mixed”, and those of Biome 13 were coded as “open”. The second CT model (CT2) was built to classify modern faunas by biogeographic realm (six categories), and the third CT model (CT3) was built to classify the modern faunas by biome (11 categories). The CT models were then used to predict the vegetative cover, biogeographic realm, and biome of the five paleofaunas. As five is the maximum number of selected ecoregions from within the same biome and biogeographic realm, terminal leaves with fewer than five faunas were pruned in order to avoid overfitting the CT models. The predictive models as well as the relationships between specific ecological variables and the modern faunas at the biome and biogeographic level were used to further test all three of our predictions.

RESULTS

Correspondence Analysis

Four axes of the CA collectively account for two-thirds of the total variation within the dataset (Table 3) and are discussed herein. Axis one (CA1) has an eigenvalue of 0.065 (27.4% of the total variation) and generally represents a continuum from closed to open habitats (Table 3; Figure 2A). The Arboreal, Frugivorous, Size 3, Scansorial, and Size 2 variables have the greatest negative correlation on CA1 (Figure 2B); higher proportions of species in these ecological attribute categories are found in mammalian communities from closed habitats (Kay et al., 1997b). The Fossorial, Semi-fossorial, Size 0, Graminivorous, and Granivorous variables have highest positive correlation on CA1 (Figure 2B); higher proportions of species in four of these ecological attribute categories (Fossorial, Semi-fossorial, Graminivorous, and Granivorous) are found in open habitats (Guo et al., 1997; Kinlaw, 1999; Gibson, 2009; Davidson et al., 2012). The Australasia and Afrotropic biogeographic realms have the broadest ranges on CA1, whereas the Nearctic and Palearctic biogeographic realms have the narrowest ranges; all of the Nearctic faunas and all but three of the Palearctic faunas have positive CA1 scores. Faunas from Biomes 1, 2, 3, and 4 mostly have negative scores on CA1; those from Biomes 5, 7, 9, and 10 largely plot near the origin on CA1; and those from Biomes 8, 12, and 13 have negative scores on CA1. Among paleofaunas, Quebrada Honda has a positive score on CA1, while La Venta, Santa Cruz, Tunguiroica, and Rümikon have negative scores on CA1.

Axis two (CA2) has an eigenvalue of 0.042, accounts for 17.5% of the total variation, and generally represents a continuum of faunas with many to relatively few large species (Table 3; Figure 2A). The Size 6, Size 5, Size 4, Semi-aquatic, and Carnivorous variables have the greatest negative correlation on CA2 (Figure 2B); higher proportions of species in these ecological attribute categories are found in mammalian communities from closed habitats (Kay et al., 1997b). The Fossorial, Semi-fossorial, Size 0, Graminivorous, and Granivorous variables have highest positive correlation on CA2 (Figure 2B); higher proportions of species in four of these ecological attribute categories (Fossorial, Semi-fossorial, Graminivorous, and Granivorous) are found in open habitats (Guo et al., 1997; Kinlaw, 1999; Gibson, 2009; Davidson et al., 2012). The Australasia and Afrotropic biogeographic realms have the broadest ranges on CA2, whereas the Neotropic and Nearctic biogeographic realms have the narrowest ranges on CA2. All Australasian faunas and all but one Nearctic fauna have positive scores on CA2; all but six faunas from the

<table>
<thead>
<tr>
<th>Axis</th>
<th>Eigenvalue</th>
<th>Total variation</th>
<th>Cumulative variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.065</td>
<td>27.4%</td>
<td>27.4%</td>
</tr>
<tr>
<td>2</td>
<td>0.042</td>
<td>17.5%</td>
<td>45.0%</td>
</tr>
<tr>
<td>3</td>
<td>0.029</td>
<td>12.1%</td>
<td>57.0%</td>
</tr>
<tr>
<td>4</td>
<td>0.023</td>
<td>9.6%</td>
<td>66.5%</td>
</tr>
<tr>
<td>5</td>
<td>0.015</td>
<td>6.2%</td>
<td>72.7%</td>
</tr>
<tr>
<td>6</td>
<td>0.013</td>
<td>5.5%</td>
<td>78.2%</td>
</tr>
<tr>
<td>7</td>
<td>0.011</td>
<td>4.6%</td>
<td>82.8%</td>
</tr>
<tr>
<td>8</td>
<td>0.009</td>
<td>3.7%</td>
<td>86.5%</td>
</tr>
<tr>
<td>9</td>
<td>0.008</td>
<td>3.2%</td>
<td>89.7%</td>
</tr>
<tr>
<td>10</td>
<td>0.005</td>
<td>2.1%</td>
<td>91.9%</td>
</tr>
<tr>
<td>11</td>
<td>0.004</td>
<td>1.7%</td>
<td>93.5%</td>
</tr>
<tr>
<td>12</td>
<td>0.004</td>
<td>1.5%</td>
<td>95.0%</td>
</tr>
<tr>
<td>13</td>
<td>0.003</td>
<td>1.2%</td>
<td>96.2%</td>
</tr>
<tr>
<td>14</td>
<td>0.002</td>
<td>1.0%</td>
<td>97.2%</td>
</tr>
<tr>
<td>15</td>
<td>0.002</td>
<td>0.8%</td>
<td>98.1%</td>
</tr>
<tr>
<td>16</td>
<td>0.002</td>
<td>0.7%</td>
<td>98.8%</td>
</tr>
<tr>
<td>17</td>
<td>0.002</td>
<td>0.6%</td>
<td>99.4%</td>
</tr>
<tr>
<td>18</td>
<td>0.001</td>
<td>0.6%</td>
<td>100.0%</td>
</tr>
<tr>
<td>19</td>
<td>&lt;0.001</td>
<td>&lt;0.1%</td>
<td>100.0%</td>
</tr>
<tr>
<td>20</td>
<td>&lt;0.001</td>
<td>&lt;0.1%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Afrotropics and Indo-Malaya have negative scores on CA2. Faunas from Biomes 1 and 3 largely have positive scores on CA2; those from Biomes 2, 4, 7, 8, 12, and 13 largely plot toward the origin of CA2; and those from Biome 9 have negative scores on CA2. The four South American paleofaunas have negative scores on CA2, while Rümikon has a positive score on CA2.

