Genetic Characterization of Solanum pimpinellifolium and Solanum lycopersicum var. cerasiforme Tomato Populations in Ecuador

Natalia Morales Palacio(1), Rafael Morales Astudillo(2), Yoamel Milián García(3), Adriana Artiles Valor(4) and Georgina Espinosa López(5)

(1) National University of Loja, Phone: (593) 0998885879, nataliamorales88@yahoo.es
(2) National University of Loja, Agricultural and Livestock Sector.
(3) University of Havana, Biology Faculty,
(4) Fisheries Research Centre, Food Ministry Institution.
(5) University of Havana, Biology Faculty,
Correspondence to Natalia Morales Palacio and Georgina Espinosa López

Abstract

The cultivated tomato (Solanum lycopersicum L.) is one of the most important vegetable worldwide. Cultivars have been bred using wild species genetic resources. The wild tomato species S. pimpinellifolium and S. lycopersicum var. cerasiforme are widely distributed in Ecuador. These species have been used in tomato breeding programs. We characterized them with microsatellite markers, using 23 microsatellite loci in 61 individuals of S. lycopersicum var. cerasiforme from three locations, and 25 microsatellite loci in 37 plants of S. pimpinellifolium from six locations. Morona Santiago (S. lycopersicum var. cerasiforme) and Manabí (S. pimpinellifolium) were the most variable populations. Populations of S. lycopersicum var. cerasiforme grouped into two clusters: one with individuals from Zamora Chinchipe, and the other with Morona Santiago’s and Islas Galápagos’ individuals. The first cluster of the S. pimpinellifolium population individuals of the El Oro province and Islas Galápagos, while a second cluster included individuals from Manabí and the third cluster included individuals from northern Peru and Loja province. Imbabura had genotypes from the three clusters. The fact that Islas Galápagos was clustered with any continental

(1)(2) National University of Loja, Small city “Guillermo Falconí Espinosa” La Argelia, Loja-Ecuador.
(3)(5) University of Havana, Biology Faculty 25 # 455e/ 2 and 4 Vedado, Havana, Cuba PC 10400.
population supports the hypothesis that the plants were introduced in this island from the continent.

**Key Words**
Wild, microsatellite, intrapopulation and interpopulation variability, tomato.

**Introduction**

The general knowledge of biodiversity and of wild and cultivated related species is of great importance for its conservation and sustainable use (Azurdia, 1996), especially to assure their availability for phyto-improvers (Meyrat, 2001).

The cultivated tomato (*Solanum lycopersicum*) is a very important economic vegetable, and in that way the wild parents are the genes’ source that would allow for improving productivity, adaptation, high nutritional facts quality, resistance to pathogens, etc., in genetic modified organisms by introgression or retro-crossover. However, before using plants for that aim, it’s absolutely necessary to know their conserved genetic variation in its natural habitat. That is why attention should be paid to wild species in order to describe them and to characterize them in the best way (IPGRI, 2000).

The primary origin of the cultivated tomato (*Solanum lycopersicum*) and related wild species is located in a long line of the Andes Mountains at Ecuador and Peru. Ecuador has especially been known as a mega diverse country; this country is custodian of a considerable part of the genetic variation of wild species. This knowledge is particularly important to follow up conservations and to assure that they are still available for phyto-improvers (Meyrat, 2001), taking into account even more that wild related tomatoes species have individual traits that could be potentially useful in culture improvement (Koenig et al., 2013). In fact, those traits have been
used by researchers of other countries (Cocaliadis, 2013). However, in Ecuador, despite the fact that tomato cultures are most of the time the only source of economic income for many agriculture-workers, the provided seeds which were improved for their biotic and abiotic conditions came from European countries. This fact provoked the uncontrolled incidence of pests and diseases, and subsequently the indiscriminate use of pesticides, most of them very hazardous, with 30-35 applications per culture cycle (INIAP, 2001). This lack of rational management points to a high risk for both consumers as well as the agriculture-workers. Also it provokes a mismatch of biological steady states with a loss in effectiveness of pesticides due to pest resistance and thus, an increase in production costs (Valarezo, 2002).

