The age of chocolate: a diversification history of *Theobroma* and Malvaceae

James E. Richardson [1,2]*, **Barbara A. Whitlock** [3], **Alan W. Meerow** [4] and **Santiago Madriñán** [5]

[1] *Programa de Biología, Universidad del Rosario, Bogotá, Colombia,* [2] *Tropical Diversity Section, Royal Botanic Garden Edinburgh, Edinburgh, UK,* [3] *Department of Biology, University of Miami, Coral Gables, FL, USA,* [4] *United States Department of Agriculture—ARS—SHRS, National Clonal Germplasm Repository, Miami, FL, USA,* [5] *Laboratorio de Botánica y Sistemática, Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia*

Dated molecular phylogenies of broadly distributed lineages can help to compare patterns of diversification in different parts of the world. An explanation for greater Neotropical diversity compared to other parts of the tropics is that it was an accident of the Andean orogeny. Using dated phylogenies, of chloroplast *ndh*F and nuclear DNA WRKY sequence datasets, generated using BEAST we demonstrate that the diversification of the genera *Theobroma* and *Herrania* occurred from 12.7 (11.6–14.9 [95% HPD]) million years ago (Ma) and thus coincided with Andean uplift from the mid-Miocene and that this lineage had a faster diversification rate than other major clades in Malvaceae. We also demonstrate that *Theobroma cacao*, the source of chocolate, diverged from its most recent common ancestor 9.9 (7.7–12.9 [95% HPD]) Ma, in the mid-to late-Miocene, suggesting that this economically important species has had ample time to generate significant within-species genetic diversity that is useful information for a developing chocolate industry. In addition, we address questions related to the latitudinal gradient in species diversity within Malvaceae. A faster diversification rate is an explanation for the greater species diversity at lower latitudes. Alternatively, tropical conditions may have existed for longer and occupied greater areas than temperate ones meaning that tropical lineages have had more time and space in which to diversify. Our dated molecular phylogeny of Malvaceae demonstrated that at least one temperate lineage within the family diverged from tropical ancestors then diversified at a rate comparable with many tropical lineages in the family. These results are consistent with the hypothesis that Malvaceae are more species rich in the tropics because tropical lineages within the family have existed for longer and occupied more space than temperate ones, and not because of differences in diversification rate.

Keywords: Andes, chocolate, latitudinal gradient, Malvaceae, phylogenetic niche conservatism, *Theobroma*

