Regionalization of Cacao (*Theobroma cacao* L.) in Colombia

Carlos E. González-Orozco * and Allende Pesca*

Corporación Colombiana de Investigacion Agropecuaria (AGROSAVIA), Centro de Investigación La Libertad, Villavicencio, Colombia

A detailed description and up to date biogeographic regionalization of cultivated cacao (*Theobroma cacao* L.) are lacking in Colombia. Here we propose a new biogeographical regionalization of cultivated cacao for Colombia. We used spatial partitioning, geospatial mapping of macro/microclimate variables and a hierarchical area taxonomy classification to describe, define and propose the biogeographical regions of cultivated cacao. The cacao regions were identified from distributions of 4,974 cocoa producing farms across Colombia. Our proposed regionalization comprises four regions (north-eastern, north-western, south-western, east), 31 sub-regions and 54 provinces. Solar radiation, precipitation, and soil temperature seasonality best explained the biogeographical regions. Rivers networks helped explain the differences at the sub-regional and provinces level. Our results indicate that biogeography is a strong indicator of cacao's agricultural expansion across the different growing regions in Colombia. This up-to-date biogeographical classification could be a useful tool for agricultural planning of cacao in Colombia. Particularly, the baseline information provided might be of use on the development of denominations of origin for cacao.

Keywords: agriculture, biogeography, chocolate, cocoa, neotropics, northern Andes

INTRODUCTION

A biogeographic regionalization is a tool to categorize geographic areas that often represent biotas (Morrone et al., 2022). Biogeographic classifications can be found in a wide range of mono crops usually at large scales in temperate regions. However, biogeographic regionalization's of important emerging crops at finer scales are less common in tropical latitudes. This emerging field, part of agriculture biogeography (Katinas and Crisci, 2018), is poorly explored. The fact that crops are placed in certain places by humans signify that man made agriculture requires empirical knowledge of the conditions where plantations could do best. Despite ideal conditions, environmental suitability of crops in the most productive areas may vary over time due to climate change. However, we do not know that what extent intuitive human decision could be related to biogeography. Here, we propose a biogeographical regionalization of cultivated cacao in Colombia. This is the first of its kind hence could be applied to define the characteristics of sites where cacao grow making it a potential tool to decision makers in relation to denominations of origin.

Cacao is one of the top priority crops in Colombia (Rodriguez-Medina et al., 2019). Cultivated cacao provides an ideal case study for agriculture biogeography because it is widely distributed across a variety of bioregions, climates, and ecosystems. In terms of human value, cacao represents a strong part of Colombia's food cultural history (Patiño, 2002). Despite its economic and human importance, cocoa producing regions (here after cacao regions) in Colombia have not been

systematically defined, described, and classified using a biogeographic approach. In this paper we applied a regionalization approach to propose the biogeographic regions of cultivated cacao (*Theobroma cacao* L.) in Colombia.

The cartography of crops is often classified as agroclimatic maps or agroecological zones (Fischer et al., 2021). They are gridded maps and often involve information about crop yield and climate. Their objective is to map the most suitable areas for agriculture. These classifications are useful to allocate suitable crop regions. However, these thematic maps do not necessarily represent a biogeographic context. Including biogeography in crop mapping has advantages because it represents the environmental conditions in the sites where the crops grow or originated. These are key variables to understand agricultural production and adaptation. Unfortunately, approaches to develop biogeographic regionalization's of crops are rare in the literature of tropical countries (González-Orozco, 2021a,b; González-Orozco and Porcel, 2021; González-Orozco et al., 2021).

Some studies in sub-tropical and tropical latitudes show how regions of cacao, classified by climate and soil suitability, will suffer the effects of climate change. For instance, cacao farms in the Brazilian amazon (Igawa et al., 2022) where reduction in precipitation and an increase in temperature may affect cacao production. Läderach et al. (2013) demonstrated that future climate change will affect cacao farming in Africa's largest cocoa producer countries. Santosa et al. (2018) showed that productions of cacao in Java-Indonesia fluctuates under climate variability. One strategy to mitigate the potential effects of climate change is to generate simple maps of the cacao regions in relation to spatial variation of climate and soils. In south America, there are not enough case studies showing how country level bioregionalizations can be of help to improve our understanding of climate threats on cacao regions.

Colombia does not have a biogeographically centered classification of cacao regions. Our study will help to fill up this knowledge gap. Currently, there are two classification schemes of cultivated cacao: (1) a cacao aptitude map generated by the National Rural Agricultural Planning Unit or SIPRA-UPRA in 2019; and (2) the edaphoclimatic cacao map generated by CORPOICA (former Agrosavia) in 2005 (Garcia Lozano et al., 2005). These agronomical classifications of cultivated cacao use thematic layering, biophysical properties, crop productivity-yield data, socioeconomic and socio-ecosystemic data. SIPRA-UPRA (2019) provide a sub-national zonification of cacao showing areas that are high, mid, low and not viable for cropping commercial cacao (https://sipra.upra.gov.co/). Garcia Lozano et al. (2005) on the other hand developed an edaphoclimatic classification including the eco physiological requirements and agronomical aspects of cultivated cacao. Both classifications are technically correct hence accepted and considered essential for the development of the cacao industry in Colombia. However, neither of them consider the biogeographical context as a criterion to characterize the cacao regions. Ignoring biogeography in a crop classification creates knowledge gaps. For example, questions such as why certain cacao qualities differ from one site to another remain unanswered. The application of biogeography to map agricultural zones of cacao is an

improvement because unorthodox variables such as rivers and history and a spatial component are used to understand the current distribution of cacao producing zones. Likewise, this study provides an independent view, and complementary to the traditional cacao aptitude and edaphoclimatic mapping. Particularly, assigning, classifying, and describing the different crop regions provide novel cartographical information. This will help to establish denominations of origin which is a key feature in economic agriculture.

MATERIALS AND METHODS

Cacao Information and Spatial Data

Cacao is widely distributed with ∼213,165 hectares planted in Colombia (Agronet, 2022) (**Supplementary Table 7**). Cacao plantations can be found from sea level to 1,200 meters (Antolinez et al., 2020; Nuñez-Rodríguez et al., 2021). The cacao farms are in 31 of the 32 departments of Colombia. This area covers 568 municipalities reaching an average cocoa production of 70,205 tons in a year as reported by the National Cacao Association Fedecacao (2021). The size of the traditional cacao farm is between 0.5 and 100 hectares. In average, each producer cultivates an area of 3 hectares. This local production generates a subsistence income for 54,341 farmers homes and provide jobs for 167,000 Colombian farmers (MADR, 2021).

The spatial data used in this study covers 95% of the cacao producing regions mentioned in **Supplementary Table 7**. This study uses locations of 4,974 farms that cultivate cacao across Colombia (**Figure 1A**). The occurrences were sourced from research projects, regional or national agricultural institutions or agencies. A full list of projects and agencies that were involved in the data collection of the cacao farms is available in the acknowledgment section. Our analysis is not done at the cacao varieties level. However, our sample includes 100% of the existing commercial cacao varieties. They are considered as special cacaos of fine aroma and flavor (ICCO-International Cocoa Organization, 2020). Among others, the most common set of clones used across Colombia are Agrosavia clones TCS01, TCS06, TCS13 TCS19; FEDECACAO clones FEAR5, FSA12, FSA13, FTA2, FLE2, FLE3, FSV41, FSV155, FEC2; Casa Luker clon Luker40; Compañía Nacional de Chocolates clones CNCH12 y CNCH13, and universal clones such as the CCN51, ICS1, ICS39, ICS60, ICS95, TSH565, EET8. Most of the cases those clones are planted in agroforestry systems with edible species such as plantain, citrus, and woody trees. They are often randomly distributed in the plots, but more recently structured mixed systems are predominant.

