



Máster Interuniversitario Oficial

Universidad Politécnica de Madrid


Universidad Politécnica de Barcelona

Universidad Politécnica de Valencia



Instituto de Conservación y Mejora de la Agrodiversidad Valenciana

COMAV



**Análisis de la variabilidad genética de
Solanum habrochaites en su lugar de origen,
un recurso infrautilizado para la mejora del tomate.**



Carmen Hermida Carrera

Director/a: Dra. Alicia Sifres Cuerda

RESUMEN

En este trabajo se ha realizado un amplio estudio de la variabilidad genética y la estructura poblacional de la especie silvestre *Solanum habrochaites* (anteriormente conocida como *Lycopersicon hirsutum*) en todo su rango de distribución natural.

Esta especie, relacionada con el tomate, es una fuente de recursos genéticos muy importante para la mejora de este cultivo, ya que presenta resistencia a gran cantidad de enfermedades y plagas, tolerancia al frío y caracteres de calidad del fruto.

S. habrochaites habita la cara oeste de los Andes, a elevadas altitudes, desde el centro de Ecuador hasta el centro de Perú. Para estudiar la variabilidad genética de esta especie se han analizado 91 entradas colectadas en todo su rango de distribución que se agruparon en siete áreas o zonas geográficas para simplificar el estudio.

Se han utilizado dos tipos de marcadores: nueve loci microsatélite y seis combinaciones de AFLPs (Amplified Fragment Length Polymorphisms). Los marcadores microsatélite, al ser marcadores codominantes permiten obtener información directa acerca de la heterocigosidad e indirectamente acerca del grado de alogamia de las poblaciones. Los marcadores AFLP, aunque son dominantes, proporcionan una gran cantidad de información ya que analizan un gran número de loci a lo largo de todo el genoma.

Se observó que la estructura poblacional de esta especie está altamente correlacionada con la diferenciación geográfica como mostraron el análisis molecular de la varianza y los diferentes métodos de agrupamiento.

El análisis de coordenadas principales para AFLPs mostró la existencia de una clina de norte a sur del rango de distribución para la variación genética. Las entradas de los diferentes grupos geográficos se organizaron en el primer eje de norte a sur encontrándose una clara separación entre ellos. Los grupos pertenecientes a la zona central del área de distribución fueron los que presentaron una mayor variación genética, mostrando, además, una mayor heterocigosidad, calculada para los loci microsatélites.

Las poblaciones situadas en los extremos, al norte y el sur del área de distribución, mostraron una menor diversidad y una extremadamente baja heterozigosidad.

Utilizando los datos morfológicos de pasaporte, no observamos las formas *typicum* y *glabratum* descritas por Muller. Al parecer, éstas constituirían las formas extremas de una variación continua de norte a sur relacionada con el grado de pubescencia. Además, estas diferencias en la pubescencia no estarían asociadas con el grado de homocigosis como describía este autor.

Aunque se encontró una alta correlación entre las distancias climáticas y genéticas, la mayor correlación se encontró entre éstas últimas y las distancias geográficas. Las principales diferencias genéticas se encuentran entre las entradas de los diferentes valles, probablemente debido al aislamiento geográfico de las poblaciones.

Nuestros resultados apoyan la hipótesis para la evolución de *S. habrochaites*, propuesta por otros autores con anterioridad, que sugiere que el centro de origen de la especie estaría en su zona central de distribución. A partir de estas poblaciones ancestrales autoincompatibles se derivarían el resto. Conforme las poblaciones migraron hacia el norte y el sur de su área de distribución la autoincompatibilidad sería sustituida por autocompatibilidad como resultado de la selección favorecida por la presencia de pocos individuos, debido al efecto fundador o por la escasez de insectos polinizadores.

SUMMARY

This paper has conducted an extensive study of genetic variability and population structure of the wild species *Solanum habrochaites* throughout its natural distribution range.

This species is related to tomatoes and it is a source of genetic resources important to the improvement of this crop, as it is resistant to many pests and diseases, cold tolerance and fruit quality characters.

S. habrochaites inhabits the west side of the Andes at high altitudes from central Ecuador to central Peru. To study the genetic variability of this species we have been analyzed 91 entries collected throughout their range of distribution. They were grouped into seven geographical areas or zones to simplify the study.

We have used two types of markers: nine microsatellite loci and six combinations of AFLP (Amplified Fragment Length Polymorphisms). With microsatellite markers, codominant markers, it is possible to obtain direct information on the heterozygosity and indirectly on the extent of allogamy populations. AFLP markers, though dominant, provide a wealth of information and analyzing a large number of loci throughout the genome.

It was noted that the population structure of this species is highly correlated with geographical differentiation as shown by molecular analysis of variance and different methods of grouping.

The principal coordinates analysis for AFLPs showed the existence of a cline from north to south distribution range for genetic variation. The entries in the various geographical groups were organized in the first north-south axis founding a clear separation between them. The groups belonging to the central distribution area were those who showed more genetic variation, showing also a higher heterozygosity, calculated for microsatellite loci.

The populations at the ends, north and south of the range, showed reduced diversity and an extremely low heterozygosity.

Using morphological data for passports, we did not see the forms *glabratum* and *typicum* described by Muller. Apparently, they constitute the extreme forms of a continuous variation from north to south related to the degree of pubescence. Moreover, differences in the pubescence are not associated with the degree of homozygosity as described by this author.

Although there was a high correlation between climate and genetic distances, the highest correlation was found between the latter and geographical distances. The major genetic differences are among the entries in the different valleys, probably due to geographical isolation of populations.

Our results support the hypothesis for the evolution of *S. habrochaites*, proposed previously by other authors, suggesting that the center of origin of the species would be at its central distribution. From these ancestral self-incompatible populations would result from the rest. As populations migrated to the north and south, its range of self-incompatibility would be replaced by self-compatibility as a result of selection favored by the presence of few individuals, due to a founder effect or by a shortage of pollinating insects.

1.	Introducción	2
1.1.	Importancia económica del tomate	2
1.2.	Valor nutricional del tomate.	4
1.3.	Taxonomía del género <i>Solanum</i> .	4
1.4.	Centro de Origen y diversificación.	7
1.5.	<i>Solanum habrochaites</i> . Importancia y principales características de la especie.	9
1.5.1.	Descripción de la planta.	9
1.5.2.	Hábitat.	11
1.5.3.	Empleo en la mejora del tomate.	11
1.6.	Los marcadores genéticos en la conservación y uso de los recursos fitogenéticos.	13
1.7.	Estudios de variabilidad en <i>Solanum habrochaites</i> .	14
2.	Objetivos	17
3.	Material y Métodos	19
3.1.	Material Vegetal.	19
3.2.	Caracterización molecular.	25
3.2.1.	Tipos de marcadores moleculares utilizados.	25
3.2.2.	Toma de muestras, extracción y cuantificación del ADN.	25
3.3.	Marcadores AFLP.	25
3.3.1.	Fundamento.	25
3.3.2.	Protocolo.	26
3.3.2.1.	Digestión.	27
3.3.2.2.	Ligación.	27
3.3.2.3.	Amplificación.	27
3.4.	Marcadores SSR.	29
3.4.1.	Fundamento.	29
3.4.2.	Marcadores Microsatélites Seleccionados.	30
3.4.3.	Protocolo.	31
3.4.3.1.	Amplificación.	31
3.5.	Separación de los productos amplificados.	31

3.6. Análisis de los datos.	33
4. Resultados y discusión.	38
4.1. Marcadores AFLPs.	38
4.1.1. Análisis de Coordenadas Principales y Diversidad Genética.	38
4.1.2. Análisis Cluster.	40
4.1.3. Estructura poblacional.	43
4.2. Marcadores SSRs.	46
4.2.1. Análisis de Coordenadas Principales.	47
4.2.2. Análisis Cluster.	49
4.2.3. Estructura poblacional.	50
4.3. Diferenciación Regional.	53
4.4. Análisis Climático.	56
4.5. Datos de Pasaporte.	59
4.6. Discusión	60
5. Conclusiones.	65
6. Bibliografía.	68
7. Anexo.	79

Introducción

- En tu tierra - dijo el principito - los hombres cultivan cinco mil rosas en un mismo jardín... Y no encuentran lo que buscan...
- No lo encuentran...- respondí.
- Y, sin embargo, lo que buscan podrían encontrarse en una sola rosa o en un poco de agua...
- Seguramente - respondí.

Antoine de Saint- Exupéry "*El principito*"

1. Introducción

1.1. Importancia económica del tomate

El tomate es la hortaliza con mayor importancia económica en el mundo, con una superficie cultivada de casi 5 millones de hectáreas y una producción de 126 millones de toneladas en 2007 (Tabla 1.1). El principal continente productor es Asia con un 56.7% de la producción, seguido del continente Europeo con un 16,24% (Figura 1.1, FAOSTAT 2009).

Tabla 1.1. Superficie cultivada de tomates en el mundo en el año 2007 (Fuente: FAOSTAT, 2009).

Área geográfica	Superficie cultivada (en miles de ha)
España	55
Europa	594
Unión Europea	300
América del Norte y Central	327
América del Sur	141
Asia	2.827
África	660
Total mundial	4.834

A nivel mundial España participa con un 1,14% del total, ocupando a nivel de la unión Europea, el segundo lugar tras Italia. Italia, España, Grecia, Portugal y Francia son, en orden descendente, los mayores productores de tomate de la Unión Europea (UE27). Una gran parte de la producción en estos países se destina a la industria de transformación, además tiene un elevado consumo interno. Holanda, Polonia y Rumanía les siguen en producción y representan un 12,03% del total de la UE. Holanda es una de las más importantes fuentes de suministro de tomate fresco de la UE (Tabla 1.2).

En España, las comunidades autónomas con mayor nivel de producción son Andalucía y Extremadura, que suman un 68.51% de la producción española, seguidas de la Comunidad de Murcia, Canarias, Castilla la Mancha, Aragón y Navarra con un 23.27% de la producción. La Comunidad Valenciana se encuentra en octavo lugar con un 2.85% del total (Anuario de Estadística Agroalimentaria 2008).

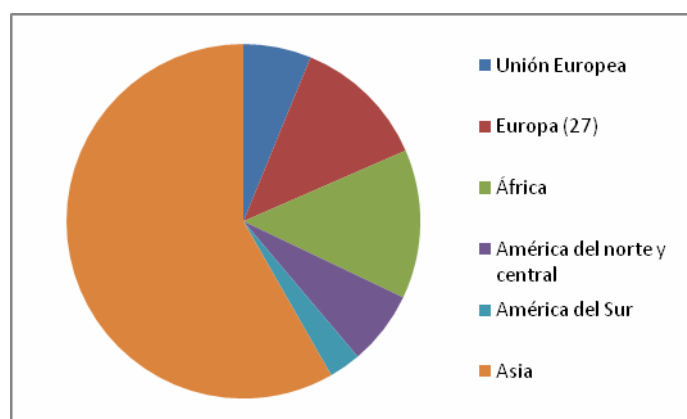


Figura 1.1. Distribución de la producción de tomates por continentes en el año 2007 (Fuente: FAOSTAT 2009).

Tabla 1.2. Distribución de la producción de tomates en la Unión europea por países en 2007 (Fuente: FAOSTAT 2009).

Área Geográfica	Producción (en miles de ton)
Italia	6025
España	3610
Grecia	1450
Portugal	1000
Francia	750
Holanda	690
Polonia	652
Rumanía	555
Total Unión Europea (27)	15764

La diversidad de cualquier producto es uno de los mayores atractivos de este frente al consumidor. El tomate tiene una gran riqueza varietal pudiéndose distinguir variedades con distinto aspecto exterior (forma, tamaño, color) e interior (sabor, textura, dureza) o variedades destinadas para uso industrial o en fresco. Las preferencias por un tipo determinado de tomate varían en función del país, tipo de población, uso al que se destina, etc.

El tomate se ha empleado como organismo modelo en estudios genéticos y de desarrollo (Tanksley & McCouch 1997). Es una hortaliza ideal para investigaciones fisiológicas, celulares, bioquímicas y de genética molecular. Es de fácil cultivo, tiene un ciclo de vida corto y es fácilmente manejable en horticultura. Se puede, además, cultivar *in vitro* y regenerar, haciendo posible su transformación genética (Gisbert *et al.* 2000). Han sido descritos numerosos marcadores del tomate. Existe un mapa de tomate público, en la página <http://sgn.cornell.edu/>, que posee una gran cantidad de marcadores

moleculares descritos en diferentes localizaciones cromosómicas de gran ayuda para la investigación del género y la Mejora Genética de la especie.

1.2. Valor nutricional del tomate

La calidad interna del tomate incluye aspectos organolépticos (dulzor, acidez, aroma, textura, perfume, firmeza) y nutritivos (vitaminas, minerales, azúcares).

El tomate contiene azúcares simples que le confieren un ligero sabor dulce y algunos ácidos orgánicos que le otorgan el sabor ácido característico. El tomate es una fuente importante de ciertos minerales (como el potasio y el magnesio) y sobre todo de sustancias con actividad antioxidante (vitaminas, pigmentos y compuestos fenólicos). De su contenido en vitaminas destacan la B1, B2, B5 y la vitamina C. Los carotenoides son uno de los principales grupos de pigmentos, como el licopeno (pigmento que da el color rojo característico al tomate) (Roselló y Nuez 2006).

1.3. Taxonomía del género *Solanum*

Desde la introducción del tomate en Europa en el siglo XVI fue reconocida su estrecha relación con el género *Solanum* por parte de los botánicos. Tournefort, en el siglo XVII, fue el primero en considerar al tomate en un género propio, *Lycopersicon*, basándose en el número de lóculos de los frutos. Sin embargo, un siglo después, Linneo clasificó a los tomates en el género *Solanum* en su libro '*Species Plantarum*' y describió *S. Lycopersicon* L. y *S. peruvianum* L., aunque su contemporáneo Miller describió el género *Lycopersicon* Mill. basándose en la clasificación de Tournefort (Peralta & Spooner 2000). Numerosos autores clásicos y modernos estuvieron de acuerdo con Miller y Tournefort (Müller 1940, Luckwill 1943, Correl 1958, D'Arci 1991, Rick *et al.* 1979, 1988; Warnock 1988).

Dos trabajos taxonómicos completos son los de Müller (1940) y Luckwill (1943). Müller consideró el género *Lycopersicon* como independiente y lo dividió en dos subgéneros, *Eulycopersicon* C. H. Müll. con dos especies de frutos anaranjados a rojizos y pilosos y *Eriopersicon* C. H. Müll. con cuatro especies de frutos verdes y pubescentes. Además describió una nueva especie *L. glandulosum* C. H. Müll. y dividió

a *L. peruvianum* en dos variedades y cuatro subespecies. Tres años después, Luckwill (1943) adoptó las mismas categorías supraespecíficas pero reconoció diferentes taxones intraespecíficos y cinco especies en el subgénero *Eriopersicon* describiendo a *L. pissisi* Phil. como una nueva especie. Estos tratamientos quedaron obsoletos debido al considerable aumento de especies y razas colectadas con posterioridad (Rick 1971, 1991; Taylor 1986).

Una relación de las diferentes clasificaciones se muestra en la tabla 1.3.

Rick (1960, 1979) propuso una clasificación basada en relaciones de cruzabilidad y aspectos morfológicos. Reconoció nueve especies silvestres de tomate y las clasificó en dos complejos. El complejo *Esculentum* incluyó siete especies, mayoritariamente autocompatibles (AC) y que cruzaban con facilidad con el tomate cultivado. Tres especies tienen frutos no pilosos y pigmentados y las demás tienen frutos verdes y pubescentes. En el complejo *Peruvianum* incluyó las especies autoincompatibles (AI) *S. peruvianum* y *S. chilense*, con frutos verdes y pubescentes que rara vez cruzan con las anteriores.

En la actualidad se acepta la clasificación de Child (1990), quien, basándose en caracteres morfológicos, incluyó a los tomates en *Solanum* subgénero Potatoe (G. Don) D'Arcy, sección *Lycopersicum*, subsección *Lycopersicum*, y los dividió en tres series *Lycopersicon* Child., *Neolycopersicon* (Correll) Child. y *Eriopersicon* (C. H. Müll.) Child. Esta clasificación está respaldada por estudios posteriores basados en datos de caracterización molecular, tanto de DNA cloroplástico como nuclear (Palmer y Zamir 1982, Miller & Tanksley 1990, Spooner *et al.* 1993).

Tabla 1.3. Taxonomía del tomate y especies silvestres relacionadas. Extraído de Peralta *et al.* 2008.

Müller 1940	Luckwill 1943	Rick 1979	Child 1990	Peralta <i>et al.</i> 2008
			Sección <i>Lycopersicon</i>	
Subgénero	Subgénero	Complejo Esculentum	Subsección <i>Lycopersicon</i>	Sección <i>Lycopersicon</i>
<i>Eulycopersicon</i>	<i>Eulycopersicon</i>		Serie <i>Lycopersicon</i>	Grupo <i>Lycopersicon</i>
<i>L. esculentum</i>	<i>L. esculentum</i>	<i>L. esculentum</i>	<i>S. lycopersicum</i>	<i>S. lycopersicum</i>
<i>L. pimpinellifolium</i>	<i>L. pimpinellifolium</i>	<i>L. pimpinellifolium</i>	<i>S. pimpinellifolium</i>	<i>S. pimpinellifolium</i>
		<i>L. cheesmaniae</i>	<i>S. cheesmaniae</i>	<i>S. cheesmaniae</i>
				<i>S. galapagense</i>
			Serie <i>Neolycopersicon</i>	Grupo
		<i>L. pennellii</i>	<i>S. pennellii</i>	Neolycopersicon
				<i>S. pennellii</i>
		<i>L. hirsutum</i>	Serie <i>Eriopersicon</i>	Grupo <i>Eriopersicon</i>
		<i>L. chmielewskii</i>	<i>S. hisutum</i>	<i>S. habrochaites</i>
		<i>L. parviflorum</i>	<i>S. chmielewskii</i>	
			<i>S. neorickii</i>	<i>S. huaylasense</i>
Subgénero <i>Eriopersicon</i>	Subgénero <i>Eriopersicon</i>	Complejo Peruvianum		<i>S. corneliomulleri</i>
<i>L. peruvianum</i>	<i>L. pissisi</i>	<i>L. peruvianum</i>	<i>S. peruvianum</i>	<i>S. peruvianum</i>
<i>L. chesmaniae</i>	<i>L. peruvianum</i>	<i>L. chilense</i>	<i>S. chilense</i>	<i>S. chilense</i>
<i>L. hirsutum</i>	<i>L. chesmaniae</i>			Grupo Arcanum
<i>L. glandulosum</i>	<i>L. hirsutum</i>			<i>S. arcanum</i>
	<i>L. glandulosum</i>			<i>S. chmielewskii</i>
				<i>S. neorickii</i>
			Subsección <i>Lycopersicoides</i>	Sección <i>Lycopersicoides</i>
			<i>S. lycopersicoides</i>	<i>S. lycopersicoides</i>
			<i>S. sitiens</i>	<i>S. sitiens</i>
			Subsección <i>Juglandifolium</i>	Sección <i>Juglandifolia</i>
			<i>S. juglandifolium</i>	<i>S. juglandifolium</i>
			<i>S. ochranthum</i>	<i>S. ochranthum</i>

El estudio de Spooner y Peralta (2001) con *S. peruvianum* posiciona a *S. juglandifolium* y a *S. ochantrum* dentro de la sección *Juglandifolium* y a *S. lycopersicoides* con *S. sitiens* dentro de la subsección *Lycopersicoides*. A la patata la enmarcan en la sección Petota dándoles tratamiento de clados hermanos.