When both CA1 and CA2 are considered together, the fossil faunas plot as follows: La Venta falls outside the range of all modern biogeographic realms and biomes; Quebrada Honda and Santa Cruz plot within the range of the Afrotropic biogeographic realm and Biome 9 (flooded grasslands and savannas); Tinguiririca plots within the ranges of the Afrotropic and Indo-Malaya biogeographic realms and Biome 7 (tropical and subtropical grasslands, savannas, and shrublands); Rümikon plots within the ranges of Biomes 1, 2, 3, 4, 7, and 10 of the Australasia, Afrotropic, Neotropic, and Palearctic biogeographic realms (Figure 2A).

Axis three (CA3) has an eigenvalue of 0.029 and accounts for 12.0% of the total variation (Table 3, Figure 3A). The Size 6, Insectivorous, Graminivorous, Size 1, and Ramivorous variables have the greatest correlation on the negative side of CA3; the Granivorous, Semi-fossorial, Carnivorous, Omnivorous and Lignivorous variables have the greatest correlation on positive side of CA3 (Figure 3B). CA3 generally discriminates among bioge-
graphic realms. All of the Nearctic and all but one of the Indo-Malayan faunas plot on the positive side of CA3; all of the Afrotropical and all but three of the Australasian faunas plot on the negative side of CA3. The Australasia and Palearctic biogeographic realms have the broadest ranges on CA3, and the Afrotropic, Neotropic, and Indo-Malaya biogeographic realms have the narrowest ranges on CA3. La Venta, Quebrada Honda, Rümikon, and Tinguiririca have positive scores of CA3, and Santa Cruz has a negative score on CA3.

Axis four (CA4) has an eigenvalue of 0.023 and accounts for 9.6% of the total variation (Table 3). No major trends are discernable among the biomes and biogeographic realms on CA4. The Ramivorous, Semi-fossorial, Graminivorous, Size 2, and Size 4 variables have the greatest correlation on the negative side of CA4; the Size 0, Semi-aquatic, Scansorial, Lignivorous, and Size 6 variables have the greatest correlation on the positive side of CA4; (Figure 3B). La Venta, Quebrada Honda, Rümikon, and Tinguiririca have positive scores on CA4, while Santa Cruz has a negative score. The Palearctic and Neotropic biogeographic realms have the broadest ranges on CA4, and Afrotropic and Nearctic biogeographic realms have the narrowest range on CA4.

On CA3 and CA4, Quebrada Honda, Santa Cruz, and Tinguiririca do not plot within the range of any biogeographic realm or biome. La Venta plots within the ranges of the Australasia biogeographic realm and Biome 10 (montane grasslands and shrublands), while Rümikon plots within the ranges of the Afrotropic, Neotropic, and Palearctic biogeographic realms and all biomes (Figure 3A).
Linear Regression

An ordinary least squares regression of MAP on CA1 score results in a predictive equation with a negative slope and an $r^2$ value of 0.61 (Figure 4). Based on the CA1 scores of the paleofaunas and the 95% confidence interval of the regression equation, MAP is estimated at 1,917–2,035 mm for La Venta, 1,210–1,286 mm for Santa Cruz, 1,115–1,214 mm for Rümikon, 1,079–1,186 mm for Tinguiririca, and 877–1,032 mm for Quebrada Honda. Regressions of MAP on CA2, CA3, and CA4 resulted in $r^2$ values of 0.01, 0.01, and 0.04, respectively.

Hierarchical Cluster Analysis

When all five paleofaunas are included in the HCA, the four South American paleofaunas form an exclusive cluster that is the second most basal cluster of the dendrogram (Figure 5); Rümikon is the basal branch within a large cluster of 51 ecoregions from 10 biomes (mostly Biomes 10 and 12, represented by nine ecoregions each), from the Afrotropic (n=20), Palearctic (n=19), Indo-Malaya (n=11), and Nearctic (n=1) biogeographic realms. When analyzed individually using HCA (i.e., without other fossil localities), the position of Rümikon is the same. In contrast, the positions of the four South American paleofaunas are distinct. La Venta plots among modern faunas, as the basalmost branch in a cluster of seven Australasian ecoregions and three Indo-Malayan ecoregions; five of these ecoregions represent Biome 1, one represents Biome 3, and one represents Biome 10. Quebrada Honda plots as the basalmost branch among a small cluster of four Afrotropical ecoregions from Biome 9. Santa Cruz does not cluster with any modern faunas due to high proportions of species in the Size 4, Scansorial, and Size 5 ecological variables. Tinguiririca also does not cluster with any modern faunas, the result of high proportions of species in the Terrestrial, Ramivorous, Graminivorous, and Size 3 variables.

Similarity Percentage

La Venta is most similar to Biomes 2, 1, and 7 (dissimilarity = 0.26, 0.29, 0.29, respectively; Appendix 9) and most dissimilar to Biomes 13, 12, and 8 (dissimilarity = 0.51, 0.48, 0.45, respectively). Collectively, the Size 1, Ramivorous, Arboreal, and Semi-fossorial variables account for most of the dissimilarity between La Venta and these latter biomes.

The Quebrada Honda fauna is most similar to Biomes 9, 7, and 2 (dissimilarity = 0.30, 0.31, 0.31, respectively; Appendix 9). The Quebrada Honda fauna is most dissimilar to Biomes 1, 12, and 13 (dissimilarity = 0.47, 0.46, 0.40, respectively). Collectively, the Arboreal, Size 1, Size 4, and Ramivorous variables account for most of the dissimilarity between Quebrada Honda and these latter biomes.
The Santa Cruz fauna is most similar to Biomes 2, 9, and 7 (dissimilarity = 0.36, 0.39, 0.41, respectively; Appendix 9). The Santa Cruz fauna is most dissimilar to Biomes 12, 13, and 1 (dissimilarity = 0.53, 0.52, 0.49, respectively). Collectively, the Ramivorous, Size 1, Size 4, and Arboreal variables account for the most dissimilarity between Santa Cruz and these latter biomes.