The improved commercial varieties in the world all came from wild species collected from their natural habitats (Peralta et al., 2006). In the breeding improvement selection process, some very important genes from the agronomic point of view have been relegated. Those genes could potentially be used to induce specific resistance for each country’s conditions. That is why is urgent and vital to study the wild Ecuadorian populations variation as well as the relation among them. The potential uses of their agronomic characteristics which are of great interest include: moisture tolerance, insects’ and fungi resistance, improvement of fruit color, soil salinity tolerance, resistance to drought, high content of lycopene, etc. (Nuez, 1995).

More evolved wild forms of foodstuff with red or yellow fruit at maturity that could be found in wild form in Ecuador are: *S. cheesmaniae* and *S. pimpinellifolium*; and, in semi-wild form: *S. lycopersicum* var. *cerasiforme*; while in Peru there are *S. pimpinellifolium* and *S. lycopersicum* var. *cerasiforme*.

The fact that in some indigenous communities, especially in the Amazon and in Ecuadorian and Peruvian coasts, have found *S. lycopersicum* var. *cerasiforme* (big
edible fruit), seems to support that tomatoes were used for human consumption in both countries and migrated to Mexico where the Spaniards know the biggest fruit known today. In any case, the highest variability of wild species relatives of cultivated tomato, are in these two countries, all compatible in direct sexual crossings with tomato (S. lycopersicum) with different difficult degrees.

The wild Peruvian tomatoes, especially S. pimpinellifolium, have been widely studied with different techniques such as: allozymes (Rick et al., 1977, 1978), nuclear DNA sequences (Caicedo and Schall, 2004), AFLPs (Spooner et al., 2005) and microsatellite markers (Sifres et al., 2007), this last is user friendly and a highly polymorphic genetic markers (Krishnasamy et al., 2011). However, studies on tomato populations in Ecuador are scarce. Recently, the morphologic variation and the resistance to a fungus (Fusarium oxysporum) has been accomplished in 146 accessions of four wild related tomatoes species as well as 20 accessions of cultivated tomatoes from a very important germoplasm collection in Ecuador (Morales et al., 2014).

Zuriaga et al. (2009) determined the genetic and climatic variability for S. pimpinellifolium in 247 accessions. Those authors concluded that the climatic differences between Ecuador and Peru could be one of the main causes of genetic variation. In Ecuadorian’s accessions they found seven exclusive alleles in S. pimpinellifolium species. Regardless, the mixed spatial patterns of S. lycopersicum, S. pimpinellifolium and S. lycopersicum var. cerasiforme support the tight genetic relation among them.

For the aforementioned reasons, the aim of this work is to genetically characterize populations of two wild species of Solanum genus related to the cultured tomato by means of microsatellite loci.
Materials and Methods

Vegetal Material Samples

Molecular diversity was estimated in two species of Solanum genus, section and group Lycopersicon (Solanum pimpinellifolium, Solanum lycopersicum var. cerasiforme). Depending on real availability, between two and ten seeds (half sibs) of each accession were germinated in controlled conditions in a greenhouse at 2134 msnm. The genetic diversity of S. lycopersicum var. cerasiforme was estimated in 61 individuals collected from three different localities (Islas Galápagos, Morona Santiago and Zamora Chinchipe) belonging to 46 accessions of the germoplasm bank from the Biotechnology Center of the National University of Loja-Ecuador (UNL, abbreviations from Spanish) collected in Ecuador. The genetic diversity of S. pimpinellifolium was estimated in 37 individuals, collected from five Ecuadorian localities (El Oro, Islas Galápagos, Imbabura, Loja and Manabí) and in 11 individuals from Peru for a total of 40 accessions (figure 1).

DNA Extractions

For each accession, genomic DNA was obtained from 100 mg of young and fresh leaves by the CTAB method. DNA was stored at -20° C in a buffer TE solution (Tris-HCl EDTA) at 10 ng/µl concentration until it was used for PCR.