Con la reciente descripción de *S. galapagense* S. C. Darwin et Peralta (Darwin *et al.* 2003) como una especie diferente a partir del antiguo taxón *L. chesmaniae f. minor* se reconocieron diez especies de tomate silvestre. Peralta, Knapp y Spooner, (2005), utilizaron las mismas accesiones que Peralta y Spooner (2001) pero basándose en la morfología como atributo independiente y llegaron a las mismas conclusiones. Además reconocen cuatro especies a partir de *S. peruvianum* L. sensu lato: *S. peruvianum* sensu estricto, *S. corneliomuelleri* J. F. Macbr, *S. arcanum* Peralta y *S. huaylasense* Peralta. En la actualidad se consideran 13 especies de tomate silvestre incluyendo al tomate cultivado (*S. lycopersicum*) y a sus especies silvestres (Peralta *et al.* 2008, p. 13). Divide la sección en cuatro grupos: el grupo *Lycopersicon* (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* (L. Riley) Fosberg y *S. galapagense*), el grupo *Neolycopersicon* (*S. pennellii*), el grupo *Eriopersicon* (*S. peruvianum* s. str., *S. corneliomuelleri*, *S. huaylasense*, *S. habrochaites* S. Knapp et D. M. Spooner y *S.*

chilense) y el grupo Arcanum (*S. arcanum*, *S. chmielewskii* (C. M. Rick *et al.*) D. M. Spooner *et al.* y *S. neorickii* D. M. Spooner *et al.*).

Zuriaga *et al.* (2009b) tras sus análisis con marcadores AFLPs y dos secuencias genéticas nucleares han diferenciado tres grupos en la sección *Lycopersicon*: el grupo *Lycopersicon*, formado por *S. pimpinellifolium*, *S. lycopersicum*, *S. cheesmaniae*, y *S. galapagense*; el grupo *Arcanum*, constituido por *S. chmielewskii*, *S. neorickii*, *S. arcanum* y *S. huaylasense*, y el grupo *Eriopersicon* formado por *S. peruvianum* y *S. chilense*. *Solanum pennellii* y *S. habrocahites* no fueron incluidas en ningún grupo, pero resultaron ser las más cercanas a la especie *S. lycopersicoides* utilizada como fuera de grupo.

La mayoría de taxónomos (y recientemente los mejoradores) están adoptando *Solanum* como el nombre genérico de los tomates; ver http://www.sgn.cornell.edu/help/about/solanum_nomenclature.html

Con esto, la familia Solanaceae contiene entre 3.000 y 4.000 especies repartidas en 90 géneros (Knapp *et al.* 2004) siendo *Solanum* el más numeroso con 1.500 especies estimadas (Bohs 2007). Este género engloba a los tomates y a sus especies silvestres relacionadas dentro de la sección *Lycopersicon*.

1.4. Centro de Origen y diversificación del tomate.

La historia de la domesticación del tomate (*Solanum lycopersicum* L.) es curiosa. Los tomates silvestres se extienden por la costa oeste de América del Sur. Sin embargo, no hay evidencias de que en la época precolombina los indios suramericanos conociesen la existencia del tomate. En la región Andina no existió un nombre para el tomate, tampoco formaba parte de su cultura y no se han encontrado restos arqueológicos del mismo (Harlan 1992, Esquinas-Alcázar *et al.* 1995). La domesticación tuvo lugar en México donde los tomates silvestres sí son conocidos, aunque sólo se encuentran plantas silvestres en el sur. Los indios mexicanos sí le dieron un nombre al tomate en la lengua Nahuatl (azteca), le llamaron 'jitomatle'. El tomate estaba bien integrado en su cultura, lo cultivaban, comercializaban con él y preparaban gran cantidad de platos. Sin embargo, nuestra palabra 'tomate' deriva de la palabra india 'tomatl' que hace referencia a otra hortaliza cultivada del género *Physalis*. Esto quiere

decir que los aztecas usaron la palabra 'jitomatl' para el tomate y que *Physalis* fue domesticada anteriormente, derivando posteriormente el tomate como otro tipo de 'tomatl' (Harlan 1992). El tomate se introdujo en Europa por primera vez a mediados del siglo XVI (Kalloo 1991). Las primeras variedades debieron ser de color amarillo ya que en Europa se le denominaba *pomodoro* o *pomme d'or* y todavía se le denomina así en Italia. Su aceptación fue lenta, ya que el tomate pertenece a la familia de la belladona y se pensaba que podía ser venenoso. La toxicidad de la belladona, perteneciente a la familia de las Solanáceas, está producida por alcaloides. El alcaloide predominante en el tomate es la tomatina, que se encuentra en elevada concentración en el follaje y en el fruto verde y que se degrada a componentes inertes al madurar el fruto, pero incluso a grandes dosis es mucho menos peligroso que los alcaloides de la hierbamora, la belladona y otras especies (Rick 1978).

Como ya se ha dicho, todas las especies silvestres del tomate tienen su centro de origen en el oeste de América del Sur. Están distribuidas a lo largo de la costa y en los Andes, desde Ecuador hasta el norte de Chile y con una especie endémica de las islas Galápagos. Los tomates silvestres crecen en una gran variedad de hábitats, desde casi el nivel del mar hasta los 3300 metros de altitud (Rick 1973, Taylor 1986), aunque se han encontrado ejemplares sobre los 3900 metros. Estos hábitats incluyen desde la árida costa Pacífica hasta las montañas méxicas de los andes. Las poblaciones silvestres de tomate crecen a diferentes alturas en estos angostos valles, están aisladas geográficamente y están adaptadas a determinados microclimas y condiciones del suelo. Esta diversidad geográfica, de hábitats y de climas ha contribuido a la diversidad de las especies silvestres de tomate (Warnock 1988). Se apoya la teoría de *Solanum lycopersicon* como el ancestro de los tomates cultivados en todo el mundo. Todas las especies relacionadas con él son diploides ($2n = 24$) (Rick *et al.*, 1979) (aunque puede existir alguna forma tetraploide) y su sistema reproductivo varía desde la alogamia autoincompatible a algamas facultativas o autógamas (Rick 1963). La mayoría de las especies de tomate silvestre se pueden cruzar con el tomate cultivado (aunque a veces con dificultad) y son de gran importancia en los programas de Mejora como fuentes de resistencia frente a enfermedades o plagas o caracteres agronómicos de interés (Stevens & Rick 1986).

Tras la difusión del tomate por el mundo perdió gran cantidad de diversidad a causa de la importación de sólo algunos individuos. Esto provocó un cuello de botella

que redujo la diversidad genética de la especie, limitando, a su vez, el número de alelos (Rick 1976). Este hecho llevó asociado un proceso de deriva genética. Posteriormente se produjo una nueva reducción de variabilidad debido a procesos de selección natural y selección artificial. De forma natural cada cultivo se adaptó a las condiciones de suelo, clima, plagas o patógenos naturales de la zona donde se cultivaba. Artificialmente el agricultor hizo una selección en base a diferentes atributos distintos en cada zona, tamaños, colores, formas, etc. Actualmente al observar gran variedad de formas, colores o tamaños se podría pensar erróneamente que hay mucha variabilidad, pero los análisis genéticos demuestran que hay muy poca diversidad genética. La mayor parte de caracteres están codificados por muy pocos genes. Estos genes muchas veces actúan como genes reguladores y pequeñas variaciones en lugares concretos pueden provocar un cambio muy amplio. Esto explicaría la diversidad aparente.

Debido a esta estrecha base genética del tomate cultivado, las especies silvestres de tomate suponen una fuente importante de genes de interés para la mejora del mismo.

1.5. *Solanum habrochaites*. Importancia y principales características de la especie

1.5.1. Descripción de la planta

Las plantas de *S. habrochaites* son arbustos decumbentes o enredaderas perennes de hasta 6 m de longitud, escasa a densamente pubescentes. Las formas menos pubescentes están cubiertas con largos tricomas que pueden ser glandulares. Presentan simpodios de tres hojas. Éstas son imparipinadas, de 7 a 30 cm de longitud con 3 a 5 pares de foliolos y varios foliolillos con márgenes aserrados. Presentan inflorescencias de 10 a 30 cm de longitud con 20 a 30 flores (Figura 1.2). Las flores son hermafroditas, actinomorfas y pentámeras. Son muy vistosas, de color amarillo intenso, con corolas divididas de 2 a 4 cm de diámetro. Presentan el cono estaminal en forma de tubo, dehiscencia longitudinal y estigma normalmente exerto. Los frutos, de 1 a 1.5 cm de diámetro, son globosos, biloculares y de color verde pálido a intenso ya sean densa o escasamente pubescentes (Figura 1.3). Las semillas son ovaladas de 2 a 3 mm de longitud y de color marrón oscuro.



Figura 1.2. Inflorescencias de *S. habrochaites*.



Figura 1.3. Detalle de frutos de *S. habrochaites* con diferentes niveles de pubescencia.

Müller (1940) subdividió la especie en dos taxones, las formas *typicum* y *glabratum*, en función de su morfología y sistema de reproducción. *S. habrochaites* f. *typicum* se caracteriza por tener tallos, hojas y frutos muy pilosos y un estigma muy exerto. Con flores de gran tamaño y con corolas con divisiones menos profundas que el tomate cultivado, frutos verdes con rayas oscuras y las semillas de color marrón y cubierta lisa. Es la forma más variable y de distribución más amplia. Es en gran medida autoincompatible (Rick y Chetelat 1991) aunque también pueden autofecundarse (Rick *et al.* 1979). Por el contrario, describió que la forma *glabratum*, localizada en Ecuador, se reconoce por la escasa pubescencia de hojas y tallos y por poseer una corola más pequeña (Müller 1940). Estas flores, más pequeñas tienen tendencia a autofecundarse. Además, esta forma cruza en ambas direcciones y sin dificultad con el tomate.

Rick *et al.* (1979), sin embargo, no siguieron esta nomenclatura, ya que encontraron entradas con características morfológicas intermedias entre ambas formas y entradas con escasa pubescencia que no presentaban las flores características de la

forma *glabratum*. Por estos motivos las dividieron en entradas ecuatorianas y peruanas, sin asignarlas a ningún taxón.

1.5.2. Hábitat

La especie *S. habrochaites* tiene una extensa distribución geográfica desde el sur de Ecuador (región de Loja, latitud 4°S) hasta el centro-sur de Perú (Taylor 1986). Ocupa las zonas de mayor altitud del género. Se encuentra habitualmente en valles húmedos y elevados, entre 500 y 3300 metros de altitud (Rick 1973), en regiones con un rango de temperaturas medias de 10 a 21°C (Warnock 1991). Aunque habita las mismas latitudes que *S. pimpinellifolium* tienen diferentes preferencias en cuanto a altitud. Mientras que esta última prefiere las áreas costeras, *S. habrochaites* crece a grandes altitudes en los valles andinos, excepto en Ecuador donde se han encontrado, en lugares próximos a la costa, ejemplares de ambas especies (Rick *et al.* 1979).

Se ha descrito que la forma *glabratum* se encuentra en la zona norte de distribución de la especie, desde el suroeste de Ecuador hasta la frontera norte de Perú. *S. habrochaites f. typicum* es más variable que *f. glabratum* y comúnmente se encuentra a grandes altitudes en el norte y centro de Perú (Taylor 1986).

1.5.3. Empleo en la mejora del tomate

Estudios moleculares en *Solanum* han demostrado que una gran parte de la variabilidad del género se presenta en las especies silvestres autoincompatibles y que aproximadamente el 17% de la variabilidad alélica presente en *S. habrochaites* no se encuentra en *S. lycopersicon* (Miller & Tanksley 1990).

Esta especie es de gran importancia en la Mejora Genética Vegetal ya que presenta resistencias a plagas de gran importancia en el cultivo del tomate como son, entre otras: *Tuta absoluta* (Leite *et al.* 2001), *Tetranychus urticae* (Guo *et al.* 1993), *Spodoptera exigua* (Lin *et al.* 1987), *Heliothis Zea* (Dimock & Kennedy 1983), *Plusia chalcites*, *Heliothis armígera*, *Trialeurodes vaporarium* (Juvik *et al.* 1982), *Aphis spp* (Williams *et al.* 1980) *Myzus persicae* (Kenedy & Yamamoto 1979) y *Lyriomiza spp.* (Webb *et al.* 1971).

La resistencia a enfermedades causadas por insectos se debe principalmente a la secreción por los tricomas de determinados compuestos químicos (Kalloo 1993). Un compuesto fenólico, el ácido clorogénico, presente en los tricomas glandulares inhibe el crecimiento de los insectos. Otros compuestos asociados con la resistencia a insectos son los sesquiterpenos (Zhang *et al.* 2008). Las hojas de *S. habrochaites f. glabratum* contienen zingibereno, un sesquiterpeno asociado a una resistencia temporal al escarabajo de la patata de Colorado (Carter *et al.* 1989). Incluso se ha patentado un repelente para artrópodos cuyo principio activo es 2-undecanona procedente de esta especie (Witting-Bissinger 2008).

Algunos caracteres morfológicos como el grosor del córtex en los tallos de *S. habrochaites* también juegan un papel importante contra algunos artrópodos (Quiros *et al.* 1977).

Así mismo, presenta resistencias a diversas enfermedades causadas por hongos como *Phytophthora infestans* y *Alternaria solani* (Foolad *et al.* 2008), a bacteriosis como *Corynebacterium michiganense* y *Pseudomonas tomato* (Kalloo 1991) y diferentes virosis como TSWV (Soler *et al.* 2000), TMV, TLCV, TYLCV (Vidavsky y Czosnek 1998) y a los nematodos *Globodera pallida* (Kalloo 1991) y diferentes especies del género *Meloidogyne* (Cortada *et al.* 2009).

Existen en el mercado portainjertos de tomate de diferentes casas comerciales procedentes de cruces interespecíficos con *S. habrochaites* resistentes a diferentes virosis, bacteriosis, hongos y nematodos (Cortada *et al.* 2008).

También se ha encontrado en esta especie tolerancia al frío y a las heladas (Paterson *et al.* 1987), así como caracteres de interés en la mejora de la calidad, como un elevado contenido en azúcares y ácidos orgánicos (Bucheli *et al.* 1999, Galiana-Balaguer *et al.* 2006).

En la actualidad se dispone de líneas de introgresión (IL), que son líneas *quasi* isogénicas homocigotas, en las que está representado todo el genoma de la entrada de *S. habrochaites* LA1777 en el fondo genético de tomate. Éstas han sido desarrolladas por el grupo de Steve Tanksley en la Universidad de Cornell y están a disposición pública en el Tomato Genetics Resource Center (TGRC) de la Universidad de California en Davis. Estas líneas facilitan en gran medida la identificación y el uso de nueva variación

genética encontrada en esta especie silvestre (Zamir 2001), además de ser de gran utilidad en el mapeo de genes de interés (Ronen *et al.* 1999) y QTLs (quantitative trait loci) (Monforte & Tanksley 2000).

1.6. Los marcadores genéticos en la conservación y uso de los recursos fitogenéticos

Los marcadores moleculares permiten obtener una gran cantidad de información sobre la diversidad genética y las relaciones filogenéticas en el germoplasma utilizado por el mejorador (Tanksley & McCouch, 1997). La aplicación de diversos métodos estadísticos (análisis multivariante de componentes principales, análisis de agrupamiento, etc.) a la información obtenida permite calcular distancias genéticas y representaciones gráficas que pueden ser interpretadas en clave filogenética.

Para que un marcador sea considerado un buen marcador utilizable en mejora tiene que ser polimórfico, codominante, no epistático, neutro e insensible al medio (García-Mas *et al.* 2000).

Los marcadores morfológicos se utilizaron durante años hasta la década de los 60, siendo generalmente fenotipos de fácil interpretación visual. Se identificaron algunos marcadores morfológicos ligados a genes de importancia económica, pudiendo ser utilizados en programas de mejora. Los marcadores morfológicos controlados por un sólo locus pueden ser usados como marcadores genéticos si su expresión es temprana, reproducible en cualquier ambiente y están estrechamente ligados con un carácter de interés agronómico.

Con el desarrollo de marcadores isoenzimáticos y de otros tipos de proteínas (marcadores bioquímicos) se amplió el número de marcadores genéticos disponibles. Los isoenzimas suelen presentar una herencia codominante, no están habitualmente influenciados por el medio y tienen una interacción nula o no identificable con otros marcadores. A principios de los 80, la utilización de enzimas de restricción permitió el análisis del polimorfismo de fragmentos de restricción del ADN (RFLPs) y ya más recientemente, el desarrollo del método de amplificación en cadena de segmentos de ADN utilizando una ADN polimerasa (PCR) facilitó la obtención de nuevos tipos de marcadores. Más tarde, combinando ambas técnicas, se han desarrollado los marcadores

AFLPs (Amplified Fragments Length Polymorphism) (Vos et al. 1995). El descubrimiento y estudio de diversas clases de secuencias repetitivas de ADN, llamadas mini y microsátélites (SSRs), ha aportado otra rica fuente de polimorfismo genético. Los marcadores de ADN representan un avance cuantitativo al añadir a las cualidades de los isoenzimas la de poder encontrarse en un número prácticamente ilimitado. Además pueden acceder a un rango de variabilidad mucho mayor, que abarca tanto las regiones de ADN codificantes (incluida la variación debida a la degeneración del código genético) como las que no lo son (intrones, zonas de ADN repetitivo, pseudogenes, etc.). El polimorfismo más simple que cabe esperar es la sustitución en una secuencia de un nucleótido por otro, estos marcadores son SNP (Single Nucleotide Polimorfism) y están siendo muy utilizados en Mejora por su elevado número y su amplia distribución en el genoma. La mejora cualitativa y cuantitativa aportada por los marcadores de ADN ha sido crucial para que se hayan convertido en una herramienta enormemente potente para el análisis genético y el estudio de la diversidad de las poblaciones silvestres para su utilización en el uso y conservación de los recursos fitogenéticos.

1.7 Estudios de variabilidad en *Solanum habrochaites*

Se han realizado numerosos estudios de variabilidad en el género *Solanum*, pero solamente unos pocos estudian con detalle la variabilidad dentro de la especie *S. habrochaites*, a pesar del gran potencial que tiene *Solanum habrochaites* como parental donador de caracteres de interés agronómico en programas de mejora.

El trabajo más exhaustivo fue el realizado por Rick *et al.* en 1979. Analizaron entradas de todo el rango de distribución de la especie conocido hasta ese momento. Realizaron una caracterización morfológica importante de la colección, estudiando el sistema de reproducción de las entradas y relacionando estos datos con el lugar de colecta de cada una de ellas. Completaron su estudio con una caracterización molecular, utilizando isoenzimas. Observaron que las entradas de los distintos valles son considerablemente diferentes. Las poblaciones que presentan mayor variabilidad son las del centro del rango de distribución de la especie, es decir el norte de Perú, en las provincias de Piura, Cajamarca y La Libertad. Estas entradas son en su mayoría autoincompatibles. Conforme nos desplazamos hacia el norte y el sur las poblaciones presentan menor variabilidad y tienden a la autocompatibilidad. Estos autores sugieren

que las poblaciones centrales representan los biotipos más antiguos a partir de los cuales derivaron el resto. Al migrar las especies hacia el norte y el sur en su amplio rango de distribución las especies autoincompatibles fueron sustituidas por autocompatibles, ya fuera debido al efecto fundador, a la ventaja de fijar uno o pocos genotipos prósperos, a la dificultad para obtener el tamaño poblacional necesario o a la escasez de insectos polinizadores.

