Edaphoclimatic Descriptors of Wild Tomato Species (*Solanum* Sect. Lycopersicon) and Closely Related Species (*Solanum* Sect. Juglandifolia and Sect. Lycopersicoides) in South America

Gabriela Ramírez-Ojeda[1], Iris Edith Peralta[2,3], Eduardo Rodríguez-Guzmán[4], Jaime Sahagún-Castellanos[1], José Luis Chávez-Servia[5], Tulio Cecilio Medina-Hinostroza[6], Jorge Rodrigo Rijalba-Vela[7], Leopoldo Pompeyo Vásquez-Núñez[7] and Juan Enrique Rodríguez-Pérez[1]*

[1]Crop Science Department, Horticulture Institute, Chapingo Autonomous University (UACh), Chapingo, Mexico, [2]Agronomy Department, Agricultural Sciences Faculty, National University of Cuyo (UNCUYO), Mendoza, Argentina, [3]Scientific Technological Center CONICET, Argentine Institute for Arid Zones Research, Mendoza, Argentina, [4]Agronomy Department, University Center for Biological and Agricultural Sciences, University of Guadalajara (UdG), Zapopan, Mexico, [5]Interdisciplinary Research Center for Integral Regional Development Oaxaca Unit, National Polytechnic Institute (IPN), Oaxaca, Mexico, [6]Directorate of Genetic Resources and Biosafety, Ministry of Environment, Lima, Peru, [7]Biological Sciences Faculty, National University Pedro Ruiz Gallo (UNPRG), Lambayeque, Peru

Wild species related to cultivated tomato are essential genetic resources in breeding programs focused on food security to face future challenges. The ecogeographic analysis allows identifying the species adaptive ranges and most relevant environmental variables explaining their patterns of actual distribution. The objective of this research was to identify the diversity, ecological descriptors, and statistical relationship of 35 edaphoclimatic variables (20 climatic, 1 geographic and 14 edaphic variables) from 4,649 accessions of 12 wild tomato species and 4 closely related species classified in *Solanum* sect. Lycopersicon and clustered into four phylogenetic groups, namely "Lycopersicon group" (*S. pimpinellifolium*, *S. cheesmaniae*, and *S. galapagense*), "Arcanum group" (*S. arcanum*, *S. chmielewskii*, and *S. neorickii*), "Eriopersicon group" (*S. habrochaites*, *S. huaylasense*, *S. corneliomulleri*, *S. peruvianum*, and *S. chilense*), "Neolycopersicon group" (*S. pennellii*); and two phylogenetically related groups in *Solanum* sect. Juglandifolia (*S. juglandifolium* and *S. ochranthum*), and section Lycopersicoides (*S. lycopersicoides* and *S. sitiens*). The relationship between the climate and edaphic variables were determined by the canonical correlation analysis, reaching 89.2% of variation with the first three canonical correlations. The most significant climatic variables were related to humidity (annual evapotranspiration, annual precipitation, and precipitation of driest month) and physicochemical soil characteristics (bulk density, pH, and base saturation percentage). In all groups, ecological descriptors and diversity patterns were consistent with previous reports. Regarding edaphoclimatic diversity, 12 climate types and 17 soil units were identified

among all species. This approach has promissory applications for biodiversity conservation and uses valuable genetic resources related to a leading crop.

INTRODUCTION

Latin America and the Caribbean are regions rich in biodiversity, hosting nearly 60% of the world's biological diversity (UNEP-WCMC, 2016). Within this region, Mesoamerica is recognized as one of the main centers of origin, diversification, domestication, and biological plant diversity of various species of agricultural interest and animal consumption (Fortuny-Fernández et al., 2017). The complex evolutionary history, phylogenetics, geology, biogeography, and climatic variability are some factors that enhance the diversity in this area (UNEP-WCMC, 2016). This condition is essential to ensure food, socioeconomic, and cultural sovereignty for sustainable development and offers a large number of ecosystem services (FAO et al., 2019).

In this sense, tomato (*Solanum lycopersicum* L.) is one of the most cultivated vegetables due to its wide distribution and environmental adaptation in warm, subtropical, and tropical regions with nutritional and commercial importance worldwide (Peralta et al., 2008; Ramírez-Ojeda et al., 2021a). Regarding the place of origin and diversification of tomato, Peru is considered the center of origin with two transitions that involve tomato diversification process; the first one in South America, from wild species *S. pimpinellifolium* L. to a partially domesticated species *S. lycopersicum* L. var. *cerasiforme* (SLC); the second transition occurred in Mesoamerica from SLC to the completely domesticated species *S. lycopersicum* L. var. *lycopersicum*. However, new findings indicate that the origin of SLC may be prior to its domestication since many typical characteristics of tomatoes grown in South America come from this species; SLC is subsequently considered to have been lost or declined once the partially domesticated forms extended to the north (Razifard et al., 2020).

Wild species related to cultivated tomatoes are essential genetic resources in breeding programs focused on food security to face future challenges. Therefore, it is of strategic importance to study the climatic and edaphic factors that help to understand their current distribution patterns, as well as to establish the best indicators predicting possible effects of climate change and natural or anthropic environmental alterations. This is why it is necessary to undertake national and regional strategies for the conservation and use of cultivated and wild tomato genetic resources (Sandoval-Ceballos et al., 2021).

Based on an integrative taxonomy, which includes multiple evidences, the classification of wild tomatoes and their wild relatives was proposed: *Solanum* section Lycopersicon (Mill.) Wettst. comprises cultivated tomato (*S. lycopersicum* L.) and 12 wild tomato species: *S. arcanum* Peralta, *S. cheesmaniae* (L. Riley) Fosberg, *S. chilense* Dunal, *S. chmielewskii* (C. M. Rick, Kesicki, Fobes, and M. Holle), D. M. Spooner, G. J. Anderson,

and R.K. Jansen, *S. corneliomulleri* J. F. Macbride, *S. galapagense* S. C. Darwin and Peralta, *S. habrochaites* S. Knapp and D. M. Spooner, *S. huaylasense* Peralta, *S. neorickii* D. M. Spooner, G. J. Anderson and R. K. Jansen, *S. pennellii* Correll, *S. peruvianum* L., and *S. pimpinellifolium* L. Four phylogenetically related *Solanum* species are also considered in the present study: *S. juglandifolium* Dunal, *S. ochranthum* Dunal (*Solanum* sect. Juglandifolia (Rydb.) A. Child), *S. lycopersicoides* Dunal, and *S. sitiens* I. M. Johnston (*Solanum* section Lycopersicoides (A. Child) Peralta) (Peralta et al., 2008; Causse et al., 2016; Tropicos.org, 2021).

Evidence of phylogenetic relationships of these species have been studied in detail by Peralta et al. (2008), who have proposed a six-group classification of wild tomatoes and phylogenetically closely related species: Section Lycopersicon: "Lycopersicon group" (*S. pimpinellifolium*, *S. cheesmaniae*, and *S. galapagense*), "Arcanum group" (*S. arcanum*, *S. chmielewskii*, and *S. neorickii*), "Eriopersicon group" (*S. habrochaites*, *S. huaylasense*, *S. corneliomulleri*, *S. peruvianum*, and *S. chilense*), "Neolycopersicon group" (*S. pennellii*); and two outgroups: Section Juglandifolia (*S. juglandifolium* and *S. ochranthum*) and Section Lycopersicoides (*S. lycopersicoides* and *S. sitiens*). This classification has been verified by molecular, genomic, and transcriptomic evidence of wild tomatoes (Rodríguez et al., 2009; Aflitos et al., 2014; Pease et al., 2016) and recently has been used for ecogeographic studies with satisfactory results (Ramírez-Ojeda et al., 2021a).

By considering plant genetic resources as the biological foundation for maintaining and improving crop productivity (Kantar et al., 2015), wild tomato species constitute an important gene pool due to the presence of genes with tolerance and resistance to biotic and abiotic factors (Arellano-Rodríguez et al., 2013; Cervantes-Moreno et al., 2014; Nosenko et al., 2016; Razali et al., 2018; Dinh et al., 2019) with potential use for breeding programs. Additionally, several questions arise about these gene pools, such as current distribution, population dynamics *in situ* or *ex situ*, and how are they used directly or as sources of genes to generate new varieties that respond to current and future basic problems of tomato cultivation (for example, climate change, diseases, pests), including the contribution of genes capable of conferring a greater nutritional–nutraceutical quality to new varieties (Chávez-Servia et al., 2011; Hernández-Bautista et al., 2014).

Identification of variables that derive in adaptation and speciation processes requires a large amount of field data of significant variables in natural populations. Recent developments and the use of remote sensing technologies, as well as a great availability of environmental information derived from Geographic information systems (GIS), have made it possible

to identify patterns of species environmental variations at different scales (Nakazato et al., 2010). These tools and the availability of databases, with passport information of specimens collected in natural areas, allow for verification of the presence of species in a geographic range, as well as possible ecological descriptors, that is, to describe in detail the environmental conditions associated with the distribution of natural populations (Nakazato et al., 2010; Sánchez-González et al., 2018; Vilchez et al., 2019; Ministerio del Ambiente, 2020; Ramírez-Ojeda et al., 2021a).

One way to identify the adaptive ranges and most relevant variables that determine species distribution of valuable genetic resources is through ecogeographic studies, focusing on collection, conservation, characterization, documentation, and use of these resources (Parra-Quinajo et al., 2012; Pease et al., 2016), with the purpose of describing and explaining spatial patterns and processes involved in biodiversity distribution through time and space (Martiny et al., 2006; Tofalo et al., 2013; Délices et al., 2019). Ecogeographic studies of plant genetic resources allow the identification of the adaptive ranges of the species and the most relevant environmental variables that define their distribution (Parra-Quinajo et al., 2012; Ramírez-Ojeda et al., 2021a). Through ecogeographic studies, it is also possible to predict the environmental characteristics of the accession sites (Steiner and Greene, 1996) from ecological descriptors obtained through GIS tools using the geographical location and environmental variables (Lobo-Burle et al., 2013; Sánchez-González et al., 2018; Ramírez-Ojeda et al., 2021a, 2021b).

Currently, several information sources about geographical distribution of tomato species can be found in public databases (GBIF, 2021; Solanaceae Source, 2021; TGRC, 2021), conservation programs and gene banks (Córdoba-Téllez and Molia-Moreno, 2006; Florido et al., 2009; Magallanes-López et al., 2020), and genetic resources baseline studies (Ministerio del Ambiente, 2020), as well as some studies on geographic distribution patterns and ecological and climatic descriptors of wild tomato species (Peralta et al., 2008; Chetelat et al., 2009; Nakazato et al., 2010; Grandillo et al., 2011; Gonzá lez et al., 2013; Vilchez et al., 2019; Ramírez-Ojeda et al., 2021a). However, information regarding edaphic conditions of the sites where these species are located is limited or unknown (Balaguera-López et al., 2009).

Soil, a finite and nonrenewable natural resource, is of great importance in a large number of environmental services such as food and biomass production, climate regulation, carbon fixation, water storage and filtration, biogeochemical cycles, biodiversity reserve, and human physical and cultural environment (Burbano-Orjuela, 2016). Therefore, when considering edaphic together with climatic characteristics, it allows having a better understanding of the ecological and distribution patterns of the species.

