

# Genetic and Metabolomic Differentiation of *Physalis Ixocarpa* Brot. Populations in Michoacan State, Mexico

**Ariana Martínez-Vega**

Instituto Tecnológico de Mexico

**Ernesto Oregel-Zamudio**

Instituto Politecnico Nacional

**Ignacio García**

Instituto Politecnico Nacional

**Vinicio Villalpando-Arteaga**

Instituto Tecnológico de Mexico

**Jesus Torres** (✉ [jrtorresg@ipn.mx](mailto:jrtorresg@ipn.mx))

Instituto Politécnico Nacional: Instituto Politecnico Nacional <https://orcid.org/0000-0002-8643-3696>

---

## Research Article

**Keywords:** domestication, metabolomic fingerprinting, tomate milpero, Chapala, GC-MS

**Posted Date:** June 10th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-585240/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

*Physalis ixocarpa* Brot. is a native species that is consumed in many localities of the Cienega-Chapala in Mexico's Michoacan state. These fruits are cultivated and collected into traditional maize crops. The fruits are similar to *P. Philadelphica*, but the differences are in the fruit size and organoleptic properties (flavor, sweetness). According to antecedents of domestication that this zone represents in Mexico, is possible that *P. ixocarpa* shows incipient differentiation signals in genetic structure and metabolomic fingerprinting. Our objective was find evidences of genetic and metabolomic differentiation among populations of *P. ixocarpa* in the Cienega-Chapala. We used the sequencing of the chloroplast intergenic sequences *psbJ – petA* and *trnL – rpL32*, and the metabolomic fingerprinting by GC-MS. The results showed that exist genetic differentiation ( $F_{ST}$ ) and signatures of selection (Fu's  $F_s$ ' neutrality test) among populations. Moreover the metabolomic fingerprinting showed differences among populations and an increase of aldehydes, aromatic aldehydes, ester, and alcohols related with organoleptic properties of *P. ixocarpa*. We conclude that *P. ixocarpa* is an important genetic resource with signatures of differentiation in the Cienega-Chapala, Michoacan state, Mexico that eventually could be related with domestication signatures.

## Introduction

Mexico is the center of diversity of the genus *Physalis* with around 50 native species (Zamora-Tavares et al. 2015). The Cienega-Chapala region that belongs to the Balsas Basin is an important center of domestication of species where some *Physalis* species produce edible fruits and are consumed by local people (Vargas-Ponce et al. 2016). In this region, some species of *Physalis* are cultivated inside of cornfield (milpa system), for this reason, the fruits are knowledge as "tomate milpero". One of this species that are harvested is *Physalis ixocarpa* Brot. The form of fruits is similar to domesticated husk tomatoe (*P. philadelphica*), but the principal difference between species is their size. *P. philadelphica* fruits are around five-folds bigger than *P. ixocarpa*. Thus, some authors have proposed that *P. ixocarpa* is the same specie or is the wild form of actually cultivated husk tomatoes (Morales-Contreras 2018; Ramos-López and Morales 2018; Ayala-Armenta et al. 2020). Local people highly value *P. ixocarpa* fruits because they have a better flavor than commercial husk tomatoes. This species' kilogram has a price four-folds in the markets with respect to cultivated tomatoes (Torres-García, personal observation).

*P. ixocarpa* fruits' production is not extensive. They are collected inside the cornfield and grew spontaneously (Vargas-Ponce et al. 2016). The farmers avoid eliminating these plants in the crop, leaving the fructification, physiological maturity, and seed dispersion to increase this species' seed-soil bank. For this reason, there are no formal genetic improvement programs for the development of new varieties.

However, the zone has antecedents of domestication of another important species as bean (*Phaseolus vulgaris* L.), corn (*Zea mays* L.), and squash (*Cucurbita argyrosperma* Huber) (Zizumbo-Villarreal et al. 2005; Zizumbo-Villarreal and Colunga-GarcíaMarín 2010). For these reasons, is possible that *P. ixocarpa* present signatures of selection also in some traits of human interest (domestication process), as the

organoleptic properties. These domestication processes can modify the genetic structure and metabolite profile of fruits of the populations in the same way that moderns genetic improvement programs (Gross and Olsen 2010; Moyers et al. 2018).

For studying the domestication processes, genetic analyses (sequencing, molecular markers, genomics, among others) have been used in many species with reliable results (Gross and Olsen 2010). However, using other scopes to make phenotyping has been delayed with respect to genetic scopes (García-Flores et al. 2015). The use of non-target metabolomics to elaborate metabolic fingerprinting allows determinate a higher number of traits with high levels of confidence (Sumner et al. 2003; Messerli et al. 2017; Alseekh and Fernie 2018). The use of these tools can help detect differentiation signaling in plants (Macel et al. 2010; Saito and Matsuda 2010).

For this reason, this study hypothesized that in the region of the Cienega-Chapala region, it would find evidence of differentiation in genetic population and metabolomic fingerprint in *P. ixocarpa* populations caused by the farmer's management. The objectives of this study were: 1) Determinate the genetic differentiation of *P. ixocarpa* populations using chloroplast intergenic sequences (*psbJ – petA* and *trnL – rpl32*), and 2) determinate the metabolomic differentiation using metabolomic fingerprinting by GC-MS.

## Methods

### Study Area

This study was carried in the region known as the extreme northwest Balsas-Jalisco region in Mexico's Michoacan state. This region is an area with a complex topography due to the Western Sierra Madre, where many climatic conditions can be found in small distances. The annual mean temperature ranges are around 14 to 18°C, an altitude of 1500 to 1700 mamsl, and annual precipitation of 400 to 800 mm.

The agricultural activities in this region are influenced by Chapala lake. Chapala is the biggest lake in Mexico, with an extension of 1100 km<sup>2</sup>. However, this lake has reduced its extension caused by natural and human events (as dams' construction). Such events reduced the lake extension leaving new cultivable areas where agricultural areas have been established. Actually, this zone is one of the principal zones for producing berries and avocados in the world.