INTRODUCTION

In his revision of *Theobroma,* Cuatrecasas (1964) recognized 22 species of understory trees all found in Neotropical lowland rainforests from the Amazon basin to Southern Mexico. Previous phylogenetic analyses (Whitlock and Baum, 1999; Silva and Figueira, 2005) have indicated that *Theobroma* is sister to *Herrania* Goudot, a genus of about 20 species monographed by
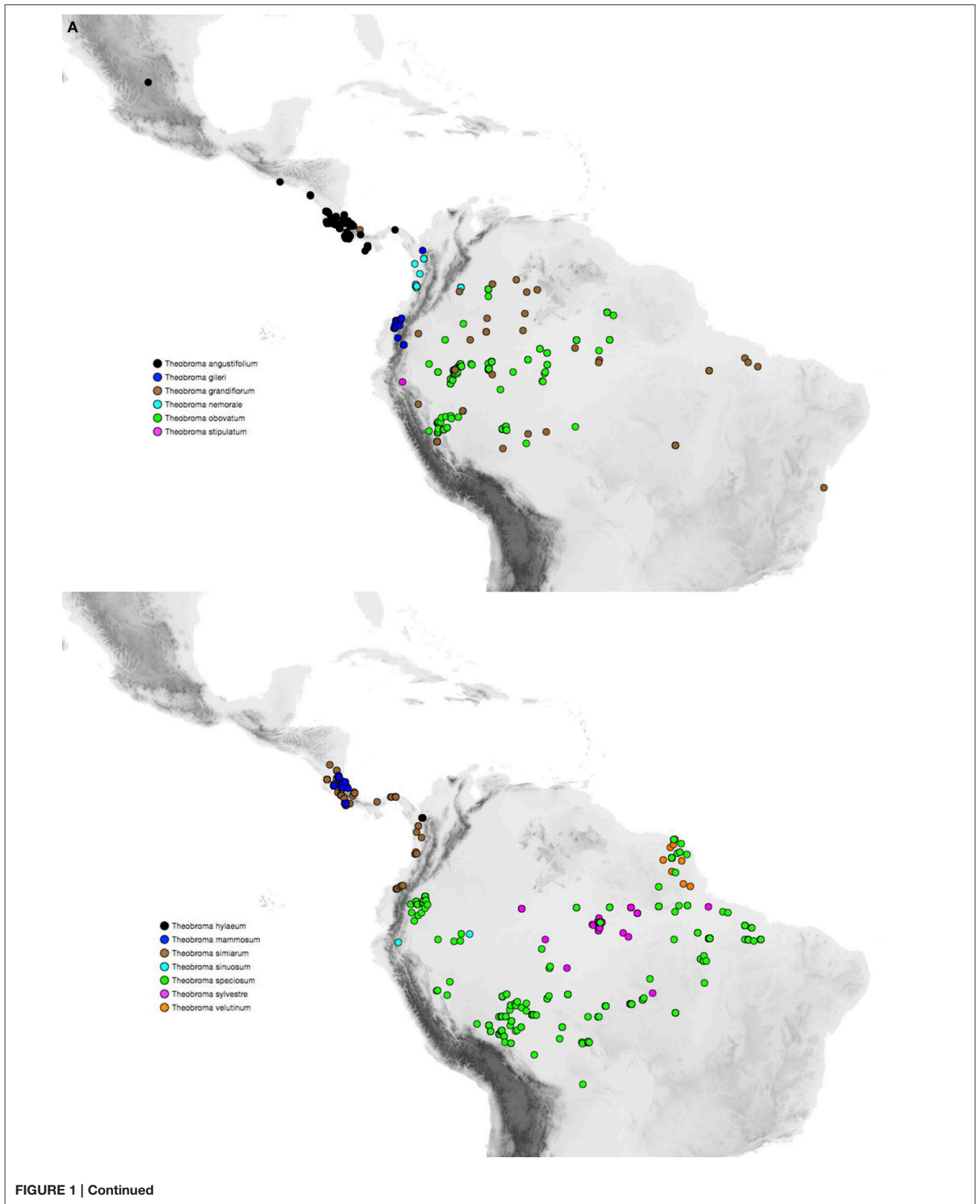
Schultes (1958), and they are both representatives of the tribe Theobromeae (Whitlock et al., 2001) along with two other genera, *Glossostemon* Desf., with one species from Arabia, and the Neotropical *Guazuma* Mill., with 2–5 species. **Figures 1A–C** indicates the distributions of species of *Theobroma* and *Herrania* with geo-referenced data taken from the Global Biodiversity Information Facility. The distribution of *Theobroma cacao* in the Neotropics is plotted separately, together with another widespread species, and this may contain some cultivated individuals. These figures show that Northwestern South America, specifically Colombia, is the most species-rich region for *Theobroma, Herrania,* and in fact all Theobromeae based on herbarium collections (26 species of Theobromeae can be found in Colombia). *Theobroma* is of great interest from a biogeographic viewpoint as it is distributed in an area that has been subject to much relatively recent geological activity. It is also of interest because it includes the economically important species *Theobroma cacao*, the source of chocolate, predicted to be a $100 billion dollar industry by 2016 with worldwide demand increasing by 2.5% a year largely driven by newly emerging markets (Markets and Markets, 2011). However, current cultivation practices face major challenges associated with the advanced age of plantations, the lack of variety of cultivated material, the low density of trees per hectare, low production, a poor comprehensive crop management strategy and fungal and viral diseases (Schnell et al., 2007; Motamayor et al., 2008). There is also a need to ensure the long-term sustainability of this industry by protecting it from the risks posed by climate change. Information on the origin and evolutionary history of *Theobroma cacao* and its relatives will assist with planning crop improvement strategies.
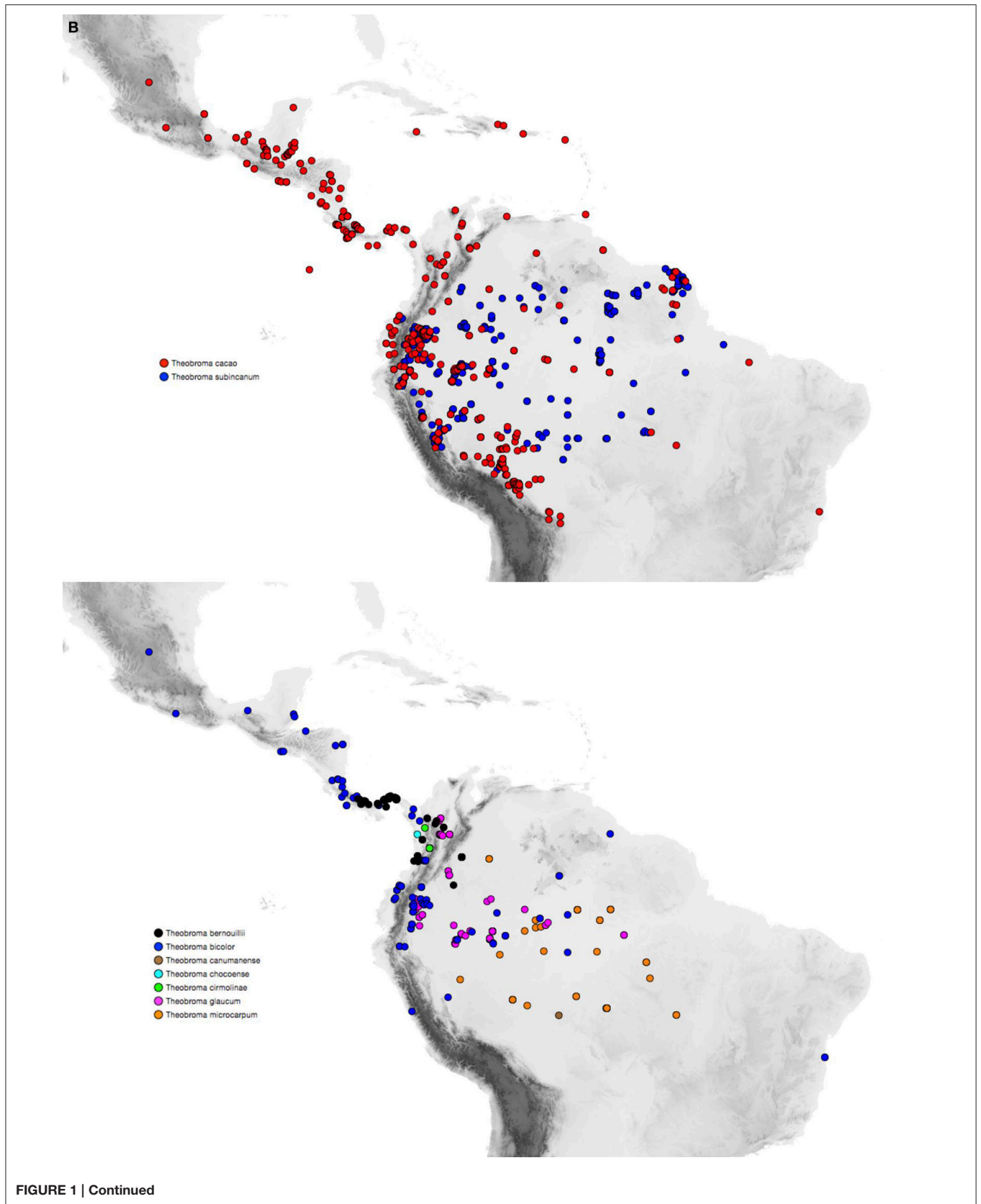
The uplift of the Andes and the bridging of the Panamanian Isthmus are two geological events that have been suggested to have had a profound impact on patterns of Neotropical plant diversification (Gentry, 1982; Burnham and Graham, 1999; Richardson et al., 2001; Knapp and Mallet, 2003; Antonelli et al., 2009; Roncal et al., 2013; Meerow et al., 2015), and may, in part, explain the greater diversity of this region in comparison with the palaeotropics. Hoorn et al. (2010) provided maps at various stage in the development of these geological systems. These events may have produced barriers to the dispersal of lineages restricted to lowland tropical forests and changed the substrate composition and fluvial systems in lowland areas (Roncal et al., 2013), facilitating diversification. The Andean Cordillera extends for 5000 km along the western coast of South America (Gregory-Wodzicki, 2000). The timing of uplift of the Andes varied from north to south and from east to west (Gregory-Wodzicki, 2000; Mora et al., 2010). The Altiplano-Puna of the Central Andes reached no more than a third of its modern elevation of 3700 by 20 Ma and no more than half its modern elevation by 10 Ma (Gregory-Wodzicki, 2000). From the middle Miocene through to the early Pliocene, elevations in the northern Eastern Cordillera of the Andes were no more than 40% of their modern values, but between two and five Ma uplift occurred at a more rapid rate reaching modern elevations by around 2.7 Ma (Gregory-Wodzicki, 2000). In Colombia the Andes divide into the Eastern, Central, and Western Cordilleras. The Western and Central
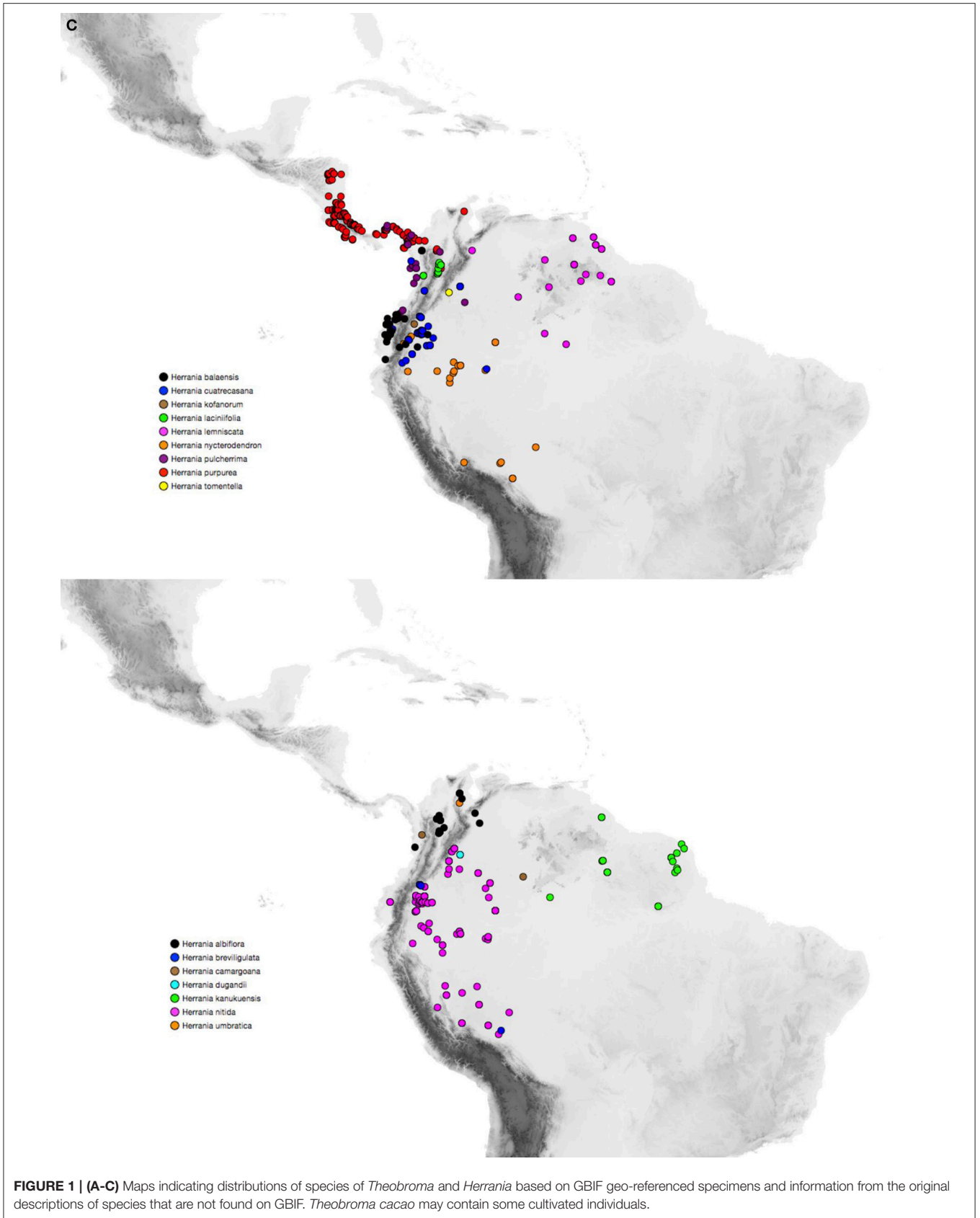
Cordilleras do not reach the northern coast of South America and therefore may not constitute barriers to dispersal for lowland-restricted organisms. The formation of the Eastern Cordillera could therefore have been crucial in erecting a montane barrier to dispersal for lowland restricted plants. The timing at which that barrier became effective in restricting migration will depend on the adaptive or dispersal capacity of individual lineages. Western Amazonia also experienced a period of submergence from the Early Miocene that resulted in the formation of an extensive wetland called the Pebas System that existed from 17 to 11 Ma (Wesselingh et al., 2002; Wesselingh, 2006; Wesselingh and Salo, 2006; Wesselingh and Ramos, 2010). This may also have acted as a barrier to dispersal of lowland wet forest restricted lineages during the period of its existence. Other potential barriers may have been the Llanos grassland ecosystem that spreads from the foothills of the Andes to the coast of Eastern Venezuela or areas of dry forest adjacent to the Andes Mountains in Colombia, e.g., to the north of Los Llanos in Arauca and Casanare or in the Inter-Andean valleys of the Magadalena and Cauca Rivers.