The Tinguiririca fauna is most similar to Biomes 7, 2, and 10 (dissimilarity = 0.36, 0.37, 0.38, respectively; Appendix 9). The Tinguiririca fauna is most dissimilar to Biomes 12, 13, and 1 (dissimilarity = 0.52, 0.51, 0.48, respectively). The Ramivorous, Arboreal, and Size 1, and Insectivorous variables account for the most dissimilarity between Tinguiririca and these latter biomes.

The Rümikon fauna is most similar to Biomes 7, 5, and 4 (dissimilarity = 0.15, 0.16, 0.17, respectively; Appendix 9). The Rümikon fauna is most dissimilar to Biomes 1, 9, and 2 (dissimilarity = 0.24, 0.23, 0.21, respectively). Collectively, the Size 3, Arboreal, Carnivorous, and Size 2 variables account for the most dissimilarity between Rümikon and these latter faunas.

Classification Trees

The CT for vegetative cover (CT1, Figure 6A) has an $r^2$ value of 0.73 and resulted in misclassification rates of 12%, 17%, and 33% for the ecoregions classified as closed, open, and mixed, respectively (Table 4.1). The Arboreal variable defines the first split in the tree; it separates 17 of the closed, 46 of the mixed, and all 24 of the open ecoregions to the left side of the tree; 57 of the closed and 35 of mixed ecoregions are separated to the right side of the tree. On the left side of the tree, the Semi-aquatic variable defines the second split; it separates four closed, three mixed, and 16 open ecoregions to the left branch and 13 closed, 43 mixed, and eight open ecoregions to the right branch. On the right side of the tree, the Size 2 variable defines the second split; it separates nine closed and 20 mixed ecoregions to the left branch and 48 closed and 15 mixed ecoregions to the right branch. In this model, La Venta is predicted to represent a closed habitat, Quebrada Honda, Santa Cruz, and Rümikon are predicted to represent mixed habitats, and Tinguiririca is predicted to represent an open habitat (Figure 6A).

The CT used to predict biogeographic realms (CT2, Figure 6B) has an $r^2$ value of 0.88 (Table 4.2). The Size 6 variable defines the first split, separating 12 of the Indo-Malayan, one of the Palearctic, and all 29 of the Afrotropical ecoregions to the left side of the tree; 13 Indo-Malayan, 25 Palearctic and all Australasian, Nearctic, and Neotropical ecoregions (N = 29, 25, and 35, respectively) are separated to the right side of the tree. On the left side of the tree, the Size 3 variable defines the second split, with 23 of the Afrotropical ecoregions...
separated to the left branch; 12 Indo-Malayan, six Afrotropical, and one of the Paleartic ecoregions are separated to the right branch. On the right side of the tree, the Size 5 variable defines the second split; 28 Palearctic, 13 Indo-Malayan, and four Nearctic are separated to the left branch, whereas 35 Neotropical, 29 Australasian, 21 Nearctic, and seven Paleartic ecoregions are separated to the right branch. In this model, La Venta, Quebrada Honda, and Tinguiririca are predicted to be within the Indo-Malaya biogeographic realm, and Rümikon and Santa Cruz are predicted to be within the Palearctic biogeographic realm (Figure 6B).

The CT used to predict biomes (CT3, Figure 7) has an $r^2$ value of 0.64 (Table 4.3). The Fruit variable defines the first split; all ecoregions from Biomes 8, 12, and 13, along with 1–17 ecoregions from all other biomes are separated to the left side of the tree, while 21 ecoregions from Biome 1 and 1–10 ecoregions from Biomes 2, 3, 4, 5, 7, 9, and 10 are separated to the right side of the tree. On the left side of the tree, the Semi-aquatic variable defines the second split; 20 of the remaining ecoregions from Biome 13 and 1–7 ecoregions from all Biomes except Biome 4 are separated to the left branch, while all 17 remaining ecoregions from Biome 4 and 1–12 ecoregions from all Biomes except Biome 1 are separated to the right branch. On the right side of the tree, the Size 0 variable defines the second split. Nine ecoregions from Biomes 1 and 7 and 1–7 ecoregions from Biomes 2, 3, 4, 7, 9 and 10 are separated to the left branch while 12 ecoregions from Biome 1 and 1–6 ecoregions from Biomes 3, 4, 5, 7, and 10 are separated to the right branch. In this model, La Venta is predicted to be within Biome 1, Santa Cruz is predicted to be within Biome 2, Quebrada Honda is predicted to be within Biome 9, Tinguiririca is predicted to be within Biome 10, and Rümikon is predicted to be within Biome 7. When the Size 0 variable is removed from the analysis, Rümikon is predicted to be within Biome 10; all other predictions are unchanged (Figure 7).
PALEOHABITAT RECONSTRUCTIONS

The five types of analyses described above result in paleohabitat interpretations that are broadly consistent with those of previous investigations (Hantke, 1984; Kay and Madden, 1997a, b; Flynn et al., 2003; Croft et al., 2008; Kälin and Kempf, 2009; Kay et al., 2012; Catena et al., 2017). Thus, they confirm our first prediction that these EDA-based paleoenvironmental reconstructions will be congruent with reconstructions based on other types of data. Nevertheless, detailed examination of the mammalian community structures of these sites reveals that they are not analogous to modern faunas from their corresponding biogeographic realm (i.e., Neotropical and Palearctic). Rather, they more closely resemble modern sites from other biogeographic realms. This confirms our second prediction and illustrates the importance of using a global comparative dataset when investigating ancient mammal community paleoecology, especially for fossil sites in biogeographic realms that have experienced significant changes in climate and/or faunal composition between the time of deposition and the Recent.