### Collection of Biological Material

The collection of the biological samples used in this study was obtained in the itinerant-local markets of the region named "tianguis". In such markets, the local producer sells their products every week. Six local markets were visited; each market is ubicated in the respective municipality center. The municipalities sampled in this study were Chavinda, Jacona, Tangamandapio, Villamar, Sahuayo, and La Barca (Fig. 1). We used the domesticated husk tomatoe (*P. philadelphica*) as a reference group.

In each market, local producers were localized and asked for their husk tomatoes' origin. The husk tomatoes were bought from farmers that cultivated and collected their own *P. ixocarpa* fruits. In La Barca's market, the seller can not identify their husk tomatoes' origin. La Barca is one of the bigger cities of the region, and the principal economic activity is intensive agriculture. However, this municipality was included because this is the east limit of the Cienega-Chapala region.

The samples were washed two times with water and soap and washed three times with distillate water. After, samples separated considering each fruit as and different individual and were lyophilized in a vacuum chamber at  $-50^{\circ}\text{C}$  for 72 h. Lyophilized samples were stored hermetically until their utilization.

### **Amplification by PCR of Intergenic Chloroplast Regions *psbJ – petA* and *trnL – rpl32***

Individual lyophilized samples were ground and used for DNA extraction using the CTAB method (Chen and Ronald 1999). As external groups, four genus species were used *P. angulata*, *P. lagascae*, *P. peruviana*, and *P. pruinosa*. Such species were obtained from the herbarium of CIIDIR-IPN, Unidad Michoacán (Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional). The quality of extraction was corroborated on 1% agarose gel in 1 % TBA buffer at 76 V for 40 min.

For the amplification of intergenic chloroplast regions, five individual of each population were taken and amplified with a set of primers designed by Shaw et al. (2005). These primers are universal for whatever plant species; besides, chloroplast regions' use establishes in a better way the phylogenetic relations among populations (Shaw et al. 2005). The two amplified regions correspond to *psbJ – petA*, and *trnL – rpl32*.

The PCR mix contained 1x of the buffer, 10 mM of dNTPs, 1  $\mu\text{L}$  of DNA, 1.5 mM of  $\text{MgCl}_2$ , 1 U of Taq (Go Taq® Flexi DNA polymerase; Promega), and 10 mM of each primer in a final volume of 15  $\mu\text{L}$ . The amplification conditions were as follows: denaturation at  $80^{\circ}\text{C}$  for 5 min followed by 30 cycles of  $95^{\circ}\text{C}$  for 1 m,  $50\text{--}65^{\circ}\text{C}$  (gradient of  $0.25^{\circ}\text{C}$  each 1 s) for 1 min, and  $72^{\circ}\text{C}$  for 4 min, followed by a final extension at  $72^{\circ}\text{C}$  for 5 min. The amplicons were sent to the Macrogen, Korea for sequencing.

### **Sequence Analyses**

Sequence analysis was performed using the software MEGA (Molecular Evolutionary Genetic Analysis) version 7 (Kumar et al. 2016). We confirmed the quality and effective size of the sequences included in the alignment before analysis. The sequences were aligned by each intergenic region, after both alignments were linked, indicating each fragment's start and end site. Each region's alignment included sequences *P. angulata*, *P. lagascae*, *P. peruviana*, and *P. pruinosa* were used as external groups. A model test for finding the best evolutive model was calculated.

Genetic diversity ( $\pi$ ) was quantified within populations based on the number of mutations. Additionally, Fu's  $F_s$ ' neutrality test was calculated for each population. We evaluated the genetic distance between populations ( $F_{ST}$ ) with a test performed with Arlequin Software v.3 (Excoffier et al. 2005). using 100 000

steps in the Markov chain and 1000 dememorization steps with a significance level of 0.05. To establish whether the genetic distance between accessions is independent of geographic location, we performed a Mantel test with 1000 permutations.

### **Phylogenetic Reconstruction**

Phylogenetic reconstruction of the different populations was made using two approaches. The first method used was constructing a phylogenetic tree using 1000 bootstrap replicates under the maximum likelihood method, gamma distribution, and invariant sites under the evolutive model of Tamura-3 parameters. Additionally, due to genetic differentiation being a continuum, and to draw the possible expansion of *P. ixocarpa* populations, a haplotype network was estimated. This network was constructed in the PopArt software (Leigh and Bryant 2015) using the TCS method (Clement et al. 2000).

### **Metabolomic Fingerprinting Using GC-MS**

Fifty milligrams of individual samples lyophilized and ground tissue were placed in amber vials (five individuals for each population). Vials were incubated at 80°C for one h. After this time, solid-phase microextraction (SPME) was used, inserting the fiber holder into vials and awaiting 10 minutes for each sample. Samples were injected into a gas chromatograph (Clarus 680, Perkin-Elmer Inc., Waltham, MA, USA), equipped with a phase capillary column: 5% diphenyl 95% dimethylpolysiloxane 30 m long, 0.32 mm i.d., 0.25  $\mu\text{m}$  film thickness, temperature limits between - 60 a 320/350°C (Elite-5 MS, Perkin-Elmer Inc., Waltham, MA, USA). Helium gas was used at a flow rate of 1 mL min<sup>-1</sup>, the flow remained constant, and there was an initial wait time of 0.5 min. The column temperature was initially maintained at 50°C for 1 min and then ramped to 250°C at 30°C/min, remaining at this temperature for a further 10 min. The temperature of the injector was 230°C. A mass spectrometer (Clarus SQ8T, Perkin-Elmer Inc., Waltham, MA, USA), with an electron impact ionization source (70 eV) in full scan mode was used. The analysis range was 40–500 *m/z*. The transfer line and ionization source temperatures were 230 and 250°C, respectively.

The feature detection, retention time correction, and peak alignment of the original chromatograms were made in XCMS Online (<https://xcmsonline.scripps.edu>) (Tautenhahn et al. 2012). To avoid false positives in the detection of metabolites, we only used metabolites with *q*-values  $\leq$  of 0.05. A Principal Component Analysis (PCA) was made to select the metabolites with significant participation in the metabolomic differentiation. The annotation of the principal metabolites was made using the NIST library using a cut-off value of 0.8.