Regionalization

A hierarchical area taxonomy classification was applied to describe the biogeographical regions (González-Orozco, 2021b). This method consists of defining the categories that species, in this case a crop, in a hierarchy using nomenclatural and biogeographical principles such as taxonomy, physical geography, spatial variation of climate and soils, topographical features. The biogeographical regions (or regionalization) are placed into categories of different hierarchical levels. In our case, the regionalization consists of three hierarchical levels:
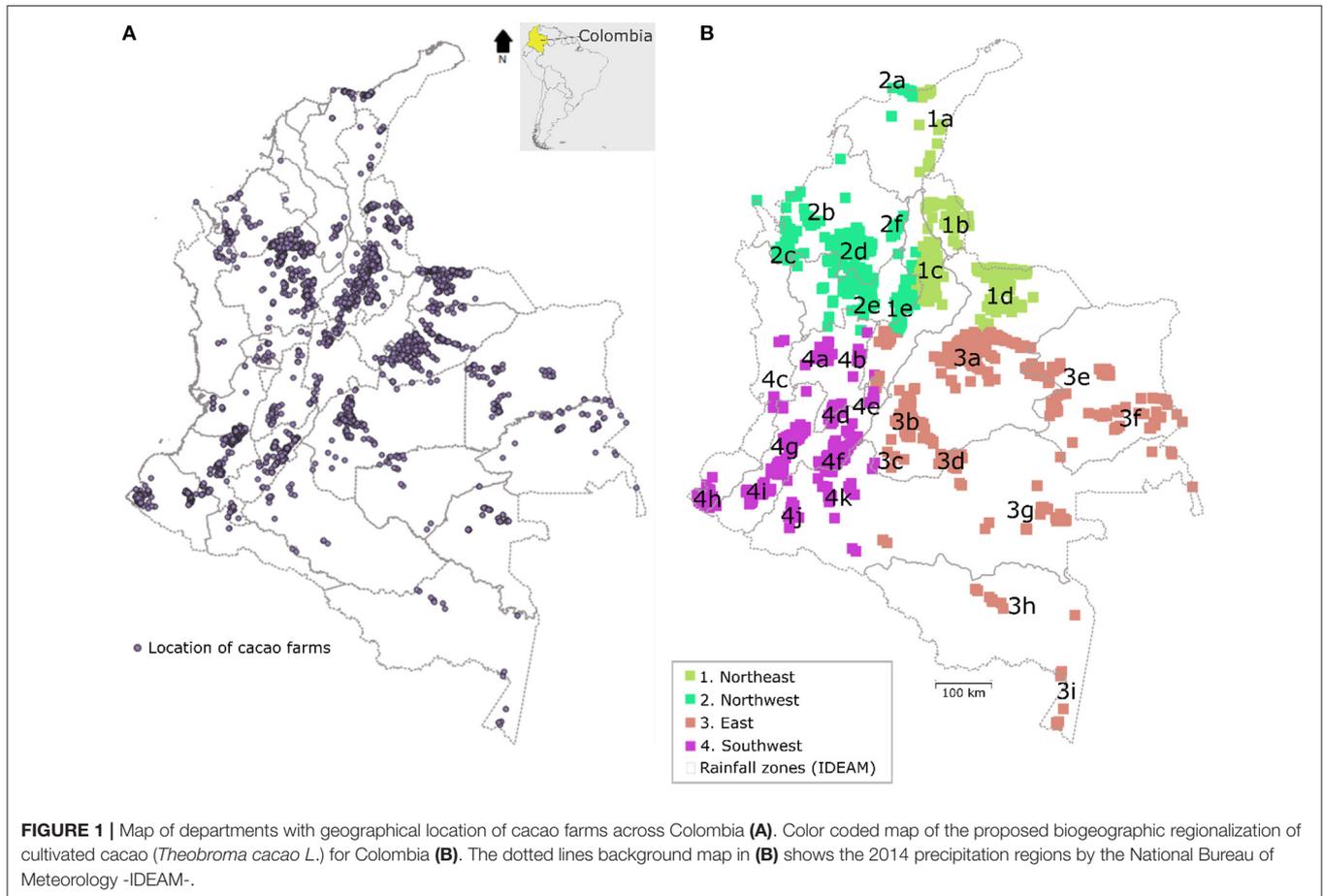
FIGURE 1 | Map of departments with geographical location of cacao farms across Colombia (A). Color coded map of the proposed biogeographic regionalization of cultivated cacao (*Theobroma cacao L.*) for Colombia (B). The dotted lines background map in (B) shows the 2014 precipitation regions by the National Bureau of Meteorology -IDEAM-.

regions, sub-regions, and provinces. Regions are the largest units and provinces the smallest. Wallace, an R-based GUI application for ecological modeling, was used to identify the regions (Kass et al., 2018). Wallace's platform works under a non-programming environment and friendly interface that allows the user to manipulate and choose the different modeling functions. In our case, we used the spatial partitioning and thinning tools. The spatial partitioning method consists of applying a random-k fold process ($k = 4$ aggregation factor) to split the dataset in spatially meaningful groups. We used 10,000 background points was applied to the 3,141 occurrences. As part of a spatial thinning method, one hundred iterations were run using 0.5 km of radius. After thinning, 1,833 occurrences were eliminated, and 3,141 cacao farms occurrences remained and were used to generate our regionalization (**Figure 1B**). Sub-regions were defined by independent sub-clusters of cacao farms within regions. The names of the sub-regions were based on the departments where they are located. At the finer spatial level, provinces were defined based on the macro and microclimate conditions that differentiated them as well as the rivers where they are distributed. Rivers are important to consider in this study because they create barriers that can be used to define geographic regions, and they provide good soils to grow cacao. A generic diagnosis, general remarks, and a detailed description of the provinces were conducted.

Climate Maps

The macro climate dataset WordlClim v.2 was used to generate the macroclimate maps of the regions (Fick and Hijmans, 2017). The macroclimate WorldClim v.2 dataset consisted of seven variables: precipitation, maximum temperature, minimum temperature, mean temperature, relative humidity, wind speed and solar radiation at 1 km^2 spatial resolution. The soil temperature bioclimatic SBIO dataset used was to generate the microclimate maps of the regions (Lembrechts et al., 2021). It consisted of eleven soil temperature variables at 1 km^2 spatial resolution and 0–5 cm depth. The microclimate soil temperature variables contained in the SBIO dataset are annual mean temperature SBIO1, mean diurnal range SBIO2, isothermality SBIO3, temperature seasonality SBIO4, maximum temperature of the warmest month SBIO5, minimum temperature of the coldest month SBIO6, temperature annual range SBIO7, mean temperature of the wettest quarter SBIO8, mean temperature of the drierst quarter SBIO9, mean temperature of the warmest quarter SBIO10 and mean temperature of the coldest quarter SBIO11. We used QGIS Desktop V. 3.10.2. (2019) to extract the macro and microclimate values for each of the 4,974 occurrences. Climate maps for each of the regions, sub-regions and provinces were generated in QGIS Development Team (2019). Regions, sub-regions, and provinces were described using the macro and microclimate maps. The river networks of Colombia were plotted

against the macro and microclimate maps. Finally, we overlapped the cacao optimum zones map (SIPRA-UPRA, 2019) and the edaphoclimatic map of cacao suitability (CORPOICA, Garcia Lozano et al., 2005) with the biogeographic classification and climate maps to corroborate and explore the presence of cacao productivity zones inside the proposed biogeographic regions. Soil types are important for cacao agriculture (León-Moreno et al., 2019). However, soil types were not included in this regionalization due to absence of a readily available digitalized soils raster layer for Colombia.