Sacks y St. Clair (1998) estudiaron la eficiencia para introgresar genes de *Solanum habrochaites* en tomate en función de la variación genética de las entradas y la variación medioambiental, cuantificando el número de semillas viables obtenido en los cruces interespecíficos. Observaron que los cruces de tomate con las entradas colectadas más al norte producían un mayor número de semillas que los cruces con las entradas del sur, con lo que la introgresión de genes de interés que estén en estas últimas será más complicada.

Ercolano *et al.* (2005) realizaron un estudio molecular empleando RAPDs. Caracterizaron únicamente 15 entradas de Ecuador y Perú. Encontraron que podían diferenciar las entradas de la forma *glabratum* y *typicum* utilizando este tipo de marcadores y observaron una mayor variabilidad en las entradas pertenecientes a la forma *typicum*.

Objetivos

2. Objetivos

- Realizar un análisis exhaustivo de la variabilidad molecular existente en la especie *S. habrochaites* en todo el rango de distribución natural de la especie.
- Estudiar la estructura poblacional de *S. habrochaites* en su rango de distribución natural.
- Comparar la utilidad de los marcadores microsatélite y AFLPs en el estudio de la variabilidad molecular de *S. habrochaites* para el manejo de las colecciones de germoplasma.

Material y Métodos

[...] cuando comparamos la multitud de razas de plantas agrícolas, culinarias, de huerta y de jardín, utilísimas al hombre en las diferentes estaciones y para diferentes fines, o tan hermosas a sus ojos, tenemos, creo yo, que ver algo más que simple variabilidad.

Charles Darwin. *“El origen de las especies”*

3. Material y Métodos

3.1. Material Vegetal

En total se analizaron 91 entradas de *S. habrochaites* de diferente origen: 36 de las entradas fueron cedidas por el Tomato Genetics Resource Center (TGRC) (entradas con nombres precedidos por LA) y 55 por el Banco de Germoplasma del Instituto de Conservación y Mejora de Agrobiodiversidad Valenciana (COMAV), colectadas por su equipo en colaboración con la Universidad Nacional de Piura, Universidad Nacional de Trujillo (Perú), Universidad Nacional de Loja y Universidad Nacional Pedro Ruíz Gallo (Ecuador).

Con el fin de facilitar el estudio de las entradas en todo su rango de distribución se dividieron en siete grupos por criterios ecogeográficos, según pertenecieran a las mismas cuencas hidrográficas o cuencas conectadas o existieran barreras geográficas que las separasen. Para ello se emplearon el Mapa Bioclimático de Ecuador y el Mapa Hidrográfico del Perú y se utilizó el programa Google Earth (v.5.0). En el caso de existir pocas entradas en esas áreas se agruparon para facilitar el análisis. Las áreas se denominaron de norte a sur zonas A, B, C, D, E, F y G. (Figuras 3.1.1, 3.1.2 y Tabla 3.1.1).

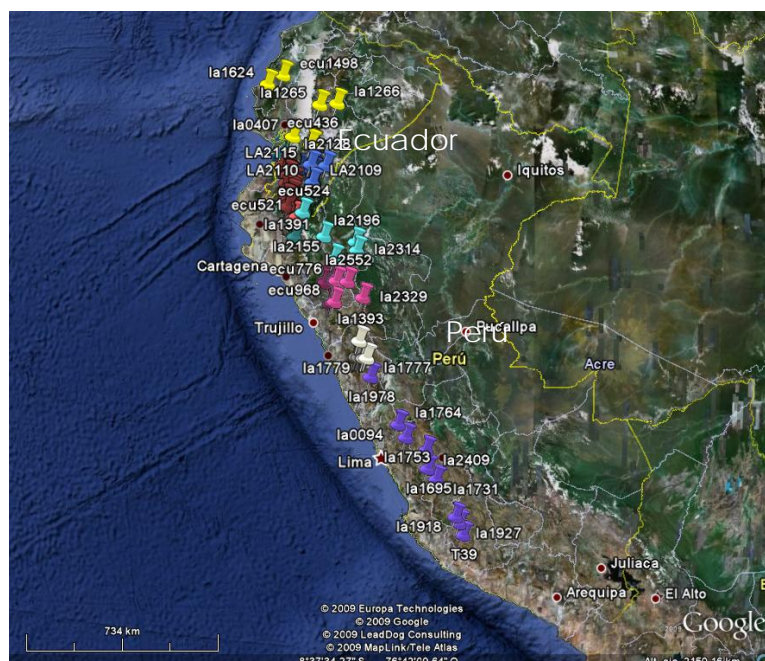


Figura 3.1.1. Fotografía satélite de Ecuador y de Perú donde se indican por colores las zonas de colecta de *S. habrochaites* y las entradas colectadas en cada zona: amarillo, zona A; azul, zona B; rojo, zona C; turquesa, zona D; rosa, zona E; blanco, zona F; morado, zona G. Fuente: Google Earth.

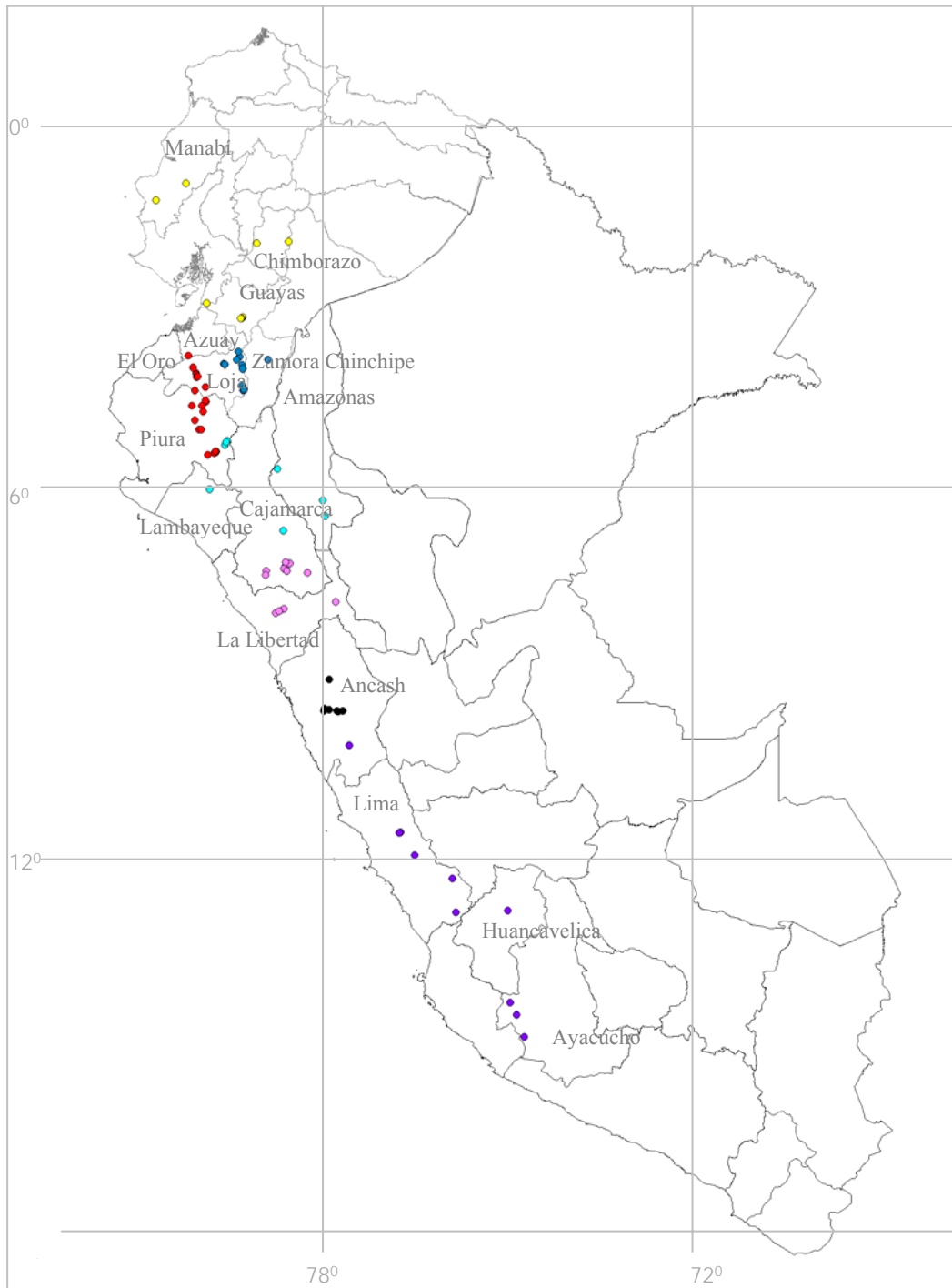


Figura 3.1.2 Detalle de mapa de Ecuador y Perú en el que se representan por colores las diferentes entradas de *S. habrochaites* colectadas en cada zona geográfica: amarillo, zona A; azul, zona B; rojo, zona C; turquesa, zona D; rosa, zona E; blanco, zona F; morado, zona G. Fuente: Programa Diva Gis.

Zona A (Centro de Ecuador) (8 entradas): en esta zona se incluyeron las entradas colectadas más al norte, en el centro de Ecuador. Entradas de Manabí, LA1624 (*glabratum*) y ECU1498; Guayas, LA0407 (*glabratum*); Chimborazo, LA1265 (*glabratum*) y LA1266 (*glabratum*); y Azuay, ECU0434, ECU0436 y ECU0950. Las entradas clasificadas como forma *glabratum* lo han sido por el TGRC.

Zona B (Noreste de Loja) (13 entradas): entradas del noreste de la provincia de Loja, en la cuenca del Río Zamora LA1252, LA2109, LA2110, LA2115, pertenecientes a la forma *glabratum*, colectadas a grandes altitudes, sobre 2000 metros, y la entrada LA2128 (*glabratum*) colectada al este de ellas y que es la única de la provincia de Zamora Chinchipe y que está colectada a 1000 metros. Se incluyeron también las entradas colectadas en la cuenca del Río Catamayo ECU0311, ECU0336, ECU0359, ECU0653, ECU0533, ECU0537, ECU0534 y ECU0536.

Zona C (Oeste de Loja-Piura) (29 entradas): entradas del oeste de la provincia de Loja en los cantones de Celica ECU0529, ECU0531, ECU0532, ECU0657; Sozoranga LA2864 (*glabratum*) y Macará LA2863. Todas estas entradas pertenecen a la cuenca del Río Chira (en la parte ecuatoriana Río Catamayo) al igual que las entradas del departamento de Piura en la provincia de Morropón cerca de Ayabaca: ECU0911, ECU0922, ECU0924, ECU0901, ECU0888 (distribuidas en altitudes desde 550 m a 2120 m). Las entradas colectadas en la cuenca del Río Piura cerca de Serran en la carretera a Huancabamba ECU0851, ECU0865, ECU0866, ECU0848 y ECU0842, ECU0521, ECU0522, ECU0540, ECU0541, ECU0542, ECU0543, ECU0544, ECU0549, ECU0652, ECU0656 y LA1739, a altitudes desde 725 a 1250m. Este grupo también incluye entradas del suroeste de Ecuador y Norte de Perú: ECU0527 (Las Lajas, El Oro) a 450m y ECU0528 (Puyango, Loja) colectada a 1250m.

Zona D (Huancabamba) (10 entradas): esta zona engloba entradas colectadas en los departamentos de Piura, Lambayeque, Cajamarca y Amazonas. Pertenecen las entradas LA1718, ECU0523, ECU0524, ECU0546, ECU0547, colectadas en la cuenca del Río Huancabamba, zona que ha sido descrita por su gran diversidad (Weigend, 2002). Se incluyeron también las entradas LA2155, LA2314, LA2196 y LA1391 por estar colectadas en lugares cercanos aunque en la cuenca del Río Marañón (desemboca en el Atlántico). La entrada colectada un poco más al sur, ECU0526, en la cuenca del Río Olmos a 230m de altitud, también pertenece a este grupo.

Zona E (Cajamarca-La Libertad) (13 entradas): se incluyeron las entradas colectadas en el sur del departamento de Cajamarca, en el valle del río Jequetepeque cerca de la ciudad de Cajamarca, colectadas a grandes altitudes alrededor de los 3000m, entradas ECU0753, ECU0755, ECU0762, ECU0776, ECU0816 y LA0386, ECU1352, ECU1353, ECU2552. También las entradas colectadas en el norte del departamento de

La Libertad en el valle del río Moche, entre Otuzco y Trujillo alrededor de los 2000m de altitud son ECU0819, ECU0968 y ECU0800 están en este grupo. La entrada LA2329 pertenece a este grupo aunque está colectada cerca de Aricapampa en La Libertad.

Zona F (Ancash) (7 entradas): entradas colectadas en el valle del río Casma, en el departamento de Ancash LA1361, LA1775, LA1777, LA1779, PE0039 y PE0043. Colectadas entre 1490 y 3300 m de altura. También se incluyó la entrada colectada en el mismo departamento más al norte LA1393, a 2800m, en la cuenca del Río Santa, cerca de Caraz.

Zona G (Sur de Perú) (10 entradas): entradas de Lima, Huancavelica y Ayacucho, de centro a sur de Perú. En el departamento de Lima: LA1978 en la cuenca del Río Fortaleza, cercano al callejón de Huaylas; en la cuenca del Río Chillón, cerca de Cantas, las entradas LA1764 y LA0094; la entrada LA1753 en la cuenca del Río Rimac, cerca de Matucana; LA1695 y LA2409 en el valle del río Cañete y LA1731 en el valle del río San Juan, ya en Huancavelica. LA1918 se colectó en la cuenca del río Palpa y LA1927 en la cuenca del río Ingenio, en Ayacucho. También hemos incluido en este grupo la entrada T0039 colectada al este de Nazca.

Tabla 3.1.1. Relación de entradas de *S. habrochaites*, origen y localización del lugar de colecta, altitud y zona asignada.

Nº de entrada	Origen	Departamento (País)	Localización	Latitud (Sur)	Longitud (Oeste)	Altitud (m)	Zona
ECU1498	COMAV	Manabí (E)	Altamira	01°02'56"	80°05'25"	456	A
LA1624	TGRC	Manabí (E)	Monta-Guayaquil	01°19'48"	80°34'30"	363	A
LA1266	TGRC	Chimborazo (E)	Pallatanga	01°59'00"	78°27'00"	3508	A
LA1265	TGRC	Chimborazo (E)	Río Chimbo	02°01'00"	78°58'00"	2494	A
LA0407	TGRC	Guayas (E)	Mirador	02°59'00"	79°46'00"	19	A
ECU0434	COMAV	Azuay (E)	Caledonias	03°12'33"	79°11'30"	1924	A
ECU0950	COMAV	Azuay (E)	Cuenca	03°13'49"	79°13'10"	1613	A
ECU0436	COMAV	Azuay (E)	Puente de Rircai	03°13'56"	79°13'05"	1560	A
ECU0336	COMAV	Loja (E)	Barrio Gualán	03°45'56"	79°15'31"	2364	B
LA2115	TGRC	Loja (E)	Pucala	03°51'00"	79°14'00"	2701	B
LA2128	TGRC	Zamora-Chinchipe (E)	Zumbi	03°53'29"	78°46'43"	820	B
ECU0311	COMAV	Loja (E)	Cachipampa	03°53'59"	79°16'55"	2262	B
ECU0359	COMAV	Loja (E)	Guadarín	03°57'34"	79°29'05"	2250	B
ECU0537	COMAV	Loja (E)	Las Chinchas	03°58'00"	79°28'00"	2119	B
ECU0653	COMAV	Loja (E)	Las Chinchas	03°58'00"	79°28'00"	2119	B
LA1252	TGRC	Loja (E)	Loja	03°59'00"	79°12'00"	2079	B
ECU0533	COMAV	Loja (E)	La Argelia	04°02'12"	79°11'20"	2448	B
ECU0534	COMAV	Loja (E)	Vilcabamba	04°18'00"	79°12'00"	2094	B
LA2109	TGRC	Loja (E)	Yangana	04°18'00"	79°12'00"	2295	B

LA2110	TGRC	Loja (E)	Yangana	04°18'00"	79°12'00"	2295	B
ECU0536	COMAV	Loja (E)	Yangana	04°21'45"	79°10'12"	1855	B
ECU0527	COMAV	El Oro (E)	Puyango Viejo	03°52'60"	80°05'60"	706	C
ECU0528	COMAV	Loja (E)	Alamor	04°01'02"	79°58'30"	1257	C
ECU0529	COMAV	Loja (E)	Quillangas	04°06'20"	79°55'45"	1866	C
ECU0531	COMAV	Loja (E)	Celica	04°07'45"	79°55'10"	1587	C
ECU0657	COMAV	Loja (E)	Celica	04°09'21"	79°54'20"	1283	C
ECU0532	COMAV	Loja (E)	Celica	04°10'00"	79°55'20"	927	C
LA2864	COMAV	Loja (E)	Sozorango	04°20'00"	79°47'00"	1038	C
LA2863	TGRC	Loja (E)	Macara	04°23'00"	79°57'00"	445	C
ECU0922	COMAV	Piura (P)	Jilili	04°33'00"	79°47'00"	1113	C
ECU0911	COMAV	Piura (P)	Sicches	04°33'53"	79°46'23"	1254	C
ECU0888	COMAV	Piura (P)	Las Lomas	04°37'52"	79°59'53"	479	C
ECU0924	COMAV	Piura (P)	Montero	04°38'00"	79°50'00"	1228	C
ECU0901	COMAV	Piura (P)	Pingola	04°43'16"	79°48'54"	1532	C
ECU0865	COMAV	Piura (P)	Frías	04°52'00"	79°57'00"	1000	C
ECU0866	COMAV	Piura (P)	Frías	04°52'00"	79°57'00"	3050	C
ECU0848	COMAV	Piura (P)	Quebrada Chilique	05°01'00"	79°53'00"	1800	C
ECU0851	COMAV	Piura (P)	Faical	05°01'00"	79°51'00"	1925	C
ECU0842	COMAV	Piura (P)	Paltashaco	05°06'00"	79°54'00"	950	C
ECU0549	COMAV	Piura (P)	Palambra	05°22'49"	79°36'50"	1117	C
LA1739	TGRC	Piura (P)	Canchaque a Serrán	05°23'00"	79°37'00"	1037	C
ECU0656	COMAV	Piura (P)	Canchaque	05°23'05"	79°37'20"	934	C
ECU0544	COMAV	Piura (P)	Canchaque	05°23'18"	79°36'48"	1182	C
ECU0540	COMAV	Piura (P)	Quinta San Juan Villa	05°23'36"	79°37'34"	915	C
ECU0541	COMAV	Piura (P)	Quinta San Juan Villa	05°23'36"	79°37'34"	915	C
ECU0521	COMAV	Piura (P)	Canchaque	05°24'18"	79°38'06"	704	C
ECU0522	COMAV	Piura (P)	Canchaque	05°24'18"	79°38'06"	704	C
ECU0652	COMAV	Piura (P)	Las Huacas-Canchaque	05°24'34"	79°38'10"	854	C
ECU0543	COMAV	Piura (P)	Las Huacas-Canchaque	05°24'45"	79°38'32"	633	C
ECU0542	COMAV	Piura (P)	Las Huacas-Canchaque	05°26'07"	79°44'37"	232	C
ECU0523	COMAV	Piura (P)	Comederos Bajos	05°12'53"	79°26'10"	2240	D
ECU0546	COMAV	Piura (P)	Comederos Bajos	05°12'53"	79°26'10"	2240	D
ECU0547	COMAV	Piura (P)	Comederos Bajos	05°12'53"	79°26'10"	2240	D
LA1740	TGRC	Piura (P)	Huancabamba	05°14'00"	79°27'00"	1990	D
LA1718	TGRC	Piura (P)	Huancabamba	05°14'00"	79°27'00"	1990	D
ECU0524	COMAV	Piura (P)	Huancabamba	05°16'48"	79°28'16"	2357	D
LA1391	TGRC	Cajamarca (P)	Bagua a Olmos	05°40'00"	79°38'00"	425	D
ECU0526	COMAV	Lambayeque (P)	Carretera Olmos- Jaén	05°59'42"	79°43'03"	213	D
LA2196	TGRC	Amazonas (P)	Río Utcabamba	06°10'30"	77°54'30"	1608	D
LA2314	TGRC	Amazonas (P)	San Francisco	06°25'00"	77°52'00"	2149	D
LA2155	TGRC	Cajamarca (P)	Bambamarca	06°39'00"	78°32'30"	2804	D
LA0386	TGRC	Cajamarca (P)	Cajamarca a Celendin	07°09'36"	78°30'00"	2683	E
ECU0816	COMAV	Cajamarca (P)	Dirección Cajabamba	07°11'05"	78°26'5"	2752	E
ECU0753	COMAV	Cajamarca (P)	Dirección San Juan	07°12'29"	78°30'17"	2966	E
ECU0755	COMAV	Cajamarca (P)	Dirección San Juan	07°13'47"	78°29'14"	3161	E
ECU0762	COMAV	Cajamarca (P)	Dirección San Juan	07°16'14"	78°31'48"	2296	E
LA1352	TGRC	Cajamarca (P)	Rupe	07°18'00"	78°48'30"	1795	E
ECU0776	COMAV	Cajamarca (P)	San Juan	07°18'17"	78°28'39"	2367	E
LA2552	TGRC	Cajamarca (P)	Las Flores	07°20'00"	78°09'00"	2371	E