In addition to directly causing diversification by splitting lowland populations as mountains rose, diversification may also have resulted indirectly from changes to lowland sediments and river systems that flank the mountains. The joining of Gondwanan and Laurasian landmasses through the formation of the Isthmus of Panama was also thought to be a key event for Neotropical biotic evolution because it allowed the interchange of terrestrial species between North and South America (Simpson, 1980). According to Coates and Obando (1996) the formation of the Isthmus of Panama did not occur in one single event, but was reportedly completed in the Middle Pliocene at around 3.4–3.1 Ma. However, recent studies indicate that the land bridge may actually have begun to form from the early Miocene (Farris et al., 2011; Montes et al., 2015). The migration history of plants and animals across the Isthmus of Panama region has been reviewed by Cody et al. (2010) who concluded that plants had a greater capacity for traversing between North and South America prior to the formation of the land bridge and more recently by Bacon et al. (2015) who re-assessed biological migrations in the light of an older isthmus closure. The role of the rise of the Andes separating the Chocó and Mesoamerican regions of the Neotropics from the Amazonian and eastern regions of South America in promoting diversification has been demonstrated in various groups of organisms including birds (Gonzalez et al., 2003; Brumfield and Edwards, 2007), primates (Cortés-Ortiz et al., 2003), insects (Arrivillaga et al., 2002), rodents (Patterson and Velazco, 2008), mammals (Patterson et al., 2012), and fish (Albert et al., 2006) but few studies have focused on lowland plants (e.g., Pirie et al., 2006; Winterton et al., 2014). The distributions of both *Theobroma* and *Herrania* make them an excellent model group to study the effects of montane uplift, the closure of the Isthmus of Panama and other geological events in the region on diversification patterns in the Neotropics.

In order to fully understand the diversification of *Theobroma* and its allies, it is necessary to place it into spatial and temporal context within the family to which it belongs. The

FIGURE 1 | Continued

FIGURE 1 | Continued

FIGURE 1 | (A-C) Maps indicating distributions of species of *Theobroma* and *Herrania* based on GBIF geo-referenced specimens and information from the original descriptions of species that are not found on GBIF. *Theobroma cacao* may contain some cultivated individuals.

circumscription of Malvales has changed markedly in recent years in the light of molecular phylogenetic studies (e.g., Alverson et al., 1999). Previously recognized families have now been sunk into a broader Malvaceae. One of these, Sterculiaceae, the former home of *Theobroma*, is polyphyletic. *Theobroma* is now placed in the tribe Theobromeae within subfamily Byttnerioideae Burnett (Whitlock et al., 2001), one of nine sub-families currently recognized within Malvaceae. Byttnerioideae includes 27 genera and 650 species (Stevens, 2001) and also includes the tribes Byttnerieae, Hermannieae, and Lasiopetaleae. Most of the nine subfamilies of Malvaceae have a predominantly tropical distribution although some have strong representation at higher latitudes. The genus *Tilia* in Tilioideae is restricted to temperate areas and Malvoideae are well-represented in both tropical and temperate zones.

The comparative evolution of temperate and tropical lineages is of great interest as it may allow us to answer questions related to the latitudinal gradient in species diversity (described in e.g., Hillebrand, 2004; Jablonski et al., 2006; Brown, 2014) that is the greater species richness at lower latitudes. Temperate lineages are those found at high latitudes (or altitudes), of which there are few in Malvaceae, and does not include those found in mid-latitudinal deserts, Mediterranean or warm regions. One explanation for this latitudinal gradient is that tropical lineages have been around for longer (Stebbins, 1974) and have occupied more space. Throughout much of the history of angiosperms global temperatures have been much warmer than modern ones. A decline in temperature was experienced throughout the course of the Tertiary creating temperate conditions at higher latitudes and biome areas would have changed in response to those changes. The fossil record has copious evidence of tropical elements at higher latitudes (e.g., London Clay Flora, Reid and Chandler, 1933) during the warmer periods of the Tertiary. As outlined by Fine and Ree (2006) tropical lineages thus occupied greater areas for longer periods of time than temperate ones, and tropical groups therefore had more time and space within which to diversify. An alternative hypothesis to explain the latitudinal gradient was outlined by Mittelbach et al. (2007) who suggested that greater diversity in the tropics is due to faster diversification rates (see also, Rolland et al., 2014). Lineages that have both tropical and temperate clades may be used to compare their age and diversification rates allowing us to determine whether either of these two hypotheses is correct. The tropical/temperate distribution of Malvaceae also permits addressing questions related to phylogenetic niche conservatism (Kerkhoff et al., 2014). Were temperate lineages derived from tropical ones? If so, how often and when did those lineages arise and did their evolution coincide with climatic changes such as Tertiary cooling?

The primary aim of the present study is to use a dated molecular phylogeny to determine the effects of the Andean uplift and the formation of the Isthmus of Panama on the temporal and spatial diversification of *Theobroma* and *Herrania*. Inability to disperse across water would result in Central/South American disjunctions being dated to after the closure of the Isthmus of Panama. Similarly, if they could not disperse over mountains with an altitude of 2000 m then west/east Andean

disjunctions would be dated to *c.* 5 Ma. Diversification may also have increased in lowland areas during periods of Andean uplift that altered the landscapes of the Amazon Basin and Chocó. We also aimed to determine the age of *Theobroma cacao* and discuss the implications for the chocolate industry. Additionally we aimed to assess the diversification history of Malvaceae throughout its range. If the latitudinal gradient is to be explained by faster diversification rates in the tropics we would expect to see higher rates in tropical lineages compared with temperate ones. Alternatively, temperate lineages may have been around for less time and occupied less space than tropical ones in which case we might expect to see temperate lineages nested within tropical ones and for both to have similar diversification rates. Few temperate lineages nested within tropical ones would be consistent with phylogenetic niche conservatism in terms of cold tolerance traits.

METHODS

Map Generation

Maps were generated that included all accessions recorded in GBIF for the species within their native range as taken from monographic treatments (we excluded accessions georeferenced outside their native rage). Additionally, all specimens which mentioned "cultivated" in the specimen description were eliminated. The map may include cultivated specimens within their native range that could not be identified through the information provided in GBIF, but these should represent only a very small percentage of the total.

Sampling

We utilized two datasets in this study. We downloaded 157 plastid *ndh*F sequences from GenBank that were derived from publications by Alverson et al. (1999), Whitlock et al. (2001) Nyffeler et al. (2005), and Wilkie et al. (2006) and aligned them automatically using ClustalW in BioEdit (Hall, 1999) and then manually be eye using Mesquite (Maddison and Maddison, 2015). Of these 137 were of Malvaceae, representing each of the currently recognized subfamilies, and 20 were outgroups from other families in Malvales, and Brassicales. We also used a previously published matrix of 23 WRKY sequences of five different orthologs from the tribe Theobromeae (Borrone et al., 2007) that included 15 individuals (**Table 1**) representing 11 species of *Theobroma*, seven species of *Herrania,* and an outgroup, *Guazuma ulmifolia*, that is sister to these genera in the *ndh*F analysis.

Phylogenetic Analysis and Molecular Dating

The *ndh*F dataset was analyzed using BEAST (Drummond and Rambaut, 2007). A fossil based calibration point was used along with a secondary calibration point that was used to constrain the age of the stem node of Malvaceae. Fossil leaves of Malvoideae from the middle-late Paleocene Cerrejón Formation in Colombia (58–60 mya) were described as a new species, *Malvaciphyllum macondicus* (Carvalho et al., 2011),

TABLE 1 | Voucher specimens or USDA-ARS MIA DNA sample numbers, GenBank accession numbers of the WRKY sequences used in the phylogenetic analyses.

Species	Voucher	WRKY3	WRKY 11	WRKY 12	WRKY 13	WRKY 14
Guazuma ulmifolia Lam.	Whitlock and Bowser 360 (GH)	168	237	191	214	260
Herrania albiflora Goudot	TC01377 [CATIE]	169	238	192	215	261
H. cuatrecasana García-Barriga	Whitlock 312 (GH)	170	239	193	216	–
H. kanukuensis Schultes	Mori 24727 (NY)	171	240	194	217	262
H. nitida (Poepp.) Schultes	TC01371 [CATIE]	172	241	195	218	263
H. nycterodendron Schultes	Whitlock 315 (GH)	173	242	196	219	264
H. purpurea (Pitt.) Schultes	Whitlock 318 (GH)	174	243	197	220	265
H. umbratica Schultes	TC01376 [CATIE]	175	244	198	221	266
Theobroma angustifolium Mociño and Sessé ex DC (1)	Whitlock 303 (GH)	176	245	199	222	267
T. angustifolium (2)	TC01368 [CATIE]	177	246	200	223	268
T. bicolor Humb. and Bonpl.	Hunter 1029 (WIS)	178	247	201	224	269
T. cacao L. cv. TSH516 (1)	TC00157 [CEPLAC]	179	248	202	225	270
T. cacao (2)	Whitlock 361 (GH)	180	249	203	226	271
T. cacao (3)	Whitlock s.n. (GH)	181	250	204	227	272
T. chocoense Cuatrec.	Whitlock 356 (GH)	182	251	205	228	273
T. gileri Cuatrec.	Whitlock 301 (GH)	183	252	206	229	274
T. grandiflorum (Willd. ex Spreng.) Schum.	Whitlock 305 (GH)	184	253	207	230	275
T. mammosum Cuatrec. and León	TC01367 [CATIE]	185	254	208	231	276
T. microcarpum C. Martius (1)	Whitlock 302 (GH)	186	255	209	232	277
T. microcarpum (2)	TC01369 [CATIE]	187	256	210	233	278
T. simiarum Donn. Smith	Whitlock 321 (GH)	188	257	211	234	279
T. speciosum Willd. ex Spreng	Hunter 1033 (WIS)	189	258	212	235	280
T. velutinum Benoist	Mori 24731 (NY)	190	259	213	236	281