The results were represented in a heatmap-bicluster. An ion matrix was constructed using the metabolites with the highest differentiation levels. The heatmap construction was made using the platform Metaboanalyst ([www.metaboanalyst.ca](http://www.metaboanalyst.ca)) (Chong et al. 2018). For the heatmap, the data were normalized and auto-scaled. The dendrograms used the Minkowski correlation as a distance function and the Ward clustering algorithm; the branches' significance was  $p \leq 0.05$ .

# Results

## Genetic Differentiation Among Populations

Each of the intergenic regions amplified by PCR had an extension of around 550 pb. The final alignment had 1023 sites. The number of transversions was higher than transitions in most of the population, except in Santiago (Table 1). The population from La Barca showed the highest number of variable sites (59); on the other hand, Jacona showed the least number (eight). The nucleotidic diversity had the same tendency as the number of variable sites, where La Barca showed the highest diversity and Jacona the lowest. The external group used, *P. philadelphica*, had 14 variables sites and one of the most reduced nucleotidic diversity.

Fu's  $F_s$  estimator reveals the type and the intensity of selection. Positive values indicate recent population bottlenecks or from overdominant selection, and negative values indicate increasing genetic diversity. In this case, these populations grow under human selection; for this reason, positive values can be considered as human selection. The neutrality tests showed that all populations studied had positive values. However, the intensity of selection was different in each population. The lowest values were observed in *P. philadelphica* (0.89) and Chavinda (1.04). The highest values were observed in Villamar population (9.32).

## Genetic Differentiation

The genetic differentiation measured using  $F_{ST}$  showed the conformation of seven groups (Table 2). The populations conformed to the first group from Chavinda, Jacona, Sahuayo, and Santiago. A second group was conformed by Jacona and Chavinda, the third by Sahuayo and Chavinda, the fourth by Santiago, Chavinda, Sahuayo. Villamar, La Barca, and commercial husk tomatoe each population was considered independent. The isolation by distance analysis showed that not exist a correlation between the geographic distance (km) and genetic differentiation ( $F_{ST}$ ) ( $R^2 = 0.01944$ ).

## Phylogenetic Reconstruction

The phylogenetic tree shows that three species used as external groups were grouped in the tree's basal branch (Fig. 2). Such species were *P. angulata*, *P. lagascae*, *P. peruviana*, and *P. pruinosa*. The most ancestral *P. ixocarpa* were Sahuayo and Villamar, followed by Chavinda, Jacona, and Santiago. However, in the recent branches, the species *P. lagascae* was arranged between the *P. ixocarpa* from La Barca and the commercial husk tomatoes (*P. philadelphica*). This result shows that the *tomates milperos* sold in La Barca is another species different from *P. ixocarpa*.

In the haplotype network, the ancestry of populations shows that Chavinda is the network center (Fig. 3). Similarly, as the phylogenetic tree, commercial husk tomatoe and La Barca's accessions appear in the network's extremes.

## Metabolomic Fingerprinting

The metabolomic fingerprinting of the different accessions of tomatoes milperos using GC-MS showed detection of 552 metabolites with ranges of q-values  $\leq 0.05$ . The PCA reduced the number of metabolites at 34 in the two first components. Such principal components explained the 80% of total variation among samples. In the Upper side of the heat map, the dendrogram shows the population's grouping considering the 34 most important metabolites; such dendrogram has two principal branches (Fig. 4). La Barca, Sahuayo, and Villamar were grouped in the left branch. Inside of this branch, two sub-branches can be observed, Sahuayo and Villamar are considered sister groups, while the La Barca population was excluded.

The second branch (right side of the heatmap) contains two sub-branches, one of which is conformed by Santiago and Chavinda; both populations did not show statistical differences in the metabolomic fingerprint. The other sub-branches are formed for Jacona, and the most divergent population in this sub-branch are the samples that correspond to commercial husk tomatoes (*P. philadelphica*).

The heatmap's left side shows the grouping dendrogram of metabolites; this dendrogram is divided into two principal branches. The first branch is located in the n two principal branches. The first branch is located in the upper section of the heatmap, showed an increase in the expression of the upper section of the heatmap showed an increase in the expression of 11 metabolites. The samples from populations Santiago, Chavinda, and Jacona showed an increase in such metabolites' expression. On the other hand, the samples from La Barca showed a significant reduction in the expression. The *P. philadelphica* fingerprint showed intermediate values with respect to other populations.

The second branch (the downside of the dendrogram) showed an inverse tendency in metabolites' expression. While La Barca, Sahuayo, and Villamar show a significant increase in such metabolites' expression, *P. philadelphica* showed intermediate expression levels. Otherwise, the populations from Santiago, Chavinda, and Jacona showed a reduction in expression with respect to other populations.

## Discussion

The populations of *P. ixocarpa* in the region Cienega-Chapala in Michoacan state, Mexico, have evidence of genetic and metabolomic differentiation. These patterns can be explained by the continuum selection process driven by the region's farmers. The characteristics under selection are the organoleptic traits as the flavor, acidity, and sweetness of the fruits. This selection process on this species was independent of the domestication process' of *Physalis philadelphica* and *Physalis lagascae* that are other species consumed in central Mexico. This study provides new evidence for studying a poorly studied species, representing an important genetic resource.

*P. ixocarpa* populations showed a high degree of genetic differentiation due to the high level of genetic diversity in this zone's species. The region Cienega Chapala belongs to the Balsas River Basin, which is considered the center of origin and domestication of many species with high importance in the actuality like *Zea mays* L., *Cucurbita argyrosperma* Huber, *Phaseolus vulgaris* L., among others (Casas et al. 2007; Zizumbo-Villarreal and Colunga-GarcíaMarín 2010; Zizumbo-Villarreal et al. 2005). Moreover, the

molecular markers used (intergenic chloroplast regions *psbJ* – *petA* and *trnL* – *rpL32*) are considered variables highly, for this reason are used for phylogenetic studies (Shaw et al. 2005).