PCR and Electrophoresis

From a total of 44 microsatellite primer pairs previously described by Smulders et al. (1997), and taking into consideration the authors’ scale (highest polymorphisms in samples and quality of bands), 26 pairs were selected, 23 of them amplified for all accessions of S. lycopersicum var. cerasiforme (LE20592, LE2A11, LEATPACAA,
LEATPACAb, LECAB9, LECHI3, LECHSOD, LEDIH4RE, LEE8, LEEF1Aa, LEEF1Ab, LEHMG2A, LEILV1B, LELAT59G, LELEUZIP, LEMDDNa, LEMDDNb, LENIA, LERBCS3B, LESODB, LESSF, LEWIPIG, LPHFS24), and 25 pairs amplified for *S. pimpinellifolium* (LE20592, LE2A11, LEATPACAA, LEATPACAb, LECAB9, LECHI3, LECHSOD, LEDIH4RE, LEE8, LEEF1Aa, LEEF1Ab, LEHMG2A, LEILV1B, LELAT59G, LELEUZIP, LEMDDNa, LEMDDNb, LENIA, LERBCS3B, LESODB, LEWIPIG, LPHFS24, LE21085, LEGAST1, LESSRPSPGb).

PCR amplification was performed in a 15 µl total volume, containing 1µl of genomic DNA template (10 ng), 1.5 µl of 10X buffer (1X), 1 µl of DNTPs (0.1 mM) (BIOSYNTHESIS), 0.5 units of Taq polymerase (INVITROGEN), 1.5 µl of each of the two primers (1 µM) and 2 mM of MgCl$_2$, with exception of LELEUZIP loci (1.5 mM MgCl$_2$) and LEDIH4RE, LELAT59G, LESSF (2.5 mM MgCl$_2$). PCR conditions employed were previously described by Smulders et al. (1997). For some primers, annealing temperature was reduced to 50° C (LE21085, LEATPACAA, LECAB9, LEDIH4RE, LEILV1B, LESSRPSPGb) and the total cycles increased to 35 (LESSRPSPGb).

PCR reactions were carried out in an iCycler thermocycler from Bio Rad. Amplification products were checked in 2% agarose gels stained in ethidium bromide after the horizontal electrophoresis ran out in an UVITEC BTS-20.LM a 365 nm transiluminator.

**Microsatellite Polymorphism’s Detection**

The samples were prepared by adding 4 µl of Blue Juice 1X (INVITROGEN) to each amplified product. Samples were subsequently denatured at 94° C for 4 min, followed by quenching on ice. Amplified products were separated by PAGE vertical electrophoresis (6% polyacrylamide, 8 M urea, Tris-borate buffer 1X) for 60-90 min at
75-W constant current, using a TECHNE S3S electrophoresis system. The DNA bands were visualized by silver staining (CIP, 1998) and allele sizes were recorded in base pairs and estimated by comparison to a commercial 30-330 bp marker (INVITROGEN).

**Data Analysis**

In order to detect possible mistakes in reading or interpreting data related to the presence of null alleles or wrong assignments of allele sizes, generated genotypes with primer sets were analyzed with Micro-Checker program (Van Oosterhout et al., 2004). This allowed the adjustment of allele frequencies and genotypes.

The program GenAlEx 6.1 (Peakall and Smouse, 2006) was used to estimate means of genetic variability parameters: Nei’s heterozygosity, number of alleles per locus (Na), effective number of alleles (Ne), number of private alleles (NPA), unbiased expected heterozygosity (UHe), and its standard error deviation (SE-UHe).

The statistical package program FSTAT, version 2.9.3 (Goudet, 2002) was used to calculate allelic richness, the observed heterozygosity (Ho), intraspecific diversity (Hs), total diversity (Ht), and disequilibrium linkage.

To make inference of population structure and to identify mixed individuals based on their genotypes, the Structure 2.2 program (Pritchard et al., 2000; Falush et al., 2007) was used. “Admixture Model” was selected for that aim in the following conditions: 500 000 burn-in iterations and 1 000 000 Markov and Montecarlo repetitions (MCMC) to obtain the estimated parameters. Seven K values were evaluated and the number of clusters (k) were estimated as the maximized number to ∆k parameter (Evanno et al., 2005) with Structure Harvester statistic package (Earl et al., 2012).