Abbreviations: CATIE, Centro de Agronómico Tropical de Investigación Enseñanza, Costa Rica; CEPLAC, CEPLAC/CEPLAC, Itanubo, Brazil. Only the last three digits are shown. All accession numbers begin with EF640. For example, the Guazuma ulmifolia WRKY03 GenBank Accession No. is EF640168

that can be assigned to the clade Eumalvoideae because of distal and proximal bifurcations of the costal secondary and agrophic veins that is a synapomorphy for this clade. Eumalvoid leaves and bombacoid pollen found in formations of the mid to late Paleocene of Colombia indicate that representatives of Malvoideae and Bombacoideae were present in neotropical forests at that time. The *Malvaciphyllum macondicus* fossil was used to constrain the age of the node representing the stem of eumalvoideae (indicated in **Figure 2**) using a log normal distribution, as recommended for fossil calibrations by Ho (2007) and Ho and Phillips (2009), with an offset of 60 mya (based on the older age estimate for the fossil according to Carvalho et al., 2011) and a mean of 1. This approach biases in favor of an older age estimate for this node. For the secondary calibration a normal prior distribution, as recommended for secondary calibrations by Ho (2007) and Ho and Phillips (2009), with a mean of 91.85 mya and standard deviation of 0.1 mya was assigned to the stem node of Malvales. This was based on the age with 95% confidence interval of this node derived from a dated phylogeny of all angiosperms that utilized numerous fossils (Magallón and Castillo, 2009).

For the *ndh*F analysis an XML (eXtensible Mark-up Language) input file was generated in the Bayesian Evolutionary Analysis Utility software (BEAUti) version v.1.6.2 (part of the BEAST package). The best performing evolutionary model was identified

under two different model selection criteria, the hierarchical likelihood ratio test (hLRT) and the Akaike information criterion (Akaike, 1974) as implemented in MrModelTest (Nylander, 2004). Both selection criteria indicated that a General Time Reversible (GTR) with site heterogeneity being gamma distributed and with invariant sites model was optimal. A relaxed clock uncorrelated lognormal minimal distribution was chosen based on the assumption of the absence of a strict molecular clock. To specify informative priors for all the parameters in the model, the Yule tree prior was used since it is recommended as being appropriate for species-level phylogenies (Ho and Phillips, 2009). The XML file was run in BEAST software version v.1.4.8 (Drummond and Rambaut, 2007). Five runs were performed with the MCMC chain length set to 10,000,000, to screen every 10,000 and sample every 10,000 trees. The resulting log file was imported into Tracer (Rambaut and Drummond, 2007) to check whether effective sample sizes (ESS) values were adequate for each parameter. LogCombiner and TreeAnnotator v1.6.2 (also part of the BEAST package) were also used to remove burn-ins and combine tree files and to produce the maximum clade credibility (MCC) tree that has the maximum sum of posterior probabilities on its internal nodes and summarizes the node height statistics in the posterior sample. MCC files were visualized using FigTree version 1.3.1 (Rambaut, 2009).
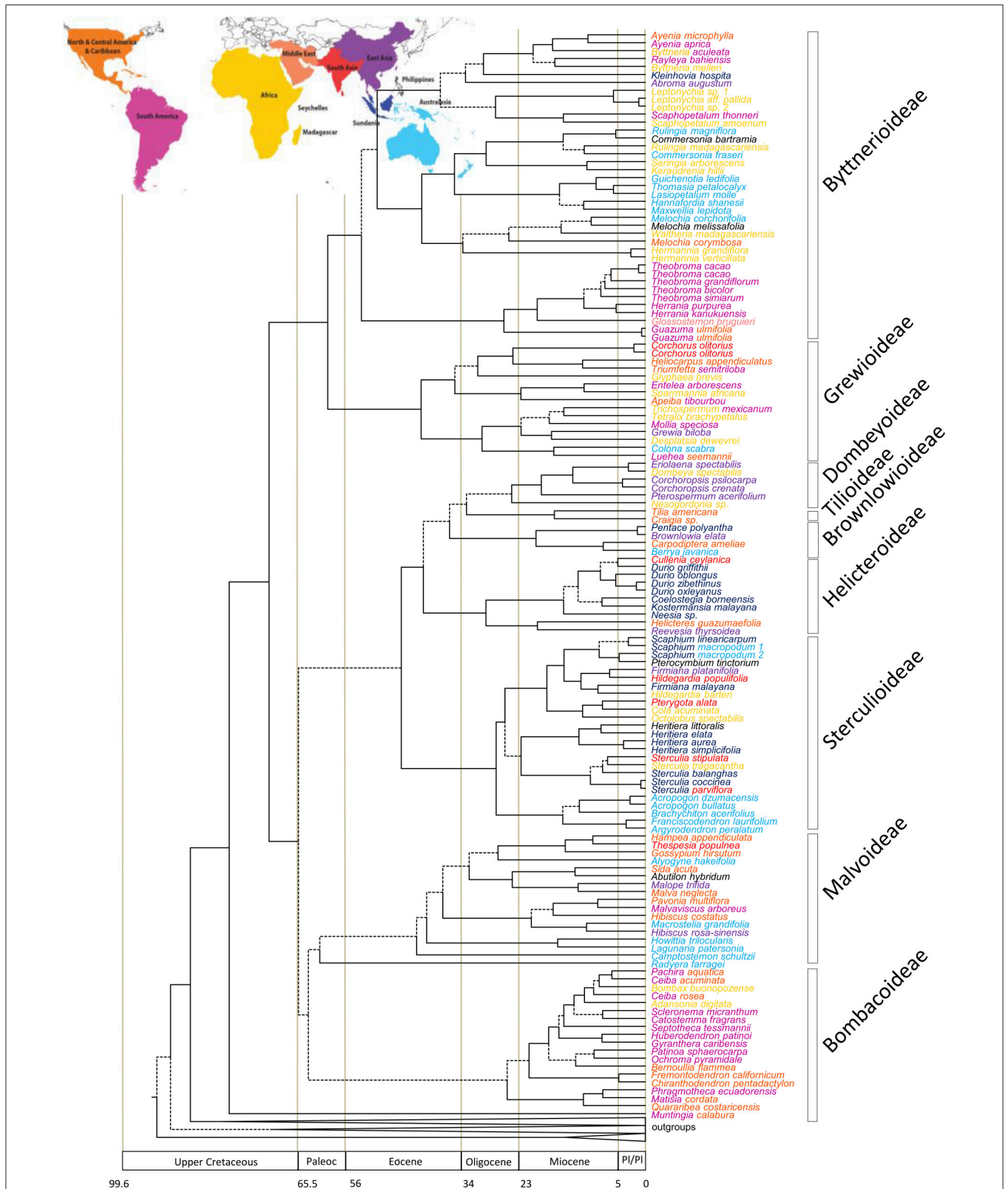
FIGURE 2 | Maximum clade credibility tree resulting from the BEAST analysis of the *ndh*F data. Broken lines represent branches with <0.95 *pp*-values. Names color coded according to the map at top of this figure. If the species is found in two areas the genus name and specific epithet are colored to match the two areas.

The BEAST package was also used to analyze the WRKY dataset employing a secondary calibration using a normal distribution based on the age of the split between *Theobroma* and *Herrania* derived from the *ndh*F analysis [11.56 (4.0–20.9 [95% HPD]) Ma] with a standard deviation of 4.5 that was chosen so that 95% of the distribution fell within the 95% confidence intervals of the age based on the *ndh*F analysis. Conditions of the analysis were identical to those for the *ndh*F analysis except that only three runs of 10 million generations were run as this was sufficient to achieve adequate ESS values for this dataset.

We calculated diversification rates using the simple estimator of Kendall (1949) and Moran (1951) where $SRln = ln(N) - ln(N_0)/T$ (where N = standing diversity, N_0 = initial diversity, here taken as = 1, and T = inferred clade age). We used this equation which was that of Magallon and Sanderson (2001) in the absence of extinction ($\in = 0.0$) and under a high relative extinction rate ($\in = 0.9$). Standing diversity was based on information from the Angiosperm Phylogeny Website (Stevens, 2001; http://www.mobot.org/MOBOT/research/APweb/), our current understanding of the phylogeny and numbers of species in each clade although we acknowledge that these numbers may change in the light of new sequence data and taxonomic studies.