Ordination and Statistical Analyses

To test the spatial patterns of climate and soil temperature for the regions of cacao across Colombia, a data ordination was performed using a singular value decomposition based PCA analyses with the seven climate and eleven soil temperature variables extracted from the cacao farms plot presence points representing the cacao regions. The PCA was carried our using the ClustVis tool (Metsalu and Vilo, 2015). Outputs from the axes 1 were spatially mapped using QGIS Development Team (2019) as conducted in previous studies that identify biogeographical and crop regions (González-Orozco et al., 2014, 2020). The loading values were used to generate the PCA Axes 1 visualization maps. This way, we were able to quantify and test the differences between the proposed regions based on the spatial patterns of the climate and soil temperature variables. Hence, the PCA maps were used to validate the regionalization obtained with the spatial partitioning classification. Descriptive statistics such as mean and standard deviation were calculated for the regions (**Supplementary Tables 1, 2**), sub-regions (**Supplementary Tables 3, 4**), and provinces (**Supplementary Tables 5, 6**).

TABLE 1 | Proposed classification of biogeographical regions, sub-regions, and provinces of cultivated cacao in Colombia (River names are in brackets).

Regions	Sub-regions	Provinces
North-east	1a. Guajira-Cesar	1.North (Ancho, Corual), 2. East (Riecito, Tucuy)
	1b. Norte de Santander	3. North (Catatumbo), 4. Central (Tibú, Nuevo presidente, Sadinata), 5. East (Zulia, Pamplonita)
	1c. Santander	6. North (Lebrija), 7. Central (Chucurí, Oponcito, Cascajales), 8. Southeast (Suarez)
	1d. Boyacá-Arauca	9. West (Cobaría, Arauca, Viejo, Banadía), 10. Central (Caranal, Cusay, Ele), 11. East (Arauca, Brazo Gaviotas, Bayonero), 12. South (Ariporo, Casanare)
	1e. Santander-Cundinamarca-Boyacá	13. Southwest (Opón, Quiratá, Guayabito, Horta), 14. West (Minero), 15. Northwest (Chirche)
North-west	2a. Magdalena	16. North (Piedras, Guachaca, Buriticá, Don Diego)
	2b. Cordoba-Sucre	17. Central (Sinú, San Jorge)
	2c. Antioquia-Chocó	18. West (Grande, Apartadó, Chigorodó), 19. North (Curbaradó)
	2d. Antioquia-Bolívar	20. Northeast (Cauca, Nechí, Porce, Bagre)
	2e. Antioquia-Caldas	21. Southeast (San Bartolome, Nus, Naré, Samaná norte), 22. Northeast (La Miel)
	2f. Bolívar	23. South (Magdalena, Boque)
East	3a. Casanare	24. North (Pauto), 25. Central (Cravo sur, Charte, Uneté, Cusiana), 26. South (Tua, Upía)
	3b. Meta	27. South (Humadea, Ariari, Guejar)
	3c. Meta	28. Southwest (Guayabero)
	3d. Guaviare	29. Central (Guaviare)
	3e. Vichada	30. North (Tomo, Caño Urimica, Múco), 31. East (Vichada, Tuparro), 32. West (Uva o Brazo de Amanavén)
	3f. Guainía	33. North (Guaviare), 34. Central (Caño Bocón, Inírida), 35. South (Guanía), 36. East (Atabapo)
	3g. Vaupes	37. East (Vaupes)
	3h. Amazonas- Caquetá	38. Central (Caquetá, Caguán)
	3i. Amazonas	39. South (Putumayo, Amazonas)
South-west	4a. Caldas-Risaralda	40. South (Cauca, Chinchiná, Campo alegre)
	4b. Tolima	41. North (Magdalena, Gualí)
	4c. Chocó- Valle del Cauca	42. North (Rijorodó), 43. Central (Calima, Dagua, Anchicayá)
	4d. Tolima	44. South (Tetuan, Amoya, Saldaña)
	4e. Tolima-Cundinamarca	45. East (Quebrada de bajas, Cuinde), 46. South (Sumapaz)
	4f. Huila	47. East (Magdalena, Bache, Bateas), 48. Central (Yaguará, Yaguaracito)
	4g. Cauca	49. North (Ovejas, Palo), 50. Central (Timbío, Guachicono, Tambo)
	4h. Nariño	51. West (Mira, Nulpe)
	4i. Cauca-Nariño	52. South (Patía, Mayo, Sambingo, Guáitara)
	4j. Cauca	53. Southeast (Caquetá)
	4k. Caquetá	54. West (La Yuza, Bodoquero)

Provinces numbers correspond to the numbered panels in Figure 3.

RESULTS

Regionalization

Four biogeographical regions (northeast, northwest, eastern and southwest), 31 sub-regions and 54 provinces of cultivated cacao are proposed (**Figures 1B**, **3**; **Table 1**). We demonstrated that the biogeography of cultivated cacao is being influenced by macro/micro-climate, elevation, and river systems (**Supplementary Figure 1**). Landscape features such as alluvial plains (<800 m.a.s.l.), and slopes of inter Andean dry or humid valleys with elevations generally between 600 and 1,200 m.a.s.l. also determined the geography of the cacao regions. In the following section, a general diagnosis, remarks, and a detailed description of the provinces is provided and discussed.

Climate Maps

Macro and microclimate maps showed different patterns for each hierarchical level (**Supplementary Figures 2, 3**). Solar radiation is macroclimate variables that best differentiate the four regions. Soil temperature seasonality on the other hand is an effective microclimate variable to determine differences within regions and provinces. The specific patterns of macro and microclimate will be discussed in the next section.

Ordination Analysis

The ordination analysis supported the regionalization shown in **Figures 1**, **3**. Axis 1 explained 99.8% of the variation in climate suggesting areas of moderate solar radiation as stable cacao regions. Axis 1 explained 82.3% of the variation

in soil temperature seasonality, suggesting areas of low to moderate seasonality as ideal cacao growing regions (**Figure 2**; **Supplementary Figure 5**). Solar radiation and soil temperature seasonality were the variables that best explained the spatial variation of macro and microclimate across the regionalization (**Figure 2**). The clusters of the climate PCA (**Figure 2A**) suggest that cacao regions in the northern of Colombia are exposed to the largest amount of solar radiation. The clusters of the soil temperature PCA (**Figure 2B**) suggest that cacao regions in the southern Andean part of Colombia are exposed to the largest amount of soil temperature seasonality as well as some regions on the eastern plains and inner valleys of the Andean Mountain ranges.

DISCUSSION

The ordination analyses proved that there are spatial differences in climate between regions, sub-regions, and provinces (**Figure 2**) suggesting that cacao plantations could be affected depending on the biogeography where they are located. Although we have not proved the potential effect on the quality of cocoa, our regionalization helps to identify the geographic differences at national and regional scales.