The program ARLEQUIN version 3.11 (Excoffier et al., 2005) was used to calculate $F_{ST}$ values among populations.
Nei’s (1983) Individual genetic distances were obtained by Population program (Arnason et al., 1995). The Neighbor-joining algorithm was used by MEGA 5 program for tree construction (Tamura et al., 2011).

Results and Discussion
Different populations of *S. lycopersicum* var. *cerasiforme* and *S. pimpinellifolium* were evaluated to estimate genetic variability and differentiation among them. Hardy–Weinberg equilibrium was not calculated because each accession was collected in distant sites (some meters or dozens of kilometers) through Ecuadorian or Peruvian territories respectively, and in that case, those samples did not form a population itself. Despite this fact, the Structure program, which presupposes the existence of genetic equilibrium, was used. All loci are in linkage disequilibrium for the total samples and for each population separately. This result indicates that each locus offers independent information to determine the genetic structure of different populations in both species.

Genetic Variability
Genetic diversity, expressed as allelic richness it is shown in table 1. For *S. lycopersicum* var. *cerasiforme*, the population of Morona Santiago presents more allelic richness in relation to the Zamora Chinchipe population, as well as more alleles, despite a smaller sample size. The number of private alleles is much higher as were the heterozygosity values. However, its standard deviation is lower. Data from Islas Galápagos individuals was not used for genetic variability analysis due to small sample size.

In the case of the other species, *S. pimpinellifolium*, the most variable population was Manabí. It records the highest values for genetic parameters such as: allele number,
effective number of allele, many private alleles and also a high heterozigosity, and the least standard deviation. However, allelic richness is slightly low, despite the higher sample size.

On the other hand, the population from Loja has the highest allelic richness, despite the small sample size, although allele quantity is somewhat lower and the standard deviation of heterozygosity is the highest one.

The genetic variability parameters from El Oro, Islas Galápagos and Imbabura are slightly lower than those reported for Manabí. However, Galápagos and Imbabura present the lowest values of allelic richness and private alleles respectively.

The results found in this study are in agreement with those reported by Zuriaga et al., 2009. Those authors found in Islas Galápagos, using microsatellite techniques, identical and indistinguishable plants amongst each other, and also clustered with some individuals of *S. pimpinellifolium* collected from the Ecuador’s coast (Provinces of Los Ríos and Guayas) that registered no genetic diversity and slow bootstrap values (Nuez et al., 2004). That is why it is very probable that the population from Islas Galápagos could be a recent colonization from the continental Ecuador and a bottleneck effect is accomplished, and subsequently, a loss of diversity and selection directed to autogamy.

The Peruvian population, despite of their highest sample size in relation to the others, is however the least variable, with the lowest heterozygosity, the absence of private alleles, and the lowest number of effective alleles.

Literature shows that Peruvian populations of *S. pimpinellifolium* have a high number of private alleles that do not exist in Ecuadorian populations (Zuriaga et al., 2009). The contradictory results of this study could be explained by the fact that the
Peruvian sample could be constituted by inbreeding related individuals collected in adjacent areas that do not represent the existent variability in Peru.

In a general sense, *S. pimpinellifolium* presents a scattered distribution (Rick et al., 1977) and records a limited level of allogamy (Rick, 1979). In such conditions, the plants have difficulties for a cross pollination and because of that, they follow the auto pollination route (Zuriaga et al., 2009). This leads to a low heterozygosity, decrease of diversity, and a strong selection to autogamy.

Zuriaga et al. (2009) studying 119 accessions of *S. pimpinellifolium* from Peru and 94 from Ecuador found the highest variability at the north of Peru (highest mean number of alleles by locus, private alleles, and heterozygosity). However, the same authors report that in both regions south and north of Ecuador there is a lot of genetic variation and geographic groups have high diversity. Similarly in this work, populations of Manabí and El Oro are outstanding for being the most variable in relation to the other Ecuadorian populations and even more in relation to the Peruvian populations.