Due to the limited edaphic information available regarding optimal characteristics for development of wild tomato species, the aim of the present work was to study ecological descriptors associated with soil characteristics and their relationship and the statistical association with climatic variables. Likewise, it was also

analyzed whether the classification of wild tomatoes is related to the edaphoclimatic descriptors and supports the proposed groups of species.

MATERIALS AND METHODS

Database

Initial database consisted of 12,131 accessions of 12 wild tomato species and 4 phylogenetically related species. Of these, 7,482 accessions were eliminated due to atypical data, repeated records, or accessions with little geographic precision and outside natural areas identified according to the altitude and ecological ranges reported (Peralta et al., 2008; Grandillo et al., 2011; Ministerio del Ambiente, 2020). The final 4,649 accessions database came from scientific reports, articles (Sotomayor et al., 2019; Razifard et al., 2020), international plant repositories (Tomato Genetic Resource Center, Global Biodiversity Information Facility, Solanaceae Source) (GBIF, 2021; Solanaceae Source, 2021; TGRC, 2021), and new accessions collected in 2018–2019 in Peru (Ministerio del Ambiente, 2020). The distribution of 16 species is shown in **Figure 1**. The species distribution is shown in Figure A1 in the **Supplementary Material**. It should be noted that S. lycopersicum was not included because its wide distribution would not reflect a natural but artificial distribution due to anthropic dispersal as a cultivated or ruderal species.

Environmental Information

For the statistical analysis and ecological descriptors, an environmental information system with 900 m spatial resolution was built with 35 variables (**Table 1**). Nineteen bioclimatic variables were obtained from WorldClim version 2.1 from period 1970 to 2000 (Fick and Hijmans, 2017). Annual evapotranspiration (ET) was calculated from the sum of monthly values reported by Trabucco and Zomer (2019). Altitude (Alt), a geographic variable, was obtained with an elevation model from WorldClim (Fick and Hijmans, 2017). Alt was analyzed together with climatic variables due to the strong influence on the definition of climates. Finally, 14 edaphic variables obtained from the Harmonized World Soil Database version 1.1 (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009) were used.

Edaphoclimatic diversity patterns were identified from climate types corresponding to world climatic classification proposed by Beck et al. (2018) with the Köppen–Geiger system and soil units from the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009) (**Table 2**).

Canonical Correlation Analysis and Ecological Descriptors

A selection of climatic and edaphic variables was made in order to identify a strong linear dependence (collinearity) between more than two explanatory variables. For this purpose, Pearson's correlations were obtained, between variables, eliminating one

FIGURE 1 | Geographic distribution of 12 wild tomato (*Solanum* sect. Lycopersicon) and 4 closely related species (*Solanum* sect. Juglandifolia, and Lycopersicoides). In parenthesis, number of accessions after each species name.
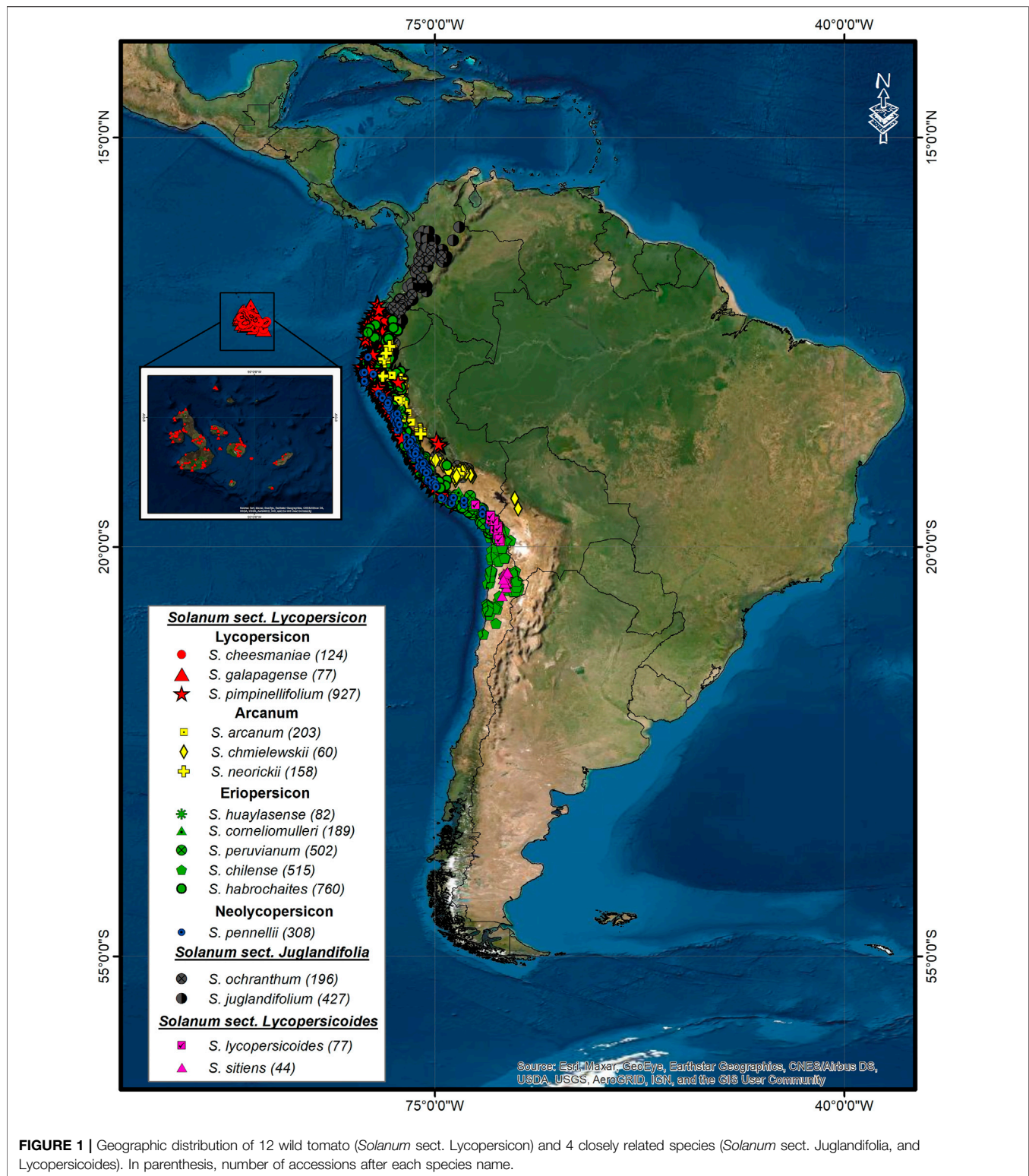
of each pair whose absolute coefficient was greater than 0.90. The conserved variable was the one that showed the highest number of correlations with other variables, and therefore, the lowest number of non–linearly associated variables was maintained.

With the selected variables, a canonical correlation analysis was carried out to identify the relationship between the group of climatic variables and the group of edaphic variables. All statistical analyses were performed using SAS

TABLE 1 | Climatic, geographic, and edaphic variables used in the canonical correlation analysis and ecological descriptors.

Climatic variables
WorldClim variables (1970–2000): Annual mean temperature (**Bio1**, ˚C), mean diurnal range (**Bio2**, ˚C), isothermality (**Bio3**, Bio2/Bio7 × 100), temperature seasonality (**Bio4**, standard deviation × 100), maximum temperature of the warmest month (**Bio5**, ˚C), minimum temperature of coldest month (**Bio6**, ˚C), temperature annual range (**Bio7**, Bio5-Bio6), mean temperature of wettest quarter (**Bio8**, ˚C), mean temperature of driest quarter (**Bio9**, ˚C), mean temperature of warmest quarter (**Bio10**, ˚C), mean temperature mean of coldest quarter (**Bio11**, ˚C), annual precipitation (**Bio12**, mm), precipitation of wettest month (**Bio13**, mm), precipitation of driest month (**Bio14**, mm), precipitation seasonality (**Bio15**, coefficient of variation), precipitation of wettest quarter (**Bio16**, mm), precipitation of driest quarter (**Bio17**, mm), precipitation of warmest quarter (**Bio18**, mm), and precipitation of coldest quarter (**Bio19**, mm) (Fick and Hijmans, 2017). Annual evapotranspiration (**ET**, mm) (Trabucco and Zomer, 2019).

Geographic variables
Altitude (**Alt**, masl) (Fick and Hijmans, 2017)

Edaphic variables
Percentage of gravel (**GR**, %), sand (**SA**, %), silt (**SI**, %), and clay (**CL**, %), bulk density (**BD**, kg/dm^3), organic carbon (**CO**, %), pH, cation exchange capacity (**CEC** cmol/kg), base saturation (**BS**, %), calcium carbonate (**CaCO$_3$**, %), total exchangeable bases (**TEB**, cmol/kg), calcium sulfate (**CaSO$_4$**, %), salinity (**SAL**, dS/m), and sodium (**SOD**, %) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009)

TABLE 2 | Climate types and soil units used to determine edaphoclimatic diversity patterns.

Climate type
Af (tropical and rainforest), Am (tropical and monsoon), Aw (tropical and savannah), BWh (arid, desert, and hot), BWk (arid, desert, and cold), BSh (arid, steppe, and hot), BSk (arid, steppe, and cold), Csa (temperate, dry summer, and hot summer), Csb (temperate, dry summer, and warm summer), Csc (temperate and dry and cold summer), Cwa (temperate, dry winter, and hot summer), Cwb (temperate, dry winter, and warm summer), Cwc (temperate, dry winter, and cold summer), Cfa (temperate, no dry season, and hot summer), Cfb (temperate, no dry season, and warm summer), Cfc (temperate, no dry season, and cold summer), Dsa (cold, dry summer, and hot summer), Dsb (cold, dry summer, and warm summer), Dsc (cold, dry summer, and cold summer), Dsd (cold, dry summer, and very cold winter), Dwa (cold, dry winter, and hot summer), Dwb (cold, dry winter, and warm summer), Dwc (cold, dry winter, and cold summer), Dwd (cold, dry winter, and very cold winter), Dfa (cold, no dry season, and hot summer), Dfb (cold, no dry season, and warm summer), Dfc (cold, no dry season, and cold summer), Dfd (cold, no dry season, and very cold winter), ET (polar and tundra), and EF (polar and frost) (Beck et al., 2018)

Soil units
AC (Acrisol), **AL** (Alisol), **AN** (Andosol), **AR** (Arenosol), **AT** (Anthrosol), **CH** (Chernozem), **CL** (Calcisol), **CM** (Cambisol), **FL** (Fluvisol), **FR** (Ferralsol), **GL** (Gleysol), **GR** (Greysem), **GY** (Gypsisol), **HS** (Histosol), **KS** (Kastanozem), **LP** (Leptosol), **LV** (Luvisol), **LX** (Lixisol), **NT** (Nitisol), **PD** (Podzoluvisol), **PH** (Phaezem), **PL** (Planosol), **PT** (Plinthosol), **PZ** (Podzol), **RG** (Regosol), **SC** (Solonchak), **SN** (Solonetz), and **VR** (Vertisol) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009)

software (Statistical Analysis System) version 9.3 (SAS Institute, 2011).

Regarding ecological descriptors, these were calculated for each variable and each species (12 wild tomato and 4 phylogenetically related species) with the methodology proposed by Steiner and Greene (1996). Ecological descriptors were determined by vectors calculated with the geographic coordinates of each accession and the punctual value of each variable extracted with GIS.