Description of the Regions

The proposed regionalization was developed using biogeographical principles (spatial distribution of species, geographic barriers, rivers, and climate patterns) and the

FIGURE 2 | Geospatial analysis of climate **(A)** and soil temperature **(B)** variables for the regions of cacao in Colombia based on a PCA for Axes 1.

FIGURE 3 | Map of regions/sub-regions (middle map) of cultivated cacao in Colombia, and a close view of the 54 provinces with corresponding rivers (panels numbered 1–54).
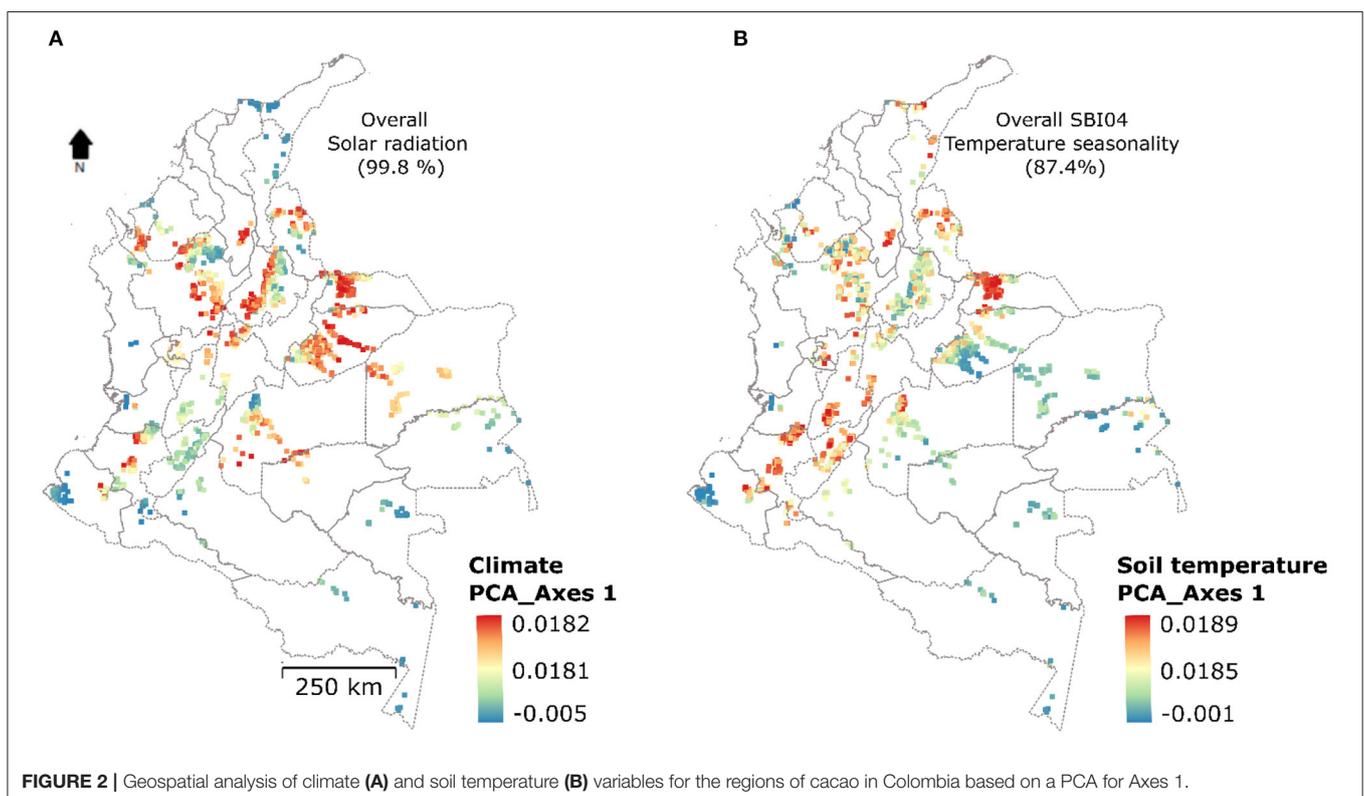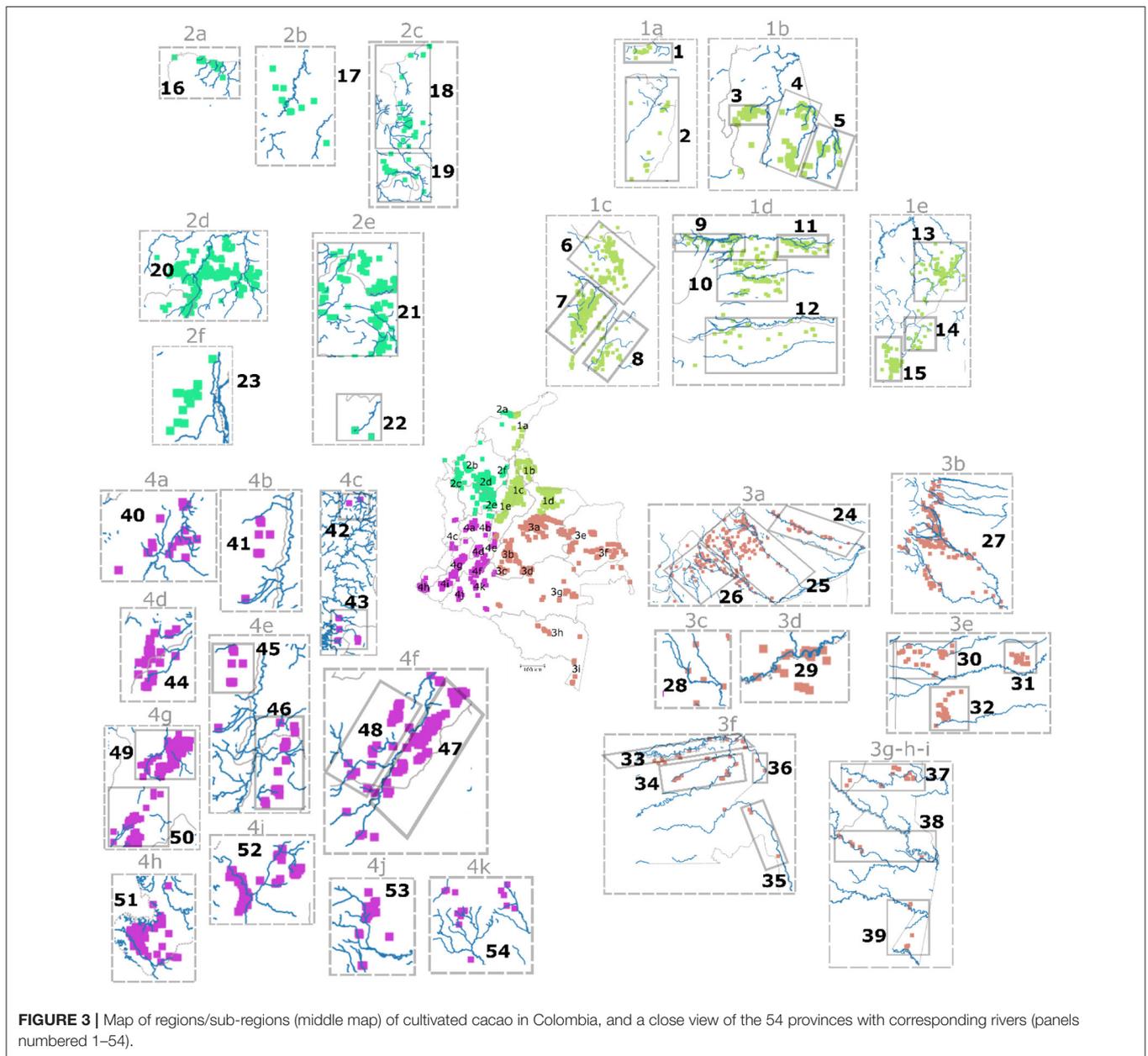
latest microclimate soil dataset which differs from all the previous cacao regionalization schemes in Colombia (**Supplementary Figure 4**; SIPRA-UPRA, 2019; and CORPOICA, Garcia Lozano et al., 2005), which mainly uses agro-climatological and cacao productivity information. Being the first of its kind, this biogeographical regionalization provides insightful information with greater geographical details and spatial accuracy. Soil and climate conditions have a strong implication on cacao production and quality (Guilherme et al., 2017). The use of climate and soil temperature bioclimatic conditions in agroclimatic zoning can be applied to determine regional denominations of origin as previously tested in other crops (Santos et al., 2020). For instance, in the