The fact that Ecuadorian populations are more variable than the Peruvian ones could be explained because both the climate and geography can influence the mutational selection that randomly occurs, leading to more variation of *S. pimpinellifolium*. In that way, while in the Peruvian coast the annual precipitation record is less than 700 mm and substantial climatic changes are reported, in some regions of Ecuador more than 2800 mm are recorded and a more uniform climate through the entire year is observed (Zuriaga et al., 2009). On the other hand, there are the Andes laying in the Ecuadorian land which has allowed the species re-diversification in a way that the country has been recognized as a mega diverse land. On the other hand, the Peruvian geography extends from the borders of Ecuador to the south, with a wide
thicket and brush land, and then a sandy desert and a slight geographic climatic differentiation (Zuriaga et al., 2009).

**Genetic Differentiation Among Population**

The collection site of each individual used in this work is represented at the end of the branches in figure 2, which identify the units of each group. The analysis for *S. lycopersicum* var. *cerasiforme* populations showed the existence of two clusters: Zamora Chinchipe and Morona Santiago/Islas Galápagos; while for *S. pimpinellifolium*, the tree did not show good resolution.

The Combination Analysis of Pritchard et al. (2000) is an individual-based and model-based approach. It uses a method based in each individual’s information to evaluate the genetic relations and to identify populations, without a priori assumptions about the number of existent populations, arbitrarily defined as entirely and together collected individuals in the same geographic location.

A potential problem with individual-based methods is that they still can yield uncertain results if genetic differentiation among populations is not substantial. In addition, the performance and reliability of individual-based methods has not been thoroughly evaluated. Thus, it seems useful and prudent to use both individual-based and population-based methods (Evanno et al., 2005).

Therefore, the population’s structure generated with 23 loci and 61 individuals belonging to *S. lycopersicum* var. *cerasiforme* species (shown in figure 3) is in agreement with the genetic Nei’s (1983) distances tree (figure 2). In both cases individuals are separating into two groups.

The first cluster belongs to the individuals collected in Zamora Chinchipe at the south-east of the Ecuadorian Amazon, with a mountain orography that can be
distinguished from the rest of the provinces of the same Amazon region. This cluster shows an assignment coefficient of Morona Santiago’s population less than 2%.

The second cluster belongs to individuals from Morona Santiago, which is limited by the south to Zamora Chinchipe province and is located at the south Center of the Amazon. This population presents an assignment coefficient to Zamora Chinchipe population around 2%. This cluster includes the individuals collected in Islas Galápagos, whose assignation coefficient of Morona Santiago is between 5 and 60%, which supports the hypothesis that the species was introduced to the Island from the continent.

Thus, the STR markers allowed the differentiation of the individuals at the level of population for the species *S. lycopersicum var. cerasiforme*.

The exploratory tree based in Nei’s (1983) genetic distances does not allow detection of a clear differentiation for each population of the *S. pimpinellifolium* species (figure 2); however the results of the Structure program run with 25 STR loci and 48 individuals of *S. pimpinellifolium* species (figure 4) were more resoluteness. Three clusters are observed. One of them includes the individuals collected from El Oro southwest of the coastal line of Ecuador and also from Islas Galápagos. This indicates that the genetic base of *S. pimpinellifolium* of the Island is the same as that of the Ecuadorian continental populations. Another group consists of individuals collected in the northwest of continental Ecuador, in the Manabí province.

Finally, the last cluster is composed by individuals from the north of Peru and from Loja province in the south of Ecuador. They have similar population structure, which is expected if we consider that the species is abundant and has been collected in the boundary land between the two countries and is in agreement with Zuriaga et al. (2009). Those authors in the PCoA analysis and NJ dendrogram based in Dc
distances among different geographic clusters of \textit{S. pimpinellifolium} found that despite the fact that the clusters of the south of Ecuador and the north of Peru are nearby, both zones are quite different on terrain and climate gradients. Nevertheless, the north Peruvian cluster is together with the south Ecuadorian. Following this idea, with exception of the population from Loja, the rest of Ecuadorian populations of \textit{S. pimpinellifolium} are quite different from Peruvian populations. Similar results were obtained by Zuriaga et al. (2009) using Structure clusters, FCA, geographic dendrogram, and AMOVA. They concluded that this division is highly correlated to severe climatic differences on both country’s boundaries, which not only affects this species, but also has an important effect on the global ecosystem that characterized both countries.