RESULTS

Phylogenetic Analysis and Molecular Dating

*ndh*F–The dataset consisted of 2219 characters and 157 taxa representing 100 of the over 200 genera of the family. The ESS values for all parameters exceeded 200. The maximum clade credibility (MCC) tree from the BEAST analysis is indicated in **Figure 2** with outgroups excluded. Dates with 95% confidence intervals and posterior probabilities for all nodes are indicated in **Supplementary Figures 1**, **2** respectively, along with GenBank accession numbers. Malvaceae were strongly supported as monophyletic [posterior probability (*pp*) = 0.96] with stem and crown node ages of 78.2 (70.1–87.2 [95% HPD]) million years old (Ma) and 70.7 (63.4–78.6 [95% HPD]) Ma, respectively. Each of the subfamilies indicated in **Figure 2** was monophyletic and received >0.95 *pp*-values except Malvoideae that had a *pp*-value of 0.8, and Bombacoideae (*pp* = 0.45). The tribe to which *Theobroma* belongs, Theobromeae, had *pp* = 0.99 with stem and crown node ages of 53.4 (36.7–70.2 [95% HPD]) and 26.6 (9.6–46.0 [95% HPD]) Ma, respectively. *Theobroma* was monophyletic but with poor support and relationships within the genus were also poorly supported. The stem node of the genus was 11.6 (4.0–20.9 [95% HPD]) and the crown node was 8.4 (2.4–15.0 [95% HPD]) Ma. The two individuals of *Theobroma cacao* formed a monophyletic group that had a *pp* of 1.0 and stem and crown node ages of 6.5 (1.0–12.5 [95% HPD]), and 1.2 (0.01–3.5 [95% HPD]) Ma, respectively. Number of species, crown ages with confidence intervals and diversification rates of subfamilies and other selected clades of Malvaceae are indicated in **Table 2**.

WRKY–The dataset consisted of 23 taxa and 3987 characters. The ESS values for all parameters exceeded 200. The MCC tree from the BEAST analysis is shown in **Figure 3** and this has a topology identical to that of Borrone et al. (2007). Confidence intervals on age estimates and posterior probabilities are shown in **Supplementary Figures 3**, **4**, respectively. The genera *Theobroma* and *Herrania* were both strongly supported as monophyletic with *pp*-values of 1. *Theobroma* had stem and crown node ages of 12.7 (11.6–14.9 [95% HPD]) and 11.0 (8.6–14.3 [95% HPD]) Ma, respectively. Relationships within *Theobroma* were not strongly supported but the three individuals of *Theobroma cacao* were with stem and crown node ages for the species of 9.9 (7.7–12.9 [95% HPD]) and 0.5 (0.95–0.1 [95% HPD]) Ma, respectively. There are multiple possible trans-Andean splits in the phylogeny indicated in **Figure 3** at nodes X [3.1 (1.7–4.9 [95% HPD]) Ma] with an eastern lineage splitting from a predominantly western species (*T. gileri*), Y [8.3 (5.9–11.1 [95% HPD])], and Z [3.9 (2.4–5.9 [95% HPD]) Ma]. Some species have distributions on either side of the Andes. Trans-isthmian splits include that between the *T. mammosum* and *T. angustifolium* clade and its sister that occurred 0.7 (0.2–1.3 [95% HPD]) Ma.
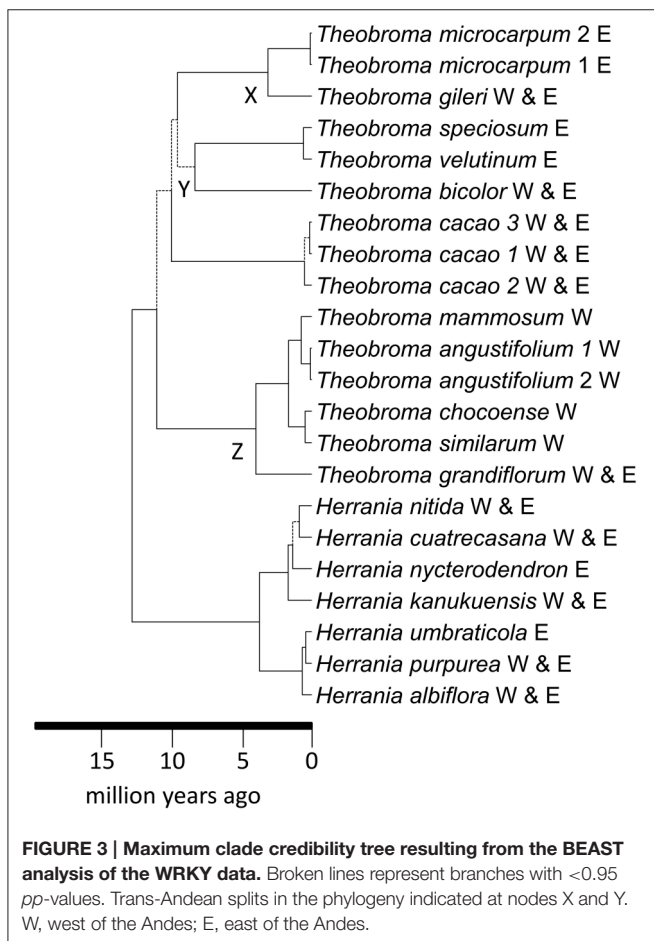
DISCUSSION

Theobromeae Biogeography

Theobroma began to diversify 11.0 (8.6–14.3 [95% HPD]) Ma, coincident with the uplift of the Northwestern Andes and resultant changes to lowland areas. **Figure 1** indicates the distributions of each of the species of *Theobroma* and *Herrania* included in the analysis. Twenty-six species of Theobromeae can be found in Colombia in the lowlands and foothills surrounding the Andes including species endemic to the Pacific coastal Chocó such as *T. chocoense* or *T. gileri*. These distributions and timings are consistent with diversification of the genus being affected by Andean uplift. This could be due to phylogenetic niche conservatism in the tribe with populations not being able to survive the cooler temperatures at higher altitudes resulting in allopatric speciation and phylogenetic splits as the mountains rose. The two trans-Andean splits in the phylogeny indicated in **Figure 3** at nodes X, Y, and Z occurred at 3.1 (1.7–4.9 [95% HPD]) Ma, 8.3 (5.9–11.1 [95% HPD]), and 3.9 (2.4–5.9 [95% HPD]) Ma, respectively. Some species have distributions on either side of the Andes. It could be that splits between groups within these species that are found on either side of the Andes have similarly old ages, e.g., *T. bicolor* or *T. cacao*, and those splits may have been caused by Andean uplift, but more samples from these species must be included to test this. The Pebas System that existed from 17 to 11 Ma does not appear to have caused any vicariance events in *Theobroma* or *Herrania* because it predates diversification within each of these genera. Possible trans-isthmian splits include that between the *T. mammosum* and *T. angustifolium* clade and its sister that occurred 0.7 (0.2–1.3 [95% HPD]) Ma. This is well after the formation of the Isthmus and therefore could have resulted from overland migration rather than long distance trans-oceanic dispersal.

TABLE 2 | Diversification rates of subfamilies of Malvaceae and selected clades in Byttnerioideae.

Subfamily	Number of species	Distribution	Crown node age (95% HPDs)	Diversification rate (Magallon$_0$)	Diversification rate (Magallon$_9$)
Malvaceae	4225	Cosmopolitan	70.7 (63.4–78.6)	0.12 (0.11–0.13)	0.09 (0.08–0.1)
Grewioideae	770	Pantropical	42.2 (25.6–60.6)	0.16 (0.11–0.26)	0.1 (0.07–0.17)
Byttnerioideae	650	Pantropical and Australia	53.4 (33.9–68.0)	0.12 (0.1–0.19)	0.08 (0.06–0.12)
Theobroma/Herrania	42	Neotropics	11.6 (4.0–20.9)	0.32 (0.18–0.93)	0.14 (0.08–0.41)
Byttnerieae	225	Byttnerieae	38.5 (22.2–56.0)	0.14 (0.1–0.24)	0.08 (0.06–0.14)
Hermannieae	270	Hermannieae	34.3 (18.8–51.8)	0.16 (0.11–0.3)	0.1 (0.06–0.18)
Lasiopetaleae	53	Australia, New Zealand, SE Asia, Madagascar	35.9 (20.6–53.5)	0.11 (0.07–0.19)	0.05 (0.03–0.09)
Sterculioideae	430	Pantropical	28.0 (13.8–46.0)	0.22 (0.13–0.44)	0.14 (0.08–0.27)
Tilioideae	50	North temperate, Central America	17.1 (2.2–33.2)	0.23 (0.12–1.78)	0.1 (0.05–0.81)
Dombeyoideae	381	Palaeotropics, Australia	25.1 (11.2–41.8)	0.24 (0.14–0.53)	0.15 (0.09–0.33)
Brownlowioideae	68	Pantropical	20.5 (5.0–37.0)	0.21 (0.11–0.84)	0.1 (0.06–0.41)
Helicteroideae	95	Pantropical	30.0 (11.7–50.4)	0.15 (0.09–0.39)	0.08 (0.05–0.2)
Malvoideae	1670	Temperate to tropical	61.2 (60.1–62.9)	0.12 (0.12–0.12)	0.08 (0.08–0.09)
Bombacoideae	120	Pantropical	26.0 (12.2–47.0)	0.18 (0.1–0.39)	0.1 (0.05–0.21)

FIGURE 3 | Maximum clade credibility tree resulting from the BEAST analysis of the WRKY data. Broken lines represent branches with <0.95 *pp*-values. Trans-Andean splits in the phylogeny indicated at nodes X and Y. W, west of the Andes; E, east of the Andes.