Subsequently, the edaphic and climatic variables were identified as significant in the canonical correlation analysis; the extreme values (maximum and minimum), the median, and the coefficient of variation (CV = (Q/Med) × 100, where Q = (Q3 − Q1)/2 (interquartile range), and Med = median) were identified.

Finally, to identify the ecological distribution patterns of every group of species, altitude, annual mean temperature, precipitation, and annual evapotranspiration were considered as climatic variables and pH, cation exchange capacity (CEC), bulk density (BD), and base saturation (BS) as edaphic variables. These variables were chosen due to the importance and influence they have on the distribution and development of the species (Ramírez-Ojeda et al., 2021b), in addition to the importance and significance that they showed in the statistical analyses.

Edaphoclimatic Diversity

Edaphoclimatic diversity was identified using GIS tools with the vector of geographic coordinates of each accession and raster images of climate types and soil units (**Table 2**). **Figure 2** shows the distribution of climate types and soil units in South America. With the resulting information, a frequency table by climate type and soil unit was obtained for each species group (6) and for each individual species (16).

Hot spot Analysis

Critical points of species abundance and areas with the greatest diversity concentration were established using ArcGIS with the "Spatial Statistics Tools" module. Spatial density maps were constructed by adding all those accessions of each species with a distance between accessions of 1 km. A distance criterion was chosen based on previous diversity studies of potato species (*Solanum* Sect. *Petota*), the sister group of tomatoes (Hijmans et al., 2002; Spooner et al., 2010). Subsequently, hot spot spatial analysis was performed with Getis-Ord Gi* statistic (Getis and Ord, 1992) to quantify the specific areas of high clustering and spatial significance for species abundance and diversity.

The hot spot analysis determines the spatial grouping of points higher (hot spot) or lower (cold spot) than the expected by a

FIGURE 2 | (A) Climate classification according to Beck et al. (2018): Af (tropical and rainforest), Am (tropical and monsoon), Aw (tropical and savannah), BWh (arid, desert, and hot), BWk (arid, desert, and cold), BSh (arid, steppe, and hot), BSk (arid, steppe, and cold), Csb (temperate, dry summer, and warm summer), Cwb (temperate, dry winter, and warm summer), Cfb (temperate, no dry season, and warm summer), and ET (polar, frost). **(B)** Soil units according to FAO/IIASA/ISRIC/ISSCAS/JRC (2009): AC (acrisol), AL (alisol), AN (andosol), AR (arenosol), CL (calcisol), CM (cambisol), FL (fluvisol), FR (ferralsol), GL (gleysol), HS (histosol), KS (kastanozem), LP (leptosol), LV (luvisol), LX (lixisol), NT (nitisol), PH (phaezem), PL (planosol), PT (plinthosol), PZ (podzol), RG (regosol), SC (solonchak), SN (solonetz), VR (vertisol).

random distribution. Significance tests were calculated using z-values (Getis and Ord, 1992).

RESULTS

Canonical Correlation Analysis and Ecological Descriptors

According to Pearson's correlation coefficients, out of the 34 edaphoclimatic variables, 19 did not present collinearity. The variables selected for subsequent statistical analyses and ecological descriptors were annual evapotranspiration, altitude, precipitation of dries month, annual precipitation, temperature annual range, isothermality, mean diurnal range, annual mean temperature, percentage of sand, silt and clay, BD, pH, organic carbon, CEC, BS, calcium carbonate CaCO$_3$), sodicity, and salinity.

The canonical correlation analysis (CCA), performed with two groups of variables (climatic and edaphic), indicated that the first three canonical correlations had values of 0.800, 0.436, and 0.415, respectively, and percentages of explanation of data variation of 71.45, 9.38, and 8.36%, respectively, with a total of 89.20%.

Likelihood ratio tests indicated that the three canonical correlations are different from zero ($p \leq 0.0001$).

Correlations between climate characteristics and their canonical variables indicated that CLIMATE1 is associated with annual evapotranspiration (ET, −0.907), annual precipitation (Bio12, −0.947), and precipitation of driest month (Bio14, −0.864). The CLIMATE2 vector is associated with altitude (Alt, −0.5164) and isothermality (Bio3, −0.5642), which roughly quantifies the "hot" (low values) and "cold" (high values) regions. The CLIMATE3 vector represents the mean annual temperature (Bio1, 0.7145) and mean diurnal range (Bio2, 0.793).

Regarding correlations of soil canonical variables, the SOIL1 vector represent BD (0.801), pH (0.658), and BS percentage (BS, 0.707). High values of the SOIL2 vector identify soils with high content of sand (0.478), pH (0.552), CaCO$_3$ concentration (0.546), and low clay content (−0.427). Finally, the SOIL3 vector does not show important correlations.

The correlations between the CLIMATE vectors with the original soil variables indicate that CLIMATE1 had correlations of importance with BD (0.641), pH (0.527), CEC (−0.452), BS (0.566), and CaCO$_3$ content (0.424). CLIMA2 and

FIGURE 3 | Scatter plot of SOIL1 vs CLIMATE1 canonical vectors from CCA with 0.80 correlation and 74% of total variation. Group 1 (red): Section Lycopersicon: "Lycopersicon group", Group 2 (yellow): "Arcanum group", Group 3 (green): "Eriopersicon group", Group 4 (blue): "Neolycopersicon group" outgroup close related species in Section Juglandifolia Group 5 (grey): and Group 6 (pink): Section Lycopersicoides. In each scatter plot first-quadrant (I) represents environments with low annual evapotranspiration (ET), annual precipitation (Bio12), precipitation of driest month (Bio14) and high pH, BD, and base saturation (BS); second-quadrant (II) represents environments with high ET, Bio12, Bio14 and high pH, BD and BS; third-quadrant (III) represents environments with high ET, Bio12, Bio14 and low pH, BD and BS; fourth-quadrant (IV) represents environments with low ET, Bio12, Bio14, and low pH, BD, and BS.

CLIMA3 did not show correlation of importance. SOIL1 with climate variables showed associations with annual evapotranspiration (ET, −0.7267), annual precipitation (Bio12, −0.7585), precipitation of driest month (Bio14, −0.6924), isothermality (Bio3, 0.5404), and temperature annual range (Bio7, 0.5117). SOIL2 and SOIL3 did not show correlation of importance.

Figure 3 shows the relationship between canonical variables CLIMATE1 and SOIL1, representing 71.45% of the total data variability and a positive correlation of both canonical variables of 0.80. This figure shows the distribution and ecological adaptation of every species regarding canonical correlations.

Table 3 and **Table 4** show the ecological descriptors of edaphic and climatic variables identified as significant in the first and second canonical correlation. These results are mostly consistent with the environmental ranges previously reported in other studies. Table A1 in **Supplementary Material** shows the ecological descriptors of the rest of the variables.

Figure 4 shows the boxplots for four climatic variables for each of the six species groups, as well as the amplitude observed for each variable.

Among the main findings, it can be observed that groups 4 (*S. pennellii*) and 6 (*S. lycopersicoides* and *S. sitiens*) are ones that contain the species that distributes in environments with the lowest availability of precipitation and evapotranspiration. Considering altitude, group 1 (*S. pimpinellifolium*, *S. cheesmaniae*, and *S. galapagense*) has the lowest average altitude, while group 6 (*S. lycopersicoides* and *S. sitiens*) has

the highest average altitude. Group 1 was located in environments with the highest mean annual temperature; by contrast, group 6 had the lowest average annual temperature. Groups 2, 3, and 5 remained in transition climatic conditions with the rest of the phylogenetic groups.

The analysis of four edaphic variables in **Figure 5** determines that group 5 (*S. juglandifolium* and *S. ochranthum*) has the lowest pH average. In all groups, BD was relatively constant, with similar values in all species. The mean BS in most of the groups was greater than 80%, except for group 5, with an average value around 40%. In general, soil characteristics in all groups of species were relatively similar, except for group 5 (*S. ochranthum* and *S. juglandifolium*) which presented an opposite trend.

Edaphoclimatic Diversity

The edaphoclimatic diversity found in 16 species is shown in **Figure 6** and **Figure 7**. Regarding, climate diversity, it was possible to identify 12 climate types of the 21 reported for Latin America by Beck et al. (2018).

Within the six phylogenetically related groups identified by Peralta et al. (2008) and used by Ramírez-Ojeda et al. (2021a), specific climate type patterns can be observed, with the same climate types occurring in different proportions within each group (**Figure 6**), confirming in most of the groups, the environmental distribution similarity between the species that make them up.

S. habrochaites has the greatest diversity (11 climate types), intermedium diversity (8 climate types) was found in *S. arcanum*,

TABLE 3 | Ecological descriptors of climatic and edaphic variables associated with the first canonical correlation (71.4%) for 12 species of wild tomato and 4 closely related species. Bio12 = annual precipitation, Bio14 = precipitation of driest month, pH = hydrogen ion concentration. *Range (maximum–minimum value), **median, ***coefficient of variation.

Group/Section	Species	Bio12 (mm)	Bio14 (mm)	ET (mm)	BD (kg/dm³)	pH	BS (%)
Lycopersicon	*S. cheesmaniae*	107–562*	0–15	187–1,125	1.1–1.4	4.3–8.5	31.0–100
		277** (21.3)***	3 (66.6)	45 (41.9)	1.3 (1.1)	7.1 (15.4)	100 (34.5)
	S. galapagense	135–546	0–15	160–1,052	1.1–1.4	4.3–8.5	31.0–100
		274 (16.0)	4 (50)	531 (30.6)	1.3 (0.4)	4.9 (22.4)	32.0 (107.8)
	S. pimpinellifolium	1–2,828	0–143	1–1,710	0.3–1.5	3.2–8.5	10.0–100
		68 (146.3)	0 (0)	43 (195.9)	1.4 (6.5)	7.6 (9.2)	100 (4.5)
Arcanum	*S. arcanum*	22–1,193	0–55	11–1,094	0.3–1.5	4.6–8.5	17.0–100
		487 (44.0)	4 (137.5)	390 (43.2)	1.3 (1.5)	6.4 (13.2)	87.0 (5.1)
	S. chmielewskii	504–1,318	4–19	429–874	1.2–1.5	5.2–8.1	19.0–100
		944 (18.2)	11 (22.7)	610 (16.9)	1.2 (2.5)	8.1 (4.3)	100 (0)
	S. neorickii	426–1,366	3–68	326–1,031	0.2–1.7	4.6–8.1	17.0–100
		817 (21.2)	18.5 (45.9)	672 (13.7)	1.2 (10.2)	5.6 (26.7)	81 (44.0)
Eriopersicon	*S. huaylasense*	128–507	0–3	73–424	1.2–1.4	5.2–7.9	21.0–100
		328 (21.5)	1 (50)	238 (23.0)	1.4 (7.9)	5.6 (12.5)	38.0 (81.5)
	S. corneliomulleri	19–434	0–2	12–354	1.1–1.5	4.2–8.1	19.0–100
		205 (59.2)	0 (0)	141 (45.7)	1.4 (6.7)	5.6 (16.9)	38.0 (86.8)
	S. peruvianum	0–534	0–3	0–427	1.2–1.5	3.2–8.6	10.0–100
		25 (97.4)	0 (0)	13 (134.6)	1.3 (4.1)	7.6 (6.3)	100 (5.0)
	S. chilense	0–355	0–1	3–275	1.0–1.5	4.2–8.6	28.0–100
		29 (68.9)	0 (0)	20 (72.5)	1.3 (2.3)	7.5 (10.6)	100 (6)
	S. habrochaites	11–2,358	0–143	8–1,682	0.3–1.7	4.3–8.5	14.0–100
		605 (42.0)	3 (266.6)	535 (43.9)	1.3 (7.6)	5.7 (15.7)	87 (35.6)
Neolycopersicon	*S. pennellii*	1–404	0–3	0–289	1.2–1.5	5.1–8.5	19.0–100
		49 (94.8)	0 (0)	33 (95.8)	1.3 (4.9)	7.9 (14.5)	100 (31.0)
Juglandifolia	*S. juglandifolium*	555–3,214	1–194	413–1,648	0.3–1.9	4.1–7.7	13.0–100
		1,895 (28.7)	60 (49.1)	1,177 (10.0)	0.9 (1.0)	5.2 (1.9)	23.0 (36.9)
	S. ochranthum	507–2,358	2–131	387–1,474	0.3–1.7	3.2–8.5	10.0–100
		1,010 (11.0)	36 (43)	814 (14.1)	1.2 (11.2)	5.6 (13.3)	45.0 (67.7)
Lycopersicoides	*S. lycopersicoides*	13–215	0–0	9–182	1.3–1.4	4.7–8.1	34.0–100
		104 (50.4)	0 (0)	82 (58.5)	1.3 (1.5)	7.5 (10.6)	100 (6.5)
	S. sitiens	8–31	0–0	9–26	1.2–1.4	6.4–7.9	88.0–100
		17 (24.2)	0 (0)	21 (16.6)	1.2 (8.5)	6.5 (10.7)	88.0 (6.8)