La Venta, Colombia

The Hierarchical Cluster Analysis (HCA) and Classification Tree (Figures 5, 6B) indicate that the ecological structure of La Venta most closely resembled those of modern mammal communities from the Indo-Malaya biogeographic realm. Even though La Venta plots within the range of Australasian faunas on Axis 3 and 4 of the Correspondence Analysis (Figure 3A), this is primarily due to its high proportion of insectivorous species and low proportion of predatory species, a pattern also noted by Croft (2006) and attributed to a lack of

<table>
<thead>
<tr>
<th>TABLE 4. Misclassification percentages of the Classification Tree analyses. 1) CT1 – vegetative cover; 2) CT2 – biogeographic realm; 3) CT3 – biome.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. CT1 - Vegetative Cover</strong></td>
</tr>
<tr>
<td>Total Correctly Incorrectly Misclassification</td>
</tr>
<tr>
<td>Closed 73 64 9 12%</td>
</tr>
<tr>
<td>Mixed 82 68 14 17%</td>
</tr>
<tr>
<td>Open 24 16 8 33%</td>
</tr>
<tr>
<td><strong>2. CT2 - Biogeographic Realm</strong></td>
</tr>
<tr>
<td>Total Correctly Incorrectly Misclassification</td>
</tr>
<tr>
<td>Australasia 29 28 0 0%</td>
</tr>
<tr>
<td>Afrotopics 29 28 1 3%</td>
</tr>
<tr>
<td>Indo-Malaya 25 21 4 16%</td>
</tr>
<tr>
<td>Nearctic 25 22 3 12%</td>
</tr>
<tr>
<td>Neotropics 35 31 4 11%</td>
</tr>
<tr>
<td>Palearctic 36 33 3 8%</td>
</tr>
<tr>
<td><strong>3. CT3 - Biome</strong></td>
</tr>
<tr>
<td>Total Correctly Incorrectly Misclassification</td>
</tr>
<tr>
<td>Biome 1 22 16 6 27%</td>
</tr>
<tr>
<td>Biome 2 10 4 6 60%</td>
</tr>
<tr>
<td>Biome 3 10 6 4 40%</td>
</tr>
<tr>
<td>Biome 4 20 16 4 20%</td>
</tr>
<tr>
<td>Biome 5 12 3 9 75%</td>
</tr>
<tr>
<td>Biome 7 16 9 7 43%</td>
</tr>
<tr>
<td>Biome 8 15 7 8 53%</td>
</tr>
<tr>
<td>Biome 9 16 12 4 25%</td>
</tr>
<tr>
<td>Biome 10 18 13 5 27%</td>
</tr>
<tr>
<td>Biome 12 16 9 7 44%</td>
</tr>
<tr>
<td>Biome 13 24 21 3 13%</td>
</tr>
</tbody>
</table>
placental carnivores (Croft, 2006; see also Croft et al., 2018a). However, La Venta differs from all Australasian faunas in CT2 (Figure 6B) in the proportion of species in mass categories Size 5 and Size 6; such species are absent in Australasian ecoregions but present in Indo-Malayan ecoregions. Therefore, we consider the paleofauna of La Venta to more closely resemble faunas of Indo-Malaya than those of Australasia in terms of overall ecological structure.

The results of the SIMPER analysis and CT3 (Appendix 9; Figure 7) indicate that the paleofauna of La Venta resembles a modern fauna from Biome 1, moist tropical to subtropical broadleaf forests. This biome is characterized by low temperature variability, high rainfall (MAP generally > 2,000 mm), and semi-evergreen to evergreen deciduous vegetation (Holdridge et al., 1971). Biome 1 ecoregions are mainly present in the modern Indo-Malaya, Neotropic, and Australasia biogeographic realms. Our MAP estimate of 1,917–2,035 mm for La Venta is consistent with those of Biome 1 ecoregions as well as Kay and Madden’s (1997a, b) estimate of 1,500–2,000 mm for the site. The classification tree for vegetative cover (CT1; Figure 6A) indicates that the paleofauna of La Venta resembles that of a modern closed vegetation habitat. Overall, the paleohabitat of La Venta is reconstructed as a humid, closed tropical forest, also consistent with the findings of Kay and Madden (1997a, b).

Quebrada Honda, Bolivia

The collective results of the CA (Figures 2A, 3A) and HCA (Figure 5) indicate that the paleofauna of Quebrada Honda had an ecological structure most similar to that of a modern mammal community from the Afrotropics. While CT2 (Figure 6B) predicts Quebrada Honda to more closely resemble faunas from Indo-Malaya, this is largely due to similar proportions of Size 3 species. Quebrada Honda is distinguished from Indo-Malayan faunas on CA2 (Figure 2A) by its high proportions of Size 4, Terrestrial, and Graminivorous species. In this analysis, faunal communities with similar proportions of species within those ecological categories are only present in modern Afrotropical faunas. We therefore consider the fauna of Quebrada Honda to have an ecological structure most similar to Afrotropical mammal communities.

The results of the HCA, SIMPER, and CT3 (Appendix 9; Figures 5, 7) indicate that the paleofauna of Quebrada Honda had a community structure similar to a modern fauna from Biome 9, flooded grasslands and savannas. This biome is characterized by tropical and subtropical climates, seasonal rainfall generally ranging from 500–1,300 mm/year, vegetation with a non-continuous canopy, and seasonal to year-round flooding (Holdridge et al., 1971; González-Jiménez, 1979). Ecoregions from Biome 9 mainly occur in the modern Afrotropic, Neotropic, and Palearctic biogeographic realms.

The estimated MAP of 877–1,032 mm for Quebrada Honda based on linear regression is within the MAP range of Biome 9 ecoregions used in this study and is also consistent with the previous MAP estimate of ~1,000 mm by Catena et al. (2017). The presence of alluvial systems and periodic of flooding events at Quebrada Honda has previously been indicated by sedimentological, paleosol, and ichnofossil analyses (MacFadden...
and Wolff, 1981; Croft et al., 2011; Catena et al., 2017) and is compatible with the presence of a freshwater turtle at the site (Acanthochelys; Cadena et al., 2015). However, other aquatic animals such as crocodylians have not yet been discovered at Quebrada Honda. The Classification Tree for vegetative cover (CT1, Figure 6A) indicates that the mammal community of Quebrada Honda had an ecological structure similar to that of a modern habitat with a mixed vegetative cover. Overall, the paleohabitat of Quebrada Honda is reconstructed as a subtropical, river-associated savanna with a relatively open, mixed vegetative cover. This is also consistent with the paleoenvironmental findings of Catena et al. (2017).