The number of transversions was higher than the number of transitions in most populations studied. This amount and type of changes are unusual because the transversions are evolutive changes less frequent than transitions. However, the intergenic regions are highly variable genomic regions that are not involved in fundamental biological processes and are neutral to selection (Kelchner 2012; Shaw et al. 2005). This high number of mutations was responsive to the higher values of genetic diversity observed due to the calculation of nucleotidic diversity ( $\pi$ ) resulting from the number of changes/number of total sites (Casillas and Barbadilla 2017).

The estimator of neutrality used (Fu's  $F_s$ ) showed differences in selection intensity, but not in the type of selection. Fu's  $F_s$  predicted recent population bottlenecks in all populations. Due that the selection of traits by farmers causes the population bottlenecks, we assume that these bottlenecks are caused by the selection process (Arias et al. 2010).

The intensity of selection was more remarkable in some populations of tomate milpero than the husk tomato (*P. philadelphica*) variety evaluated. The cause of the major impact of human selection on *P. ixocarpa* can be caused by the reduced effective population size and diversity management. Because each corn plot can be considered an independent population, and the number of individuals is reduced. Moreover, each corn plot's agricultural practices are each farmer's decision (Vargas-Ponce et al. 2016). Thus can give an idea about the high variation levels and distinct domestication degrees that can be found in this area.

The phylogenetic reconstruction using two approaches give different results. In the phylogenetic tree, the species used as external groups were separated from the *P. ixocarpa* populations as was expected. However, in *P. lagascae* this sample was grouped between *P. ixocarpa* populations. Nerveless in the case of La Barca's population, the samples were collected in the local market and not directly with the farmers. This opens the possibility that the samples correspond to another specie as *Physalis angulata* or small fruits of *P. philadelphica*. In some Jalisco regions, the consumption of *Physalis angulata* also named tomate milpero.<sup>2</sup> Another critical result of the phylogenetic trees is that *P. ixocarpa* is not the ancestor of the actual husk tomatoes (*P. philadelphica*), according to the distribution of the branch, the ancestor of *P. philadelphica* may be *P. lagascae*.

The haplotype network showed different results to the phylogenetic tree. Chavinda was ubicated in the center of the network; this indicates that it is probably that in this municipality is possible to find the most conserved genetics of populations sampled in this study. Fu's  $F_s$  result showed that this population had the lowest value with respect to other populations. This means that the selection process has not been so intense for this population. The individuals from La Barca and commercial husk tomato (*P. philadelphica*) were ubicated in the network's extremes; this confirms the grouping estimated by  $F_{st}$  and the

phylogenetic tree. The  $F_{ST}$  showed that the La Barca population did not have any genetic relationship with the other populations.

The genetic analyses demonstrated the evidence of population differentiation, probably caused by human selection. Besides, the metabolomic fingerprinting gave evidence for the traits under selection. The heatmap showed differences among populations and confirmed the differences between La Barca and *P. philadelphica* with respect to the rest of the population. The non-target metabolomics applications for the classification of food samples have been well documented (Gamboa-Becerra 2017; García-Flores et al. 2015). Thus complement the genetic scope for the identification of the differentiation process.

The dendrogram of the left side of the heatmap showed two principal branches. The metabolites present in the upper branch showed marked differences among populations; these metabolites were aldehydes, aromatic aldehydes, ester, and alcohol. These compounds are responsible of some organoleptic properties as flavor. The heatmap showed that fruits from Jacona, Chavinda, and Villamar showed over-expression of such metabolites. On the other hand, La Barca showed a down-expression in the same metabolites, and *P. philadelphica* showed intermediate values.

Local people recognize organoleptic differences between the tomatoes milperos and commercial husk tomatoes. The metabolites upper expressed in the population Jacona, and the evidence of selection confirm the differentiation process occurring in the Balsas Balsin. This region has been recognized as an important domestication center in Mexico (Arias et al. 2010; Casas et al. 2007; Zizumbo-Villarreal and Colunga-GarcíaMarín 2010; Zizumbo-Villarreal et al. 2005).

This study contributes to knowledge of a poorly studied species with a high potential to approaching. The evidence showed that tomatoes milperos (*P. ixocarpa*) cultivated in the region Cienega-Chapala, Michoacan State, Mexico, been in a continuum selection process. This differentiation process has changed some organoleptic properties, consumers of the region more appreciate these characteristics.

## Declarations

**Funding** : Consejo Nacional de Ciencia y Tecnología (CONACyT), Mexico

**Conflicts of interest/Competing interests** : Not applicable

**Availability of data and material** : Available

**Code availability**: Not applicable

## References

1. Alseekh S, Fernie AR (2018) Metabolomics 20 Years on: What Have We Learned and What Hurdles Remain? Plant J 94(6):933–942. DOI:10.1111/tpj.13950