while *S. sitiens* has the greatest climatic restriction, located only in climates BWk (arid, desert, cold). The climate type identified in most of the accession sites was associated with the 16 species was BSk (arid, steppe, and cold), and only absent in species of Lycopersicon group (S. *cheesmaniae*, S. *galapagense*, and S. *pimpinellifolium*) and in *S. juglandifolium* and *S. sitiens*. The opposite case was presented with Cwc climate (temperate, dry winter, and cold summer) present only in some areas where *S. habrochaites* was collected. *S. juglandifolium* and *S. ochranthum* share similar climatic types but were most frequently found in Cfb (temperate, no dry season, warm summer).

Diversity of soil units among wild tomato species (**Figure 7**) found 17 different soil units out of the 23 reported for Latin America in Harmonized World Soil Database from FAO/IIASA/ISRIC/ISSCAS and JRC (2009). Most frequent soils types were leptosol (LP), present in all species, regosol (except *S. huaylasense*) and less frequently acrisol (AC), present only in some *S. pimpinellifolium* accessions.

The greatest edaphic diversity was found in *S. pimpinellifolium*, with accessions in 16 of the 17 reported soil units (except VR). The opposite case was identified for species of Lycopersicoides section *S. sitiens* and *S. lycopersicoides*, with two and four soil units, respectively. Likewise in the patterns of climatic diversity described, edaphic diversity is similar within species, integrating each of the six phylogenetically related groups.

Hot spot Analysis

Areas with a high number of species and accessions were determined by hot spot analysis. **Figure 8** shows the result of hot spot analysis applied with a distance of 1 km between accessions for 4,649 accessions of 12 wild tomato and 4 phylogenetically related species. The highest concentration of species is located in two areas of Peru, one near Trujillo and Chimbote, and the second area around Lima. Likewise, a small area with high diversity is located in southern Peru and northern limit of Chile. The zone in Trujillo–Chimbote is characterized by the presence of seven species (*S. pennellii*, *S. arcanum*, *S. neorickii*, *S. huaylasense*, *S. habrochaites*, *S. pimpinellifolium*, and *S. ochranthum*). The region of high diversity around *Lima* also

TABLE 4 | Ecological descriptors of climatic and edaphic variables associated with the second canonical correlation (9.3%) for 12 species of wild tomato and 4 closely related species. Bio3 = isothermality, Sand = sand percentage, Clay = clay percentage. *Range (maximum–minimum value), **median, ***coefficient of variation.

Group/Section	Species	Alt (m)	Bio3 (°C × 100)	Sand (%)<	CaCO₃ (%)	Clay (%)
Lycopersicon	_S. cheesmaniae_	5–1,478*	58.3–74.8	33–60	0–3.1	3–37
		87** (152.4)***	65 (4.5)	43 (11.6)	2 (67.3)	28 (16.0)
	S. galapagense	4–868	59.1–73.9	33–60	0–3.1	3–37
		45 (240.0)	68.3 (4.9)	34 (14.7)	0 (0)	36 (12.5)
	S. pimpinellifolium	1–1,774	48.2–89.4	0–94	0–4.3	0–56
		92 (101.9)	65.6 (8.0)	44 (51.1)	2 (65)	17 (47)
Arcanum	_S. arcanum_	132–3,292	65.5–90.1	0–83	0–4.3	0–45
		1,767 (33.1)	87 (2.7)	54 (10.1)	0 (0)	17 (8.8)
	S. chmielewskii	1,803–3,195	73.6–86.6	25–63	0–2.0	14–31
		2,445 (11.3)	82.6 (2.9)	63 (12.6)	2 (17.5)	14 (14.2)
	S. neorickii	1,202–3,262	76.4–89.1	0–76	0–2.4	0–56
		2,230 (12.2)	84.5 (1.9)	63 (30.1)	0 (0)	12 (20.8)
Eriopersicon	_S. huaylasense_	978–3,304	80.7–90.2	25–67	0–3.5	10–28
		2,301 (18.8)	87.8 (1.1)	67 (31.3)	0 (0)	10 (85)
	S. corneliomulleri	1,018–3,097	64.6–87.4	25–80	0–4.0	4–32
		2,310 (17.0)	75.8 (3.9)	57 (16.6)	0 (0)	16 (43.7)
	S. peruvianum	2–3,191	40.5–87.8	25–94	0–21.6	2–32
		532, (128.0)	62.5 (15.3)	54 (21.2)	2.9 (55.1)	16 (31.2)
	S. chilense	0–3,995	41.6–87.2	30–96	0–21.6	1–32
		1,910 (57.9)	68.1 (10.1)	54 (24.0)	3.1 (59.6)	19 (31.5)
	S. habrochaites	40–3,692	50.1–91.0	0–94	0–4.3	0–46
		2,137 (30.6)	84.5 (5.1)	51 (32.3)	0 (0)	17 (44.1)
Neolycopersicon	_S. pennellii_	5–2,921	48.0–87.4	25–94	0–7.2	2–28
		831 (52.4)	68.5 (8.2)	54 (12)	2.9 (60.3)	16 (18.7)
Juglandifolia	_S. juglandifolium_	1,005–3,153	76.5–94.5	9–94	0–2.0	2–56
		2,195 (14.3)	89.2 (1.9)	40 (25)	0 (0)	13 (46.1)
	S. ochranthum	1,195–4,008	72.4–93.8	0–83	0–2.4	0–46
		2,742 (10.5)	85.8 (3.3)	60 (16.6)	0 (0)	12 (26.0)
Lycopersicoides	_S. lycopersicoides_	1,290–3,775	65.8–85.5	30–57	0–7.2	18–32
		2,960 (13.2)	73.1 (4.6)	50 (11)	3.1 (53.2)	20 (20)
	S. sitiens	2,276–3,330	67.9–71.6	43–69	0.3–7.2	12–28
		2,740 (5.7)	69.1 (1.2)	69 (8.6)	0.3 (1,150)	12 (25)

features seven species: _S. pennellii_, _S. neorickii_, _S. corneliomulleri_, _S. peruvianum_, _S. chilense_, _S. habrochaites_, and _S. pimpinellifolium_.

Finally, the region of high diversity on the border between Chile and Peru is home to five species: _S. pennellii_, _S. peruvianum_, _S. chilense_, _S. pimpinellifolium_, and _S. lycopersicoides_.