**Santa Cruz, Argentina**

The results of the CT2 (Figure 6B) indicate that the paleofauna of Santa Cruz had an ecological structure similar to a modern mammal community from the Palearctic. While Santa Cruz is within the range of Afrotropical faunas in the CA (Figures 2A, 3A), it differs from Afrotropical ecoregions in its low proportion of species within the Size 6 category. The results of the SIMPER and CT3 (Appendix 9; Figure 7) analyses indicate that the fauna of Santa Cruz was ecologically similar to a modern fauna from Biome 2, dry tropical to subtropical broadleaf forests. This biome is characterized by low temperature variability, predominantly deciduous trees, and strongly seasonal rainfall generally ranging from 500–2,000 mm/yr but sometimes greater (Holdridge et al., 1971). Ecoregions from Biome 2 occur mainly in the modern Neotropical and Indo-Malaya biogeographic realms and are absent in the modern Palearctic biogeographic realm.

The estimated MAP of 1,210–1,286 mm for Santa Cruz based on linear regression is within the range of MAPs from ecoregions of Biome 2 used in this study is higher than the previous MAP estimate of ~600 mm by Flynn et al. (2003). The discrepancy between these MAP estimates could be due, in part, to the higher proportion of taxa coded for Ramivory (i.e., Browse) in this analysis compared to Flynn et al. (2003). In our analysis, taxa with high-crowned teeth (notohippids, archaeohyracids, and the interatheriine Santigorothia) were coded both Ramivorous and Graminivorous (i.e., as both browsers and grazers) instead of just Graminivours (grazers) as in Flynn et al. (2003); this was done to lessen the chance of misinterpreting diet based on tooth crown height (see Croft et al., 2008). Hypsodonty (having high-crowned teeth) in ungulates is considered an adaptation to abrasive diets, such as those with high amounts silica and/or abiotic material that is inadvertently consumed in the form of dust and grit (Janis, 1988; Pérez Barbería and Bordon, 2001; Williams and Kay, 2001; Reguero et al., 2010). In modern faunas, higher relative abundances of browsing taxa (i.e., ungulates) are associated with higher annual rainfall, generally around 1,000 mm annually (Bell, 1971; Cumming, 1982; McNaughton and Georghiades, 1986). In our analysis, ecoregions with faunas that have the highest proportions of Ramivorous species (≥ 25%) have an average MAP of 1,417 and a median MAP of 1,017 mm. Mean annual precipitation estimates for Patagonian sites equivalent in age to Tinguiririca based on phytolith ecomorphology and stable carbon isotopes are quite low (ca. 200-700 mm; Dunn et al., 2015; Kohn et al., 2015). The relatively arid conditions indicated by these studies imply that high-crowned herbivores from Tinguiririca may have been open habitat specialists (i.e., pure grazers) rather than mixed feeders, as coded here.

**Tinguiririca, Chile**

The results of the CA and CT2 (Figures 2A, 3A, 6B) indicate that the paleofauna of Tinguiririca had an ecological structure most similar to that of modern mammal communities from Indo-Malaya. The results of the SIMPER and CT3 (Appendix 9; Figure 7) indicate that the paleofauna had a community structure similar to a modern fauna from Biome 10, the montane grasslands and shrublands. This biome is characterized by montane elevation generally ranging from 1,500–3,500 m, temperate, humid to semi-arid climates, and open, shrubby vegetation (Holdridge et al., 1971; North et al., 2016). Ecoregions from Biome 10 occur mainly in the modern Afrotropic, Neotropic, and Palearctic biogeographic realms; a single ecoregion from Biome 10 is present in the Indo-Malaya biogeographic realm. The estimated MAP of 1,079–1,186 mm for Tinguiririca based on linear regression is within the range of MAPs from ecoregions of Biome 2 used in this study is higher than the previous MAP estimate of ~600 mm by Flynn et al. (2003). The discrepancy between these MAP estimates could be due, in part, to the higher proportion of taxa coded for Ramivory (i.e., Browse) in this analysis compared to Flynn et al. (2003). In our analysis, taxa with high-crowned teeth (notohippids, archaeohyracids, and the interatheriine Santigorothia) were coded both Ramivorous and Graminivorous (i.e., as both browsers and grazers) instead of just Graminivours (grazers) as in Flynn et al. (2003); this was done to lessen the chance of misinterpreting diet based on tooth crown height (see Croft et al., 2008). Hypsodonty (having high-crowned teeth) in ungulates is considered an adaptation to abrasive diets, such as those with high amounts silica and/or abiotic material that is inadvertently consumed in the form of dust and grit (Janis, 1988; Pérez Barbería and Bordon, 2001; Williams and Kay, 2001; Reguero et al., 2010). In modern faunas, higher relative abundances of browsing taxa (i.e., ungulates) are associated with higher annual rainfall, generally around 1,000 mm annually (Bell, 1971; Cumming, 1982; McNaughton and Georghiades, 1986). In our analysis, ecoregions with faunas that have the highest proportions of Ramivorous species (≥ 25%) have an average MAP of 1,417 and a median MAP of 1,017 mm. Mean annual precipitation estimates for Patagonian sites equivalent in age to Tinguiririca based on phytolith ecomorphology and stable carbon isotopes are quite low (ca. 200-700 mm; Dunn et al., 2015; Kohn et al., 2015). The relatively arid conditions indicated by these studies imply that high-crowned herbivores from Tinguiririca may have been open habitat specialists (i.e., pure grazers) rather than mixed feeders, as coded here.
The Classification Tree for vegetative cover (Figure 6A) indicates that the paleofauna of Tinguiririca resembles that of a modern open vegetation habitat. Overall the paleoenvironment of Tinguiririca is reconstructed as a temperate, open shrubland; despite the differences in estimated MAP, this is consistent with the previous paleoenvironmental findings of Flynn et al. (2003) and Croft et al. (2008).