2. Arias D, Peñaloza-Ramírez J, Dorado O, Cuevas-Reyes P, Leyva E, Albarrán-Lara AL, Rangel-Altamirano G (2010) Phylogeographic Patterns and Possible Incipient Domestication of *Jacaratia mexicana* A. DC. (Caricaceae) in Mexico. *Genetic Resource Crop Evolution* 57:1227–1238. DOI:10.1007/s10722-010-9569-1
3. Ayala-Armenta QA, Tovar-Pedraza JM, Apodaca-Sánchez MA, Correia KC, Saucedo-Acosta CP, Camacho-Tapia M, Beltrán-Peña H (2020) Phylogeny and Pathogenicity of Soilborne Fungi Associated with Wilt Disease Complex of Tomatillo (*Physalis ixocarpa*) in Northern Sinaloa, Mexico. *European Journal of Plant Pathology* 157, 733–749 DOI: 0.1007/s10658-020-02030-9
4. Casas A, Otero-Arnaiz A, Pérez-Negrón E, Valiente-Banuet A (2007) In Situ Management and Domestication of Plants in Mesoamerica. *Ann Bot* 100:1101–1115. DOI:10.1093/aob/mcm126
5. Casillas S, Barbadilla A (2017) Molecular Population Genetics. *Genetics* 205:1003–1035. DOI:10.1534/genetics.116.196493
6. Chen DH, Ronald PC (1999) A Rapid DNA Minipreparation Method Suitable for AFLP and Other PCR Applications. *Plant Molecular Biology Report* 17(1):53–57. DOI:10.1023/A:1007585532036
7. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque SG, Wishart DS, Xia J (2018) MetaboAnalyst 4.0: Towards More Transparent and Integrative Metabolomics Analysis. *Nucleic Acids Res* 46(W1):W486–W494. DOI:10.1093/nar/gky310
8. Clement M, Posada D, Crandall KA (2000) TCS: A Computer Program to Estimate Gene Genealogies TCS : A Computer Program to Estimate Gene Genealogies. *Mol Ecol* 9:1657–1659
9. Excoffier L, Laval G, Schneider S (2005) Arlequin (Version 3.0): An Integrated Software Package for Population Genetics Data Analysis. *Evolutionary Bioinformatics* 1:47–50. DOI:10.1177/117693430500100003
10. Gamboa-Becerra R, Montero-Vargas JM, Martínez-Jarquín S, Gálvez-Ponce E, Moreno-Pedraza A, Winkler R (2017) Rapid Classification of Coffee Products by Data Mining Models from Direct Electrospray and Plasma-Based Mass Spectrometry Analyses. *Food Anal Methods* 10: 1359–1368. DOI:10.1007/s12161-016-0696-y
11. García-Flores M, Juárez-Colunga S, García-Casarrubias A, Trachsel S, Winkler R, Tiessen A (2015) Metabolic Profiling of Plant Extracts Using Direct-Injection Electrospray Ionization Mass Spectrometry Allows for High-Throughput Phenotypic Characterization According to Genetic and Environmental Effects. *Journal of Agricultural Food Chemistry* 2015, 63, 1042–1052. DOI:10.1021/jf504853w
12. Gross BL, Olsen KM (2010) Genetic Perspectives on Crop Domestication. *Trends Plant Sci* 15:529–537. DOI:10.1016/j.tplants.2010.05.008
13. Kelchner SA (2012) The Evolution of Non-Coding Chloroplast DNA and Its Application in Plant Systematics. *Ann Mo Bot Gard* 87(4):482–498
14. Kumar S, Stecher G, G., and Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33(7):1870–1874. DOI:10.1093/molbev/msw054

15. Leigh JW, Bryant D (2015) POPART: Full-Feature Software for Haplotype Network Construction. *Methods in Ecology Evolution* 6:1110–1116. DOI:10.1111/2041-210X.12410
16. Macel M, van dam NM, Keurentjes JJB (2010) Metabolomics: The Chemistry between Ecology and Genetics. *Molecular Ecology Resources* 10(4):583–593. DOI:10.1111/j.1755-0998.2010.02854.x
17. Messerli G, Nia VP, Trevisan M, Kolbe A, Schauer N, Geigenberger P, Chen J, Davison AC, Fernie AR, Zeeman SC (2017) Rapid Classification of Phenotypic Mutants of Arabidopsis via Metabolite Fingerprinting. *Plant Physiol* 143:1484–1492. DOI:10.1104/pp.106.090795
18. Morales-Contreras BE, Rosas-Flores W, Contreras-Esquivel JC, Wicker L, Morales-Castro J (2018) Pectin from Husk Tomato (*Physalis ixocarpa* Brot.): Rheological Behavior at Different Extraction Conditions. *Carbohyd Polym* 179:282–289. DOI:10.1016/j.carbpol.2017.09.097
19. Moyers BT, Morrell PL, Mckay JK (2018) Genetic Costs of Domestication and Improvement. *J Hered*: 103–116. DOI:10.1093/jhered/esx069
20. Ramos-López BI, Morales I (2018) Yield Analysis of *Physalis ixocarpa* Brot. Ex Hornem Varieties under Greenhouse and Field Conditions. *Ciencia Rural* 2014:1–7. DOI:10.1590/0103-8478cr20180044
21. Saito K, Matsuda F (2010) Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review Plant Biology* 61(1):463–489. DOI:10.1146/annurev.arplant.043008.092035
22. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) Tortoise and the hare II: *relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis*. *Am J Bot* 92(1):142–166. DOI:10.3732/ajb.92.1.142
23. Sumner LW, Mendes P, Dixon RA (2003) Plant Metabolomics: Large-Scale Phytochemistry in the Functional Genomics Era. *Phytochemistry*. DOI:10.1016/S0031-9422(02)00708-2
24. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G (2012) XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Anal Chem* 84(11):5035–5039. DOI:10.1021/ac300698c
25. Vargas-Ponce, O., J. Sánchez Martínez, M. del P. Zamora-Tavares, and L. E. Mares Valdivia. 2016. Traditional Management of a Small-Scale Crop of *Physalis angulata* in Western Mexico. *Genetic Resources and Crop Evolution* 1383–1395. DOI:10.1007/s10722-015-0326-3.
26. Zamora-Tavares P, Vargas-Ponce O, Sánchez-Martínez J, Dánae CT (2015) Diversity and Genetic Structure of the Husk Tomato (*Physalis philadelphica* Lam.) in Western Mexico. *Genet Resour Crop Evol* 62:141–153. DOI:10.1007/s10722-014-0163-9
27. Zizumbo-Villarreal D, Colunga-GarcíaMarín P (2010) Origin of Agriculture and Plant Domestication in West Mesoamerica. *Genetic Resource Crop Evolution* 813–825. DOI:10.1007/s10722-009-9521-4
28. Zizumbo-Villarreal D, Colunga-GarcíaMarín P, Delgado-Valerio PP, Gepts P (2005) Population Structure and Evolutionary Dynamics of Wild – Weedy – Domesticated. *Crop Sci* 1083:1073–1083. DOI:10.2135/cropsci2004.0340

## Tables

Table 1

Number of mutations, transicions, transversions and nucleotidic diversity, Fu's Fs test of neutrality of populations of *Physalis ixocarpa* and *P. philadelphica* in Michoacan state, Mexico.