Cold spots correspond to the geographical distribution of _S. ochranthum_ and _S. juglandifolium_ accessions in Colombia and Ecuador, and _S. sitiens_ in the northern region of Chile. The rest of the areas of distribution are insignificant according to the statistical criteria, assuming a random distribution.
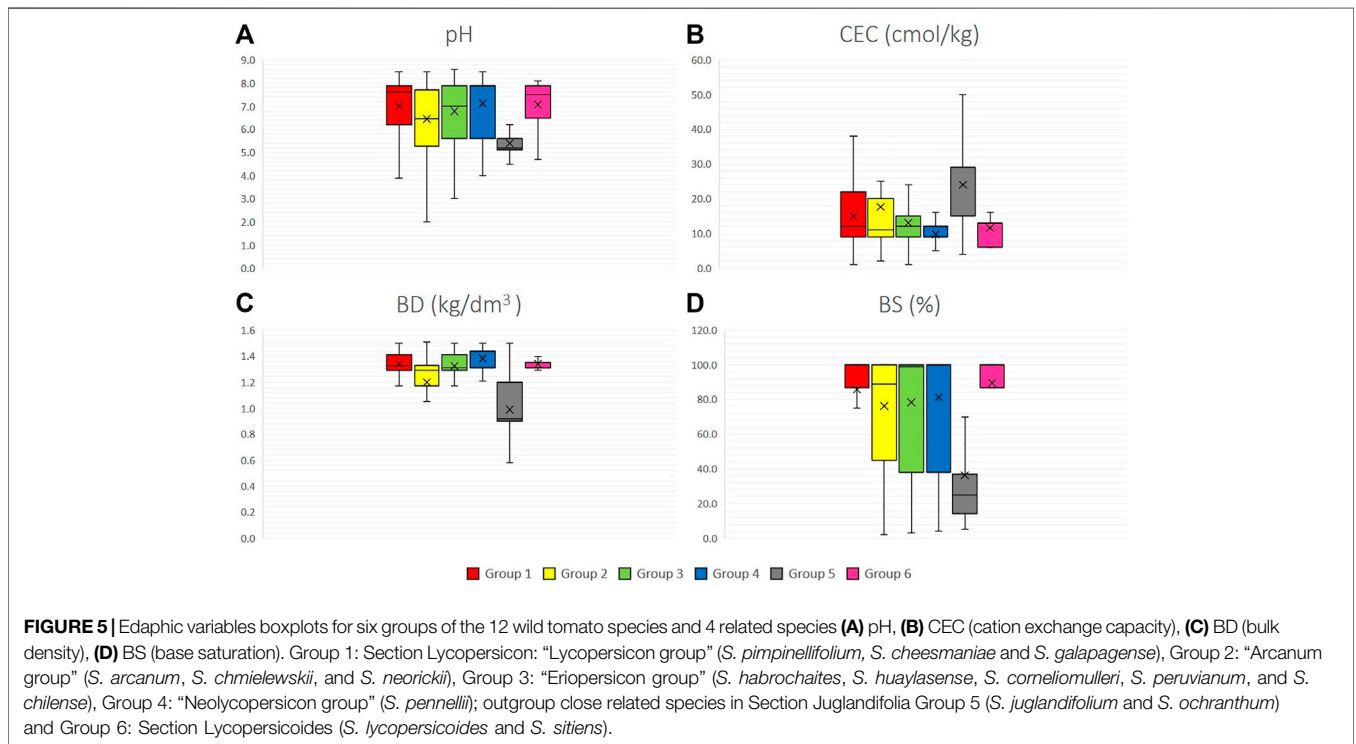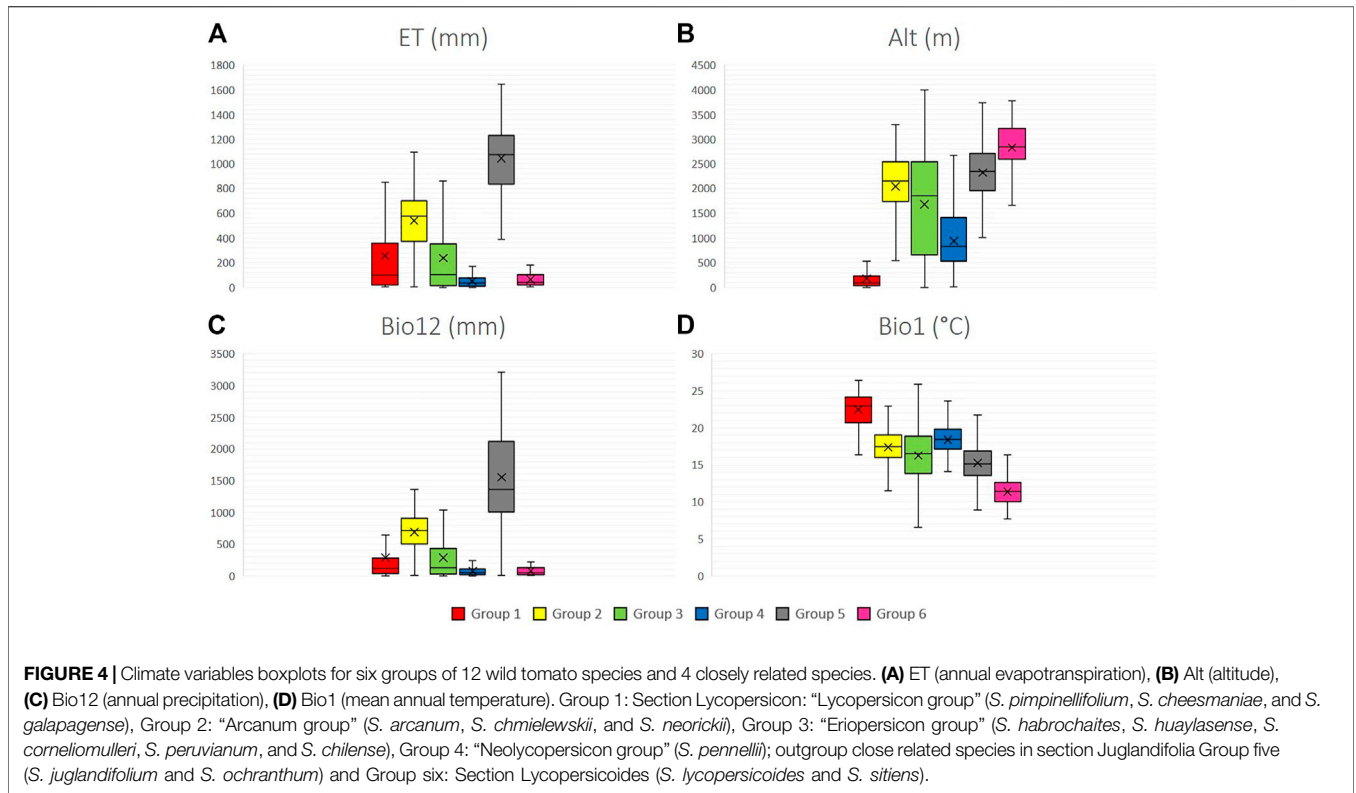
DISCUSSION

This research provides a relevant ecogeographic characterization to understand the distribution patterns of wild species that complement the phenotypic and genetic information. Characterization of genetic resources through environmental characteristics associated with accession areas and use of GIS

tools allows the identification of adaptive ranges and most relevant environmental factors affecting species distribution and ecological adaptation (Parra-Quinajo et al., 2012).
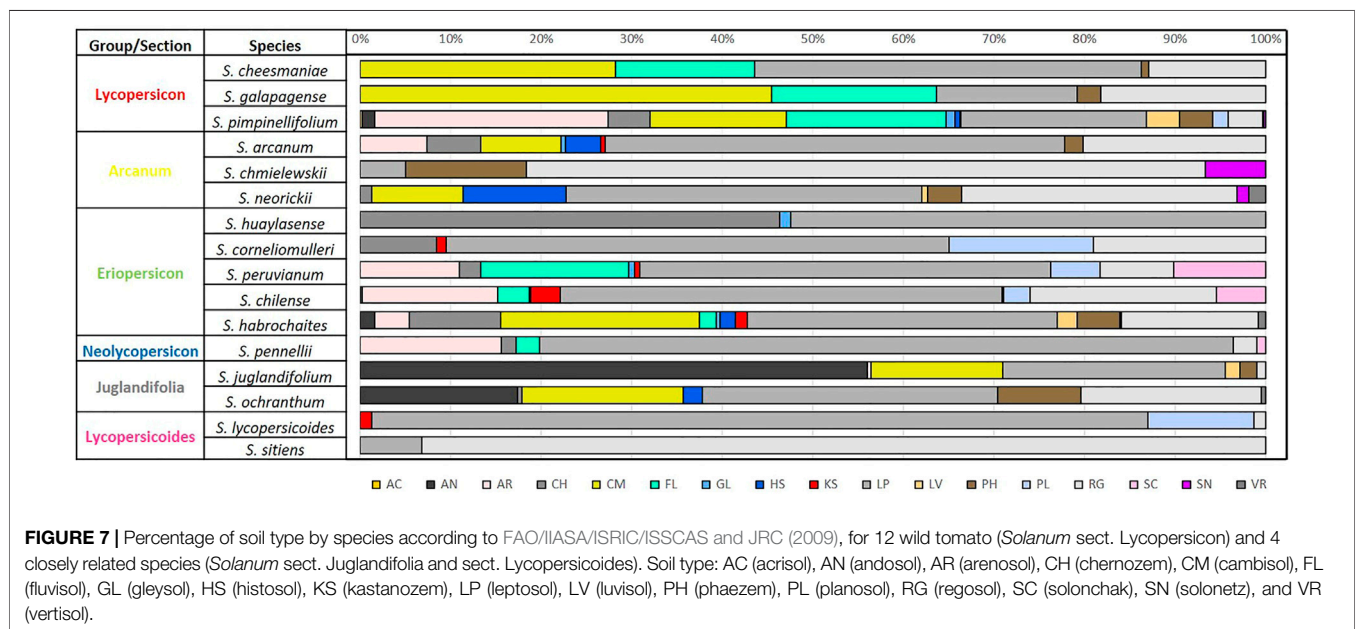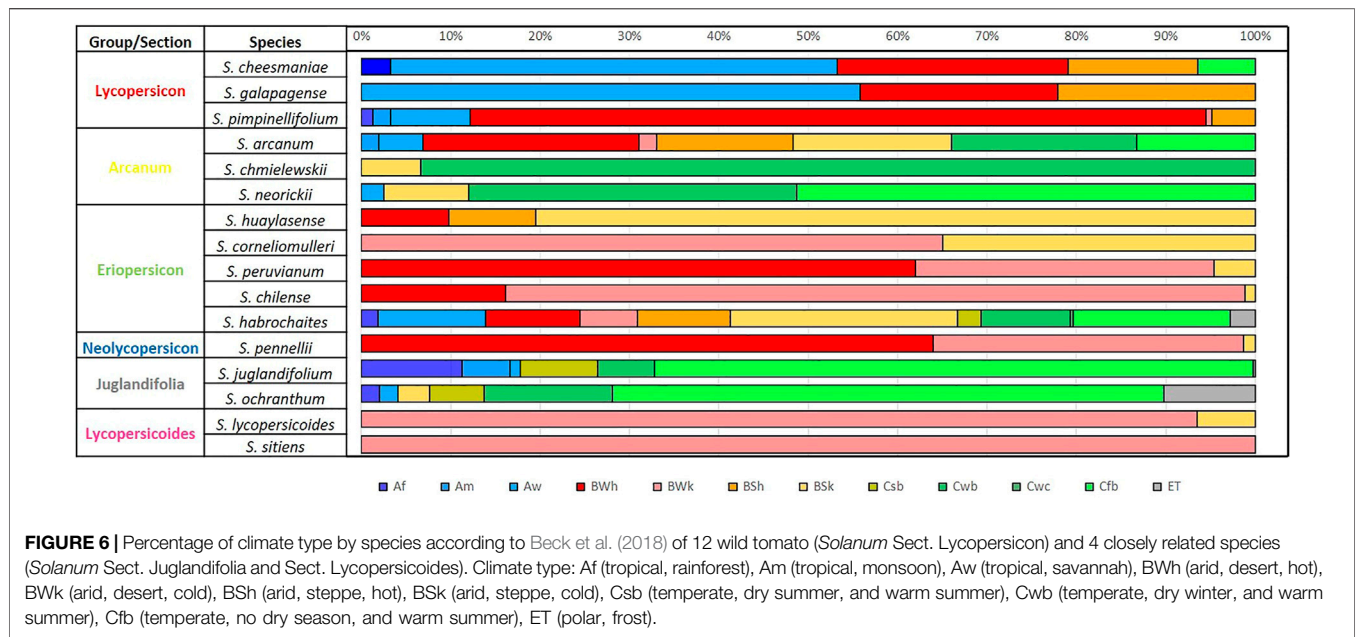
Likewise, through GIS and georeferenced information of species locations, it is possible to quantify geographical distances and distribution patterns of germplasm accession sites. From this perspective, it is likely to determine specific environmental conditions in which wild species and local varieties of crops have acquired their adaptive characters (Hijmans and Spooner, 2001). Therefore, the results obtained in this research constitute a source of updated and valuable information on the edaphoclimatic characteristics in which wild tomatoes and phylogenetically related species are distributed along its natural geographic range.

In general, geographical distribution of 16 wild species related to the cultivated tomato is wide, from Colombia through Peru, comprising Pacific coastal region to Chile and the Andean mountains, with an altitudinal range from sea level to 3,300 m (Peralta et al., 2008; Bergougnoux, 2014. However, within this distribution, there are overlapping areas between several species

FIGURE 4 | Climate variables boxplots for six groups of 12 wild tomato species and 4 closely related species. **(A)** ET (annual evapotranspiration), **(B)** Alt (altitude), **(C)** Bio12 (annual precipitation), **(D)** Bio1 (mean annual temperature). Group 1: Section Lycopersicon: "Lycopersicon group" (*S. pimpinellifolium*, *S. cheesmaniae*, and *S. galapagense*), Group 2: "Arcanum group" (*S. arcanum*, *S. chmielewskii*, and *S. neorickii*), Group 3: "Eriopersicon group" (*S. habrochaites*, *S. huaylasense*, *S. corneliomulleri*, *S. peruvianum*, and *S. chilense*), Group 4: "Neolycopersicon group" (*S. pennellii*); outgroup close related species in section Juglandifolia Group five (*S. juglandifolium* and *S. ochranthum*) and Group six: Section Lycopersicoides (*S. lycopersicoides* and *S. sitiens*).