Rümikon, Switzerland

The results of the CA, HCA (Figures 2A, 3A, and 6), and CT2 (Figure 6B) indicate the paleofauna of Rümikon had an ecological structure most similar to that of a modern mammal community from the Palearctic. The results of the SIMPER and CT3 (Appendix 9, Figure 7) indicate that the paleofauna of Rümikon had a community structure similar to a modern fauna from Biome 7, tropical and subtropical grasslands, savannas, and shrublands. This biome is characterized by semi-arid to sub-humid climates, seasonal rainfall generally ranging from 500–1,300 mm/year, and an open to mixed vegetative cover dominated by shrubs and scattered trees (González-Jiménez, 1979). Biome 7 ecoregions mainly occur in the modern Afro Arctic and Neotropic biogeographic realms and are absent from the modern Palearctic biogeographic realm. The estimated MAP of 1,115–1,214 mm for Rümikon based on the linear regression is lower than the general estimate of 1,300–1,500 mm for all fossil localities within the Swiss Upper Freshwater Molasse of Hantke (1984) but within the range of the MAPs from ecoregions of Biome 7 used in this study.

Classification Tree 1 (Figure 6A) indicates that, based on the structure of its fauna, the paleohabitat of Rümikon resembles a modern habitat with a mixed vegetative cover. Overall the paleohabitat of Rümikon is reconstructed as a temperate, mixed woodland, which is consistent with the previous interpretations of Kälin and Kempf (2009). Paleoenvironmental interpretations for units within the Upper Freshwater Molasse of Hantke (1984) but within the range of the MAPs from ecoregions of Biome 7 used in this study.

Classification Tree 1 (Figure 6A) indicates that, based on the structure of its fauna, the paleohabitat of Rümikon resembles a modern habitat with a mixed vegetative cover. Overall the paleohabitat of Rumikon is reconstructed as a temperate, mixed woodland, which is consistent with the previous interpretations of Kälin and Kempf (2009). Paleoenvironmental interpretations for units within the Upper Freshwater Molasse of Hantke (1984) but within the range of the MAPs from ecoregions of Biome 7 used in this study.

DISCUSSION

Biotic History, Contingency, and Non-Analog Faunas

Perhaps the most noteworthy result of this study is that all four of the South American paleofaunas analyzed most closely resemble modern faunas from outside of (rather than within) the Neotropic biogeographic realm in all analyses. These results confirm our second prediction and demonstrate that the ecological structure of South American paleofaunas was quite different from that of modern mammal communities of that continent. This difference is largely driven by proportions of species in two diet categories, Carnivorous and Ramivorous, and three body mass categories, Size 1, 5, and 6. With the exception of Santa Cruz, South American paleofaunas have fewer carnivorous species than any modern fauna from the Neotropic biogeographic realm; only faunas from Australasia typically have similarly low proportions of carnivorous species (cf. Croft, 2006). Conversely, the South American paleofaunas we analyzed have higher proportions of ramivorous (browsing) species than any modern Neotropical fauna used in this study; these partly or wholly ramivorous taxa are primarily large (size categories 4–6) xenarthrans (ground sloths) and ungulates (notoungulates, litopterns, astrapotheres). On average, these taxa represent of 41% of the species in these paleocommunities. In modern faunas, ramivorous species are primarily represented by extant ungulates such as deer (cervids) and peccaries (tayassuids) and several groups of rodents (e.g., cricetids and echimyids). These taxa account for only 18% of species in modern South American mammal communities, on average. The decrease in the proportion of ramivorous species in modern South American mammal faunas could be related to changes in global climate over the course of the Cenozoic. Janis et al. (2000) suggested that declining temperatures and associated lower primary productivity could account for fewer browsing species in modern North American mammal faunas compared to those of the Miocene. A similar shift could have taken place in South America. The extinction of large (> 44 kg), folivorous (browsing and/or grazing) species during the late Cenozoic (Koch and Barnoksy, 2006; Cione et al., 2009) is
Another possible, non-exclusive explanation for this difference, as most of the browsers in these paleofaunas were large-bodied taxa.

Another difference between the South American paleofaunas used in this study and modern South American faunas is the lower proportions of Size 1 species. While this could partially reflect a preservation or collection bias (see below for further discussion), this lower proportion could also reflect differences in community structure between pre-, and post-GABI faunas. In modern South American faunas, the species in the Size 1 mass category are predominantly cricetid rodents, a group that dispersed to South America by the late Miocene and today accounts for well over 400 South American species (Barbiere et al., 2016). Prior to the rise of cricetids, Size 1 ecological niches in South America were mainly occupied by metatherians (e.g., argyrolagoids; Zimicz, 2011), which were far less diverse. Most native South American rodents (caviomorphs) were larger or much larger than most modern cricetids (cf. Walton, 1997). It is certainly the case that many small species of metatherians and rodents are unrepresented in the South American fossil record, but it is unlikely that such undiscovered species rival the taxonomic diversity represented by modern cricetids. Perhaps a paucity of very small species in pre-GABI faunas is one of the reasons cricetids experienced such an explosive radiation in South America (Parada et al., 2013; Maestri et al., 2017).

On the other end of the body mass scale, with the exception of Quebrada Honda, South American paleofaunas also have higher proportions of Size 5 species than modern faunas from the Neotropical biogeographic realm. Additionally, no modern fauna within this realm has any species in the Size 1 mass category, while the species in this category account for 8% and 6% of the species of La Venta and Quebrada Honda, respectively (Appendix 10). The absence of such very large species in modern Neotropical faunas can largely be attributed to the Pleistocene megafaunal extinction (Koch and Barnosky, 2006; Cione et al., 2009; Barnosky and Lindsey, 2010; Villavivencio et al., 2015). The only faunas with similar proportions of species within the Size 5 and 6 variables are from the Afrotropics and Indo-Malaya, the two biogeographic realms that did not experience a significant megafaunal extinction event at the end of Pleistocene (~100 ka) (Barnosky et al., 2004; Sandom et al., 2014; Smith et al., 2018).