	Jacona	Santiago	Chavinda	Villamar	Sahuayo	La Barca	P. philadelphica
Number of variable sites	8	11	36	19	25	59	14
Number of transicions	1	6	18	8	12	19	5
Number. of transversions	7	5	18	11	13	40	9
Nucleotidic diversity ( $\pi$ )	14.00	26.67	31.5	38.67	36.33	52.10	16.33
Fu's Fs	6.52*	2.16*	1.04*	9.32*	1.75*	1.58*	0.89

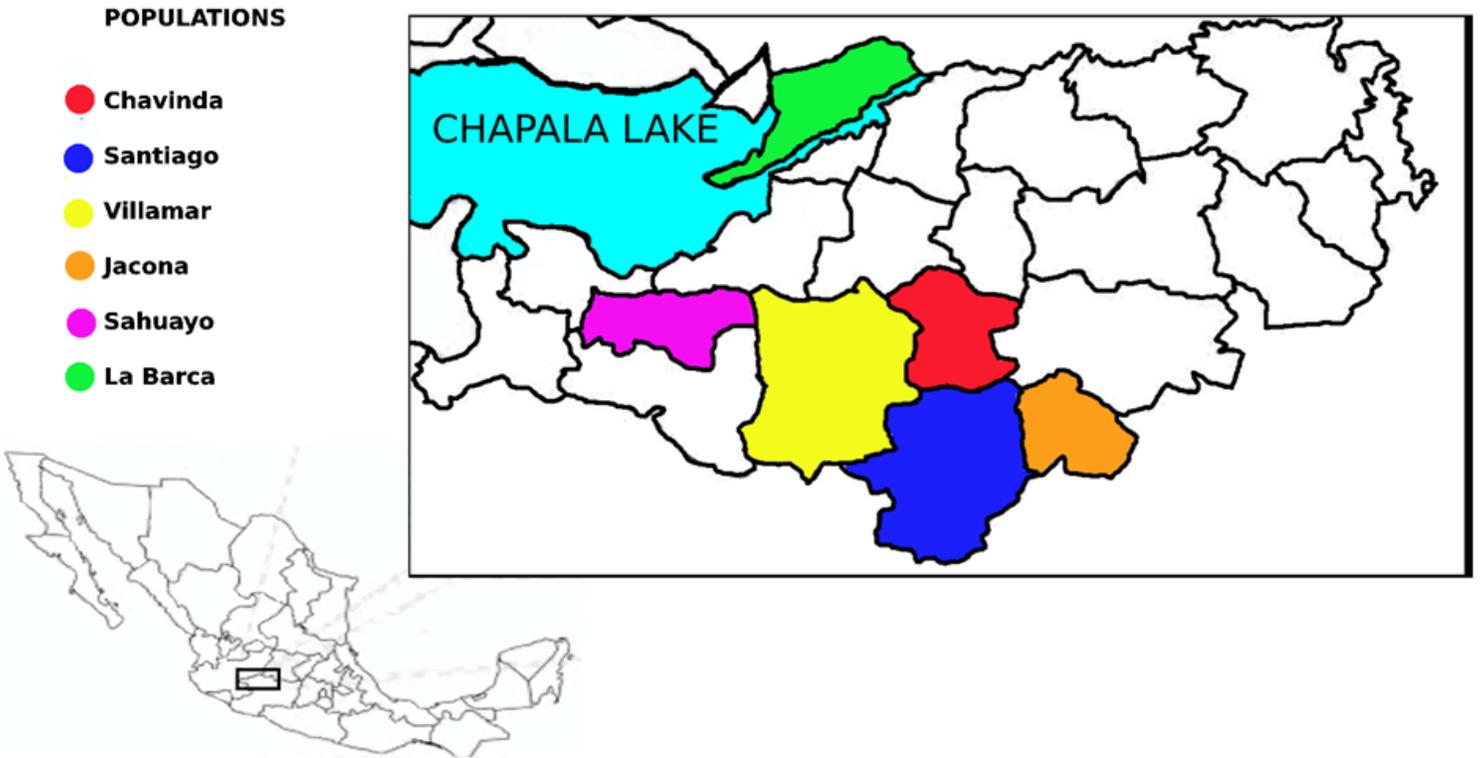
Table 2

Genetic differentiation ( $F_{ST}$ ) among populations of *Physalis ixocarpa* and *P. philadelphica* in Michoacan state, Mexico. Below of diagonal the genetic distance. Above of the diagonal the significance of  $F_{ST}$

	Jacona	Santiago	Chavinda	Villamar	Sahuayo	La Barca	Physalis philadelphica
Jacona	0	0.02*	0.24	0.02*	0.02*	0.02*	0.02*
Santiago	0.35*	0	0.13	0.03*	0.08	0.02*	0.02*
Chavinda	0.06	0.13	0	0.01*	0.06	0.02*	0.01*
Villamar	0.60*	0.57*	0.49*	0	0.02*	0.01*	0.03*
Sahuayo	0.31*	0.35*	0.18	0.39*	0	0.04*	0.02*
La Barca	0.35*	0.31*	0.29*	0.48*	0.36*	0	0.02*
Physalis philadelphica	0.68*	0.67*	0.48*	0.72*	0.57*	0.34*	0