case of cacao region in Colombia, areas of low seasonality (**Supplementary Figure 4D**) could be used by planners as sites with more stable conditions for the crop which are important currently under increasing climate change or extreme variability. Here, we map and describe in detail each region, sub-regions and provinces which are proposed as a baseline to identify the different conditions for cacao denominations of origin in Colombia.

Regionalization's are traditionally applied to wild species of plants and animals (Ebach, 2015). However, we adapted some of the biogeographic principles of regionalization to the mapping of cacao regions. Considering the geographic context, our classification should not be considered as an agronomical

classification *per-se*. The proposed biogeographical regions or regionalization represent a snapshot of the cacao regions expansion during the last 250 years. Despite many regions showed biophysical differences (**Figure 2**), we acknowledge that our regionalization does not directly measured features such as cultural and management difference. Our regionalization should not be considered as an agronomical classification. However, it can be considered as a surrogate because plantations of cacao respond to each environment accordingly and equally reflect human intervention. For instance, the eastern half of Colombia is characterized by Andean foothills, floodplains, savannas, and rainforests. This makes the eastern cacao regions fundamentally different from the Andes. Even within the same mountain's biome, the regionalization suggests a split between the northern and southern parts of the Andes. The splits are supported by the PCA maps and the geography of the cacao regions (**Figure 2**). The regions of cacao proposed here were not classified based on varieties of cacao because our consisted of developing a generic framework that serve as a baseline to further investigations.

The biogeographical regions, sub-regions, and provinces (**Figures 1**, **3**) showed differential macro and microclimate conditions (**Figure 2**; **Supplementary Figures 2, 3**) making them likely candidates to determine areas of denomination of origin for cacao. The following sections contain a discussion for each region containing a diagnosis, remarks and detailed description of the provinces as shown in **Table 1** and **Figure 3**.

NORTH-EAST Region
Diagnosis
This region covers between 6°S−11°N and 70°E−73°W. It represents areas of the dry Guajira, Andean and Orinoquia bioregions (González-Orozco et al., 2021). This region is sub-divided into four sub-regions: Guajira-Cesar, Norte de Santander, Santander, Boyacá-Arauca, Santander-Cundinamarca-Boyacá (**Table 1**; **Figure 1B**, numerals 1a–e). All sub-regions except Arauca have a bimodal precipitation trend with two rainy seasons. The macro-climate of the northern part of the region is characterized by low precipitation, high solar radiation, temperature, and wind speed (**Supplementary Figures 2A,B,G**). Soil temperature is high in the northern areas of the region (**Supplementary Figures 3A,J,K**). The macroclimate in the southern Andean part of this region is milder than the northern region. For instance, sub-regions 1d which is on the slopes of the eastern mountain range has cooler and wetter climate conditions, but soil temperatures increase further east in Arauca (**Supplementary Figure 3G**).

Remarks
As early as 1621, the town of Cúcuta in Norte de Santander was considered as the oldest area with cultivated cacao in Colombia (Patiño, 2002). In 1809, Francisco Jose de Caldas also mentioned Cucuta as one of the main cacao trading centers in the New Granada. Caldas said that the cacao from Cucuta is the best pay in the market, and it is known as the Cacao del Magdalena which is traded from Cucuta to Maracaibo along the Zulia Rive? (Caldas, 1809). Caldas mentioned that the Sogamoso and Suratá riverbanks in Santander as places where cacao grows abundantly.

Currently, Santander is the largest and one of the top three cacao producing regions of Colombia (Rodriguez-Medina et al., 2019).

Provinces
This region is sub-divided into 15 provinces (1–15 in **Figure 3**; **Table 1**). The north province (sub-region 1a) is characterized by dry and hot climate conditions (**Table 1**). The east province (sub-region 1a) is divided into two areas, one in the northern side of the Serranía de Perijá at mid elevations with cooler temperatures, and another one in the southern areas where temperatures are higher (i.e., Rioacha). The east province (sub-region 1b) is low in elevation, and it is geographically isolated by the barrier known as the Catatumbo canyon. The north area of the east province (sub-region 1b) has wetter and cooler climate and soil temperature than the southern area in Zulia which is hotter. The north and central provinces (sub-region 1b) provide milder climate conditions because they are at mid-elevations. The cacao region of Santander has two major areas: north (sub-region 1c) and south (sub-region 1e). The Sogamoso river is a geographic barrier that separates the north from the south sub-region. Overall, it shows an east to west climate and soil temperature gradient related to elevation and proximity to the Magdalena River. The lower foothills are hotter whereas mid elevations on the slopes of the mountain range present cooler conditions. The Santander sub-region 1c is sub-divided into three provinces: north, central, and southeast. The north province is the Lebría area, which shows the lowest values of soil temperature seasonality in the whole Santander cacao region. This suggests the most stable microclimate conditions in the region. The central province locates on the lower slopes of the eastern mountain range. It is composed of the Chucurí, Oponcito and Cascajales rivers. The Chucurí area shows microclimates with moderate soil temperatures compared to the rivers on the lower part of the catchment. The south-east province is a plateau on the other side of the mountain range, with even higher elevational limits than Chucurí providing the lowest soil and climate temperatures of the whole region suggesting different cacao growing conditions to all the rest of the provinces. Oponcito and Cascajales rivers show the hottest temperatures. Sub-region 1d is sub-divided into four provinces: west, central, east, and south (**Table 1**). These provinces follow a west to east temperature and rainfall gradient with hotter and drier conditions toward the east. It is known as the Cacao Arauca region. The west province runs along the riverbanks of the upper Arauca River. It is the coolest and wettest province of all because of its proximity to the foothills of the east mountain range. The central province has predominantly mid soil temperature seasonality. The east province shows higher values of soil temperature seasonality suggesting more variable microclimate conditions. The precipitation reported on the east province is the lowest in the whole region. The south province is part of the northern Casanare department where a transition from dissected non-flooding plains to flooding savannas occurs. The conditions in climate and soils for the south province are like the east province. Sub-region 1e is sub-divided into three provinces: southwest, west, and northwest. It is composed of the departments of Santander, Cundinamarca, and Boyacá. The limit of the southwest province, part of the

southern Santander department, is demarked by the Opón river, creating an isolated cacao area with wettest and hotter conditions along the Guayabito river. Horta river on the other hand has moderate microclimate like the Quiratá river. In Boyacá department, the west province is a transitional zone with moderate soil temperature gradients. The northwest province is part of Cundinamarca department and has the coldest conditions of all in the region.