Some trans-Andean species evolved after the Andes had reached a significant height, therefore ancestral species must have been able to disperse over high mountain passes, e.g.,

T. gileri or *H. cuatrecasasana* (**Figures 1**, **3**). The mode of dispersal is unknown for *Theobroma* and *Herrania*, as it is for many Neotropical trees with large indehiscent fruit. Direct dispersal by vertebrates is a possibility (Cuatrecasas, 1964), including by extinct megafauna (Janzen and Martin, 1982) or early human populations. Dispersal by water is also a possibility and pods have been observed floating in rivers (Whitlock, pers. obs.), although this is not a plausible process to account for dispersal over high passes. There are areas of lower elevation along the eastern cordillera that may have allowed migration of otherwise lowland restricted lineages. However, these areas are flanked by the desert of Tatacoa and the dry forests mentioned in the introduction that lie to the west of the Eastern Cordillera and these could have acted as barriers for dispersal of wet forest restricted taxa. In fact the age of splits could also be used to estimate ages for the development of these dry biomes that could have arisen as a result of rain shadow effects resulting from montane uplift. Splits between wet forest restricted lineages of *Manilkara* (Sapotaceae) on either side of dry cerrado and caatinga vegetation in Brazil more or less coincided with diversification of cerrado lineages restricted lineages (Armstrong et al., 2014). Timing of splits within other lineages that share similar distributions need to be used along with geological and paleontological data to reconstruct biotic and abiotic history.

We demonstrate that diversification in Theobromeae coincided with major periods of uplift of the Andean mountains and that diversity in the tribe is greatest in areas flanking the Andes. Diversification could therefore have been a direct result of allopatric speciation resulting from the rise of the mountains, as mentioned above, and/or as a result of changes in substrates (as shown for example by Savolainen et al., 2006) and fluvial patterns that occurred in lowland areas around them or other modes of speciation as reviewed by Haffer (1997). Although we have only focused on diversification rates of selected clades (**Table 2**), within which the sampling

is often low and for which there is overlap in confidence intervals for age estimates, the mean rate we report for the diversification of the *Theobroma/Herrania* clade is much greater than any of the others in Malvaceae that we highlight. This is consistent with montane uplift resulting in elevated diversification rates in comparison with those not found in montane regions.

The Age of Chocolate

According to the WRKY BEAST analysis *T. cacao* diverged from its most recent common ancestor 9.9 (7.7–12.9 [95% HPD]) Ma. We prefer to accept this date rather than that from the *ndh*F analysis because the WRKY data sampled more species and gave a better resolved tree. The possibility of retrieving a younger age for the species when all taxa in the genus are added cannot of course be ignored, but the present data indicate that *T. cacao* diverged early from the remaining lineages within the genus. The phylogenetically isolated position of *T. cacao* in *Theobroma* was also recovered by Whitlock and Baum (1999) and is further supported by its placement in its own monotypic section by Cuatrecasas (1964). Its early divergence time indicates that it may have had ample opportunity to achieve a broad natural distribution with high levels of genetic diversity, although human effects on its distribution and diversity cannot be discounted. The timing and extent of diversification within the species will require greater sampling of more individuals throughout its geographic range. Studies using microsatellite data of *T. cacao* have indicated substantial genetic diversity within wild and cultivated representatives of the species (e.g., Motamayor et al., 2008; Thomas et al., 2012; Motilal et al., 2013). The timing of diversification and extent of variability has implications for the chocolate industry as basing plantations on only a percentage of this genetic diversity means that it may be at unnecessary risk from disease and other threats such as climate change (Motamayor et al., 2008). Under-utilized wild varieties may be brought into cultivation to introduce greater genetic diversity that might protect against these risks and also introduce a wider range of flavors to the industry.

The low support and the lack of a complete species level phylogeny in the WRKY phylogeny means that we still cannot be sure what the closest ancestor of *T. cacao* is (see also Whitlock and Baum, 1999). Sampling of more individuals and more genes will be necessary to determine its closest relatives and when the species itself began to diversify. The current sample of three cultivated individuals needs to be expanded to use a phylogeographic approach to determine more precisely where and when the species originated and diversified and to complement the many studies on the population genetics of the species (e.g., Motamayor et al., 2008).

Malvaceae and the Latitudinal Gradient in Species Diversity

Most species of Malvaceae are tropical and the family thus conforms to a latitudinal gradient in species diversity. Temperate lineages of Malvaceae (or at least those groups that contain species with a temperate distribution) are nested within tropical ones, consistent with this pattern that was demonstrated by Judd et al. (1994) and Baum et al. (2004) and have stem nodes dated to the mid-Eocene (Tilioideae) or from the late Oligocene (**Figure 2**, e.g., *Malva*, *Hibiscus*). This is consistent with them having evolved from tropical progenitors as temperate climates spread as a result of climatic cooling. The fact that groups from temperate regions are found in few lineages indicates a high degree of phylogenetic niche conservatism with respect to the ability to adapt to cooler temperatures. Interestingly, diversification rates based on crown node ages of the temperate lineage Tilioideae [50 species, crown node age of 17.1 (2.2–33.2 [95% HPD])] Ma and diversification rate of 0.13 [0.07–0.57] with zero extinction and 0.1 [0.05–0.81] with a high relative extinction rate) compare favorably with many sub-families that are restricted to tropical regions (**Table 2**), e.g., Brownlowioideae [68 species, crown node age of 20.5 (5.0–37.0 [95% HPD]) Ma with a diversification rate of 0.12 [0.07–0.32] with a zero extinction rate and 0.1 [0.06–0.41] with a high relative extinction rate]. The fact that this temperate lineage has a similar diversification rate to many tropical ones (**Table 2**) is consistent with the age and area of occupancy of a lineage having been a more important factor in determining the latitudinal diversity gradient than differences in diversification rates between temperate and tropical regions. This comparison is of course limited and better sampling of Malvaceae will permit comparison with other temperate lineages in the family. Studies on other groups of organisms have yielded contrasting results with higher diversification rates in the tropics for some primates (Böhm and Mayhew, 2005), birds (Cardillo, 1999; Cardillo et al., 2005; Ricklefs, 2006; Martin and Tewksbury, 2008), amphibians (Wiens, 2007) or plants (Jansson and Davies, 2008), but no significant differences for mammals and birds (Weir and Schluter, 2007) or amphibians (Wiens et al., 2006, 2009). Our results contrast with Jansson and Davies (2008) study that indicated a latitudinal gradient in diversification rates in flowering plants. Diversification patterns through time will be better determined by dating complete species level phylogenies of lineages that have both temperate and tropical elements. Care must also be taken to decouple latitudinal effects from local or regional geological processes that might have had an enormous impact on diversification rates at regional scales. It could be argued that tropical regions have been more geologically active, e.g., Andean uplift in the Neotropics and complex tectonic activity and orogenic events in Southeast Asia. The diversification rate of *Theobroma/Herrania* that we focus on here is 0.32 (0.93–0.18) with zero extinction and 0.14 (0.08–0.41) with a high relative extinction rate based on the crown node age of this group of 11.6 (4.0–20.9 [95% HPD]) Ma that is considerably faster than that of any of the subfamilies alone (**Table 2**). The diversification of these genera coincides with Andean uplift that is consistent with this event causing this faster speciation rate. Antonelli et al. (2015) have demonstrated that the Neotropics have a higher rate of evolutionary turnover and emigration than in other parts of the tropics helping to explain the reasons for the longitudinal gradient in diversity that exists in addition to a latitudinal one.

CONCLUSIONS

The diversification of *Theobroma* and *Herrania* coincide with periods of uplift in the Northwestern Andes. The few temperate lineages of Malvaceae are nested within tropical ones having evolved as temperatures cooled during the Tertiary. Tropical lineages do not generally have faster diversification rates than temperate ones but have been around for longer and likely occupied more space consistent with age and area being more important than differences in diversification rate in explaining the latitudinal gradient in species diversity. Finally, *T. cacao* diverged from its MRCA 9.9 (7.7–12.9 [95% HPD]) Ma, and has had ample time to diversify although the timing of the onset of this diversification and within species variability requires denser sampling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fevo. 2015.00120

Supplementary Figure 1 | MCC tree from the BEAST analysis of the *ndh*F data that includes GenBank numbers, error bars on node age estimates.

Supplementary Figure 2 | MCC tree from the BEAST analysis of the *ndh*F data that includes GenBank numbers and posterior probabilities of nodes.

Supplementary Figure 3 | MCC tree from the BEAST analysis of the WRKY data that includes error bars on node age estimates and posterior probabilities of nodes.