LA1353	TGRC	Cajamarca (P)	Contumaza	07°22'00"	78°49'00"	2559	E
ECU0968	COMAV	La Libertad (P)	Quirivilca	07°55'00"	78°32'00"	3231	E
ECU0800	COMAV	La Libertad (P)	Cruce Otuzco-Trujillo	07°57'00"	78°36'00"	2714	E
ECU0819	COMAV	La Libertad (P)	Cruce Otuzco-Trujillo	07°58'48"	78°39'37"	2752	E
LA1393	TGRC	Ancash (P)	Huaraz-Caraz	09°03'00"	77°48'30"	2251	F
LA1775	TGRC	Ancash (P)	Río Casma	09°31'00"	77°53'00"	2005	F
PE0043	COMAV	Ancash (P)	Cullaspungro	09°32'00"	77°48'00"	2658	F
PE0039	COMAV	Ancash (P)	Quita	09°33'00"	77°41'00"	3609	F
LA1777	TGRC	Ancash (P)	Río Casma	09°33'00"	77°35'24"	3742	F
LA1361	TGRC	Ancash (P)	Pariacoto	09°33'00"	77°48'30"	1307	F
LA1779	TGRC	Ancash (P)	Río Casma	09°33'36"	77°39'36"	3883	F
LA2329	TGRC	La Libertad (P)	Ariacapampa	07°48'00"	77°42'00"	1853	G
LA1978	TGRC	Ancash (P)	Río Fortaleza	10°06'30"	77°29'00"	2313	G
LA1764	TGRC	Lima (P)	Río Chillón	11°30'00"	76°40'00"	2278	G
LA0094	TGRC	Lima (P)	Canta -Yangas	11°31'00"	76°41'00"	2045	G
LA1753	TGRC	Lima (P)	Surco	11°53'05"	76°26'23"	2030	G
LA2409	TGRC	Lima (P)	Río Miraflores	12°16'00"	75°50'00"	4186	G
LA1731	TGRC	Huancavelica (P)	Río San Juan	12°47'00"	74°57'00"	3638	G
LA1695	TGRC	Lima (P)	Valle Río Canete	12°48'45"	75°47'01"	2777	G
LA1918	TGRC	Ayacucho (P)	Puquio	14°16'00"	74°55'00"	2991	G
LA1927	TGRC	Ayacucho (P)	Río Ingenio	14°28'00"	74°48'42'	2784	G
T0039	COMAV	Ayacucho (P)	Huallhua	14°49'27"	74°41'05"	2630	G

País: E: Ecuador, P: Perú. En negrita: *S. habrochaites* f. *glabratum* según el TGRC.

Como fuera de grupo se incluyeron 11 entradas del mismo género que se corresponden con un representante de cada una de las especies de la sección, *S. pennellii*, *S. cornelliomuelleri*, *S. chmielewskii*, *S. neorickii*, *S. chesmaniae*, *S. pimpinellifolium*, *S. lycopersicum*, *S. arcanum*, *S. peruvianum*, *S. huaylasense* y *S. chilense* (Tabla 3.1.2).

Tabla 3.1.2. Relación de entradas de otras especies del género utilizadas como fuera de grupo.

Nº de entrada	Especie	Departamento (País)	Localización
LA0716	<i>S. pennellii</i>	Arequipa (P)	Chala
LA1283	<i>S. cornelliomuelleri</i>	Lima (P)	Santa Cruz de Laya
LA1306	<i>S. chmielewskii</i>	Ayacucho (P)	Tambo
LA1326	<i>S. neorickii</i>	Apurímac (P)	Río Pachachaca
LA1450	<i>S. chesmaniae</i>		
LA1581	<i>S. pimpinellifolium</i>	Lambayeque (P)	Punto Cuatro
LA1673	<i>S. lycopersicum</i>	Lima (P)	Nana
LA2185	<i>S. arcanum</i>	Amazonas (P)	Pongo de Rentema
LA2744	<i>S. peruvianum</i>	Tarapaca (C)	Sobraya
LA2809	<i>S. huaylasense</i>	Ancash (P)	Huaylas-Ancash
LA2884	<i>S. chilense</i>	Antofagasta (C)	Ayaviri

3.2. Caracterización molecular

3.2.1. Tipos de marcadores moleculares utilizados

Para el estudio de la variabilidad molecular se ha empleado un marcador molecular codominante, SSR (Simple Sequence Repeats) o microsatélite, y un marcador dominante, AFLP (Amplified Fragment Length Polymorphisms).

3.2.2. Toma de muestras, extracción y cuantificación del ADN

Se tomaron muestras de cinco plantas por entrada. Se pesaron aproximadamente 50 mg de tejido de hoja joven de cada una de las plantas. La extracción de ADN se realizó siguiendo una modificación del método CTAB (Bromuro de Cetiltrimetilamonio). Tras la extracción del ADN se comprobó su calidad en un gel de agarosa al 0.8% (en tampón TBE 0.5x) teñido con bromuro de etidio y se cuantificó en un espectrofotómetro Nanodrop ND-1000. Para los marcadores AFLP, tras la cuantificación, se mezclaron los ADN genómicos de cada entrada, de forma que en la mezcla hubiera la misma concentración de ADN de cada planta. En el caso de los SSRs se analizó cada una de las plantas de las diferentes entradas de forma independiente.

3.3. Marcadores AFLP

3.3.1. Fundamento

Esta técnica fue empleada por primera vez por Vos *et al.* (1995). Combina la restricción con endonucleasas y la amplificación por PCR. Su principal ventaja es la obtención de un elevado número de polimorfismos con mayor reproducibilidad que otro tipo de marcadores.

Los marcadores AFLP se basan en la amplificación vía PCR de fragmentos de ADN previamente digeridos por dos enzimas de restricción y ligados a dos adaptadores específicos de las enzimas utilizadas. La amplificación de todos los fragmentos originados tras la digestión dificultaría su interpretación e individualización debido a su elevado número. Por esta razón los cebadores empleados en la amplificación son extendidos en dos o tres bases en el extremo 3'. De esta manera el número de fragmentos obtenidos es manejable. La utilización de dos o tres bases selectivas depende del tamaño del genoma a analizar y de la cantidad de fragmentos de ADN que deseamos obtener.

En primer lugar se generan fragmentos de restricción usando dos endonucleasas que produzcan extremos cohesivos (Figura 3.3.1). Una endonucleasa reconoce una secuencia de corte poco frecuente (suele emplearse *EcoRI*), mientras que la otra endonucleasa reconoce una secuencia frecuente (suele emplearse *MseI*). Seguidamente, mediante una reacción de ligación se unen los adaptadores complementarios a los extremos generados por los cortes con las endonucleasas. A continuación se realiza una amplificación preselectiva de los fragmentos, utilizando cebadores que son copias complementarias de los adaptadores más una base adicional en el extremo 3'. Finalmente se realiza una amplificación “selectiva”, en la que los cebadores tienen la misma secuencia que los cebadores de la amplificación preselectiva, pero con dos o tres nucleótidos adicionales en el extremo 3'.

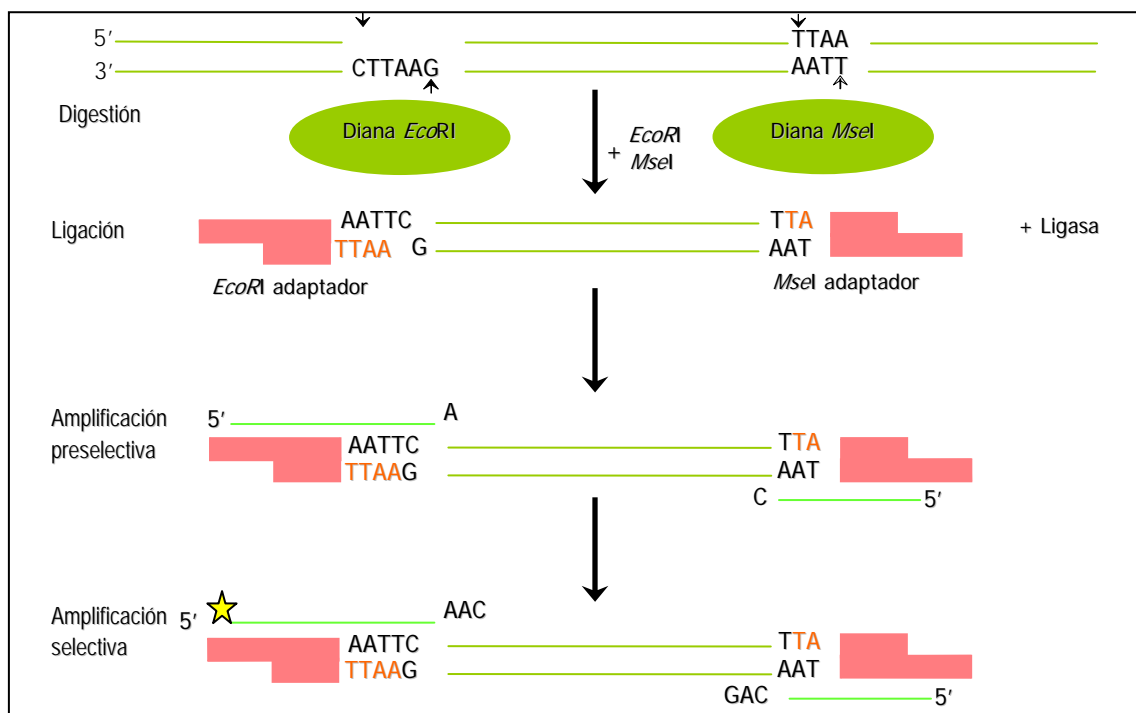


Figura 3.3.1. Esquema de la técnica AFLPs.

3.3.2. Protocolo

Las reacciones AFLP se llevaron a cabo según describe Vos *et al.* (1995). El protocolo consiste en cuatro pasos: digestión, ligación, amplificación y separación de los productos amplificados. Se utilizó un kit comercial en los pasos de digestión y ligación (Invitrogen “AFLP® Core Reagent” N° Cat.: 10482-016).

3.3.2.1. Digestión

Se realizó la digestión de 125 ng de la mezcla de ADN genómico en un volumen final de 10 μ l, con una unidad de EcoRI/MseI en tampón 1x (Tris-HCl 10mM pH 7.5, MgAc₂ 10mM y KAc 50mM). Se incubó durante 2 horas y 30 minutos a 37°C y 15 minutos a 70°C para inactivar los enzimas.

3.3.2.2. Ligación

El ADN digerido se ligó a unos adaptadores específicos de las endonucleasas empleadas. Estos adaptadores tienen la peculiaridad de que una vez ligados no pueden volver a ser digeridos por los mismos enzimas. Los adaptadores empleados fueron:

Adaptador *EcoRI*: 5'-CTCGTAGACTGCGTACC

CATCTGACGCATGGTTAA-5'

Adaptador *MseI*: 5'-GACGATGAGTCCTGAG

TACTCAGGACTCAT-5'

Para la ligación se añadió al volumen total de la reacción de digestión 9.6 μ l de la solución de ligación de adaptadores (ATP 0.4mM, Tris-HCl 10 mM pH 7.5, MgAc₂ 10mM, KAc 50 mM) y 0.4 unidades de ligasa T4, incubándose la mezcla durante 3 horas a 20°C.

3.3.2.3. Amplificación

- a) Reacción de amplificación preselectiva: para la amplificación de los fragmentos flanqueados por los adaptadores EcoRI/MseI se utilizaron dos cebadores complementarios de cada uno de los adaptadores, a los que se añadió un nucleótido adicional. Los cebadores empleados fueron:

Cebador EcoRI+A: 5'-GAC TGC GTA CCA ATT CA-3'

Cebador MseI+C: 5'- GAT GAG TCC TGA GTA AC-3'

Se amplificó, en un volumen total de 25 μ l, 5 μ l de la reacción de ligación diluida (1:10) con los cebadores anteriores (20 μ M), 100 μ M de cada dNTP, 1mM MgCl₂ y 0.8 unidades de *Taq* polimerasa en tampón de PCR (Tris-HCl 10mM pH 8.3, MgCl₂ 1.5 mM, KCl 50mM, 0.001%).

El programa de amplificación preselectiva fue el siguiente:

- 1 ciclo de 2 minutos de desnaturalización a 94°C.
 - 20 ciclos de:
 - 20 segundos de desnaturalización a 94°C.
 - 30 segundos de hibridación a 56°C.
 - 25 segundos de extensión a 72°C.
 - 1 ciclo de extensión final de 10 minutos a 60°C.
- b) Reacción de amplificación selectiva: los cebadores utilizados para la amplificación selectiva son idénticos a los de la preamplificación pero con dos nucleótidos selectivos añadidos al extremo 3'. Se emplearon 3 combinaciones de cebadores. Se utilizó el cebador *EcoRI* marcado con diferentes fluorocromos (Tabla 3.3.1).

Tabla 3.3.1. Combinación de cebadores utilizados en la amplificación selectiva y fluoróforo empleado en cada caso.

COMBINACIÓN DE CEBADORES	FLUORÓFORO
EcoRI-ACA / MseI- CTA	6-FAM
EcoRI-AGC / MseI- CAA	HEX
EcoRI- AAC / MseI- CAA	NED
EcoRI- ACC / Mse- CTA	6-FAM
EcoRI- AGG / Mse- CAT	HEX
EcoRI- AAC / Mse- CTA	NED

En un volumen total de 25 µl se amplificó 2 µl de la reacción de amplificación preselectiva diluida (1:10) con 200 µM del cebador *EcoRI* y 1 mM del cebador *MseI*, 100 µM de cada dNTP, MgCl₂ 1mM y 0.8 unidades de *Taq* polimerasa en tampón de PCR (Tris- HCl 10mM pH 8.3, MgCl₂ 1.5 mM, 0.001% glicerol). La concentración es diferente para ambos cebadores de manera que la reacción de amplificación termine cuando el cebador marcado quede totalmente consumido.

El perfil térmico de la amplificación selectiva fue el siguiente:

- 1 ciclo de 2 minutos de desnaturalización a 94°C.
- 10 ciclos de:
 - 20 segundos de desnaturalización a 94°C.
 - 30 segundos de hibridación a 66°C, disminuyendo 1°C la temperatura de cada ciclo.

- 25 segundos de extensión a 72°C.
- 20 ciclos de:
 - 20 segundos de desnaturalización a 94°C.
 - 30 segundos de hibridación a 56°C.
 - 25 segundos de extensión a 72°C.
- 1 ciclo de 30 minutos de extensión final a 60°C.

En la figura 3.3.2 se muestran los productos de la amplificación selectiva de algunas muestras de *S. habrochaites*.

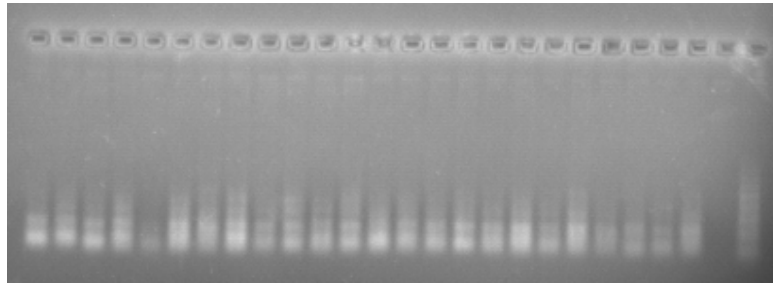


Figura 3.3.2. Gel de agarosa que muestra los productos de amplificación selectiva.

3.4. Marcadores SSR

3.4.1. Fundamento

Los microsatélites, también conocidos como SSRs, fueron descritos por primera vez por Hamada *et al.* (1982) como secuencias cortas de ADN constituidas por motivos de 1 a 6 nucleótidos repetidos en tándem. El número de repeticiones es variable y, en general, el grado de polimorfismo aumenta con la longitud total del microsatélite (Weber 1990). Debido a su excepcional variabilidad y a que se encuentran de forma abundante y uniforme a lo largo de la mayoría de genomas eucariotas, están considerados como los marcadores moleculares más potentes (Goldstein & Pollock 1997). Son marcadores idóneos para la obtención de mapas genéticos de alta densidad (Dib *et al.* 1996).

Las principales ventajas de los microsatélites es que son marcadores codominantes y altamente reproducibles. Además, en tomate se ha desarrollado un gran número de cebadores específicos para las secuencias únicas que flanquean el motivo

repetido (Smulders *et al.* 1997; Areshchenkova y Ganal 1999, 2002). También puede encontrarse una extensa colección de cebadores para amplificar microsatélites mapeados en <http://www.sgn.cornell.edu/>. El empleo de estas secuencias permite la amplificación por PCR de la región repetida y la visualización de todos los alelos posibles para un locus microsatélite.