The individuals collected from the north mountains of Ecuador, in Imbabura province, showed three genotypes.

The continental Ecuador is divided in several geographic regions, whose populations are isolated and separated by big distances and geographic accidents. If climatic variation is added to this fact, the genetic variability as well as the different populations found throughout could be easily explained. In the case of the north of Peru, the distribution of \textit{S. pimpinellifolium} is almost continuous and there is not considerably genetic differences for the collected plants (Zuriaga et al., 2009), yet at the same time the plants’ distribution is irregular (Caicedo and Schall, 2004).

In order to prove the right estimation of number of clusters (K), the Delta K statistic was used. This is based on the probability rate of the data logarithm among successive values of K, which precisely detects the highest level of hierarchic structure. Thus, for the species \textit{S. lycopersicum} var. \textit{cerasiforme}, Delta K recognizes two clusters and for \textit{S. pimpinellifolium} three of them.
A mix of two cluster’s individuals was identified with a Q predominant value of the population.

**Differentiation by $F_{ST}$ Among Populations**

The differences among populations of *S. lycopersicum* var. *cerasiforme* are illustrated in table 2. As it can be seen, the individuals from Islas Galápagos differentiated from those of Morona Santiago, and they do not differentiate from those of Zamora Chinchipe despite belonging to the same cluster as Morona Santiago.

For *S. pimpinellifolium* species (table 3), individuals from El Oro are different from all populations, as well as those from Islas Galápagos. This last population has the least value of $F_{ST}$, but the probability is less than 0.05. The population from Imbabura, does not differ from Manabí and Peru because, as it is observed in figure 4, it presents genotypes from both places. Loja and Manabí are not different from Peru.

The AMOVA for both species (table 4) showed that the highest variation among them is inside the individuals, and it is remarkable that the highest percentage of variation is among populations and individuals of *S. pimpinellifolium* in relation to *S. lycopersicum* var. *cerasiforme*.

The populations of *S. pimpinellifolium* clustered in three groups (table 4), and the fourth part of the variance is among them. In this differentiation the remarkable fact is that those are populations distributed throughout the Ecuadorian land.

Concerning *S. lycopersicum* var. *cerasiforme*, the least variation is explained because the Morona Santiago and Zamora Chinchipe province are contiguous in the Ecuadorian Amazon. However, they are included in a big area of more than 34000 Km², which represents 12.15% of the entire national territory. That is why, it is understood that populations are scatter distributed through those two province lands, and possibly the different existent cultures in both territories played an important role
in selection of the species with different criteria. This fact could explain similar variation between populations and individuals. The population from Galápagos was probably carried over by the colonists from the continent to the Islands.

It is important to remark that all variance components are highly significant (P=0).

**Conclusions**

The analyzed loci differentiate populations from species *S. lycopersicum* var. *cerasiforme* and *S. pimpinellifolium*. Those markers also allowed differentiation among individuals at the level of population.

The genetic variability average values show that the Manabí population of *S. pimpinellifolium* is the most variable. In the same way, the Morona Santiago population is the most variable of the *S. lycopersicum* var. *cerasiforme* species.

**Acknowledgements**

Thanks to the National Secretary of Higher Education, Science and Technology of the Ecuador state (Secretaría Nacional de Educación Superior, Ciencia y Tecnología del Estado Ecuatoriano SENESCYT) for granting the scholarship to the first author and consequently, the budget for the research. Also to all professionals that contributed to the establishment and maintenance of the germoplasm bank of the Biotechnology Center at National University of Loja (Centro de Biotecnología de la Universidad Nacional de Loja). Finally, to the Integrated Center of Environmental Geomatic (Centro Integrado de Geomática Ambiental, CINFA), especially to Engineer José Merino Alverca for his valuable contribution with the use of SIG software.