Supplementary Figure 4 | MCC tree from the BEAST analysis of the WRKY data that includes posterior probabilities of nodes.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Albert, J. S., Lovejoy, N. R., and Crampton, W. G. R. (2006). Miocene tectonism and the separation of cis- and trans-Andean river basins: evidence from Neotropical fishes. *J. South Am. Earth Sci.* 21, 14–27. doi: 10.1016/j.jsames.2005.07.010

Alverson, W. S., Whitlock, B. A., Nyffeler, R., Bayer, C., and Baum, D. A. (1999). Phylogeny of core *Malvales*: evidence from *ndh*F sequence data. *Am. J. Bot.* 86, 1474–1486. doi: 10.2307/2656928

Antonelli, A., Nylander, J. A. A., Persson, C., and Sanmartín, I. (2009). Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9749–9754. doi: 10.1073/pnas.081 1421106

Antonelli, A., Zizka, A., Silvestro, D., Scharn, R., Cascales-Miñana, B., and Bacon, C. D. (2015). An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Front. Genet.* 6:130. doi: 10.3389/fgene.2015.00130

Armstrong, K. E., Stone, G. N., Nicholls, J. A., Valderrama, E., Anderberg, A. A., Smedmark, J., et al. (2014). Patterns of diversification amongst tropical regions compared: a case study in Sapotaceae. *Front. Genet.* 5:362. doi: 10.3389/fgene.2014.00362

Arrivillaga, J. C., Norris, D. E., Feliciangeli, M. D., and Lanzaro, G. C. (2002). Phylogeography of the neotropical sand fly *Lutzomyia longipalpis* inferred from mitochondrial DNA sequences. *Infect. Genet. Evol.* 2, 83–95. doi: 10.1016/S1567-1348(02)00087-4

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., and Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6110–6115. doi: 10.1073/pnas.1423853112

Baum, D. A., Smith, S. D., Yen, A., Alverson, W. S., Nyffeler, R., Whitlock, B. A., et al. (2004). Phylogenetic relationships of *Malvatheca* (Bombacoideae and Malvoideae; Malvaceae sensu lato) as inferred from plastid DNA sequences. *Am. J. Bot.* 91, 1863–1871. doi: 10.3732/ajb.91. 11.1863

Böhm, M., and Mayhew, P. J. (2005). Historical biogeography and the evolution of the latitudinal gradient of species richness in the *Papionini* (Primata: Cercopithecidae). *Biol. J. Linn. Soc.* 85, 235–246. doi: 10.1111/j.1095-8312.2005.00488.x

Borrone, J. W., Meerow, A. W., Kuhn, D. N., Whitlock, B. A., and Schnell, R. J. (2007). The potential of the WRKY gene family for phylogenetic reconstruction: an example from the Malvaceae. *Mol. Phylogenet. Evol.* 44, 1141–1154. doi: 10.1016/j.ympev.2007.06.012

Brown, J. H. (2014). Why are there so many species in the tropics? *J. Biogeogr.* 41, 8–22. doi: 10.1111/jbi.12228

Brumfield, R. T., and Edwards, S. V. (2007). Evolution into and out of the Andes: a Bayesian analysis of historical diversification in *Thamnophilus* antshrikes. *Evolution* 61, 346–367. doi: 10.1111/j.1558-5646.2007.00039.x

Burnham, R. J., and Graham, A. (1999). The history of neotropical vegetation: new developments and status. *Ann. Mo. Bot. Gard.* 86, 546–589. doi: 10.2307/2666185

Cardillo, M. (1999). Latitude and rates of diversification in birds and butterflies. *Proc. R. Soc. Lond. B* 266, 1221–1225. doi: 10.1098/rspb.1999.0766

Cardillo, M., Orme, C. D. L., and Owens, I. P. F. (2005). Testing for latitudinal bias in diversification rates: an example using New World birds. *Ecology* 86, 2278–2287. doi: 10.1890/05-0112

Carvalho, M. R., Herrera, F. A., Jaramillo, C. A., Wing, S. L., and Callejas, R. (2011). Paleocene Malvaceae from northern South America and their biogeographical implications. *Am. J. Bot.* 98, 1337–1355. doi: 10.3732/ajb.1000539

Coates, A., and Obando, J. (1996). "Geologic evolution of the Central American Isthmus," in *Evolution and Environment in Tropical America*, eds J. Jackson, A. Budd, and A. Coates (Chicago, IL: University of Chicago Press), 21–56.

Cody, S., Richardson, J. E., Rull, V., Ellis, C., and Pennington, R. T. (2010). The great American biotic interchange revisited. *Ecography* 33, 1–7. doi: 10.1111/j.1600-0587.2010.06327.x

Cortés-Ortiz, L., Bermingham, E., Rico, C., Rodriguez-Luna, E., Sampaio, I., and Ruiz-García, M. (2003). Molecular systematics and biogeography of the Neotropical monkey genus, *Alouatta*. *Mol. Phylogenet. Evol.* 26, 64–81. doi: 10.1016/S1055-7903(02)00308-1

Cuatrecasas, J. (1964). Cacao and its allies: a taxonomic revision of the genus *Theobroma*. *Contrib. U.S. Natl. Herb.* 35, 379–614.

Drummond, A. J., and Rambaut, A. (2007). BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214

Farris, D. W., Jaramillo, C., Bayona, G., Restrepo-Moreno, S. A., Montes, C., Cardona, A., et al. (2011). Fracturing of the Panamanian Isthmus during initial collision with South America. *Geology* 39, 1007–1010. doi: 10.1130/G3 2237.1

Fine, P. V. A., and Ree, R. H. (2006). Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am. Nat.* 168, 796–804. doi: 10.1086/508635

Gentry, A. H. (1982). Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations or an accident of the Andean orogeny. *Ann. Mo. Bot. Gard.* 69, 557–593. doi: 10.2307/2399084

Gonzalez, M. A., Eberhard, J. R., Lovette, I. J., Olson, S. L., and Bermingham, E. (2003). Mitochondrial DNA phylogeography of the Bay Wren (Troglodytidae: *Thryothorus nigricapillus*) complex. *Condor* 105, 228–238. doi: 10.1650/0010-5422(2003)105[0228:MDPOTB]2.0.CO;2

Gregory-Wodzicki, K. M. (2000). Uplift history of the Central and Northern Andes: a review. *Geol. Soc. Am. Bull.* 112, 1091–1105. doi: 10.1130/0016-7606(2000)112<1091:UHOTCA>2.3.CO;2

Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonia: an overview. *Biodivers. Conserv.* 6, 451–476. doi: 10.1023/A:1018320925954

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *Am. Nat.* 163, 192–211. doi: 10.1086/381004

Ho, S. Y. M. (2007). Calibrating molecular estimates of substitution rates and divergence times in birds. *J. Avian Biol.* 38, 409–414. doi: 10.1111/j.0908-8857.2007.04168.x

Ho, S. Y. W., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Amazonia through time: andean uplift, climate change, landscape evolution and biodiversity. *Science* 330, 927–931. doi: 10.1126/science.1194585

Jablonski, D., Roy, K., and Valentine, J. W. (2006). Out of the Tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* 314, 102–106. doi: 10.1126/science.1130880

Jansson, R., and Davies, T. J. (2008). Global variation in diversification rates of flowering plants: energy versus climate change. *Ecol. Lett.* 11, 173–183. doi: 10.1111/j.1461-0248.2007.01138.x

Janzen, D. H., and Martin, P. S. (1982). Neotropical anachronisms: the fruits the gomphotheres ate. *Science.* 215, 19–27. doi: 10.1126/science.215.4528.19

Judd, W. S., Sanders, R. W., and Donoghue, M. J. (1994). Angiosperm family pairs - preliminary phylogenetic analyses. *Harv. Pap. Bot.* 5, 1–51.

Kendall, D. G. (1949). Stochastic processes and population growth. *J. R. Stat. Soc. B Stat. Methodol.* 11, 230–264.

Kerkhoff, A. J., Moriarty, P. E., and Weiser, M. D. (2014). The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8125–8130. doi: 10.1073/pnas.1308932111

Knapp, S., and Mallet, J. (2003). Refuting refugia? *Science* 300, 71–72. doi: 10.1126/science.1083007

Maddison, W. P., and Maddison, D. R. (2015). *Mesquite: A Modular System for Evolutionary Analysis.* Version 3.04. Available online at: http://mesquiteproject.org

Magallón, S., and Castillo, A. (2009). Angiosperm diversification through time. *Am. J. Bot.* 96, 349–365. doi: 10.3732/ajb.0800060

Magallon, S., and Sanderson, M. J. (2001). Absolute diversification rates in angiosperm clades. *Evolution* 55, 1762–1780. doi: 10.1111/j.0014-3820.2001.tb00826.x

Markets and Markets (2011). *Global Chocolate, Cocoa Beans, Lecithin, Sugar and Vanilla Market By Market Share, Trade, Prices, Geography Trend and Forecast.* Report Code: CG 1111.