FIGURE 5 | Edaphic variables boxplots for six groups of the 12 wild tomato species and 4 related species **(A)** pH, **(B)** CEC (cation exchange capacity), **(C)** BD (bulk density), **(D)** BS (base saturation). Group 1: Section Lycopersicon: "Lycopersicon group" (*S. pimpinellifolium, S. cheesmaniae* and *S. galapagense*), Group 2: "Arcanum group" (*S. arcanum*, *S. chmielewskii*, and *S. neorickii*), Group 3: "Eriopersicon group" (*S. habrochaites*, *S. huaylasense*, *S. corneliomulleri*, *S. peruvianum*, and *S. chilense*), Group 4: "Neolycopersicon group" (*S. pennellii*); outgroup close related species in Section Juglandifolia Group 5 (*S. juglandifolium* and *S. ochranthum*) and Group 6: Section Lycopersicoides (*S. lycopersicoides* and *S. sitiens*).

or regions with specific distribution such as the endemic species of the Galapagos Islands (*S. cheesmaniae* and *S. galapagense*) or hyper arid regions of northern Chile with other rare endemic

species, *S. sitiens*. Within these distribution patterns, it is also possible to identify differences and similarities between the species that conform each group, for example, the similarity

FIGURE 6 | Percentage of climate type by species according to Beck et al. (2018) of 12 wild tomato (*Solanum* Sect. Lycopersicon) and 4 closely related species (*Solanum* Sect. Juglandifolia and Sect. Lycopersicoides). Climate type: Af (tropical, rainforest), Am (tropical, monsoon), Aw (tropical, savannah), BWh (arid, desert, hot), BWk (arid, desert, cold), BSh (arid, steppe, hot), BSk (arid, steppe, cold), Csb (temperate, dry summer, and warm summer), Cwb (temperate, dry winter, and warm summer), Cwc (temperate, dry winter, and cold summer), Cfb (temperate, no dry season, and warm summer), ET (polar, frost).

FIGURE 7 | Percentage of soil type by species according to FAO/IIASA/ISRIC/ISSCAS and JRC (2009), for 12 wild tomato (*Solanum* sect. Lycopersicon) and 4 closely related species (*Solanum* sect. Juglandifolia and sect. Lycopersicoides). Soil type: AC (acrisol), AN (andosol), AR (arenosol), CH (chernozem), CM (cambisol), FL (fluvisol), GL (gleysol), HS (histosol), KS (kastanozem), LP (leptosol), LV (luvisol), PH (phaeozem), PL (planosol), RG (regosol), SC (solonchak), SN (solonetz), and VR (vertisol).

between *S. arcanum* and the species of Lycopersicon group (**Figure 6**), reflecting a wider distribution and adaptations to local sites of ecotypes (Peralta et al., 2008). These environmental characteristics reflect the ecological adaptation patterns and habitat preference of each species (Nakazato et al., 2010; Vilchez et al., 2019) (**Figures 6**, **7**, **Table A1 and A3 in Supplementary Material**). It is worth mentioning that these results also suggest a thorough revision of the proposed groups, incorporating the new passport data as well as genetic and molecular information to corroborate the belonging of each species to the phylogenetic assigned groups. The aforementioned

are under the assumption that the species are closely and genetically related and in expecting that their adaptation areas are similar.

Regarding wild tomato species and phylogenetically related species, few studies have been carried out with an ecogeographic or climatic focus. A comprehensive treatment integrates main botanical, biological, and ecological characteristics of each wild tomato and related species (Peralta et al., 2008); other studies focused on distribution of species richness and diversity through the analysis with GIS (González, 2013) and established conservation priorities (Vilchez et al., 2019); further

FIGURE 8 | Species diversity map hot spots (red) and cold spots (blue) for 12 wild tomato (*Solanum* Sect. Lycopersicon) and 4 closely related species (*Solanum* sect. Juglandifolia and sect. Lycopersicoides).
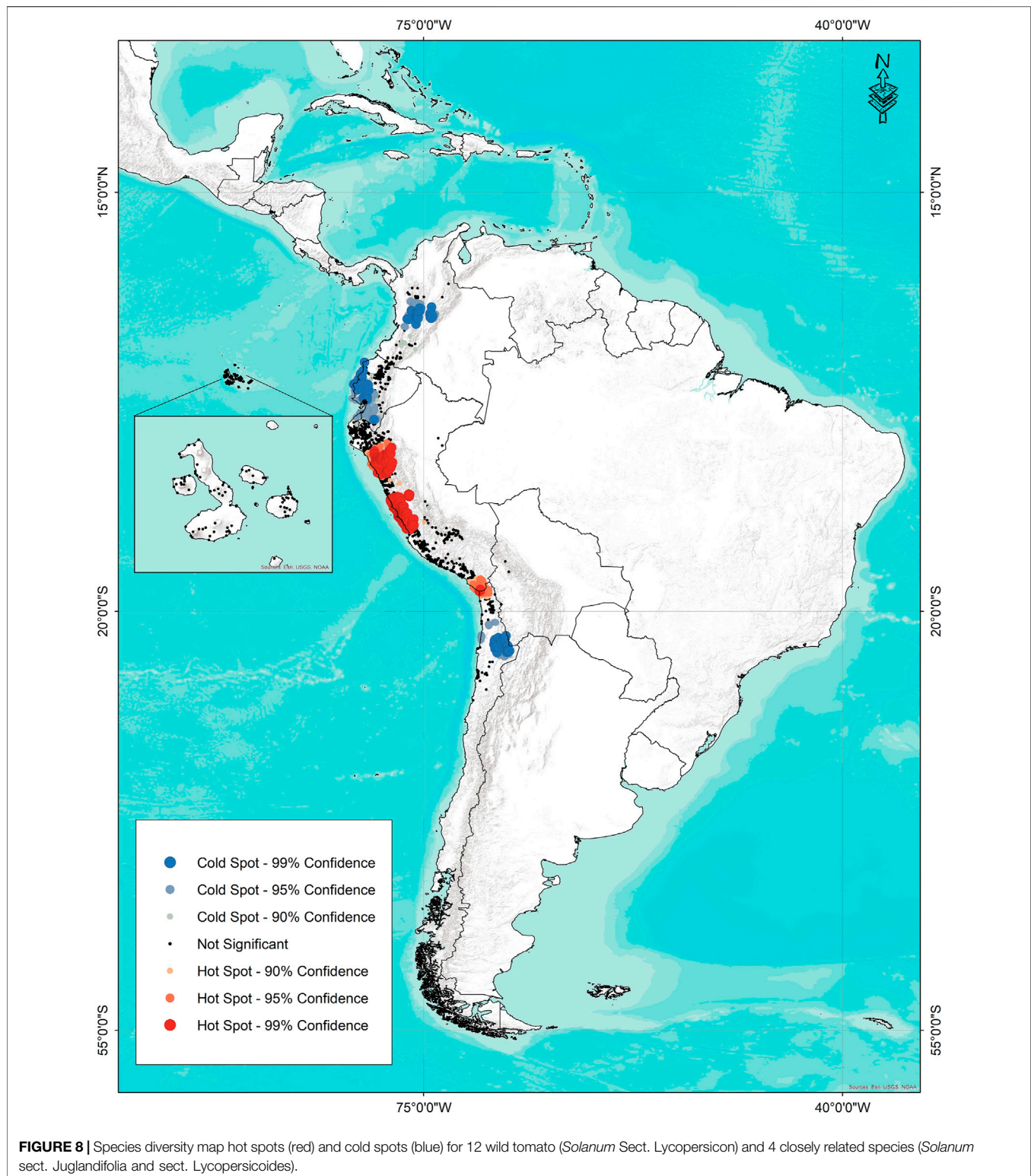
geographical and ecological characterization have been investigated in 10 tomato species determining soil and climate variables (Nakazato et al., 2010); studies have been conducted on tomato biogeography, *S. lycopersicum* var. *cerasiforme*, in its

center of origin and domestication (Délices et al. (2019); and finally climatic effects on species distribution (Lin et al., 2020) and bioclimatic characterization, and identification of ecological descriptors and patterns of climatic diversity of 12 wild

tomato and 4 closely related species (Ramírez-Ojeda et al., 2021a) have been studied. In this sense, this study complements the information available, providing information on soil characteristics that had not been analyzed.

The canonical correlation analysis satisfactorily identifies climatic variables with greatest influence on edaphic variables and vice versa, with a correlation of 0.80 representing 74% of total variation in 4,649 accessions. One main conclusion is that variables related to water availability (ET, Bio12, Bio14) have a great influence on physical (BD) and chemical soil characteristics (BS, pH). This pattern is persistent in all six groups. This relationship can be better observed in group 5 (*S. juglandifolium* and *S. ochranthum*) accessions with greater availability of annual precipitation and evapotranspiration, which present lower pH, BD, and BS than the rest of species; that is, they are located in soils with the lowest agricultural quality (**Figures 4**, **5**). This methodological approach is promising to be applied at other scales, considering the analysis at population level of each species and climatic and edaphic factors limited to smaller areas of distribution. This basis of ecogeographic characterization could incorporate information from genetic and ecological studies. A better understanding of these variables would allow the generation of projection models in different climate change scenarios (Violle and Jiang, 2009;Luebert and Weigend, 2014; Godoy-Bürki, 2016; Lin et al., 2020).

Ecological descriptors obtained, despite the incorporation of new accessions, are very similar to the ranges reported by Peralta et al. (2008) and Ramírez-Ojeda et al. (2021a) and generally identify the groups of species proposed in the classification. It is important to mention that this methodology has been widely used in the study of other species (Ruiz-Corral et al., 2008; Cerda-Hurtado et al., 2018; Sánchez-González et al., 2018; Ramírez-Ojeda et al., 2021a; 2021b). With this information, it is also possible to identify those species with tolerance to extreme conditions, for example, low and high temperatures, humidity conditions, altitude, pH, BD, and all the possible conditions when associating a species with a climate type or soil unit (**Table 3** and **Table 4**, Tables A1, A2, and A3 in **Supplementary Material**).

Edaphic diversity (**Figure 7**) tends to be more constant between species groups and sections with respect to climate diversity. In general, considering climate and soil characteristics, specific adaptation patterns for each species group can be identified: Lycopersicon group (group 1) corresponds to species with lower altitude and higher mean annual temperature; species of Juglandifolia section (group 5) are those with the highest water availability, lowest pH, BD, and base saturation; species of Lycopersicoides section (group 6) are the ones with the highest altitude, the lowest mean annual temperature, and lowest water availability, groups 4 and 6 have the lowest water availability and soils with favorable agricultural characteristics, differing by altitude. The rest of the species (groups 2 and 3) are in the transition zones with the rest of the wild tomato species. One aspect to highlight is that when combining or considering climatic and edaphic

information, it is possible to characterize in a better way the different groups, being able to better identify their differences and similarities.