**References**
Arnason AN, Schwarz CJ, and Boyer G (1995). POPAN-4: A data maintenance and
analysis system for mark-recapture data. Scientific report, Department of Computer
Science, University of Manitoba, Winnipeg, viii, pp. 267.

Azurdia C (1996). Las malezas como un reservorio genético de las plantas. Lecturas
en Recursos Fitogenéticos. Subprograma de Recursos Genéticos Vegetales
(REGEVE). Instituto de Investigaciones Agronómicas. Facultad de Agronomía.

Caicedo AL, and Schall BA (2004). Population structure and phylogeography of

Centro Internacional de la Papa (CIP) (1998). Protocolos de Laboratorio de Biología
Departamento de Recursos Genéticos. Manual de capacitación. 2ª edición (revisada

Cocaliadiis MF (2013). Desarrollo y utilización de marcadores moleculares para el
genotipado de la colección de líneas de introgresión de Solanum lycopersicoides en
el fondo genético del tomate cultivado. Tesis. Universidad Politécnica de Valencia-
COMAV.

Earl DA, and Von Holdt BM (2012). STRUCTURE HARVESTER: a website and
program for visualizing STRUCTURE output and implementing the Evanno method.
Conservation Genetics Resources vol. 4 (2) pp. 359-361 doi:10.1007/s12686-011-
9548-7.

Evanno G, Regnaut S, and Goudet J (2005).Detecting the number of clusters of
individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–
2620.

software package for population genetics data analysis. Evolutionary Bioinformatics
Online, 1, 47–60.

Version 2.2.Structure.

Goudet J (2002). FSTAT. A program to estimate and test gene diversities and
fixation indices (version 2.9.3.2). Available in:
[http://www2.unil.ch/popgen/softwares/fstat.htm].

INIAP (2001). Diagnóstico, bioecología y manejo sostenible de la negrita Prodiplosis
longifila en el Ecuador. INIAP-PROMSA-CEDEGE. Estación Experimental Portoviejo,
Informe anual, pp. 19.

Boletín de las Américas. Cali, Colombia. V. 6, No. 1, pp. 4-8.

Krishnasamy T, Narayanan M, Nagasamy N, Ayyanan G, Pandi S, Santoshkumar M,
Rajasekar R (2011). Sampling and exploitation of genetic variation exist in locally


# Tables and Figures

**Table 1.** Average values of genetic variability parameters for populations of wild species of *Solanum* genus, section *Lycopersicon*, estimated between 23 and 25 microsatellite loci (*S. lycopersicum* var. *cerasiforme* and *S. pimpinellifolium*, respectively), with statistical program GenAlEx.

<table>
<thead>
<tr>
<th>Populations</th>
<th><em>S. lycopersicum</em> var. <em>cerasiforme</em></th>
<th><em>Solanum pimpinellifolium</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morona Santiago</td>
<td>Zamora Chinchipe</td>
</tr>
<tr>
<td>N</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td>Na</td>
<td>3,043</td>
<td>2,304</td>
</tr>
<tr>
<td>Ne</td>
<td>2,415</td>
<td>1,946</td>
</tr>
<tr>
<td>RA</td>
<td>2,883</td>
<td>2,189</td>
</tr>
<tr>
<td>NAP</td>
<td>1,087</td>
<td>0,348</td>
</tr>
<tr>
<td>UHe</td>
<td>0,531</td>
<td>0,403</td>
</tr>
<tr>
<td>SE(UHe)</td>
<td>0,045</td>
<td>0,052</td>
</tr>
</tbody>
</table>

**Legend:** Number of individuals (N). Mean number of alleles by locus (Na). Effective number of alleles (Ne). Allele Richness (RA). Number of private alleles (NAP). Unbiased Expected heterozygosity (UHe). Standard deviation of expected heterozygosity (SE UHe).
**Table 2.** Pairwise differences among populations of *S. lycopersicum* var. *cerasiforme*. Inside parenthesis the probability value (P) of $F_{ST}$.