NORTH-WEST Region

Diagnosis

This region covers between 5.5°S–11°N and 77°E–73°W. It represents areas of the humid Magdalena, northern Andean-Páramo, and Chocó-Darien bioregions (González-Orozco et al., 2021). This region is subdivided into six sub-regions: Magdalena, Cordoba-Sucre, Antioquia-Choco, Antioquia-Bolivar, Antioquia-Caldas, Bolivar (**Table 1**; **Figure 1B**, numerals 2a–f). All sub-regions have a bimodal precipitation trend with two marked rainy seasons, except for the humid Magdalena which has a monomodal precipitation trend with one long rainy season and a short dry one.

Remarks

The north-west region is the only one with predominantly low values of soil temperature seasonality suggesting stable microclimate conditions (**Supplementary Figure 3D**). The Sierra Nevada de Santa Marta is known as growing good quality cacao types. The southern part of this region is known as the lower Cauca and Magdalena cacao producing region, which includes the Sinú, San Jorge valleys and Urabá. Historically, this southern region is considered as part of the 1900's cacao expansion zones of Colombia. Areas of high cacao productivity are found in the fertile plains of the Magdalena and Cauca rivers (**Supplementary Figure 1**). In 1809, Caldas mentioned that "the Cauca River in Antioquia cultivate sugar Cane, maize, and plantain but instead cacao is expanding and if locals grow it properly, this province will be opulent".

Provinces

This region is subdivided into eight provinces (16–23 in **Figure 3**; **Table 1**). The north province (sub-region 2a) is the Sierra Nevada de Santa Marta. It is characterized by low precipitation, high solar radiation, and wind speed. Soil temperature is high in the north province (**Supplementary Figure 3A**). The central province (sub-region 2b), Sinú and San Jorge River valleys, is a transitional zone from hot-dry to wet-humid conditions. The west and north provinces (sub-region 2c) borders with Chocó-Darién and are characterized by hyper humid conditions surrounded by rainforests (**Supplementary Figures 2A,C,F**). At the top end of the central mountain range is the northeast province (sub-region 2d), which is part of the northeast of Antioquia department. This province includes the Cauca River as the main catchment and has cool and wet climate conditions because of its proximity to the foothills. However, as it moves to the east on the lower Cauca valley areas, conditions become hotter and drier and temperature seasonality in the soil increases suggest less stable microclimate conditions. The southeast and

northeast (sub-region 2d–e) provinces are characterized by stable climate and soil temperatures (**Supplementary Figures 3B,C,G**) suggesting less drastic changes in microclimates. The south province (sub-region 2f) is on the eastern foothills of the Serranía de San Lucas facing the Magdalena and Roque Rivers in the south of the Bolivar department. It has dry and cool climate with mild soil temperatures and solar radiation.

EAST Region

Diagnosis

This region covers between 5.5°N to −3.8°S and 67°E to 74.5°W. It represents areas of the Savannas-Orinoquia and Imerí - Amazonian bioregions (González-Orozco et al., 2021). Overall, the climate is monomodal with a trend of 8 months of rain and four hot and dry months. This region is subdivided into eight sub-regions: Casanare, Meta, Guaviare, Vichada, Guanía, Vaupes, Amazonas-Caquetá, and Amazonas (**Table 1**; **Figure 1B**, numerals 3a–i). The macro climate of this region is characterized by low solar radiation, mid temperatures, and moderate wind speed (**Supplementary Figure 2**). Soil temperature in this part of the Amazon is mid to low. The flooding savannas of the Orinoquia show high seasonality whereas the Amazon areas present intermediate soil temperature seasonality suggesting a less variable microclimate (**Supplementary Figure 3A**).

Remarks

This region is composed of large river catchments such as Meta, Orinoco, Guaviare, and Caquetá. It could be considered as the next frontier of cacao production in Colombia (2000's-present). Although considered part of the center of origin of cacao (Baker et al., 1953), this region is unexplored for cacao productivity because it is a relatively new colonization zone for cacao (**Supplementary Figure 4**). Deforestation is one of the main impacts on forests in this region (González-Orozco et al., 2021). This region is a biodiversity hotspot hence should not be used as a non-planned expansion region of cacao unless sustainable practices are put in place (i.e. cacao agroforestry systems). The areas of best cacao productivity are on the foothills of the eastern mountain range below 600 meters above sea level (**Supplementary Figure 4**). Most of the cacao producing areas in the region are considered as small-scale farming. Sub-regions 3a–e are part of the Orinoquia eastern plains. Sub-regions 3b–d and f are transitional Orinoquia to amazon zone. Sub-regions 3g–i are considered part of the Amazonian bioregion.

Provinces

This region is subdivided into 16 provinces (24–39 in **Figure 3**; **Table 1**). The north province (sub-region 3a) occupies parts of the flooding savannas of Casanare department. From north to south precipitation decreases, and areas closer to the foothills become cooler and wetter. From west to east soil temperatures increases causing higher seasonality of soil microclimate (**Supplementary Figures 3A,D,E,I**). The soil temperature in the central province (sub-region 3a) is high suggesting more variability in the microclimate (**Supplementary Figures 3D,I,J**) than the south province where soil temperature seasonality decreases as it gets closer to the foothills. The Meta River

creates a geographic barrier between the Orinoquia and the transitional zone to the amazon. The south province (sub-region 3b) is composed of the Humadea, Ariari and Guejar rivers. The alluvial plains of the Humadea and Ariari rivers are fertile creating ideal soil fertility conditions for cacao growing (León-Moreno et al., 2019). The south province in the Ariari area is subdivided into an upper, middle, and lower zones. At the upper part of the Ariari catchment, temperatures are cooler whereas to the east conditions becomes hotter and drier (**Supplementary Figures 2C,D,E**). Soil temperature seasonality is moderate overall, but soil temperature during the dry season is high which crates more variability in the microclimate (**Supplementary Figure 3I**). The Guejar river in the southwest (sub-region 3c) of the south province is part of a rocky canyon formation, which has similar condition to the south Meta province but differs on having lower precipitation and higher soil temperatures suggesting drier and hotter microclimates. High temperatures are predominant in Guaviare's central and southwest provinces (sub-region 3d). To the northeast of Guaviare, sub-region 3e is sub-divided into three provinces: north, east, and west. The north province is strongly determined by roughed terrain and dry savannas part of the Orinoquia bioregion (i.e., Cumaribo cacao area where new plantations of cacao are replacing illegal crops). The north province is characterized by hot temperatures, mid ranges of solar radiation and high soil temperature seasonality. The east and west provinces otherwise have milder conditions than the north showing a moderate soil temperature seasonality. The south province has the largest levels of solar radiation within sub-region 3e.