Among possible uses of this approach is the identification of the germplasm with tolerance to adverse biotic and abiotic factors (Foolad and Lin., 2000; Mittova et al., 2004; Venema et al., 2005; Zhao et al., 2005; Ruiz-Corral et al., 2008; Chetelat et al., 2009; Arellano-Rodríguez et al., 2013; Ruiz-Corral et al., 2013; Cervantes-Moreno et al., 2014; Chen et al., 2015; Nosenko et al., 2016; Stam et al., 2017a; Stam et al., 2017b; Flores-Hernández et al., 2017; Razali et al., 2018; Dinh et al., 2019; Vilchez et al., 2019) with potential use for genetic breeding, identification of routes of germplasm accession, and areas of high and low diversity for use and conservation (Vilchez et al., 2019). In the information contained in **Table 3**, **Table 4**, and **Figures 4**, **5**, it is possible to identify species with extreme values that indicate tolerance or resistance to climatic and edaphic factors, with potential use as germplasm for genetic breeding.

Finally, the hot spot analysis could satisfactorily identify regions with the greatest diversity of species. These are priority areas for conservation, either due to high or low diversity. Regions identified as of great importance for conservation comprise endemism. Diversity contained in populations with few isolated individuals or with restricted distribution could be more affected by environmental and anthropic changes. This result could be explained by the quantity and geographic distance between the accessions of species studied. However, this first approximation is very useful and agrees with the diversity results obtained for wild potato species in Peru (Hijmans and Spooner, 2001).

This research determines the most important edaphoclimatic descriptors of wild tomato species and its closely related species along their natural geographic range in South America. Patterns of climatic diversity correlate with species groups and sections proposed in current classification. New edaphic characteristics analyzed in the same areas were also useful, although with less discrimination than the climatic variables. Interaction between climatic and edaphic factors allows for understanding species distribution and their adaptation patterns. Another feature to highlight is the incorporation of new data from recent collections of specimens being properly identified (Ministerio del Ambiente, 2020) that were not considered before in other studies, and thus expanding precision and reliability of these results. Most important areas for conservation of wild tomato species and related outgroups were detected. Under this premise, this contribution is promissory for further ecogeographic study of wild tomatoes and closely related species at the local population scale, especially focused *in situ* conservation reserves as well as in localities outside protected areas. Edaphoclimatic descriptors in addition with other abiotic or biotic factors could help to better estimate the species ecological niches and determine local ecotypes. Selected descriptors would be tested in models of current and future distribution considering the impact of climate change and anthropic activities along the distribution range of these valuable genetic resources.

Finally, this research can be used as a study model to replicate in other species.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Solanacea source: http://solanaceaesource.org/; Tomato Genetic Resource Center: https://tgrc.ucdavis.edu; Global Biodiversity Information Facility: https://www.gbif.org.

AUTHOR CONTRIBUTIONS

Conceptualization: GR-O, IP, and JR-P; methodology: GR-O, IP, and JR-P; software: GR-O and JR-P; validation: JR-P, JS-C, ER-G, IP, JC-S, TM-H, JR-V, and LV-N; formal analysis: GR-O, JR-P, and JS-C; data curation: IP, TM-H, JR-V, and LV-N; writing—original draft preparation: GR-O, JR-P, and IP; writing—review and editing: GR-O, JR-P, JS-C, ER-G, IP, JC-S, TM-H, JR-V, and LV-N. All authors have read and agreed to the published version of the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.748979/full#supplementary-material

REFERENCES

Aflitos, S., Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., et al. (2014). Exploring Genetic Variation in the Tomato (SolanumsectionLycopersicon) Clade by Whole-Genome Sequencing. *Plant J.* 80 (1), 136–148. doi:10.1111/tpj.12616

Arellano-Rodríguez, L. J., Rodríguez-Guzmán, E., Ron-Parra, J., Martínez-Ramírez, J. L., Lozoya-Saldaña, H., Sánchez-Martínez, J., et al. (2013). Evaluación de resistencia a *Phytophthora* en poblaciones silvestres de *Solanum lycopersicum* var. *cerasiforme*. *Rev. Mex. Cien. Agríc.* 4, 753–766. doi:10.29312/remexca.v4i5.1173

Balaguera-López, H. E., Álvarez-Herrera, J. G., Martínez-Arévalo, G. E., and Balaguera, W. A. (2009). Clay Content of Soil Influence on Yield of Tomato (*Solanum lycopersicum* L.) Crop. *Rev. Colomb. Cienc. Hort.* 3 (2), 199–209. doi:10.17584/rcch.2009v3i2.1213

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F. (2018). Present and Future Köppen-Geiger Climate Classification Maps at 1-km Resolution. *Sci. Data* 5, 180214. doi:10.1038/sdata.2018.214

Bergougnoux, V. (2014). The History of Tomato: from Domestication to Biopharming. *Biotechnol. Adv.* 32 (1), 170–189. doi:10.1016/j.biotechadv.2013.11.003

Burbano-Orjuela, H. (2016). El suelo y su relación con los servicios ecosistémicos y la seguridad alimentaria. *rcia* 33 (2), 117–124. doi:10.22267/rcia.163302.58

Cerda-Hurtado, I. M., Mayek-Pérez, N., Hernández-Delgado, S., Muruaga-Martínez, J. S., Reyes-Lara, M. A., Reyes-Valdés, M. H., et al. (2018). Climatic Adaptation and Ecological Descriptors of Wild Beans from Mexico. *Ecol. Evol.* 8 (13), 6492–6504. doi:10.1002/ece3.4106

Cervantes-Moreno, R., Rodríguez-Pérez, J. E., Rodríguez-Pérez, J. E., Carrillo-Fonseca, C., Sahagún-Castellanos, J., and Rodríguez-Guzmán, E. (2014). Tolerance of 26 Native Tomato Collections from Mexico to Nematode *Meloidogyne incognita* (KOFOID and WHITE) Chitwood. *rchsh* XX, 5–18. doi:10.5154/r.rchsh.2012.12.071

Chávez-Servia, J. L., Carrillo-Rodríguez, J. C., Vera-Guzmán, A. M., and Rodríguez-Guzmán, y. R. (2011). *Utilización actual y potencial del jitomate silvestre mexicano Subsistema Nacional de Recursos Fitogenéticos para la Alimentación y la Agricultura (SINAREFI), Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación*. México: CIIDIR-Unidad Oaxaca del Instituto Politécnico Nacional e Instituto Tecnológico del Valle de Oaxaca. Oaxaca.

Chen, H., Chen, X., Chen, D., Li, J., Zhang, Y., and Wang, A. (2015). A Comparison of the Low Temperature Transcriptomes of Two Tomato Genotypes that Differ in Freezing Tolerance: *Solanum lycopersicum* and *Solanum Habrochaites*. *BMC Plant Biol.* 15, 132. doi:10.1186/s12870-015-0521-6

Chetelat, R. T., Pertuzé, R. A., Faúndez, L., Graham, E. B., and Jones, C. M. (2009). Distribution, Ecology and Reproductive Biology of Wild Tomatoes and Related Nightshades from the Atacama Desert Region of Northern Chile. *Euphytica* 167, 77–93. doi:10.1007/s10681-008-9863-6

Córdova-Téllez, L., and Molina-Moreno, J. C. (2006). "Conservación *Ex Situ*," in *Recursos Fitogenéticos de México para la Alimentación y la Agricultura; Informe Nacional sobre el estado de los recursos fitogenéticos para la agricultura y la alimentaciónSecretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación y Sociedad Mexicana de Fitogenética, A.C: Chapingo, México*. Editors J C M. Moreno and L C Téllez.

del Ambiente, M. (2020). *Línea de base de la diversidad del tomate peruano con fines de bioseguridad*. Lima: Perú.

Délices, G., Otto, R., , O. R., Andrés Meza, P., Serna-Lagunez, R., Gamez Pastrana, R., et al. (2019). Biogeografía del tomate Solanum lycopersicum var. cerasiforme (Solanaceae) en su centro de origen (sur de América) y de domesticación (México). *Rbt* 67 (4), 1023–1036. doi:10.15517/rbt.v67i4.33754

Dinh, Q.-D., Dechesne, A., Furrer, H., Taylor, G., Visser, R. G. F., Harbinson, J., et al. (2019). High-Altitude Wild Species Solanum arcanum LA385-A Potential Source for Improvement of Plant Growth and Photosynthetic Performance at Suboptimal Temperatures. *Front. Plant Sci.* 10, 1163. doi:10.3389/fpls.2019.01163

FAO/IIASA/ISRIC/ISSCAS/JRC (2009). *Harmonized World Soil Database 1.1*. Rome, Italy.

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 (12), 4302–4315. doi:10.1002/joc.5086

Flores-Hernández, L. A., Lobato-Ortiz, R., García-Zavala, J. J., Molina-Galán, J. D., Sargerman-Jarquín, D. M., and Velasco-Alvarado, M. d. J. (2017). Parientes Silvestres Del Tomate Como Fuente De Germoplasma Para El Mejoramiento Genético De La Especie. *RevFitotecMex* 40 (1), 83–91. doi:10.35196/rfm.2017.1.83-91

Florido, M., Plana, D., Álvarez, M., Moya, C., Regla, M., and Caballero, A. (2009). Evaluación del germoplasma de tomate (Solanum L. sección Lycopersicon) conservado Ex situ en Cuba utilizando prueba In Vitro para predecir las respuestas a altas temperaturas. Cult. Trop. 30 (4), 57–61.

Foolad, M. R., and Lin, G. Y. (2000). Relationship between Cold Tolerance during Seed Germination and Vegetative Growth in Tomato: Germplasm Evaluation. jashs 125 (6), 679–683. doi:10.21273/JASHS.125.6.679

Fortuny-Fernández, N. M., Monserrat Ferrer, M., and Ruenes-Morales, M. d. R. (2017). Centros de origen, domesticación y diversidad genética de la ciruela mexicana, Spondias purpurea (Anacardiaceae). Acta Bot. Mex. 121, 7–38. doi:10.21829/abm121.2017.1289

GBIF (2021). Global Biodiversity Information Facility. . https://www.gbif.org/ (Accessed March 21, 2021).

Getis, A., and Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. Geographical Anal. 24 (3), 189–206. doi:10.1111/j.1538-4632.1992.tb00261.x

Godoy-Bürki, A. C. (2016). Efectos del cambio climático sobre especies de plantas vasculares del sur de los Andes Centrales: un estudio en el noroeste de Argentina (NOA). Ecología Austral 26 (1), 83–94. doi:10.25260/EA.16.26.2.0.11010.25260/ea.16.26.1.0.110

González, P. (2013). Geographical Distribution of Wild Tomatoes (Solanum L. Sect. Lycopersicon (Mill.) Wettst. Solanaceae). Arnaldoa 20 (2), 301–314.

Grandillo, S., Chetelat, R., Knapp, S., Spooner, D., Peralta, I., Cammareri, M., et al. (2011). "Solanum Sect. Lycopersicon," in Wild Crop Relatives: Genomic and Breeding Resources Vegetables. Editor K. Chittaranjan (Germany: SpringerBerlin/Heidelberg), 129–215. doi:10.1007/978-3-642-20450-0_9

Hernández-Bautista, A., Lobato-Ortiz, R., Cruz-Izquierdo, S., and García-Zavala, Y. J. L. (2014). Variación fenotípica, heterosis y heredabilidad en una cruza interespecífica de tomate. Interciencia 39 (5), 327–332.