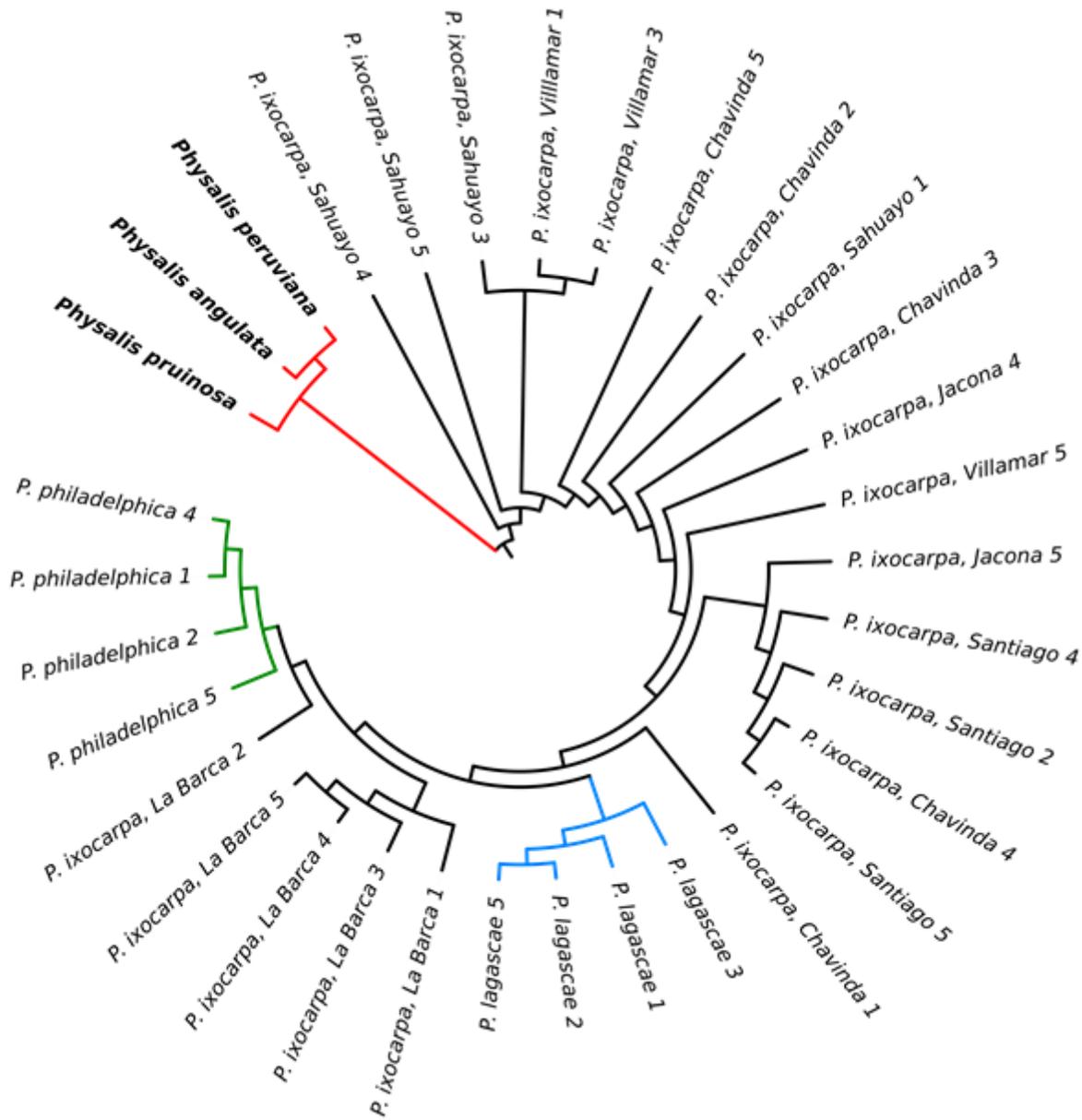
Asterisks indicate significant differences