Martin, P. R., and Tewksbury, J. J. (2008). Latitudinal variation in subspecific diversification of birds. *Evolution* 62, 2775–2788. doi: 10.1111/j.1558-5646.2008.00489.x

Meerow, A. W., Noblick, L., Salas-Leiva, D. E., Sanchez, V., Francisco-Ortega, J., Jestrow, B., et al. (2015). Phylogeny and historical biogeography of the cocosoid palms (Arecaceae, Arecoideae, Cocoseae) inferred from sequences of six WRKY gene family loci. *Cladistics* 31, 509–534. doi: 10.1111/cla.12100

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., et al. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* 10, 315–331. doi: 10.1111/j.1461-0248.2007.01020.x

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, V., et al. (2015). Middle Miocene closure of the Central American Seaway. *Science* 348, 226–229. doi: 10.1126/science.aaa2815

Mora, A., Baby, P., Roddaz, M., Parra, M., Brusset, S., Hermoza, W., et al. (2010). "Tectonic history of the Andes and sub-Andean zones: implications for the development of the Amazon drainage basin," in *Amazonia: Landscape and Species Evolution: A Look into the Past,* eds C. Hoorn and F. P. Wesselingh (Oxford: Wiley-Blackwell), 38–60.

Moran, P. A. (1951). Estimation methods for evolutive processes. *J. R. Stat. Soc. B Stat. Methodol.* 13, 141–146.

Motamayor, J. C., Lachenaud, P., da Silva, E., Mota, J. W., Loor, R., Kuhn, D. N., et al. (2008). Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE* 3:e3311. doi: 10.1371/journal.pone.0003311

Motilal, L. A., Zhang, D., Mischke, S., Meinhardt, L., and Umaharan, P. (2013). Microsatellite-aided detection of genetic redundancy improves management of the International Cocoa Genebank, Trinidad. *Tree Genet. Genomes* 9, 1395–1411 doi: 10.1007/s11295-013-0645-5

Nyffeler, R., Bayer, C., Alverson, W. S., Yen, A., Whitlock, B. A., Chase, M. W., et al. (2005). Phylogenetic analysis of the *Malvadendrina* clade (Malvaceae s.l.) based on plastid DNA sequences. *Org. Divers. Evol.* 5, 109–123. doi: 10.1016/j.ode.2004.08.001

Nylander, J. A. A. (2004). *MrModeltest v2.* Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Patterson, B. D., Solari, S., and Velazco, P. M. (2012). "The role of the Andes in the diversification and biogeography of neotropical mammals," in *Bones, Clones, and Biomes: the History and Geography of Recent Neotropical Mammals,* eds B. D. Patterson and L. P. Costa (Chicago, IL: University of Chicago Press), 351–378.

Patterson, B. D., and Velazco, P. M. (2008). Phylogeny of the rodent genus *Isothrix* (Hystricognathi, Echimyidae) and its diversification in Amazonia and the Eastern Andes. *J. Mamm. Evol.* 15, 181–201. doi: 10.1007/s10914-007-9070-6

Pirie, M. D., Chatrou, L. W., Mols, J. B., Erkens, R. H. J., and Oosterhof, J. (2006). 'Andean-centred' genera in the short-branch clade of Annonaceae: testing biogeographic hypotheses using phylogeny reconstruction and molecular dating. *J. Biogeogr.* 33, 31–46. doi: 10.1111/j.1365-2699.2005.01388.x

Rambaut, A. (2009). *FigTree v1.3.1. Computer Program.* Available online at: http://tree.bio.ed.ac.uk/software/figtree/ (Accessed March, 2012).

Rambaut, A., and Drummond, A. J. (2007). *Tracer v1.5. Computer Program.* Available online at: http://tree.bio.ed.ac.uk/software/tracer/

Reid, E. M., and Chandler, M. E. J. (1933). *The London Clay Flora.* London: British Museum of Natural History.

Richardson, J. E., Pennington, R. T., Pennington, T. D., and Hollingsworth, P. M. (2001). Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293, 2242–2245. doi: 10.1126/science.1061421

Ricklefs, R. E. (2006). Global variation in the diversification rate of passerine birds. *Ecology* 87, 2468–2478. doi: 10.1890/0012-9658(2006)87[2468:GVITDR]2.0.CO;2

Rolland, J., Condamine, F. L., Jiguet, F., and Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.* 12:e1001775. doi: 10.1371/journal.pbio.1001775

Roncal, J., Kahn, F., Millan, B., Couvreur, T. L. P., and Pintaud, J.-C. (2013). Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). *Bot. J. Linn. Soc.* 171, 120–139. doi: 10.1111/j.1095-8339.2012.01297.x

Savolainen, V., Anstett, M. C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V., et al. (2006). Sympatric speciation in palms on an oceanic island. *Nature* 441, 210–213. doi: 10.1038/nature04566

Schnell, R. J., Kuhn, D. N., Brown, J. S., Olano, C. T., Phillips-Mora, W., Amores, F. M., et al. (2007). Development of a marker assisted selection program for cacao. *Phytopathology* 97, 1664–1669. doi: 10.1094/PHYTO-97-12-1664

Schultes, R. (1958). A synopsis of the genus *Herrania. J. Arnold Arbor.* 39, 216–295.

Silva, S., and Figueira, A. (2005). Phylogenetic analysis of *Theobroma* (Sterculiaceae) based on Kunitz-like trypsin inhibitor sequences. *Plant Syst. Evol.* 250, 93–104. doi: 10.1007/s00606-004-0223-2

Simpson, G. G. (1980). *Splendid Isolation: The Curious History of South American Mammals.* Yale, MI: Yale University Press.

Stebbins, G. L. (1974). *Flowering Plants: Evolution Above the Species Level.* Cambridge, MA: Belknap.

Stevens, P. F. (2001 onwards). *Angiosperm Phylogeny Website.* Version 12, July 2012 [and more or less continuously updated since].

Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G., and van Etten, J. (2012). Present spatial diversity patterns of *Theobroma cacao* L. in the neotropics reflect genetic differentiation in pleistocene refugia followed by human-influenced dispersal. *PLoS ONE* 7:e47676. doi: 10.1371/journal.pone.0047676

Weir, J. T., and Schluter, D. (2007). The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* 315, 1574–1576. doi: 10.1126/science.1135590

Wesselingh, F. P. (2006). Miocene long-lived lake Pebas as a stage of mollusc radiations, with implications for landscape evolution in western Amazonia. *Scripta Geol.* 133, 1–17.

Wesselingh, F. P., and Ramos, M. I. F. (2010). "Amazonian aquatic invertebrate faunas (Mollusca, Ostracoda) and their development over the past 30 million years," in *Amazonia: Landscape and Species Evolution,* eds C. Hoorn and F. P. Wesselingh (Oxford: Wiley-Blackwell), 302–316.

Wesselingh, F. P., Räsänen, M. E., Irion, G., Vonhof, H. B., Kaandorp, R., Renema, W., et al. (2002). Lake Pebas: a palaeoecological reconstruction of a Miocene, long-lived lake complex in western Amazonia. *Cainozoic Res.* 1, 35–81.

Wesselingh, F. P., and Salo, J. A. (2006). A miocene perspective on the evolution of the Amazonian biota. *Scripta Geol.* 133, 439–458.

Whitlock, B. A., and Baum, D. A. (1999). Phylogenetic relationships of *Theobroma* and *Herrania* (Sterculiaceae) based on sequences of the nuclear gene vicilin. *Syst. Bot.* 24, 128–138. doi: 10.2307/2419544

Whitlock, B. A., Bayer, C., and Baum, D. A. (2001). Phylogenetic relationships and floral evolution of the Byttnerioideae (Sterculiaceae or Malvaceae s.l.) based on sequences of the chloroplast gene *ndh*F. *Syst. Bot.* 26, 420–437. doi: 10.1043/0363-6445-26.2.420

Wiens, J. J. (2007). Global patterns of diversification and species richness in amphibians. *Am. Nat.* 170, S86–S106. doi: 10.1086/519396

Wiens, J. J., Graham, C. H., Moen, D. S., Smith, S. A., and Reeder, T. W. (2006). Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. *Am. Nat.* 168, 579–596. doi: 10.1086/507882

Wiens, J. J., Sukumaran, J., Pyron, R. A., and Brown, R. M. (2009). Evolutionary and biogeographic origins of high tropical diversity in old world frogs (Ranidae). *Evolution* 63, 1217–1231. doi: 10.1111/j.1558-5646.2009.00610.x

Wilkie, P., Clark, A., Pennington, R. T., Cheek, P., Bayer, C., and Wilcock, C. C. (2006). Phylogenetic relationships within the subfamily Sterculioideae (Malvaceae/Sterculiaceae-Sterculieae) using the chloroplast gene *ndh*F. *Syst. Bot.* 31, 160–170. doi: 10.1600/036364406775971714

Winterton, C., Richardson, J. E., Hollingsworth, M., Clark, A., Zamora, N., and Pennington, R. T. (2014). "Historical biogeography of the Neotropical legume Genus Dussia: the Andes, the Panama Isthmus and the Chocó," in *Paleobotany and Biogeography: A Festschrift for Alan Graham in his 80th Year*, Chapter 14, eds W. D. Stevens, O. M. Montiel, and P. Raven (St Louis, MO: Missouri Botanical Press), 389–404.