Hijmans, R. J., and Spooner, D. M. (2001). Geographic Distribution of Wild Potato Species. Am. J. Bot. 88 (11), 2101–2112. doi:10.2307/3558435

Hijmans, R. J., Spooner, D. M., Salas, A. R., Guarino, L., and de la Cruz, J. (2002). Atlas of Wild Potatoes. Systematic and Ecogeographic Studies on Crop Genepools 10. Rome, Italy: International Plant Genetic Resources Institute.

FAO (2019). "The State of the World's Biodiversity for Food and Agriculture," in FAO Commission on Genetic Resources for Food and Agriculture Assessments. Editors J. Bélanger and D. Pilling (Rome. doi:10.4060/CA3129EN

Kantar, M. B., Sosa, C. C., Khoury, C. K., Castañeda-Álvarez, N. P., Achicanoy, H. A., Bernau, V., et al. (2015). Ecogeography and Utility to Plant Breeding of the Crop Wild Relatives of sunflower (Helianthus Annuus L). Front. Plant Sci. 6, 841. doi:10.3389/fpls.2015.00841

Lin, Y.-P., Lu, C.-Y., and Lee, C.-R. (2020). The Climatic Association of Population Divergence and Future Extinction Risk of Solanum Pimpinellifolium. AoB PLANTS 12, plaa012. doi:10.1093/aobpla/plaa012

Lobo-Burle, M., Torres-Cordeiro, C. M., Fonseca, J. R., de Melo, L. A. M. P., de Belem-das, R. N. A., and Abadie, T. (2013). Characterization of Germplasm According to Environmental Conditions at the Collecting Site Using GIS: Two Case Studies from Brazil. Plant Genet. Resour. Newsl. 135, 1–11.

Luebert, F., and Weigend, M. (2014). Phylogenetic Insights into Andean Plant Diversification. Front. Ecol. Evol. 2 (27), 1–17. doi:10.3389/fevo.2014.00027

Magallanes-López, A. M., Martínez-Damián, M. T., Sahagún-Castellanos, J., Pérez-Flores, L. J., Marín-Montes, I. M., and Rodríguez-Pérez, J. E. (2020). Calidad Poscosecha De 40 Poblaciones De Tomate (Solanum Lycopersicum L.) Nativas De México. agro 54 (6), 779–796. doi:10.47163/agrociencia.v54i6.2184

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., et al. (2006). Microbial Biogeography: Putting Microorganisms on the Map. Nat. Rev. Microbiol. 4 (2), 102–112. doi:10.1038/nrmicro1341

M. Causse, J. Giovannoni, M. Bouzayen, and M. Zouine (Editors) (2016). The Tomato Genome (Berlin, Germany: Springer). doi:10.1007/978-3-662-53389-5

Mittova, V., Guy, M., Tal, M., and Volokita, M. (2004). Salinity Up-Regulates the Antioxidative System in Root Mitochondria and Peroxisomes of the Wild Salt-Tolerant Tomato Species Lycopersicon Pennellii. J. Exp. Bot. 55 (399), 1105–1113. doi:10.1093/jxb/erh113

Nakazato, T., Warren, D. L., and Moyle, L. C. (2010). Ecological and Geographic Modes of Species Divergence in Wild Tomatoes. Am. J. Bot. 97 (4), 680–693. doi:10.3732/ajb.0900216

Nosenko, T., Böndel, K. B., Kumpfmüller, G., and Stephan, W. (2016). Adaptation to Low Temperatures in the Wild Tomato speciesSolanum Chilense. Mol. Ecol. 25 (12), 2853–2869. doi:10.1111/mec.13637

Parra-Quinajo, M., Iriondo, J. M., and Torres, E. (2012). Review. Applications of Ecogeography and Geographic Information Systems in Conservation and Utilization of Plant Genetic Resources. Span. J. Agric. Res. 10 (2), 419–429. doi:10.5424/sjar/2011102-303-11

Pease, J. B., Haak, D. C., Hahn, M. W., and Moyle, L. C. (2016). Phylogenomics Reveals Three Sources of Adaptive Variation during a Rapid Radiation. Plos Biol. 14 (2), e1002379. doi:10.1371/journal.pbio.1002379

Peralta, I. E., Spooner, D. M., and Knapp, S. (2008). Taxonomy of Wild Tomatoes and Their Relatives (Solanum Sect. Lycopersicoides, Sect. Juglandifolia, Sect. Lycopersicon; Solanaceae). Syst. Bot. Monogr. 84, 1–186.

Ramírez-Ojeda, G., Orozco-Gutiérrez, G., RuizSanchez, E., and Ruiz-Sánchez, E. (2021b). EDAPHOCLIMATIC DIVERSITY and ECOLOGICAL DESCRIPTORS of Guadua BAMBOO SPECIES (Poaceae: Bambusoideae) IN MEXICO. Ijaeb 06 (3), 228–244. doi:10.35410/IJAEB.2021.5641

Ramírez-Ojeda, G., Peralta, I. E., Rodríguez-Guzmán, E., Chávez-Servia, J. L., Sahagún-Castellanos, J., and Rodríguez-Pérez, J. E. (2021a). Climatic Diversity and Ecological Descriptors of Wild Tomato Species (Solanum Sect. Lycopersicon) and Close Related Species (Solanum Sect. Juglandifolia Y Sect. Lycopersicoides) in Latin America. Plants 10, 855. doi:10.3390/plants10050855

Razali, R., Bougouffa, S., Morton, M. J. L., Lightfoot, D. J., Alam, I., Essack, M., et al. (2018). The Genome Sequence of the Wild Tomato Solanum Pimpinellifolium Provides Insights into Salinity Tolerance. Front. Plant Sci. 9 (1402), 1–21. doi:10.3389/fpls.2018.01402

Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., et al. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. Mol. Biol. Evol. 37 (4), 1118–1132. doi:10.1093/molbev/msz297

Rodríguez, F., Wu, F., Ané, C., Tanksley, S., and Spooner, D. M. (2009). Do potatoes and Tomatoes Have a Single Evolutionary History, and what Proportion of the Genome Supports This History? BMC Evol. Biol. 9 (9), 191. doi:10.1186/1471-2148-9-191

Ruiz Corral, J. A., Durán Puga, N., Sánchez González, J. d. J., Ron Parra, J., González Eguiarte, D. R., Holland, J. B., et al. (2008). Climatic Adaptation and Ecological Descriptors of 42 Mexican Maize Races. Crop Sci. 48 (4), 1502–1512. doi:10.2135/cropsci2007.09.0518

Ruiz-Corral, J. A., Sánchez-González, J. J., Hernández-Casillas, J. M., Willcox, M. C., Ramírez-Ojeda, G., Ramírez-Díaz, J. L., et al. (2013). Identificación de razas de maíz adaptadas a condiciones deficientes de humedad mediante datos biogeográficos. Rev. Mex. Cienc. Agric. 4 (6), 829–842. doi:10.29312/remexca.v4i6.1152

Sánchez González, J. d. J., Ruiz Corral, J. A., García, G. M., Ojeda, G. R., Larios, L. D. l. C., Holland, J. B., et al. (2018). Ecogeography of Teosinte. PLoS ONE 13, e0192676.. doi:10.1371/journal.pone.019267

Sandoval-Ceballos, M. G., Kalungwana, N. A., Griffin, J. H. C., Martínez-Guerra, G., Ramírez-Ramírez, I., Maldonado-Peralta, R., et al. (2021). The Importance of Conserving Mexico's Tomato Agrodiversity to Research Plant Biochemistry under Different Climates. Plants People Planet., 1–7. doi:10.1002/ppp3.10218

SAS Institute (2011). SAS/STAT User's Guide: Software Version 9.3. Cary, NC, USA: Statistical Analysis System Institute.

Solanaceae Source (2021). A Global Taxonomic Source for the Nightshade Family. Available at: http://solanaceaesource.org/(Accessed March 18, 2021).

Sotomayor, D., Vilchez, D., and Zorrilla, C. (2019). Dataset on Occurrences of Lycopersicum Species from Peru. Figshare. Dataset. doi:10.6084/m9.figshare.9880073.v1

Spooner, D. M., Gavrilenko, T., Jansky, S. H., Ovchinnikova, A., Krylova, E., Knapp, S., et al. (2010). Ecogeography of Ploidy Variation in Cultivated Potato (Solanum Sect. Petota). Am. J. Bot. 97 (12), 2049–2060. doi:10.3732/ajb.1000277

Stam, R., Scheikl, D., and Tellier, A. (2017b). The Wild Tomato speciesSolanum Chileneshows Variation in Pathogen Resistance between Geographically Distinct Populations. PeerJ 5 (5), e2910. doi:10.7717/peerj.2910

Stam, R., Silva-Arias, G. A., Tellier, A., Scheikl, D., Hörger, A., Stephan, W., et al. (2017a). Subsets of NLR Genes Drive Adaptation of Tomato to Pathogens during Colonisation of New Habitats. bioRxiv, 210559. doi:10.1101/210559

Steiner, J. J., and Greene, S. L. (1996). Proposed Ecological Descriptors and Their Utility for Plant Germplasm Collections. *Crop Sci.* 36 (2), 439–451. doi:10.2135/cropsci1996.0011183X003600020037x

TGRC (2021). *Tomato Genetics Resource Center*. Available at: https://tgrc.ucdavis.edu/(Accessed March 18, 2021).

Tofalo, R., Perpetuini, G., Schirone, M., Fasoli, G., Aguzzi, I, Corsetti, A., et al. (2013). Biogeographical Characterization of *Saccharomyces cerevisiae* Wine Yeast by Molecular Methods. *Front. Microbiol.* 4 (166), 1–13. doi:10.3389/fmicb.2013.00166

Trabucco, A., and Zomer, R. J. (2019). Global High-Resolution Soil-Water Balance. *Figshare. Fileset.* doi:10.6084/m9.figshare.7707605.v3

Tropicos.org. Missouri Botanical Garden (2021). Available at: https://tropicos.org/home. (Accessed March 17, 2021).

Unep-Wcmc (2016). *El estado de la biodiversidad en América Latina y el Caribe*. Cambridge, Reino Unido: UNEP-WCMC.

Venema, J. H., Linger, P., Heusden, A. W., Hasselt, P. R., and Brüggemann, W. (2005). The Inheritance of Chilling Tolerance in Tomato (Lycopersiconspp). *Plant Biol.* 7 (2), 118–130. doi:10.1055/s-2005-837495

Vilchez, D., Sotomayor, D. A., , D., and Zorrilla, C. (2019). *Ex Situ* Conservation Priorities for the Peruvian Wild Tomato Species (*Solanum* L. Sect. Lycopersicum (Mill.) Wettst). *Ecol. Apl.* 18 (2), 171–183. doi:10.21704/rea.v18i2.1335

Violle, C., and Jiang, L. (2009). Towards a Trait-Based Quantification of Species Niche. *J. Plant Ecol.* 2 (2), 87–93. doi:10.1093/jpe/rtp007

Zhao, L., Qiu, C., Li, J., Chai, Y., Kai, G., Li, Z., et al. (2005). Investigation of Disease Resistance and Cold Tolerance of *Solanum Lycopersicoides* for Tomato Improvement. *HortSci* 40 (1), 43–46. doi:10.21273/HORTSCI.40.1.43

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.