Taken together, the ecological characteristics of South American paleofaunas suggest that they are not closely analogous to modern Neotropical faunas. Although the idea of non-analog faunal associations was originally developed in the context of Pleistocene mammal associations (Graham, 2005; Soligo and Andrews, 2005; Semken et al., 2010), it has subsequently been used more generally to refer to faunas that are not closely similar to any modern fauna. The distinctive nature of South American paleofaunas is clearly evident in Figure 5, in which they form a cluster distinct from all modern faunas. This cannot reasonably be interpreted as a sampling bias common to fossil faunas in general, given that: (1) thorough sampling was the primary criterion used to select the fossil faunas analyzed; and (2) the other (non-South American) paleofauna included in the analysis plots separately, among modern faunas. Rather, this cluster of South American paleofaunas is distinguished by its unique ecological characteristics (mainly, a low proportion of Size 1 species and a high proportion of ramivorous species). In this sense, South American paleocommunities can be accurately characterized as broadly non-analog compared to those present today. This conclusion, which is based on the entire mammal community, parallels similar conclusions arising from study of the early Miocene carnivore/predator guild of South America (Croft et al., 2018a). Our findings are also concordant with the idea that non-analog vegetational communities were present in southern South America during the Eocene and Oligocene (Dunn et al., 2015; Kohn et al., 2015). As more and better data ecological data are gathered about South American paleofaunas and localities, we suspect that many other non-analog aspects of South American ecological communities will come to light.

Changes in Global Climate

Changes in climate over the course of the Cenozoic have changed the distributions of biomes and thus faunal communities within biogeographic realms (Kutzbach et al., 1997; Gates, 2003; Hardy, 2003; Peñuelas and Boada, 2003; Ortiz-Jaureguizar and Cladera, 2006; Rowan et al., 2019). The faunal resemblance of Rümikon to Biome 7 (tropical and subtropical grasslands, savannas, and shrublands), a biome absent from the modern Palearctic biogeographic realm, is almost certainly due in large part to the underlying climatic history of Central Europe. In this region, global cooling immediately following the Mid-Miocene Climatic Optimum led to decreased precipitation and more open habitats (Hanke, 1984; Van der Meulen and Daams, 1992; Kälin and Kempf, 2009). Modern
central Europe is generally humid and forested, primarily composed of Biomes 4 and 5 (Kottek et al., 2006; Peel et al., 2007). The modern humid conditions in central Europe are largely a result of the relatively recent (~5–3 Ma) closure of the Central American Seaway and the subsequent strengthening of the Atlantic Gulf Stream (Raymo et al., 1989; Murdock et al., 1997; Haug and Tiedemann, 1998; Cederbom et al., 2004). Geographic shifts in biomes have occurred during the Cenozoic in all biogeographic realms, and no single modern biogeographic realm includes all biomes; one to three biomes are missing from each. Thus, restricting an EDA of a paleofauna to a modern comparative dataset from the same biogeographic realm could result in an inaccurate reconstruction due to a priori limits on the breadth of biomes to which it is being compared.

**Fidelity of the Fossil Record**

The paleohabitat reconstructed for Rümikon based on this EDA differs from those inferred for the South American paleofaunas. This confirms our third prediction of no systematic, sampling bias among paleofaunas that results in spurious similarities between all five paleofaunas. Rümikon differs from the South American sites partly in its higher proportion of species within the Size 1 category (Appendix 10). As noted previously, South American faunas are characterized by relatively few species of this size (3–10%), lower than that of any modern faunas used in this analysis (Appendix 10). By contrast, 35% of species at Rümikon are in the Size 1 category, a proportion within 4% of the mean for most other biogeographic realms (all except Indo-Malaya, at 20%). The relatively low proportion of Size 1 species in the South American paleofaunas could be the result of a preservation bias against small species, a common phenomenon in the fossil record (e.g., Wolff, 1975; Damuth, 1982; Kidwell and Flessa, 1995). However, it seems unlikely that such a bias would affect all four South American paleofaunas relatively uniformly but not Rümikon, particularly considering the geographic and temporal breadth of the South American faunas and the diversity of depositional environments they represent. As noted above, the low proportion of Size 1 species in middle Cenozoic South American paleofaunas is more likely due to the continent’s peculiar biogeographical history than a systematic bias in preservation or collection methods.

In a similar vein, the congruence between the analytical results obtained by including and excluding the Size 0 variable indicates that – even if a preservation or collection bias against very small (≤ 10 g) species is a common feature of Cenozoic mammal paleofaunas – it is unlikely to have a major effect on reconstructing community structure and paleohabitat. For the four South American paleofaunas, reconstructed habitats were the same with and without the Size 0 variable in all analyses. Nearly the same was true for Rümikon; excluding the Size 0 variable did not affect the overall habitat reconstruction, though the paleofauna was grouped with Biome 10 rather than Biome 7 in one analysis (the Classification Tree for biome prediction, CT3). However, this had no effect on the overall paleohabitat reconstruction, as Rümikon lacks Size 0 species, and Rümikon was also found to have a similar faunal composition to Biome 7 in the SIMPER analysis (Appendix 9).

**Scales of Comparison**

When performing a paleoecological analysis such as this, an important factor to consider is the geographic area represented by both the modern and fossil faunas (Kovarovic et al., 2018). Ideally, the areas of the two should be similar, though this can be challenging to achieve in practice due to a variety of factors. For example, modern species lists typically derive from national parks and other protected areas, which vary substantially in size; these most often represent gamma diversity (the diversity of a landscape or geographic area) rather than alpha diversity (diversity within a single community, typically < 10 km²; Whittaker, 1977; Stoms and Estes, 1993) and can differ from one another in area by several orders of magnitude (cf. Kay and Madden, 1997). The area represented by a fossil list can also vary depending on several factors; limited outcrop exposure can lead to underestimates of geographic area, while time-averaging can have the opposite effect, particularly in the context of habitats that shifted spatially through time. As discussed extensively by Croft (2013), it has not yet been determined how these different spatial scales can affect ecological diversity analyses of fossil faunas. Similarly, it is not possible to know the degree to which our findings are the result of using regional faunas as opposed to more restricted ones (e.g., Kay et al., 2012; Spradley et al., 2019). These approaches are likely complimentary to one another; studies based on more restricted geographic areas may be more suitable for incorporating finer-scale habitat and climate data, while those based on larger areas may be more appropriate for revealing general or large-scale patterns.
CONCLUSIONS AND FUTURE DIRECTIONS

Accurate paleohabitat reconstructions are essential for understanding the evolution and ecologies of paleofaunas. Our study is the first EDA of South American paleofaunas using a comparative dataset of modern temperate and tropical faunas from multiple biogeographic realms. The paleohabitat reconstructions for the South American paleofaunas indicate that modern mammalian faunas from other continents represent better ecological analogues for these mammalian paleocommunities than those from within the Neotropical biogeographic realm. The similarities between Rümikon, a European fossil fauna, and faunas from a biome not currently represented in the Palearctic biogeographic realm further underscores the importance of using global comparative datasets for accurate paleohabitat reconstructions, even in areas that have not experienced faunal changes of the same magnitude as in South America.