En el género *Solanum* las repeticiones de dinucleótidos son las más abundantes, principalmente (TA)_n, aunque también se encuentra abundancia de trinucleótidos, principalmente (AAG)_n, (AAT)_n y (ACC)_n.

3.4.2. Marcadores microsatélite seleccionados

Se seleccionaron 8 microsatélites de la bibliografía existente (Smulders *et al.* 1997 y SolGenomics) que ya habían sido probados y presentaban polimorfismo en esta especie (Sifres *et al.* 2003): LE20592, LE2A11, LELEUZIP, LECAB9, LEGAST1, LEGTOM5, LPHFS24, LEWIPIG; y se añadió uno de la base de datos de SolGenomics Network (<http://sgn.cornell.edu/>): SSR115 (Tabla 3.3.2).

Tabla 3.3.2. Marcadores microsatélite empleados, nombre del locus, cebadores directo y reverso, tipo de repetición, tamaño del alelo en *S. lycopersicon* según Smulders *et al.* 1997 y SolGenomics, temperatura de hibridación y fluoróforo empleado en cada caso.

NOMBRE LOCUS	CEBADORES	REPETICIÓN	TAMAÑO	T _m	FLUORÓFORO
LE2A11	5'AATTTGTAAGGAGAAGACGG3' 5'TCATATTCTTCACACCAAAGG3'	(ATCT) ₅₋₁	157pb	55°C	6-FAM
LE20592	5'CTGTTTACTTCAAGAAGGCTG3' 5'ACTTTAACTTTATTATTGCCACG3'	(TAT) ₁₅₋₁ (TGT) ₄	166pb	57°C	6-FAM
LEGAST1	5'ATCTCTATTGTTTTCGACTCG3' 5'TCTGTTGTTGCTGCTGCTC3'	(TA) ₁₂ (TG) ₄	143pb	55°C	NED
LEGTOM	5'AAAGATAAAGCATGAAATGAA3' 5'GGAGTTGAGATAAAGTGAAGA3'	(TA) ₁₀	181pb	50°C	NED
LECAB9	5'TTTATTATCCAGAAGCCTTC3' 5'CCTCACATTTAAACAAATTGC3'	(TA) ₆ (CA) ₃	118pb	55°C	NED
LELEUZIP	5'GGTGATAAATTGGGAGGTTAC3' 5'CGTAACAGGATGTGCTATAGG3'	(AAG) ₆₋₁ TT(GAT) ₇	105pb	55°C	6-FAM
LEWIPIG	5'GAGTCAAAGTTTGTCTCACATC3' 5'CTCTTCTGAACCTTGCTTTGAG3'	(CT) ₈₋₁ (AT) ₄	254pb	48°C	NED
LPHFS24	5'TTGGATTTACAAGTTCGATGT3' 5'GCATTTGACTTGATAGCAGTC3'	(TA) ₆	149pb	50°C	HEX
SSR115	5'CACCCTTTATTTCAGATTCTCT3' 5'ATTGAGGGTATGCAACAGCC3'	(AT) ₁₆	211pb	50°C	HEX

3.4.3. Protocolo

3.4.3.1. Amplificación

El volumen final de cada reacción de PCR (12 μ l) estaba compuesto por 12 ng de ADN genómico, tampón de PCR 1x, 10 μ M de cada dNTP, 10 μ M de cada cebador y 1 unidad de *Taq* polimerasa.

La reacción de PCR se llevó a cabo según describe Smulders *et al.* (1997). El perfil térmico fue el siguiente:

- 1 ciclo inicial de desnaturalización a 94°C durante 3 minutos.
- 30 ciclos con las siguientes condiciones:
 - Desnaturalización durante 45 s a 94°C.
 - Hibridación durante 1 min 45 s a la temperatura adecuada para cada pareja de cebadores (ver tabla 3.3).
 - Extensión durante 45 s a 72°C.
- 1 ciclo final de elongación de 30 min a 72°C.

La amplificación se comprobó en un gel de agarosa al 1% (Figura 3.3.3).

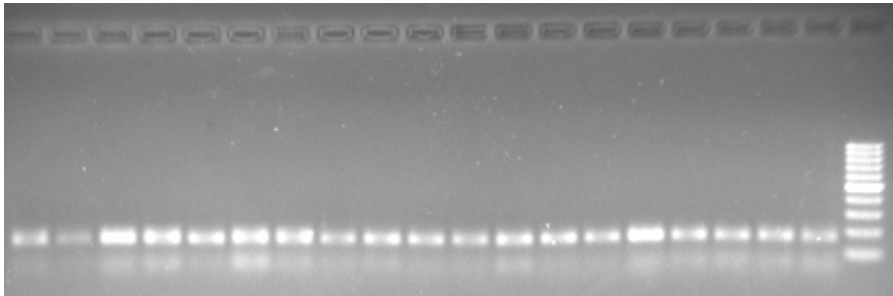


Figura 3.3.3. Gel de agarosa que muestra la amplificación de un locus microsatélite en diferentes entradas de *S. habrochaites*.

3.5. Separación de los productos amplificados

Para ambos marcadores se utilizó un secuenciador capilar ABI PRISM® 3100 Genetic Analyzer, basado en un sistema de análisis de ADN con marcadores fluorescentes que utiliza 4 capilares operando en paralelo. Los fluoróforos son excitados por un láser de forma que son detectados por un sistema de lectura que permite detectar 5 fluoróforos diferentes en una misma reacción. Así se resuelve el producto de la amplificación de hasta cuatro cebadores marcados con fluoróforos que emitan en

distinta longitud de onda más el fluoróforo correspondiente al patrón de pesos moleculares.

Las muestras se prepararon en 12 μ l de formamida, con 0.5 μ l de cada una de las tres reacciones diluidas y 0.25 μ l de marcador de peso molecular GeneScan 400 HD marcado con el fluoróforo ROX, que emite en el color rojo.

Los fragmentos amplificados se visualizaron en forma de picos mediante el programa de análisis GeneScan (v.3.1.2) (Figuras 3.3.4 y 3.3.5).

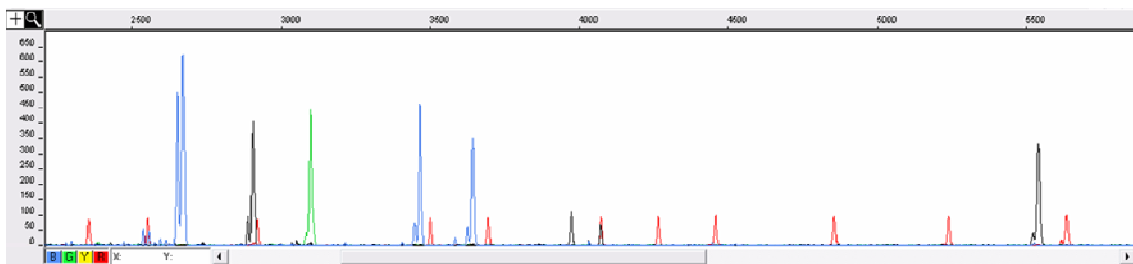


Figura 3.3.4. Electroferograma obtenido para seis microsatélites diferentes en una muestra concreta.

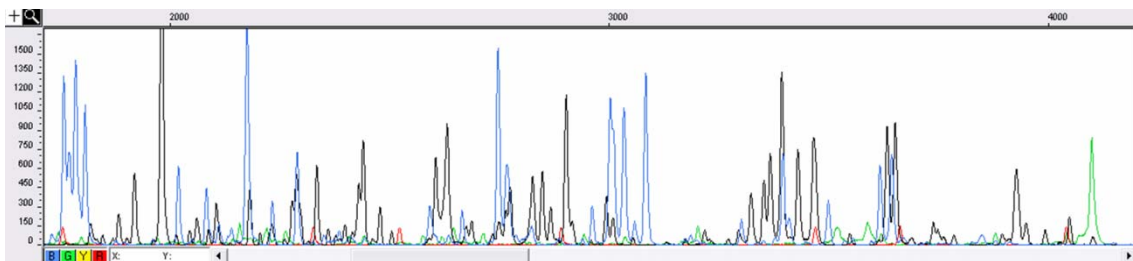


Figura 3.3.5. Electroferograma obtenido para tres combinaciones de AFLPs en una muestra concreta.

Para facilitar la interpretación de los resultados en el caso de AFLPs los picos se convirtieron en bandas utilizando el programa Genographer (v.2.1.4) (Figura 3.3.6). Los marcadores SSR se analizaron utilizando el programa Genotyper (v.2.0) (Figura 3.3.7).

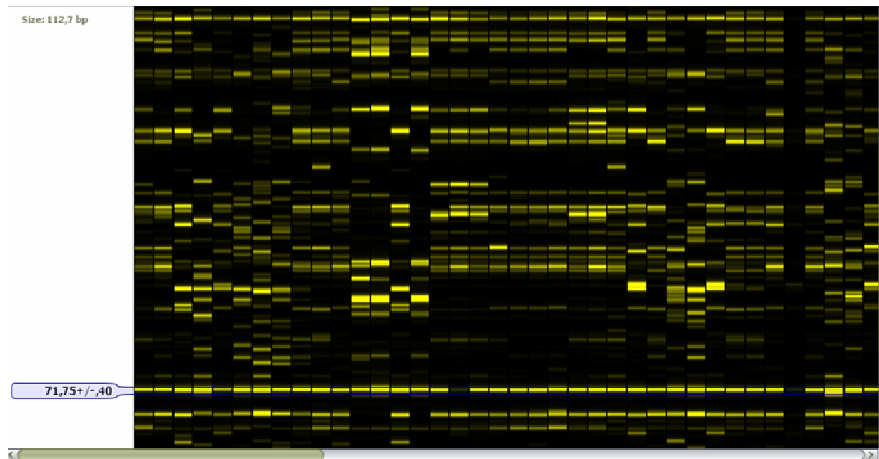


Figura 3.3.6. Imagen del programa Genographer tras importar los archivos GeneScan de algunas muestras de AFLP de *S. habrochaites* cuyos cebadores estaban marcados con el fluoróforo NED (amarillo).

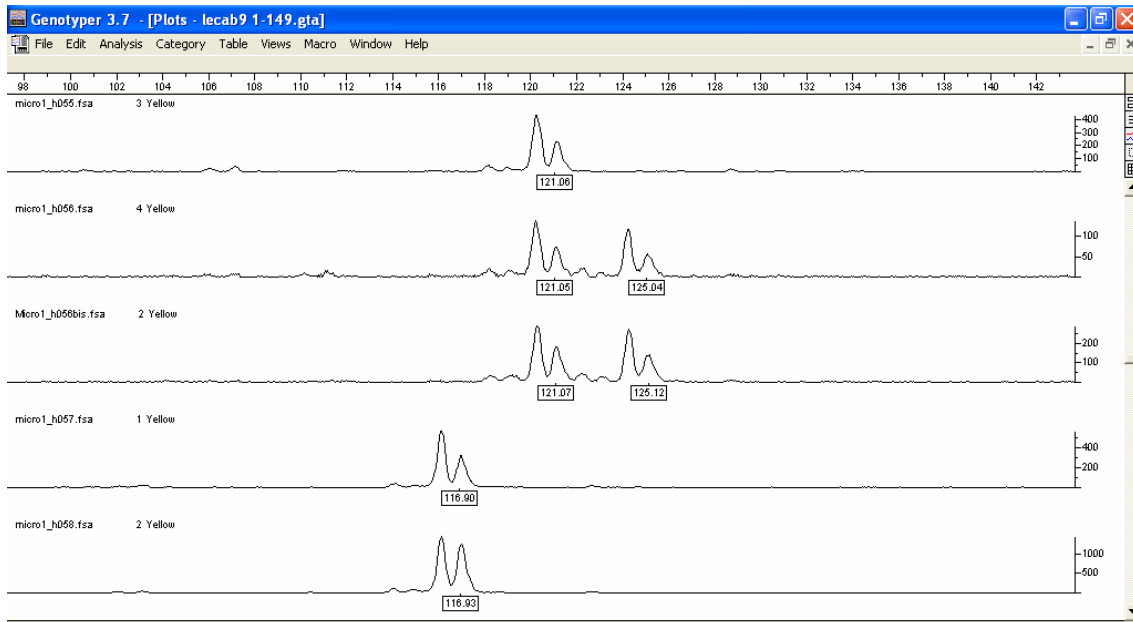


Figura 3.3.7. Imagen del programa Genotyper tras importar los archivos GeneScan de algunas muestras microsatélites de *S. habrochaites* cuyos cebadores estaban marcados con el fluoróforo NED (amarillo).

3.6. Análisis de los datos:

En el caso de los marcadores AFLPs:

Los fragmentos amplificados por los marcadores AFLP fueron evaluados empleando el programa Genographer (v.2.1.4). Las bandas del gel se transformaron en una matriz de (1) y (0), es decir, de presencia o ausencia de banda. A partir de la matriz de 1 y 0 se estimó el coeficiente de similitud de Dice (Nei y Li, 1979):

$$S_{ij} = 2a / (2a + b + c)$$

donde S_{ij} es la similitud entre dos individuos i y j , a es el número de bandas compartidas, b y c , son el número de bandas amplificadas exclusivamente en i y j respectivamente.

Se realizó un análisis cluster empleando el método Neighbour-Joining utilizando el paquete informático PHYLIP (v.3.69) (Felstein 2008). Para comprobar la robustez de los nodos se realizó un remuestreo de la matriz de 1 y 0 con 1000 muestras y un reemplazo del 35% de los datos. Los dendrogramas se visualizaron con el programa TreeViewPPC (v.1.6.6) (Page 2001).

Con esta matriz de similitudes también se realizó un análisis de coordenadas principales (ACoP) utilizando el programa Ginkgo Multivariate Analysis System (v.1.6.0) (Bouxin 2005).

Para cada grupo geográfico se calculó el número de loci polimórficos y la diversidad genética de Nei (1973),

$$D_k = 1 - \sum p_i^2,$$

donde p_i es la frecuencia del i -ésimo alelo en un locus k . Para ello se utilizó el programa PopGene (v.1.32) (Yeh et al. 1997).

En el caso de los marcadores SSR:

Se calcularon las frecuencias alélicas para cada locus en cada una de las zonas utilizando el programa Genetix (v.4.04). A partir de las frecuencias alélicas obtenidas para cada uno de los loci microsatélites se obtuvo la matriz de distancias empleando el coeficiente de Cavalli-Sforza y Edwards (1967).

$$D_C = (2 / \pi r) \sum_j^r \sqrt{2 \left(1 - \sum_i^{m_j} \sqrt{x_{ij} y_{ij}} \right)}$$

donde x_{ij} e y_{ij} son las frecuencias del alelo i del locus j en las poblaciones x e y , respectivamente, m_j es el número de alelos que presenta el locus j y r es el número de loci analizados.

Se construyó un dendrograma con el método Neighbour-Joining con ayuda del programa Populations (v.1.2.28) (Langella 2002). Con el mismo programa se realizó bootstrap para saber cuál era la robustez de los nodos.

Con esta matriz de distancias también se realizó un análisis de coordenadas principales (ACoP) utilizando el programa Ginkgo Multivariate Analysis System (v.1.6.0) (Bouxin 2005).

La diversidad genética intrapoblacional para cada grupo geográfico se obtuvo con el programa Genetix (v.4.04) (Belkhir *et al.* 1996). Se calculó el número medio de alelos por locus (A), la heterozigosidad observada (H_o), la heterozigosidad esperada (H_e & Nei 1973) y la proporción de loci polimórficos, en los que el alelo más frecuente aparece por debajo de 95%.

En ambos casos, la estructura de las poblaciones se comprobó utilizando el programa STRUCTURE (Pritchard 2000). Dada una serie de individuos se asumen K poblaciones y los individuos son asignados a cada una de las poblaciones. Se calcularon desde $K=1$ hasta $K=9$ poblaciones. El número final de poblaciones (K) se definió cuando las probabilidades alcanzaron un máximo para AFLPs y un estado de meseta para SSRs. Además, se calculó la probabilidad de que cada entrada perteneciera a cada una de las poblaciones con el mismo programa.

La varianza molecular dentro de cada zona y entre zonas o grupos geográficos se calculó mediante un análisis AMOVA con el programa Arlequin 3.1 (Excoffier *et al.* 2005).

Con el fin de establecer las relaciones genéticas entre las distintas regiones geográficas se realizó un ACoP y un análisis cluster. Como coeficientes de distancias entre grupos se utilizó el índice de diferenciación G_{st} (Nei 1973) para AFLPs y el índice F_{st} (Weir & Cockerman 1984) para SSRs.

Para comprobar si las distancias geográficas estaban correlacionadas con las genéticas se calcularon las distancias geográficas entre cada par de entradas con el método del Arco del Círculo Máximo a partir de las coordenadas geográficas.

Para estudiar la relación entre el clima y la diferenciación genética de *S. habrochaites* se utilizaron diferentes parámetros de temperatura y precipitaciones. Éstos se extrajeron de la base de datos de la página web del programa Diva-Gis (<http://www.diva-gis.org/gData>). Se obtuvieron los valores para cada una de las entradas y se realizó un análisis de la varianza para comprobar si las medias entre las diferentes zonas eran significativamente diferentes. Con esta matriz de datos estandarizada se realizó un análisis de componentes principales (ACP) con el programa Ginkgo (v1.6.0) y se representaron las entradas en función de éstas.

Se calculó el coeficiente de correlación de Pearson entre las matrices de distancias genéticas para cada marcador y las distancias geográficas y climáticas. Para ello se realizó un test de Mantel (1967) con ayuda del programa Ginkgo (v.1.6.0).

Resultados y Discusión

4. Resultados y Discusión

4.1. Marcadores AFLPs

En el análisis de la colección con marcadores AFLP se obtuvo un gran número de fragmentos amplificados para cada combinación (Tabla 4.1.1).

Se amplificaron un total de 923 bandas, incluyendo los representantes de cada una de las especies de la sección, de las cuales 887 fueron polimórficas, aproximadamente un 96% del total. En la especie *S. habrochaites* se amplificaron un total de 617, con un promedio de 103 bandas por combinación, de las cuales 561 fueron polimórficas, lo que representa un 91% del total aproximadamente, con un promedio de 94 bandas polimórficas por combinación.

Tabla 4.1.1. Relación de bandas obtenidas para cada combinación de cebadores selectivos, para todas las especies y para *S. habrochaites*.

Combinación de cebadores selectivos	Todas las especies		<i>S. habrochaites</i>	
	Número de bandas amplificadas	Número de bandas polimórficas	Número de bandas amplificadas	Número de bandas polimórficas
Eco ACA Fam - Mse CTA	200	187	148	129
Eco AGC Hex - MSE CAA	103	100	43	35
Eco AAC Ned - Mse CAA	173	163	137	125
Eco ACC Fam - Mse CTA	186	183	114	111
Eco AGG Hex - Mse CAT	95	92	60	55
Eco AAC Ned - Mse CTA	166	162	115	106
Total	923	887	617	561

4.1.1. Análisis de Coordenadas Principales y Diversidad Genética.

El Análisis de Coordenadas Principales obtenido para los marcadores AFLP empleando la distancia de Nei (1979) se representa en la figura 4.1.2. La primera coordenada, que explica una gran cantidad de la variación (56%), separa claramente las entradas del Norte de las del Sur de su área de distribución, mientras que en la zona central del diagrama aparecen las entradas de la zona D, cercana a Huancabamba. En este grupo se observa una mayor dispersión de las entradas. Es en esta zona donde hemos encontrado una mayor diversidad genética (Tabla 4.1.2), lo que concuerda con lo obtenido por Rick *et al.* (1979) que observaron una gran variabilidad morfológica en esta zona donde el polimorfismo electroforético tenía un máximo.