The Amazonian cacao provinces start from the southern part of the Guaviare River (see south province in 3d,e; **Figure 1B**). Overall, the climate is characterized by high precipitation, solar radiation, moderate temperature, but low solar radiation (**Supplementary Figure 2**). Microclimates becomes more stable than the northeast areas showing moderate to low soil temperatures (**Supplementary Figure 3**). The far eastern corner of Colombia is the most remote sub-region (3f), which is subdivided into four provinces: north, central, south, and east. The north province runs along the Guaviare riverbanks. The central province is composed of the upper catchments of the Inírida river. The south province is part of the Guanía river. The east province is part of the Atabapo river. All the other amazon sub-regions (3g–i) and corresponding provinces (east, central and south) are characterized by moderate to low soil temperatures and high precipitation suggesting stable microclimates of rainforests (**Figure 1B**).

SOUTH-WEST Region
Diagnosis
This region covers between 5.5°N−0.3°S and 74.6°E−78.8°W. It represents areas of the Pacific coast and Andean- Paramo bioregions (González-Orozco et al., 2021). The precipitation trend is bimodal in the Andean region and monomodal on the western Pacific coastal region which is the rainiest in the country. This region is subdivided into eleven sub-regions (**Table 1**; **Figure 1B**, numerals 4a–k). The climate of the region is characterized by mid to low precipitation and

temperatures in the Andean areas, and high precipitation on the pacific and amazon foothills (**Supplementary Figure 2**). Soil temperature is low on the Andean region of the western mountain range, whereas the Andean areas in the eastern mountain range and the pacific–amazon foothills have moderate temperatures (**Supplementary Figure 3A**). Moderate soil temperature seasonality is observed on the lowlands of the eastern mountain ranges (**Supplementary Figure 3D**).

Remarks
This region is composed of the southwest Pacific, Macizo Colombiano, alto Cauca, alto Magdalena, amazon foothills, Patía depression, and central pacific biogeographic regions. The Macizo Colombiano is the only Andean region where the three mountain ranges join. As well as Cucuta in the north-eastern region, the departments of Huila and Valle del Cauca are also considered one of the oldest cacao regions in the country. The geographical location of these departments in the Andean mountains serves as a corridor between the amazon foothills and the Pacific region. This aspect is one of the possible explanations to explain the importance of the cacaos del Magdalena region that Caldas mentioned in 1800. Cacao may have been dispersed by humans from the center of origin in the amazon region to Cucuta in the north of the country along the Magdalena River. The area of current high cacao productivity is the sub-region 4f in Huila department (**Supplementary Figure 4**).

Provinces
This region is subdivided into 14 provinces (40–54 in **Figure 3**; **Table 1**). The south province (sub-region 4a) is a transitional zone that has cool climate and moderate soil temperatures in the lower elevations of the Cauca River. The north province (sub-region 4b) in Tolima is part of the Magdalena River depression, which is characterized by high temperatures, mid precipitation, and high soil temperatures (**Supplementary Figures 2, 3**). The north and central provinces (sub-region 4c) are part of the hyper humid Pacific regions of Choco and Valle del Cauca, characterized by low soil temperature, roughed topography, and mostly surrounded by rainforest. In the alto Magdalena, the south province (sub-region 4d) is part of the eastern face of the central mountain range following the foothills of the Magdalena River. This area has low precipitation and moderate soil temperatures. The east province (sub-region 4e) is isolated on the internal face of the eastern mountain range connecting the upper tableland and the Magdalena depression. Hence, this province is an inner valley of the upper Magdalena River where low precipitation and mid to high soil temperatures are observed. In Tolima and Cundinamarca departments, provinces east and south share similar conditions with moderate seasonality in the soil microclimate and moderate to low temperatures. Further south, part of the 4f sub-region, is the well-known cacao region of Huila along the upper Magdalena valley (**Figure 1B**). On the east side of the Magdalena River, the east province is divided from north to south into three cacao zones. Bache river is the lowest in elevation, hottest and drier than all the rest of Huila cacao provinces. In the same province, the area of Quebrada Bateas further east toward the foothill's conditions are cooler. On the same valley, Campoalegre and Rivera rivers creates an area

of hotter temperatures making it different than its neighboring zone. Further south, toward the towns of Gigante and Garzon along the Magdalena River, soil temperatures and climate are milder all the way to Timaná. In 1809, Caldas mentioned that "the town of Timaná was one of the main cacao producing region in New Granada". The seasonality of soil temperature in this region is high when closer to the Magdalena River and become moderate further away from the riverbed. The central province on the western side of the Magdalena River behaves similarly to the east province. In the Cauca department, the north province (sub-region 4g) has high soil temperature creating drier and hotter conditions in the Palo River. In the same province, cooler temperatures are observed toward the mid slopes of the Ovejas river. Despite close geographic proximity of areas within the north province, these climate conditions make those zones contrasting for cacao growing. The climate conditions in the central province (sub-region 4g) are wet climate and cooler soil temperatures due to higher elevations. The south province (sub-region 4i) is known as the Patía depression where temperatures in the soil are the highest, but a significant day and night climate variability occurs creating microclimates of the hottest and driest conditions of all areas in sub-region 4g. In 1809, Caldas mentioned that the Patía valley was the main cacao growing region in the most southern latitude of the New Granada. In Cauca and Caquetá departments, the southeast (sub-region 4j) and west (sub-region 4k) provinces are part of the eastern Andean range amazon foothills. In this region, the climate is moderate but high precipitation and humidity are typical of the amazon rainforest. In Nariño department, the west province (sub-region 4h) is part of the Pacific region closer to the ocean. It has conditions of high humidity and precipitation, but low seasonality. The south province (sub-region 4i) in Nariño department is part of the western Andean range foothills, which provide unique conditions due to high values of high soil climate seasonality.

CONCLUSIONS

We proposed four regions (north-eastern, north-western, south-western, east), 31 sub-regions and 54 provinces of cacao in Colombia. Our proposed regionalization showed that each cacao region, sub-region, and province have their own signature which can be used to develop cacao denominations depending on its geographic origin. Solar radiation, precipitation, and soil temperature seasonality and isothermality were the variables that best helped to explain the variability of spatial cacao regions in Colombia. Even within provinces, we found biogeographic differences that could determine the quality of cocoa depending on its biogeography. Our regionalization represent ∼250 years of cacao expansion and cultivation in Colombia.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, under request.

AUTHOR CONTRIBUTIONS

CG-O designed the study, assemble spatial data, prepared the manuscript, and analyzed the climate data. AP assembled spatial data and discussed results. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsufs.2022.925800/full#supplementary-material

REFERENCES

Agronet. (2022). *Red de información y comunicación del sector Agropecuario Colombiano*. Available online at: https://www.agronet.gov.co/estadistica/Paginas/home.aspx?cod=1

Antolinez, E., Almanza, P., Baraona, A., Polanco, E., and Serrano, P. (2020). Current state of cocoa plantation: A review of its main limitations. *Ciencia y Agricultura* 17, 1–11. doi: 10.19053/01228420.v17.n2.2020.10729

Baker, R. E. D., Cope, F. W., Holliday, P. C., Bartley, B. G., and Taylor, D. J. (1953). *The Anglo-Colombian Cacao Collecting Expedition. A Report on Cacao Research at the Imperial College of Tropical Agriculture*, St. Augustine, 8–29. Available online at: EurekaMag.com

Caldas, F. J. (1809). *Relación territorial de la provincia de Pamplona. Seminario del Nuevo Reino de Granada 13*. Santafé: Biblioteca Nacional de Colombia.

Ebach, M (2015). *Origins of Biogeography: The Role of Biological Classification in Early Plant and Animal Geography*. Dordrecht; Heidelberg; New York, NY; London: Springer.