We envision several ways in which the accuracy and precision of ecological diversity analyses could be improved in the future, besides using a global comparative dataset. The incorporation of smaller-scale modern faunal communities from within the predicted biomes and biogeographic realms, such as those in nature reserves and natural parks, could be used to increase the precision of the paleohabitat reconstructions through the use of faunal lists based on alpha (locality), rather than gamma (regional) data and species richness. The integration of modern temperature and elevation data from the habitats of the smaller-scale faunal communities could also further refine the habitat reconstructions to a local level and be used to test hypotheses regarding the effects of local environmental differences, the spatial dynamics, diversification, and provinciality of paleofaunas. This method and this global comparative dataset can also provide essential habitat context when other paleoenvironmental indicators are not present and can be used in conjunction with other methods, such as paleopedology and ichnology, to create robust, multidisciplinary paleoenvironmental interpretations.

ACKNOWLEDGEMENTS

We thank K. Kovarovic and one anonymous reviewer for comments and suggestions that significantly improved and expanded upon the scope of this paper. We thank M. Sameh and B. Townsend for inspiring some of the ideas and techniques used in this study. D. Su for input and insight, and R. Patel for support and for assistance with programming R. This research was supported by Case Western Reserve University and the National Science Foundation (EAR 0958733 and EAR 1423058 to DAC).

REFERENCES


APPENDIX 1.

Faunal lists for the five fossil localities included in this study and ecological categorizations of extinct species. Extinction symbols are placed next to name representing the highest taxonomic rank for each group. The faunal list for La Venta is from Croft (2016); see Kay et al. (1997) for additional details. The faunal list for Quebrada Honda is from Croft (2016) with modifications based on Engelman et al. (2017, 2018), Brandoni et al. (2018), and McGrath et al. (2018); see also Croft and Anaya (2006), Croft, (2007, 2016), Croft et al. (2011), Engelman and Croft (2014), and Engelman et al. (2015). The faunal list for Santa Cruz is from Kay et al. (2012); see Vizcaíno et al. (2012) for additional paleobiological information. The faunal list for Tinguiririca is from Croft (2016) with modifications based on Wyss et al. (2018); for additional details about specific taxa, see Wyss et al. (1994), Croft et al. (2003, 2008), Flynn et al. (2003), Reguero et al. (2003), Bertrand et al. (2012), and Bradham et al. (2015). The faunal list for Rümikon is from the NOW database (http://www.helsinki.fi/science/now/), which was downloaded on November 01, 2017. See Table 2 for body mass categories. Diet abbreviations: Ca, Carnivorous; Fr, Frugivorous; Gr, Graminivorous (grass consumer/grazer); In, Insectivorous; Li, Lignivorous (bark consumer); Om, Omnivorous; Ra, Ramivorous (branch and leaf consumer/browser); Se, Granivorous (seed consumer). Locomotion abbreviations: Ar, Arboreal; Fo, Fossorial; Sa, Semi-aquatic; Sc, Scanorial; Sf, Semi-fossorial; Te, Terrestrial.

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 2.

Faunal lists for the 29 Australasian ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 3.

Faunal lists for the 29 Afrotropical ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 4.

Faunal lists for the 25 Indo-Malayan ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda
APPENDIX 5.

Faunal lists for the 25 Nearctic ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 6.

Faunal lists for the 35 Neotropical ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 7.

Faunal lists for the 36 Palearctic ecoregions used in this study. Faunal lists are from the WWF Wild Finder Database (https://www.worldwildlife.org/publications/wildfinder-database).

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 8.

Ecological attributes for extant species included in this study. Diet abbreviations: Ca, Carnivorous; Fr, Frugivorous; Gr, Graminivorous (grass consumer/grazer); In, Insectivorous; Li, Lignivorous (bark consumer); Om, Omnivorous; Ra, Ramivorous (branch and leaf consumer/browser); Se, Granivorous (seed consumer). Locomotion abbreviations: Ar, Arboreal; Fo, Fossorial; Sa, Semi-aquatic; Sc, Scansorial; Sf, Semi-fossorial; Te, Terrestrial.

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda

APPENDIX 9.

Similarity Percentage results comparing the fossil faunas to the modern biomes. Fauna abbreviations: LV, La Venta; QH, Quebrada Honda; RU, Rümikon; SC, Santa Cruz; TG, Tinguiririca. Diet abbreviations: Ca, Carnivorous; Fr, Frugivorous; Gr, Graminivorous (grass consumer/grazer); In, Insectivorous; Li, Lignivorous (bark consumer); Om, Omnivorous; Ra, Ramivorous (branch and leaf consumer/browser); Se, Granivorous (seed consumer). Locomotion abbreviations: Ar, Arboreal; Fo, Fossorial; Sa, Semi-aquatic; Sc, Scansorial; Sf, Semi-fossorial; Te, Terrestrial.

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda
APPENDIX 10.

Percentages of species in each ecological variable within each paleofauna (at top) and ecoregion (in alphanumerical order). Mean Annual Precipitation (MAP) is also included, based on Fick and Hijmans (2017) for the ecoregions and estimates for paleofaunas from a linear regression of MAP on CA1 score (see Figure 4). Fauna abbreviations: LV, La Venta; QH, Quebrada Honda; RU, Rümikon; SC, Santa Cruz; TG, Tinguiririca. Locomotion abbreviations: Ar, Arboreal; Fo, Fossorial; Sa, Semi-aquatic; Sc, Scansorial; Sf, Semi-fossorial; Te, Terrestrial. S0-S6 refer to body mass categories 0 through 6, respectively. Dietary abbreviations: Ca, Carnivorous; Fr, Frugivorous; Gr, Graminivorous; In, Insectivorous; Li, Lignivorous; Om, Omnivorous; Ra, Ramivorous; Se, Granivorous.

All appendices are available in one zipped electronic file at https://palaeo-electronica.org/content/2020/2909-south-america-mammal-eda