## Figures



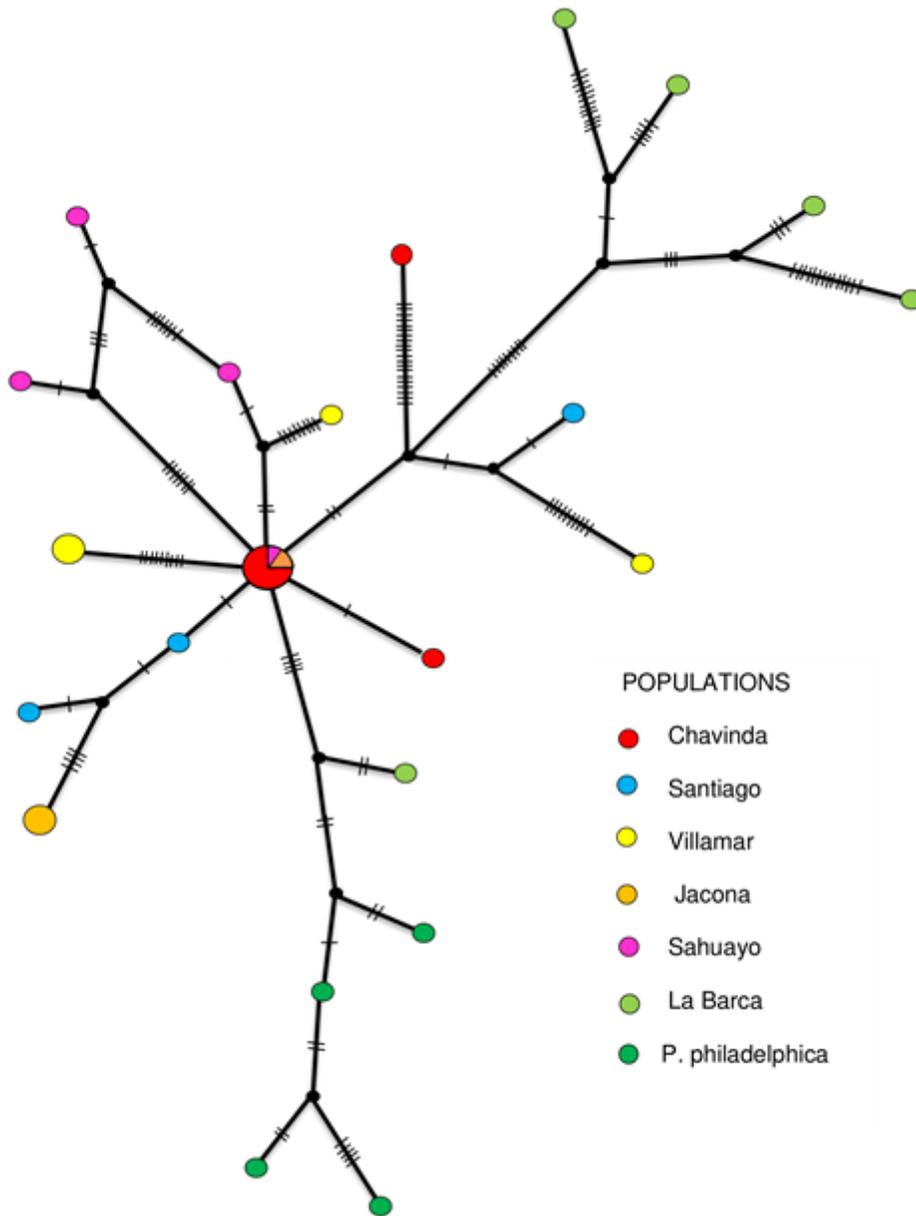
**Figure 1**

Cienega-Chapala Region, Mexico. Colors represent the municipalities where accessions of *Physalis ixocarpa* Brot. were collected.



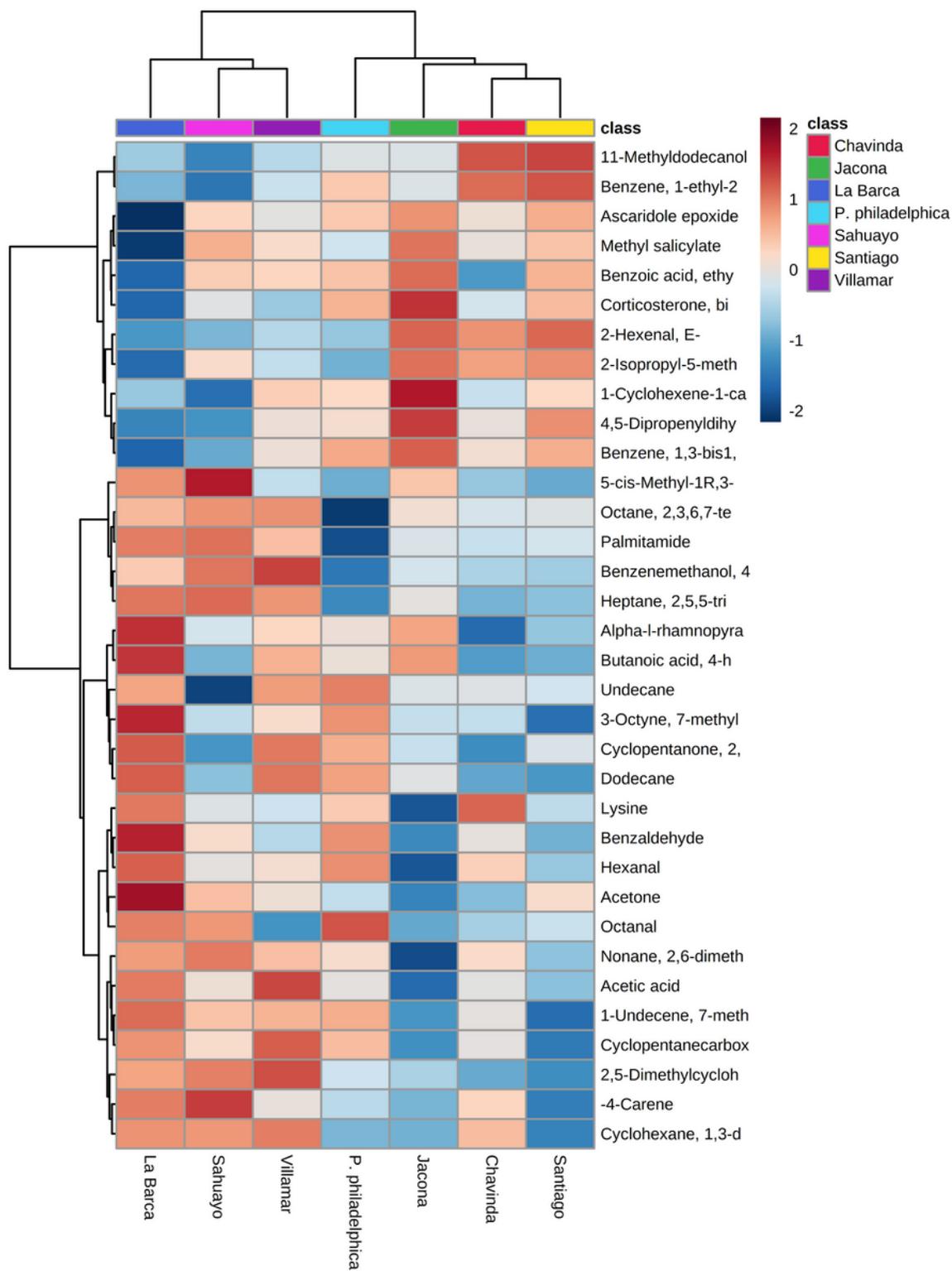
**Figure 2**

Phylogenetic tree of *Physalis* species from the Cienega of Michoacán. The tree was reconstructed using 1000 repetitions of bootstrap under maximum likelihood methods. The red branches indicate the external groups with which our samples were compared. The branches in blue represent a sample of herbarium (*Physalis lagascae*). Green branches point to commercial tomatoes (*Physalis philadelphica*).



**Figure 3**

Haplotype network (using sequencing of chloroplast intergenic regions psbJ – petA and trnL – rpL32 ) of the *Physalis ixcoarpa* Brot. populations in the Cienega-Chapala Region, Michoacan state, Mexico. Size of the circles represents the number of individuals.



**Figure 4**

The heatmap was constructed with the 12 metabolites that showed  $q$  and  $p$  values  $\leq 0.05$ . Colors represent the abundance of metabolites; the blue color indicates down-expression and red color over-expression. The metabolites are clustered according to their Minkowski correlation as a distance function, and the Ward clustering algorithm, the significance of the branches were of  $p \leq 0.05$ .