Las entradas colectadas más al norte, en las áreas A y B aparecen entremezcladas, pero podemos observar dos subgrupos con las entradas de cada área más cercanas entre sí. En las entradas del área C se observan dos subgrupos claramente diferenciados pertenecientes a las entradas colectadas más al norte o al sur en esta área (Figura 4.1.1). Las entradas de las zonas E y F, colectadas más al sur en las provincias de Cajamarca y La Libertad, aparecen agrupadas con respecto a estas dos coordenadas. También las entradas colectadas en la zona más al sur, zona G, aparecen agrupadas aunque separadas de las anteriores con respecto a la componente dos, que explica apenas un 9% de la variación.

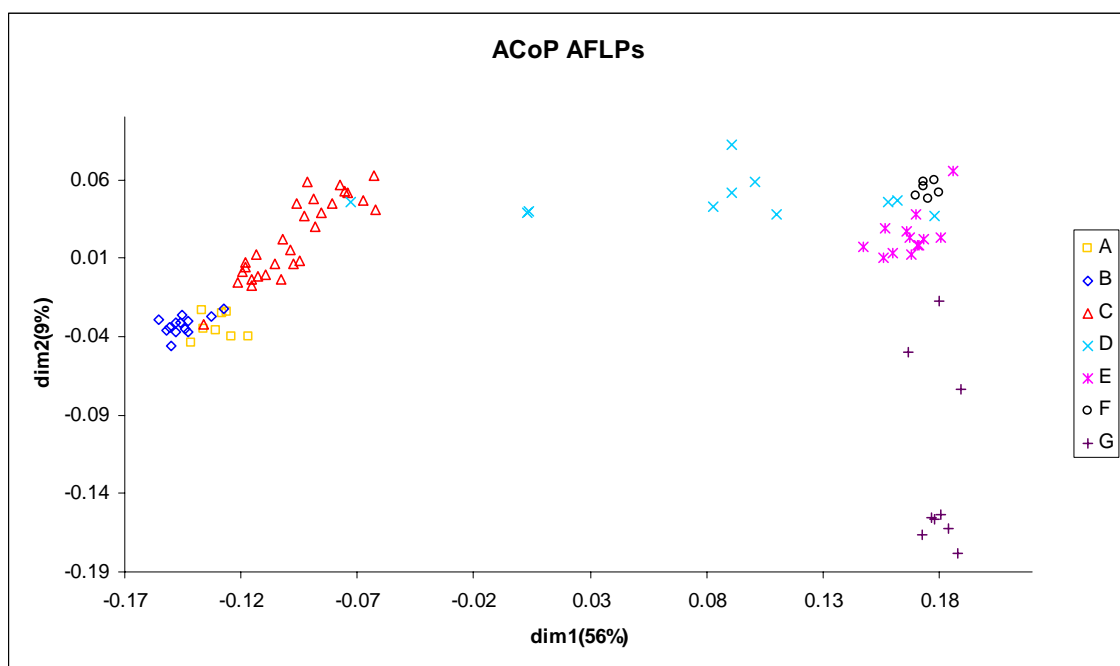


Figura 4.1.1. Análisis de Coordenadas Principales obtenido con los marcadores AFLP en el que se representan con diferentes símbolos las entradas pertenecientes a cada una de las zonas de colecta.

Las zonas centrales del rango de distribución fueron las más diversas, desde el noreste de Loja hasta La Libertad, alcanzando un máximo en la región de Huancabamba. Las regiones del extremo norte y sur mostraron una menor diversidad genética (tabla 4.1.2).

Tabla 4.1.2. Zonas, número de individuos (n), diversidad genética de Nei (1973) (D) y su varianza (entre paréntesis), proporción de *loci* polimórficos en los que el alelo más frecuente está por debajo del 95% (P).

Poblaciones	AFLPs		
	n	D	P
Centro de Ecuador (A)	8	0.12 (0.03)	0.26
Noreste de Loja (B)	13	0.05 (0.02)	0.16
Oeste de Loja-Piura (C)	29	0.11 (0.03)	0.44
Huancabamba (D)	11	0.15 (0.04)	0.46
Cajamarca-La Libertad (E)	13	0.12 (0.03)	0.42
Ancash (F)	6	0.05 (0.02)	0.14
Sur de Peru (G)	9	0.10 (0.03)	0.32
Total	89	0.19 (0.03)	0.91

4.1.2. Análisis Cluster

El dendrograma obtenido incluyendo las entradas pertenecientes a todas las especies de la sección separó claramente las entradas de *S. habrochaites* del resto de especies. En la figura 4.1.2 se representa el dendrograma en el que se ha utilizado como fuera de grupo la entrada de la especie *S. pennelli* porque es la más cercana a *S. habrochaites* entre todas las especies silvestres de la sección.

En este dendrograma puede observarse una clara separación en dos clusters de las entradas colectadas en las zonas más al norte del área de distribución de la especie de las colectadas más al sur con un valor de bootstrap del 100%. Las entradas LA1391 y ECU0526 procedentes de Cajamarca y Lambayeque respectivamente, aparecen siempre agrupadas y separadas del resto (bootstrap 100%). Estas entradas son las únicas colectadas en estas latitudes a altitudes muy bajas entre 200 y 500m.

En un mismo cluster aparecen agrupadas las entradas colectadas a latitudes más bajas desde Huancabamba hasta Nazca. Las entradas de la zona G, situada más al sur, aparecen separadas del resto de entradas, en un único cluster. En este cluster se agrupan las entradas colectadas en las provincias de Ayacucho con un valor de bootstrap muy alto, del 100%. Las entradas de las zonas E y F, colectadas en Cajamarca, La Libertad y Ancash aparecen en un cluster entremezcladas y separadas del resto de entradas del sur con un valor de bootstrap del 59%. Las entradas de la zona D, que hemos llamado zona de Huancabamba, aparecen en este cluster aunque separadas del resto. A pesar de ser más cercanas geográficamente a las entradas del norte aparecen en este cluster con las

entradas colectadas en el sur de distribución de la especie, lo cual confirma que existe una barrera geográfica que separa esta zona del resto.

En el segundo cluster (100%) se agrupan las entradas colectadas más al norte del área de distribución. Por un lado se agrupan todas las entradas de la zona B, cercanas a Loja, en el mismo cluster que las entradas colectadas más al norte en el centro de Ecuador o zona A. Las entradas colectadas en Loja-Piura, o zona C, aparecen agrupadas en este segundo cluster aunque las ramas no tienen un valor de bootstrap alto.

En general, podemos decir que las diferentes entradas de *S. habrochaites* se agrupan en función de su lugar de colecta, viéndose una gran separación entre las entradas colectadas más al norte de su área de distribución y las colectadas más al sur. Además se observa una diferenciación de las entradas colectadas en la Depresión de Huancabamba. Algunos autores han sugerido que esta Depresión representa una frontera biogeográfica importante para taxones de la zona andina (Ayers 1999). Otros autores han propuesto que la zona Amotape-Huancabamba en los Andes del norte de Perú y en el extremo sur de Ecuador es un área con una biodiversidad excepcional y posiblemente una zona fitogeográfica propia (Weigend 2002). Esto estaría de acuerdo con la existencia de tres zonas biogeográficas diferenciadas en esta región central para *S. habrochaites*. Esto se confirma con la diversidad genética encontrada, que es la mayor entre todas las áreas de colecta, a pesar de contar con un menor número de entradas (Tabla 4.1.2).

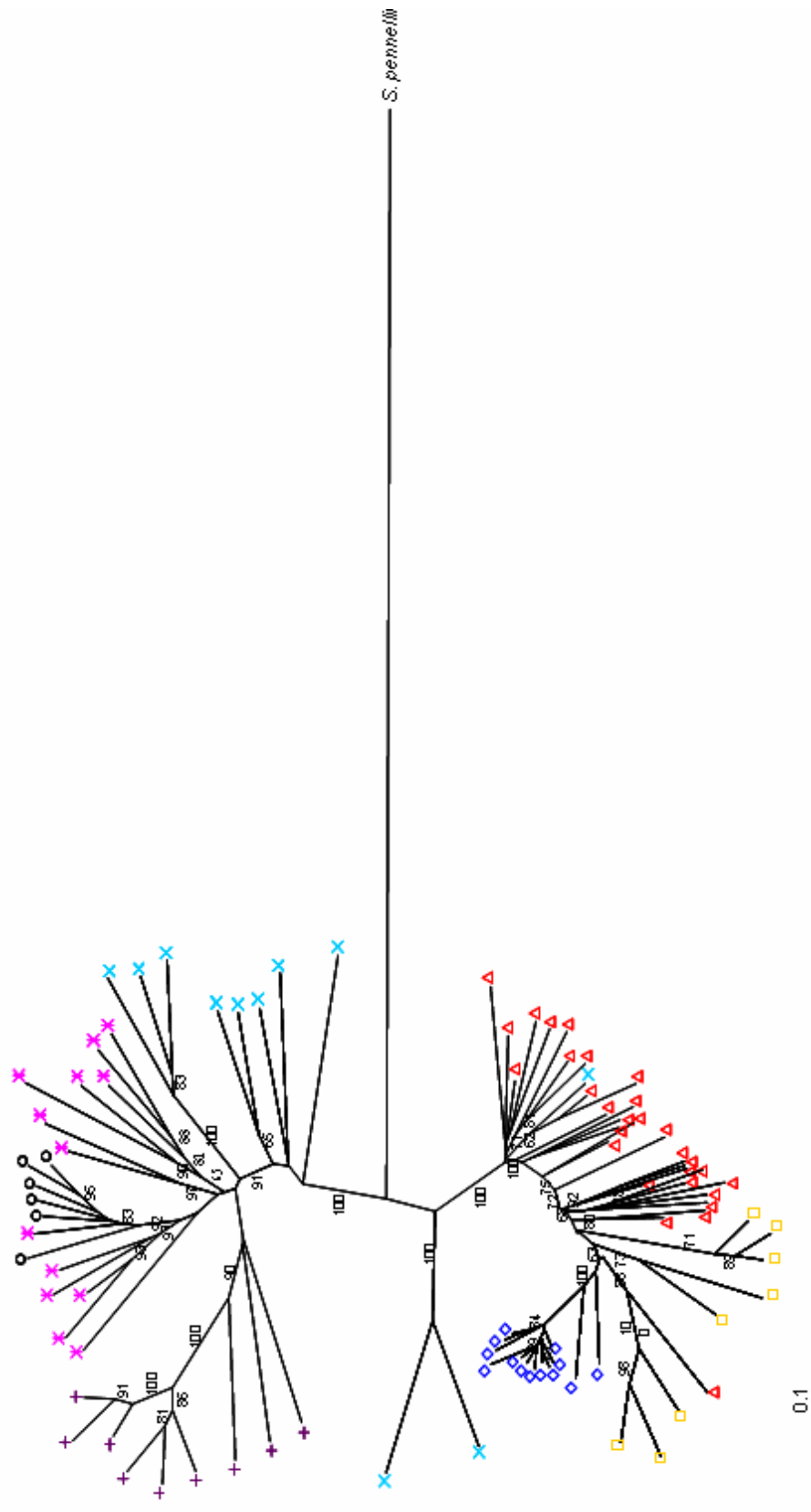


Figura 4.1.2. Dendrograma obtenido para AFLPs empleando el coeficiente de similitud de Dice (1979) y el método Neighbour-Joining para las distintas entradas de *S. habrochaites*. Se indican los valores de *bootstrap* en %.

Amarillo: zona A, azul: zona B, rojo: zona C, turquesa: zona D, rosa: zona E, blanco: zona F y morado: zona G.

4.1.3. Estructura Poblacional

La estructura de las poblaciones se comprobó utilizando el programa STRUCTURE (Pritchard 2000). Este programa informático usa una aproximación bayesiana de clustering para identificar subpoblaciones y para asignar individuos a estas poblaciones en base a sus genotipos. Dada una serie de individuos se asumen K poblaciones y los individuos son asignados a cada una de las poblaciones. Se calcula la verosimilitud para cada K , que va aumentando con el número de poblaciones. Se asume el valor de K tal que su verosimilitud alcanza un máximo o una fase de meseta.

En el caso de AFLPs se observa un máximo para $K=7$ lo que se corresponde con el número de grupos geográficos que anteriormente habíamos sugerido (Figura 4.1.3).

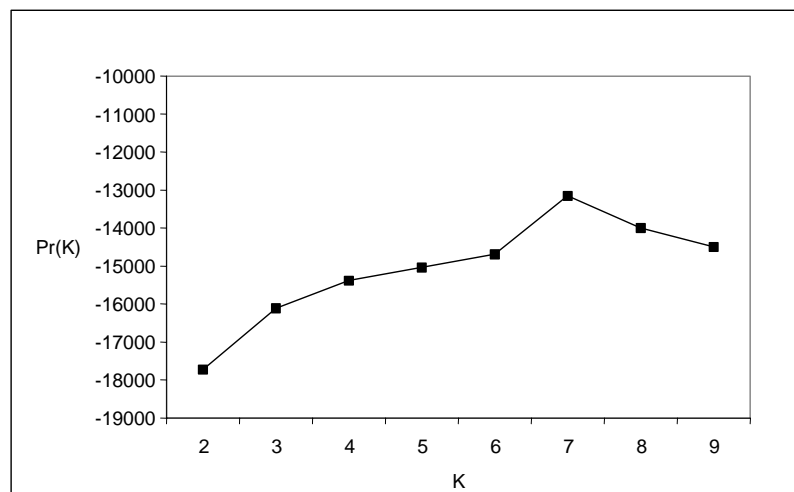


Figura 4.1.3. Representación de la probabilidad de que existan K poblaciones, $Pr(K)$, en función del número estimado de poblaciones, K .

Al representar el dendrograma obtenido con estos datos AFLP junto a la probabilidad de que cada uno de los individuos pertenezca a uno u otro grupo (Figura 4.1.4) observamos que la mayoría de asignaciones de entradas a grupos geográficos sí se corresponden con su agrupación genética según el análisis realizado. Éste análisis asume que las poblaciones están en equilibrio de Hardy-Weinberg lo cual no es cierto. Las entradas pertenecientes a la zona más al sur, G, está definida claramente como una población. Las áreas que nosotros hemos llamado E y F podrían formar una única población, aunque con una subpoblación clara que sería las entradas pertenecientes a la

zona F. Las entradas pertenecientes a Huancabamba formarían una población aunque con independencia de las entradas ECU0526 y LA1391 que formarían una población según los datos obtenidos con este tipo de marcadores. Las entradas pertenecientes a la zona C formarían dos subpoblaciones, las colectadas en la cuenca del río Piura claramente diferenciadas y las colectadas más al norte en la cuenca del Chira y Ayabaca que podrían pertenecer también a la población ecuatoriana A, cuyas entradas están claramente diferenciadas. Las entradas que hemos asignado a la zona geográfica B formarían una única población aunque dos de las entradas podrían pertenecer también a la población A.



Figura 4.1.4. Dendrograma obtenido con los datos AFLP, empleando el coeficiente de similitud de Dice y el método Neighbour-Joining. En él se representan en diferentes colores las diferentes agrupaciones según criterios ecogeográficos. En el gráfico paralelo se representa la probabilidad de que cada uno de que cada entrada pertenezca a una población u otra definida con el programa Structure considerando que existen 7 poblaciones definidas. Amarillo: zona A, azul: zonaB, rojo: zona C, turquesa: D, rosa: zona E, gris: zona F y morado: zona G.

4.2. Marcadores Microsatélite

El número de alelos encontrado en las entradas de *S. habrochaites* para cada locus microsatélite fue variable, desde 26 alelos en el caso de LE20592 hasta 5 alelos para LPHFS24 con frecuencias alélicas variables (Tabla 4.2.1).

Tabla 4.2.1. Relación del número de alelos y tamaño de los mismos encontrados para cada locus microsatélite.

NOMBRE <i>LOCUS</i>	NÚMERO TOTAL ALELOS	NÚMERO ALELOS <i>S.habrochaites</i>	TAMAÑO ALELOS (pb) en <i>S. habrochaites</i>
LE2A11	11	6	164, 166, 168, 170, 178, 184
LE20592	36	26	148, 151, 154, 157, 160, 162, 164, 167, 170, 173, 176, 179, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 220, 223, 229
LECAB9	8	8	115, 117, 119, 121, 123, 125, 127, 129
LEGAST1	14	11	118, 120, 126, 128, 129, 130, 132, 134, 136, 138, 140
LEGTOM5	16	16	171, 172, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187
LELEUZIP	8	7	93, 96, 99, 102, 105, 109, 112
LEWIPIG	10	8	249, 250, 251, 253, 254, 255, 256, 257
LPHFS24	9	5	139, 145, 146, 149, 153
SSR115	14	8	207, 208, 209, 211, 213, 215, 217, 219

En algunos de los loci microsatélites se han encontrado alelos que estaban presentes únicamente en las zonas más al norte de distribución de la especie y alelos exclusivos de las zonas más al sur de la zona de colecta, como es el caso de LE20592 (Figura 4.2.1) o que se encuentran con una frecuencia mucho mayor en una de las áreas de distribución como es el caso de LEGAST1 (Figura 4.2.2). En el caso de LE20592 cabe destacar que la zona C comparte alelos con las zonas norte y sur.

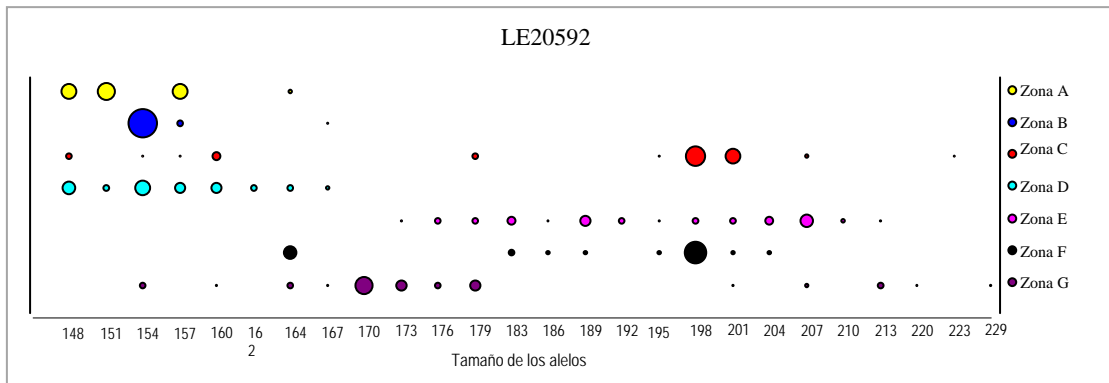


Figura 4.2.1. Frecuencias alélicas de cada alelo para el locus SSR LE20592, en diferentes colores aparecen las frecuencias para cada una de las zonas. El área del círculo es proporcional a la frecuencia del alelo.