Fedecacao. (2021). *Federacion Nacional de Cacaoteros. Año cacaotero 2020-2021, el de mayor producción de cacao en la historia de Colombia*. Available online at: https://www.fedecacao.com.co/post/a%C3%B1o-cacaotero-2020-2021-el-de-mayor-producci%C3%B3n-de-cacao-en-la-historia-de-colombia

Fick, S. E., and Hijmans, R. J. (2017). Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi: 10.1002/joc.5086

Fischer, G., Nachtergaele, F. O., van Velthuizen, H. T., Chiozza, F., Franceschini, G., Henry, M., et al. (2021). *Global Agro-Ecological Zones V4- Model Documentation*. Rome: FAO.

Garcia Lozano, J., Romero Carrascal, M., and Ortiz, L. A. (2005). *CORPOICA-Evaluación edafoclimática de las tierras del trópico bajo y medio colombiano para el cultivo de cacao*. Bogota: CONVENIO CORPOICA-MINAGRICULTURA N° 034/2003.

González-Orozco, C.E., Thornhill, A.H., Knerr, N.J., Laffan, S.W., and Miller, J.T. (2014). Biogeographical regions and phytogeography of the eucalypts. *Divers. Distrib.* 20, 46–58. doi: 10.1111/ddi.12129

González-Orozco, C. E. (2021a). Biogeographical regionalisation of Colombia: a revised area taxonomy. *Phytotaxa* 484, 247–260. doi: 10.11646/phytotaxa.484.3.1

González-Orozco, C. E. (2021b). Biogeographical regions of the genus *Cinchona* L. (Rubiaceae- Cinchoneae). *Novedades Colomb.* 16, 135–156. doi: 10.47374/novcol.2021.v16.1987

González-Orozco, C. E., and Porcel, M. (2021). Two centuries of changes in Andean crop distribution. *J. Biogeogr.* 48, 1972–1980. doi: 10.1111/jbi.14126

González-Orozco, C. E., Porcel, M., Alzate-Velásquez, D. F., and Orduz, J. (2020). Extreme climate variability weakens a major tropical agricultural hub. *Ecol. Indic.* 111, 106015. doi: 10.1016/j.ecolind.2019.106015

González-Orozco, C. E., Porcel, M., Rodriguez-Medina, C., and Yockteng, R. (2021). Extreme climate refugia: a case study of wild relatives of cacao (*Theobroma cacao*) in Colombia. *Biodivers. Conserv.* 31, 161–182. doi: 10.1007/s10531-021-02327-z

Guilherme, A. H. A. L., Araujo, Q. R., Sodré, G. A., Valle, R. R., Souza Jr, J.O., Ramos, E. M. L. S., et al. (2017). Cacao quality: highlighting selected attributes. *Food Rev. Int.* 33, 382–405. doi: 10.1080/87559129.2016.1175011

ICCO-International Cocoa Organization. (2020). *34th Special Session, the International Cocoa Council has amended the Annex "C" of the ICA, 2010 in application of Rule 24 Reconsideration of a decision of the Administrative Rules of the ICCO*. Available online at: https://www.icco.org/revision-of-annex-c-of-the-international-cocoa-agreement-ica-2010/ (accessed May 05, 2022).

Igawa, T.K., Toledo, P., and Md., Anjos, L.J.S. (2022). Climate change could reduce and spatially reconfigure cocoa cultivation in the Brazilian Amazon by 2050. *PLoS ONE* 17, e0262729doi: 10.1371/journal.pone.0262729

Kass, J. M., Vilela, B., Aiello-Lammens, M. E., Muscarella, R., Merow, C., and Anderson, R. P. (2018). *Wallace*: a flexible platform for reproducible modelling of species niches and distributions built for community expansion. *Methods Ecol. Evol.* 9, 1151–1156. doi: 10.1111/2041-210X.12945

Katinas, L., and Crisci, J.V. (2018). Agriculture biogeography: an emerging discipline in search of a conceptual framework. *Progr. Phys. Geogr.* 42, 513–529. doi: 10.1177/0309133318776493

Läderach, P., Martinez-Valle, A., Schroth, G., and Castro, N. (2013). Predicting the future climatic suitability for cocoa farming of the world's leading producer countries, Ghana and Côte d'Ivoire. *Clim. Change* 119, 841–854. doi: 10.1007/s10584-013-0774-8

Lembrechts, J., Van den Hoogen, J., Aalto, J., Ashcroft, M., De Frenne, P., Kemppinen, J., et al. (2021). Global maps of soil temperature. *Global Chang. Biol.* 28, 3110–3144. doi: 10.1111/gcb.16060

León-Moreno, C.E., Rojas-Molina, J., and Castilla-Campos, C. E. (2019). Características fisicoquímicas de los suelos con cacao (*Theobroma cacao* L.) en Colombia: ¿Están adecuados para mejorar la productividad? *Agron. Colomb.* 37, 28–38. doi: 10.15446/agron.colomb.v37n1.70545

MADR (2021). *Ministerio de Agricultura y Desarrollo Rural de Colombia. Cadena de cacao. Direccion de Cadenas Agricolas y Forestales*. Available online at: https://sioc.minagricultura.gov.co/Cacao/Documentos/2021-03-31%20Cifras%20Sectoriales.pdf

Metsalu, T., and Vilo, J. (2015). Clustvis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* 43, 566–570. doi: 10.1093/nar/gkv468

Morrone, J. J., Escalante, T., Rodríguez-Tapia, G., Carmona, A., Arana, M., et al. (2022). Biogeographic regionalization of the neotropical region: new map and shapefile. *Anais Acad. Bras. Ciênc.* 94, 1–5. doi: 10.1590/0001-3765202220211167

Nuñez-Rodríguez, J., Mendoza-Ferreira, O., González-Verjel, M., Carvajal-Rodríguez, J., and Carrero-Carmona, D. (2021). Altitudinal displacement of cocoa-producing areas in the department of Norte de Santander due to the effects of climatic variations. *Aibi Revista De investigación, administración E ingeniería* 8, 217–224.

Patiño, V. M. (2002). *Historia y dispersión de los frutales nativos del neotrópico*. Cali: Centro Internacional de Agricultura Tropical, Cali. CO, 665.

QGIS Development Team (2019). *QGIS Geographic Information System*. Open-Source Geospatial Foundation. Available online at: http://qgis.Org (accessed May 22, 2022).

Rodriguez-Medina, C., Caicedo Arana, A., Sounigo, O., Argout, X., Alvarado Alvarado, G., and Yockteng, R. (2019). Cacao breeding in Colombia, past, present and future. *Breed. Sci.* 69, 373–382. doi: 10.1270/jsbbs.19011

Santos, M., Fonseca, A., Fraga, H., Jones, G.V., and Santos, J.A. (2020). Bioclimatic conditions of the Portuguese wine denominations of origin under changing climates. *Int. J. Climatol.* 40, 927–941. doi: 10.1002/joc.6248

Santosa, E., Sakti, G. P., Fattah, M.Z., Zaman, S., and Wahjar, A. (2018). Cocoa production stability in relation to changing rainfall and temperature in East Java, Indonesia. *J. Trop. Crop. Sci.* 5, 6–17. doi: 10.29244/jtcs.5.1.6-17

SIPRA-UPRA (2019). *Aptitud Cacao. Escala 1:100.000*. Available online at: https://sipra.upra.gov.co/ (accessed May 22, 2022).