<table>
<thead>
<tr>
<th>Populations</th>
<th>Islas Galápagos</th>
<th>Morona Santiago</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morona Santiago</td>
<td>0.200</td>
<td><em>(0.000+-0.000)</em></td>
</tr>
<tr>
<td></td>
<td><em>(0.000+-0.000)</em></td>
<td></td>
</tr>
<tr>
<td>Zamora Chinchipe</td>
<td>0.015</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td><em>(0.288+-0.029)</em></td>
<td><em>(0.000+-0.000)</em></td>
</tr>
</tbody>
</table>
Table 3. Pairwise differences among six populations of *S. pimpinellifolium*. On parenthesis the values of probability of $F_{ST}$.

<table>
<thead>
<tr>
<th>Populations</th>
<th>El Oro</th>
<th>I. Galápagos</th>
<th>Imbabura</th>
<th>Loja</th>
<th>Manabí</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Galápagos</td>
<td>0.121</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009+-0.009)</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imbabura</td>
<td>0.228</td>
<td>0.305</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009+-0.009)</td>
<td>(0.000+-0.000)</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loja</td>
<td>0.515</td>
<td>0.612</td>
<td>0.325</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000+-0.000)</td>
<td>(0.000+-0.000)</td>
<td>(0.009+-0.009)</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Manabí</td>
<td>0.283</td>
<td>0.374</td>
<td>-0.052</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000+-0.000)</td>
<td>(0.000+-0.000)</td>
<td>(0.882+-0.035)</td>
<td>(0.018+-0.012)</td>
<td>*</td>
</tr>
<tr>
<td>Perú</td>
<td>0.430</td>
<td>0.472</td>
<td>0.092</td>
<td>0.538</td>
<td>0.143</td>
</tr>
<tr>
<td></td>
<td>(0.000+-0.000)</td>
<td>(0.000+-0.000)</td>
<td>(0.144+-0.036)</td>
<td>(0.000+-0.000)</td>
<td>(0.018+-0.012)</td>
</tr>
</tbody>
</table>
Table 4. Molecular variance estimations among local populations and among Structure’s clusters in two wild species of *Solanum* genus, *Lycopersicon* section, to microsatellite loci by means of AMOVA generated by GenAlEx. All differences were highly significant.

<table>
<thead>
<tr>
<th></th>
<th><em>S. lycopersicum</em> var. cerasiforme</th>
<th><em>Solanum pimpinellifolium</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of loci</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>No. of populations</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Among populations</td>
<td>18 %</td>
<td>25 %</td>
</tr>
<tr>
<td>Among individuals</td>
<td>19 %</td>
<td>30 %</td>
</tr>
<tr>
<td>Inside individuals</td>
<td>79.30 %</td>
<td>45 %</td>
</tr>
</tbody>
</table>
Figure 1. Geographical collecting locations for two wild species of *Solanum, Lycopersicon* section: *Solanum pimpinellifolium* and *Solanum lycopersicum* var. *cerasiforme*. 
Figure 2. Exploratory and unrooted tree constructed by the Neighbor-joining method based on Nei’s (1983) genetic distances for two species of Solanum genus, estimated by means of microsatellite data Population program. (A) *S. lycopersicum* var. *cerasiforme*, two clusters: (B) *S. pimpinellifolium*, three clusters. At the ends, bars indicate the origin of each individual.
Figure 3. Population Structure estimated by means of 23 loci in the species *S. lycopersicum* var. *cerasiforme*, by Structure program. Each individual is represented by a horizontal bar and each color remarks the main clusters. (IG) Islas Galápagos; (MS) Morona Santiago Province; (ZCh) Zamora Chinchipe Province.
Figure 4. Population structure estimated by means of 25 loci for the species *S. pimpinellifolium*. Each individual is represented by a horizontal bar and each color remarks the main clusters. (Oro) El Oro province; (Gal) Islas Galápagos; (Ma) Manabí province; (Imb) Imbabura province; (Loj) Loja province; (Per) Peru.