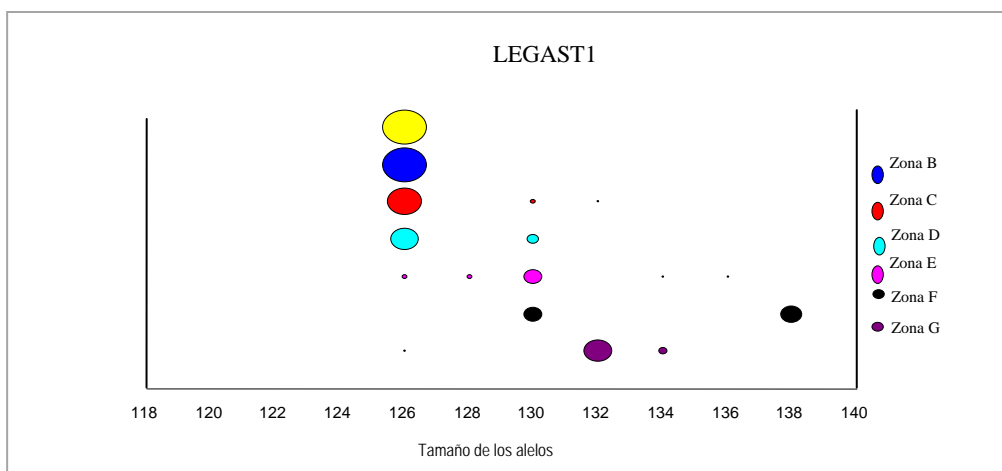


Figura 4.2.2. Frecuencias alélicas de cada alelo para el locus SSR LEGAST1, en diferentes colores aparecen las frecuencias para cada una de las zonas. El área del círculo es proporcional a la frecuencia del alelo.

Estos datos estarían de acuerdo con los resultados de Rick *et al* (1979) que encontraba alelos para isoenzimas preferentemente en la región norte o sur del área de distribución de la especie y mezcla de alelos en la zona centro lo que apoya la idea de que el origen de la especie estaría en esta zona centro y los individuos habrían migrado hacia el norte y el sur perdiendo alelos en el proceso.

4.2.1. Análisis de Coordenadas Principales y Diversidad Genética

El Análisis de Coordenadas Principales obtenido para los marcadores SSRs empleando la distancia de Cavalli-Sforza y Edwards (1967) se representa en la figura 4.2.3. Las dos primeras coordenadas, que explican una variación del 22% y 17%,

respectivamente, separan las entradas del norte y sur de distribución pero no tan claramente como los AFLPs. En cuanto a las entradas de la zona Norte, las entradas de los grupos A y C aparecen entremezcladas y separadas de las de la zona B. Las entradas de las zonas D y E situadas más al centro del área de distribución aparecen entremezcladas en el centro del diagrama, mientras que en la zona derecha del diagrama aparecen las entradas de las zonas F y G, colectadas en el sur, claramente separadas del resto.

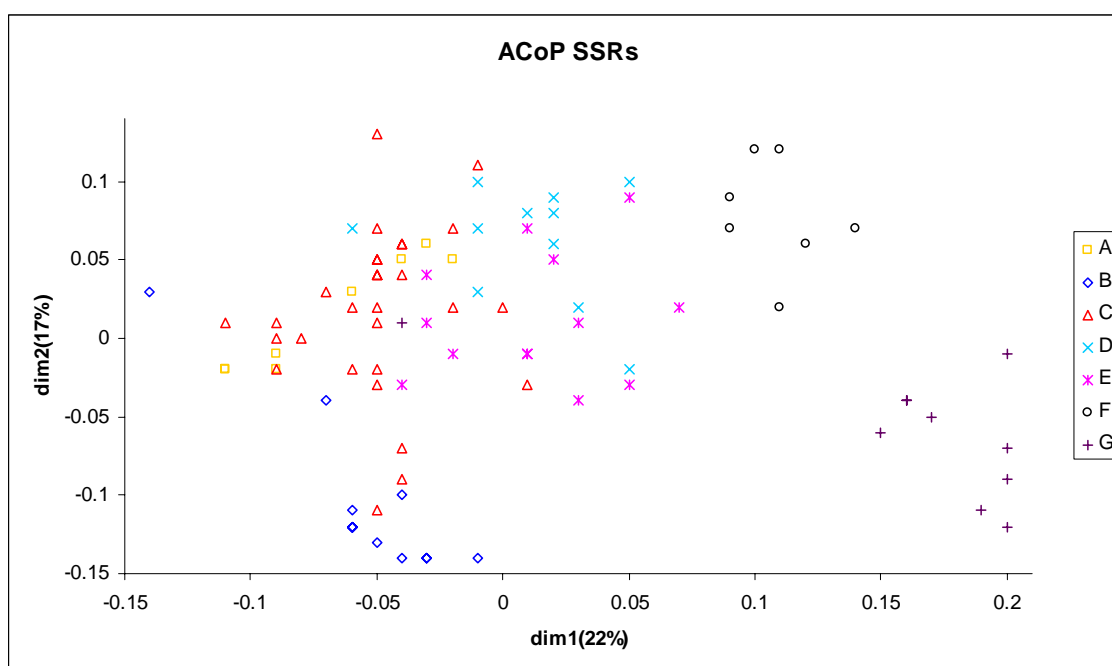


Figura 4.2.3. Análisis de Coordenadas Principales obtenido con para marcadores microsatélites en el que se representan con diferentes símbolos las entradas pertenecientes a cada una de las zonas de colecta.

Se confirma que la zona centro es la C y D pero hay más intercambio hacia el norte.

Si observamos las heterocigidades medias para cada una de las entradas en las diferentes zonas de colecta vemos que en las zonas centrales, C y D, aparecen las más altas, al igual que los polimorfismos, mientras que es muy baja en las entradas ecuatorianas y en las colectadas más al sur en el departamento de Lima (Tabla 4.2.2). Esto estaría de acuerdo con lo observado por Rick *et al.* (1979) que describió una mayor heterocigidad, asociada a una mayor alogamia, en la zona centro de distribución que tendía hacia la autocompatibilidad al alejarse hacia el norte y el sur.

Tabla 4.2.2. Número de individuos (n), heterozigosidad observada (Ho) y su varianza (entre paréntesis), heterocigosidad esperada (He) y su varianza (entre paréntesis), proporción de *loci* polimórficos en los que el alelo más frecuente está por debajo del 95% de las veces (P) y número medio de alelos por locus (A).

Población	SSRs				
	n	Ho	He	P	A
Zona A: Centro de Ecuador	69	0.03 (0.001)	0.37 (0.10)	0.67	3.0
Zona B: Noreste de Loja	120	0.05 (0.002)	0.28 (0.03)	0.89	3.1
Zona C: Oeste de Loja-Piura	271	0.34 (0.036)	0.62 (0.63)	1	6.9
Zona D: Huancabamba	105	0.29 (0.023)	0.57 (0.03)	1	5.2
Zona E: Cajamarca- La Libertad	111	0.30 (0.040)	0.55 (0.08)	0.89	5.6
Zona F: Ancash	66	0.30 (0.063)	0.40 (0.06)	0.78	3.2
Zona G: Sur de Peru	83	0.10 (0.008)	0.43 (0.06)	0.89	4.9
Zona A: Centro de Ecuador	825	0.24 (0.014)	0.66 (0.06)	1	10.6

Además, las heterozigosidades observadas son muy bajas con respecto a las heterozigosidades que se esperarían en el supuesto de que la población se encontrase en equilibrio Hardy-Weinberg, lo que significa que en todas las poblaciones existe cierto grado de autogamia o cruces entre plantas hermanas.

El número medio de alelos por locus más alto se observa en las zonas centrales, C y D, como era previsible, y también en la zona E, un poco más al sur.

4.2.2. Análisis Cluster

El dendrograma obtenido para los marcadores microsatélites empleando la distancia de Cavalli Sforza y Edwards (1967) y el método Neighborg-Joining se muestra en la figura 4.2.4. Aunque en este caso los valores de bootstrap no son tan altos como en el caso de AFLPs, las entradas también se agrupan según su procedencia geográfica. Las entradas colectadas más al norte de su área de distribución, las zonas que llamamos A y B están agrupadas en dos clusters bien diferenciados. También las entradas colectadas en la zona sur del rango de distribución de la especie aparecen en clusters diferenciados, por un lado las entradas colectadas en la zona más al sur, zona G, y por otro lado las entradas colectadas en la zona F aunque agrupadas con las de la zona E, un poco más al norte. Las entradas de la zona C aparecen, igual que ocurría en AFLPs en dos clusters uno correspondiente a la cuenca del río Chira y otro a la cuenca del río Piura. En este caso las entradas de la zona de Huancabamba, o zona D, aparecen mayoritariamente en un solo cluster aunque mezcladas con las de la zona C, algunas entradas que habíamos incluido en esta zona y que se colectaron más al sur no se agrupan con ellas.

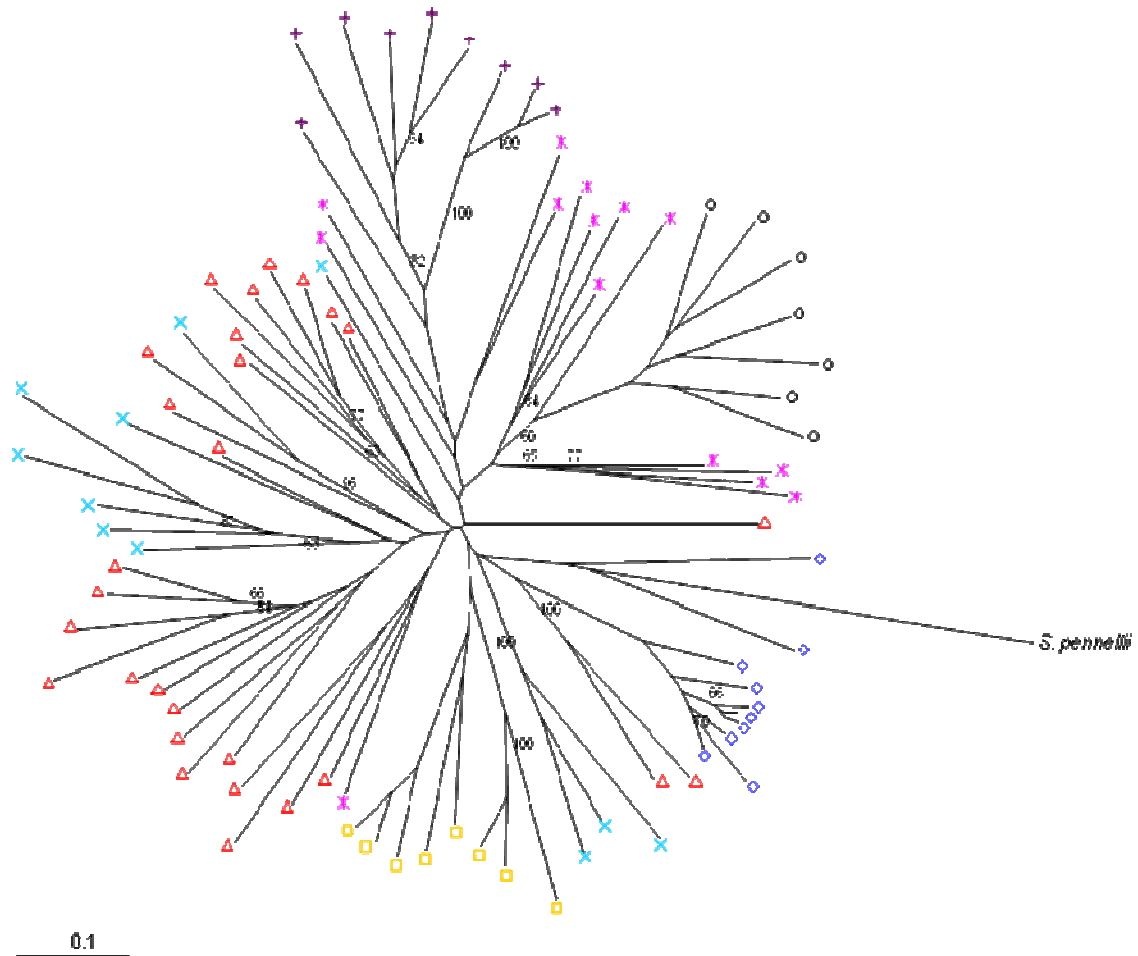


Figura 4.2.4. Dendrograma obtenido para microsatélites empleando la distancia de Cavalli-Sforza y Edwards (1967) y el método Neighbour-Joining para las distintas entradas de *S. habrochaites* y grupos geográficos. Se indican los valores de bootstrap en %. Amarillo: zona A, azul: zona B, rojo: zona C, turquesa: zona D, rosa: zona E, negro: zona F y morado: zona G.

4. 2. 3. Estructura Poblacional

Al igual que hicimos con AFLPs, estudiamos la estructura de las poblaciones con este tipo de marcadores utilizando el programa Structure. Al representar la verosimilitud de que existan K poblaciones frente al número de poblaciones total (Figura 4.2.5) no se observa un máximo, como en el caso de AFLPs sino que la probabilidad va aumentando con el número de poblaciones y llega a una meseta, por lo que asumimos 7 poblaciones.

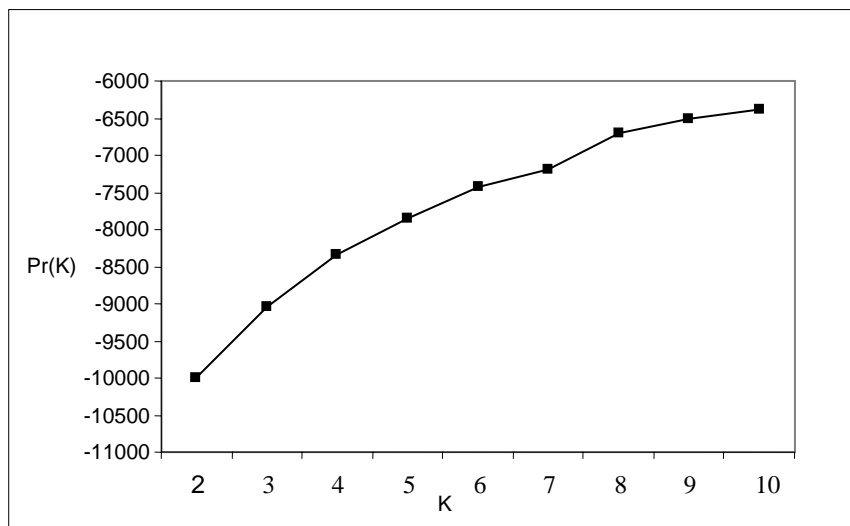


Figura 4.2.5. Representación de la probabilidad de que existan K poblaciones, $Pr(K)$, en función del número estimado de poblaciones, K .

Si, asumiendo 7 poblaciones, representamos la probabilidad de que cada planta pertenezca a cada una de las poblaciones no obtenemos una agrupación geográfica tan clara como en el caso de AFLPs. Las entradas colectadas en las zonas más al sur, E, F y G sí que están bien diferenciadas en tres poblaciones como habíamos sugerido, pero el resto de grupos geográficos no formarían poblaciones definidas. Las entradas de la zona A en el centro de Ecuador estaría formada por dos subpoblaciones. Las entradas de la zona B al este en la provincia de Loja sí que formarían una población. No se separarían de forma clara las entradas de la zona de Huancabamba de las entradas del sur de Ecuador y norte de Perú, zona C, sino que estas entradas de Huancabamba formarían una población con las entradas de sur de Piura más cercanas a éstas (Figura 4.2.6).

En general, podemos decir que los marcadores microsatélites no separan tan claramente las entradas según su procedencia geográfica como en el caso de AFLPs debido, probablemente, a que el número de loci analizados es mucho menor, no obstante son de gran utilidad para saber si hay intercambio de material genético entre las poblaciones y para conocer el grado de alogamia de las plantas en cada una de ellas.

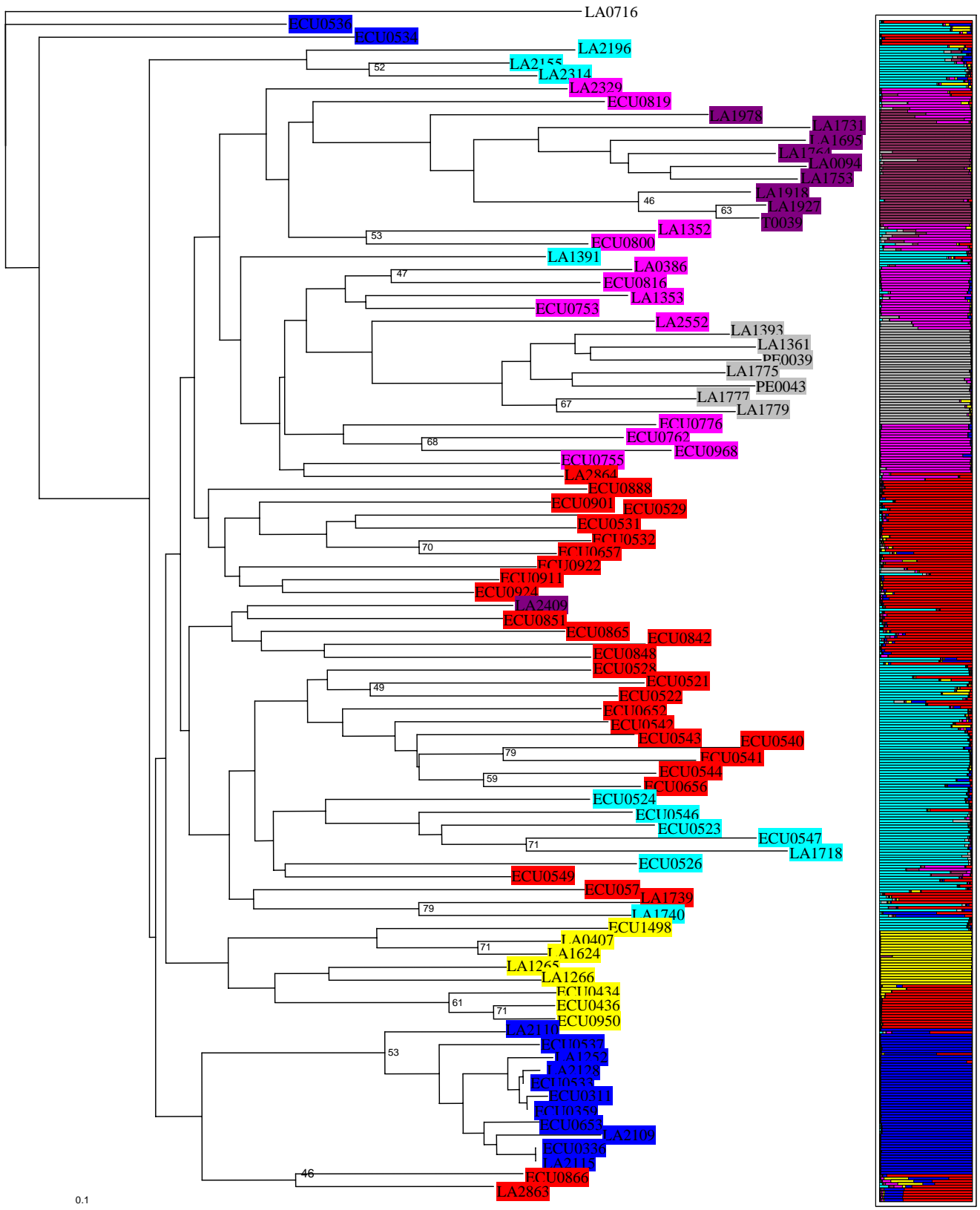


Figura 4.2.6. Dendrograma obtenido con los datos SSR, empleando la distancia de Cavalli-Sforza y Edwards (1967) y el método Neighbour-Joining, en él se representan en diferentes colores las diferentes agrupaciones según criterios ecogeográficos. En el gráfico paralelo se representa la probabilidad de que cada una de las plantas de cada entrada pertenezca a una población u otra definida con el programa Structure considerando que existen 7 poblaciones diferenciadas. Amarillo: zona A, azul: zona B, rojo: zona C, turquesa: D, rosa: zona E, gris: zona F y morado: zona G.

4.3. Diferenciación Regional

El análisis molecular de la varianza (AMOVA) para microsatélites mostró diferencias significativas entre los distintos grupos ecogeográficos ($p < 0.001$). De la variación genética total, el 31.1% fue debida a diferencias entre grupos, el 29.5% se atribuyó a diferencias entre las entradas dentro de cada grupo y el 39.41% de la variación se encontró entre entradas. Todos los componentes de la varianza fueron altamente significativos.

Dado que el 31.1% de la variación es debida a diferencias entre grupos, para estudiar la relación entre los distintos grupos geográficos, se realizó un ACoP y un análisis cluster, tanto para los marcadores AFLP como para los marcadores microsatélite.

En el caso de los marcadores AFLP se utilizó el coeficiente de diferenciación genética entre grupos, G_{st} (Nei, 1973), como medida de las distancias genéticas entre las áreas ecogeográficas. En el ACoP, las dos primeras coordenadas explicaron el 92% de la variación entre grupos (Figura 4.4.1). En él puede observarse la separación clara de las áreas peruanas y ecuatorianas, apareciendo entre ambos grupos las entradas del centro de distribución.

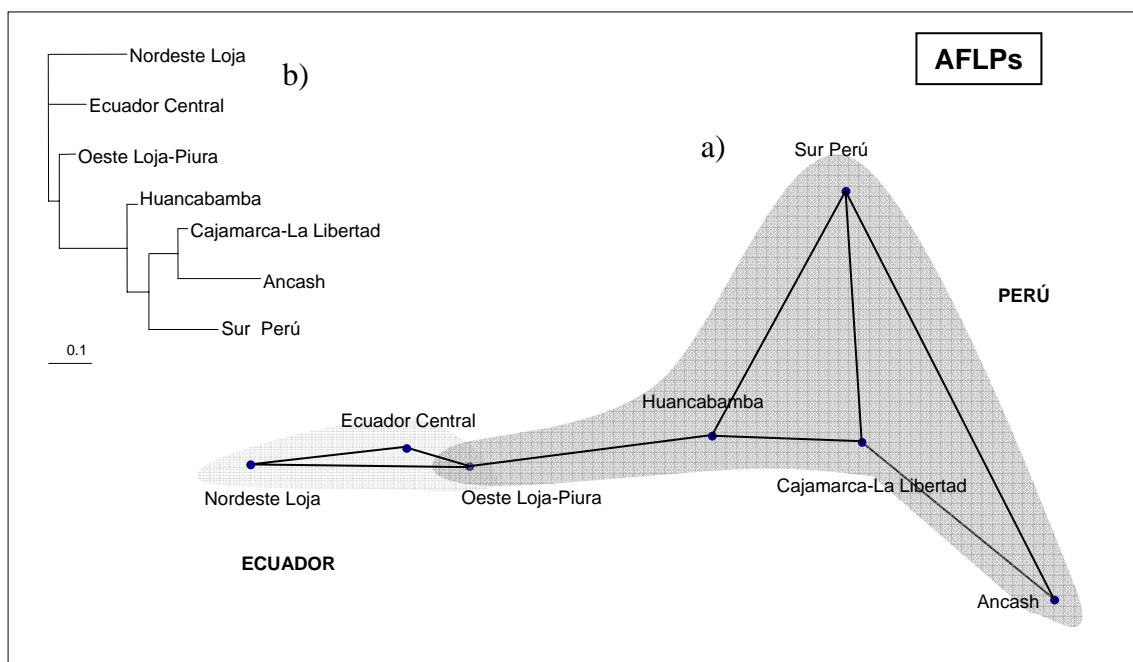


Figura 4.3.1. Análisis PCoA basado en el índice de distancias G_{st} (Nei 1973) para AFLPs (a) y dendrograma obtenido con el método Neighbor-Joining basado en el mismo índice de distancias (b).

El análisis cluster mostró la agrupación de las áreas peruanas, mientras que las ecuatorianas aparecieron separadas y la zona fronteriza entre ambas agrupaciones. La zona C, del oeste de Loja-Piura, aparece agrupada con las entradas del sur.

En el caso de los SSRs se utilizó el coeficiente F_{st} entre grupos (Weir and Cockerman 1984) como estima de las distancias entre áreas ecogeográficas. En el ACoP, las dos primeras coordenadas explicaron el 77% de la variación entre grupos (figura 4.4.2). Al igual que con AFLPs, en este diagrama los grupos ecogeográficos peruanos y ecuatorianos aparecen separados, pero en este caso los grupos ecuatorianos muestran mucha mayor diferenciación genética que en el caso anterior, aunque es difícil comparar ambos diagramas ya que en este caso la diferenciación se calcula a partir de la heterocigosidad.

El análisis cluster agrupó también los grupos ecuatorianos y peruanos, pero en este caso el grupo del sur de Ecuador y norte de Perú se agrupó con los grupos Ecuatorianos.

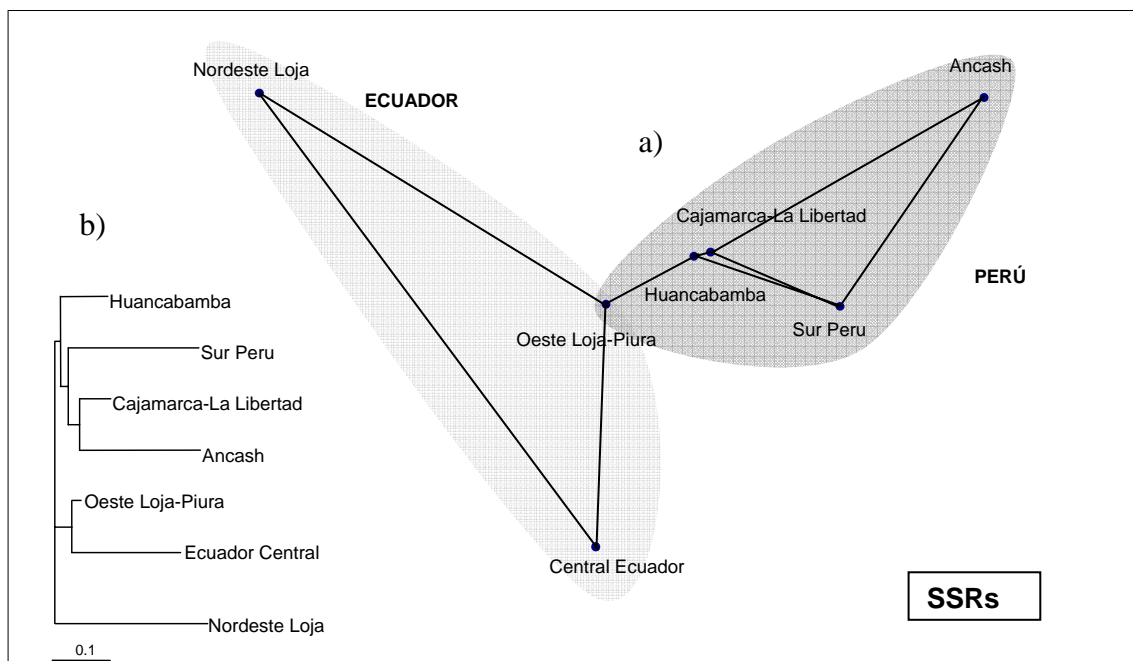


Figura 4.3.2. Análisis PCoA basado en el índice de distancias F_{st} (Weir & Cockerman 1984) para SSRs (a) y dendrograma obtenido con el método Neighbor-Joining basado en el mismo índice de distancias (b).

Con ambos tipos de marcadores los grupos ecuatorianos y los peruanos están bien diferenciados lo que apoya la agrupación por criterios ecogeográficos. Las diferencias entre las representaciones para ambos tipos de marcadores son, probablemente debidas a la diferente naturaleza de ambos tipos de marcadores, que como ya hemos dicho, es dominante en el caso de AFLPs y codominante en el caso de los microsatélites. Además el número de loci analizados en el primer caso es mucho mayor. Otra causa de las diferencias entre las representaciones puede ser que el cálculo de los AFLPs se realiza a partir de la diversidad y el de los SSRs se calcula a partir de la heterozigosidad.

Se calcularon las distancias geográficas entre cada pareja de entradas empleando el método del arco del círculo máximo (Great Circle Distance). Para cuantificar la correlación entre las distancias genética y geográfica se llevaron a cabo diferentes tests de Mantel. El coeficiente de correlación entre las distancias genéticas y geográficas fue de 0.69 para AFLPs y de 0.58 para microsatélites.

4.4. Análisis climático

En otras especies del género como *S. pimpinellifolium* se ha descrito influencia del clima en la variación genética (Zuriaga et al. 2009). Con el fin de estudiar si existía asociación entre el clima y la diferenciación poblacional en *S. habrochaites* se extrajeron diferentes parámetros de clima de la página web <http://www.worldclim.org/>. En ésta aparecen datos de 19 variables climáticas diferentes con una resolución de 2.5 minutos de arco, unos 5 Km². Se obtuvieron datos de seis de estos parámetros climáticos de cada uno de los lugares concretos de colecta para cada una de las entradas. Se analizaron los siguientes parámetros: temperatura media anual, rango de temperatura anual, temperatura media diurna, precipitación anual, estacionalidad en las precipitaciones y estacionalidad en las temperaturas. En la tabla 4.4.1 se presentan las medias y desviaciones estándar de estos parámetros para cada una de las zonas geográficas.

Tabla 4.4.1. Relación de temperaturas y precipitaciones para cada zona geográfica, entre paréntesis aparecen las desviaciones estándar.

	Temperatura media anual (°C)	Rango de temperatura media diaria (°C)	Precipitación anual (mm)	Precipitación estacional (mm)	Temperatura estacional (mm)	Rango de temperatura anual (mm)
Centro de Ecuador (A)	18.7 (5.4)	10.1 (0.9)	904 (343)	70 (24)	4.62 (2.63)	11.9 (0.5)
Noreste de Loja (B)	17.0 (2.1)	10.8 (0.8)	1154 (306)	108 (8)	3.2 (0.96)	12.5 (0.9)
Oeste de Loja-Piura (C)	20.8 (2.9)	12.5 (0.4)	811 (322)	110 (15)	4.65 (2.75)	14.1 (0.5)
Huancabamba (D)	17.8 (3.8)	12.2 (0.9)	796 (239)	57 (36)	4.99 (4.04)	14.5 (1.1)
Cajamarca-La Libertad (E)	13.3 (1.5)	15.2 (0.4)	676 (178)	71 (16)	7.433 (1.88)	18.2 (0.3)
Ancash (F)	11.6 (5.3)	14.2 (1.1)	396 (247)	97 (14)	9.7 (2.6)	18.2 (1.3)
Sur de Peru (G)	10.6 (3.2)	14.9 (1.1)	398 (233)	123 (26)	11.88 (2.78)	19.1 (1.6)

Se realizó un análisis de la varianza para todos estos parámetros observándose que las medias eran significativamente diferentes para todos los parámetros entre el norte y el sur del área de distribución natural de la especie.

Algunos parámetros, como la precipitación anual o la temperatura media anual mostraron un claro gradiente en dirección norte a sur. Las precipitaciones varían desde 1500mm en Ecuador central hasta 200mm en el sur de Perú mientras que las temperaturas varían de 25°C en Ecuador hasta menos de 7°C en algunas regiones del sur de Perú. Lo más destacable es la gran diferencia en la estacionalidad de ambos parámetros, mientras que en el sur de Perú el rango de temperaturas diario puede llegar

a 15°C y el rango de temperaturas anual a los 19°C, en Ecuador estas diferencias entre las máximas y las mínimas no es tan grande, y los rangos en ambos casos son de 10°C. En lo que respecta a las precipitaciones la estacionalidad varía desde los 40mm en el norte de Perú hasta los 146mm en el sur de Perú. En general, podemos decir que en la zona norte del área de distribución de *S. habrochaites* el clima es más uniforme a lo largo del día y del año, mientras que en el sur se producen grandes variaciones.

Con los datos del clima estandarizados se realizó un Análisis de Componentes Principales para comprobar cuáles de estos parámetros explicaban la mayor parte de la variación entre los lugares de colecta de cada una de las entradas (Figura 4.4.1). Las dos primeras componentes explican el 79% de la variación climática. La primera dimensión que explica el 54% de la variación, depende principalmente de la temperatura, tanto del rango de temperaturas como la estacionalidad y está negativamente correlacionada con la temperatura media y la precipitación anuales. Esta componente separa las entradas del norte del área de distribución de las del sur, observándose un gradiente de norte a sur (Figura 4.4.2). La segunda dimensión, que explica el 25% de la variación, está correlacionada negativamente con la temperatura y precipitación estacional y la temperatura media anual y de forma positiva se correlaciona con la temperatura media diurna, el rango anual de temperatura y con la precipitación anual. Las entradas de los diferentes grupos aparecen entremezcladas para esta segunda componente.

Además, a partir de estos datos climáticos se obtuvo una matriz de distancias euclídeas. Se realizó un test de Mantel entre esta matriz y las diferentes matrices de distancias basadas en datos moleculares y distancias geográficas. La correlación entre las distancias geográficas y el clima fue de 0.68, mientras que fue de 0.58 y 0.42 entre éste y la matriz de distancias obtenida para AFLPs y SSRs, respectivamente. En ambos casos las correlaciones fueron menores que para las distancias geográficas, ya que se encuentran poblaciones de diferentes valles en las mismas regiones climáticas. Esto indicaría que posiblemente las diferencias entre las poblaciones se deben más al aislamiento de las entradas de los diferentes valles que a la adaptación a los diferentes ambientes.

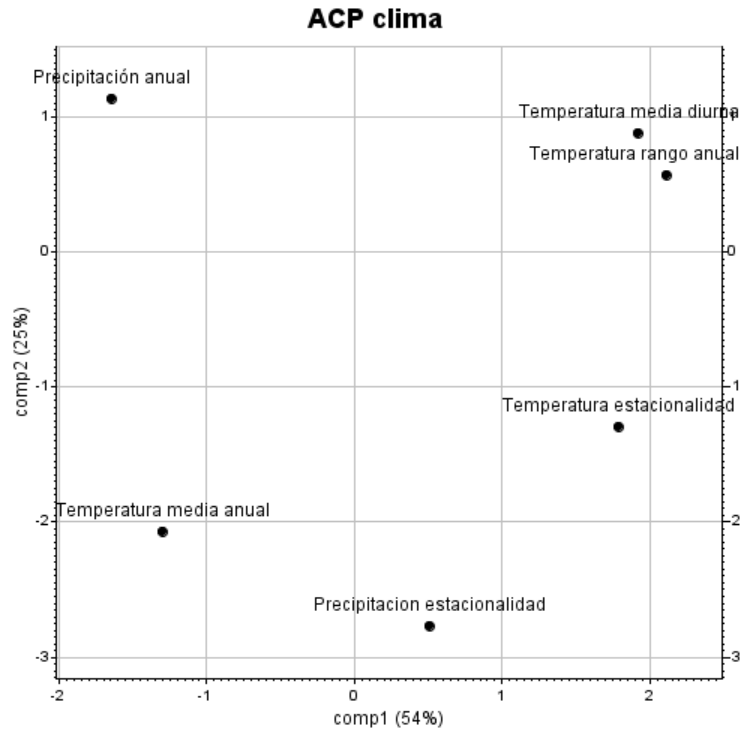


Figura 4.4.1. Análisis de Componentes Principales. Representación de los diferentes parámetros climáticos en los dos ejes principales para la variación climática.

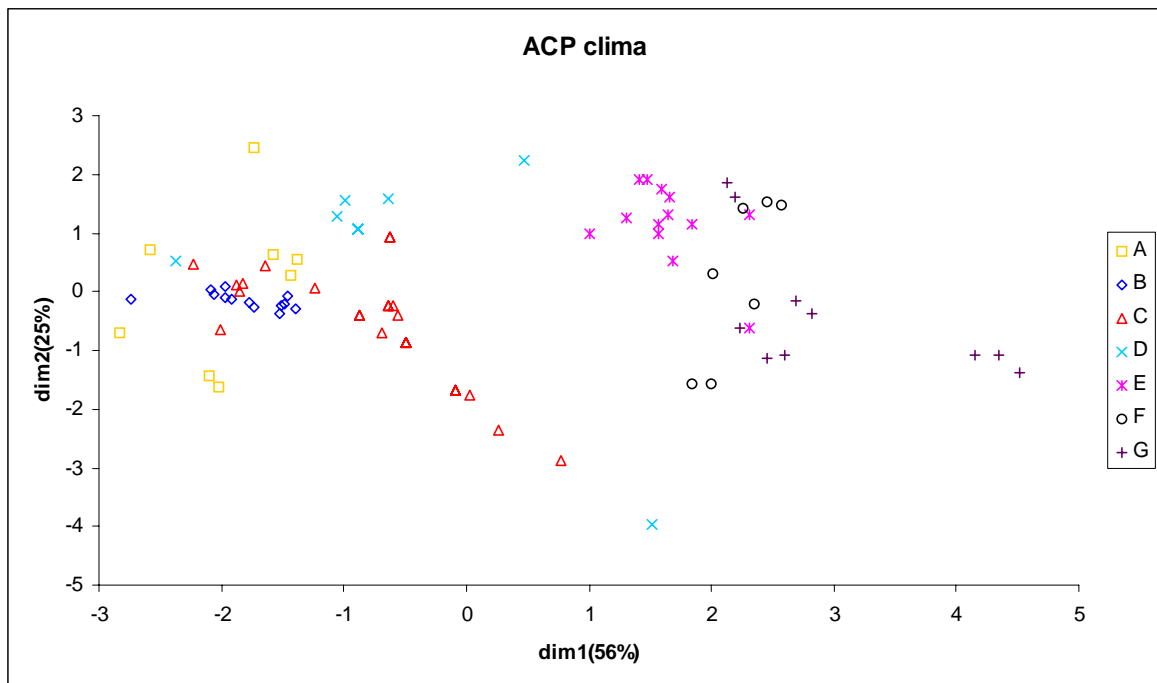


Figura 4.4.2. Representación de las entradas en los dos componentes principales para la variación climática. Se representan con diferentes símbolos las entradas pertenecientes a cada una de las zonas de colecta.

4.5. Datos de Pasaporte

Se hizo una revisión de los datos morfológicos de pasaporte de las entradas que lo disponían. Estos datos fueron tomados en el momento en que se colectaron las entradas en su lugar de origen (tabla 4.5.1).

Teniendo en cuenta estos datos se puede decir que, en cuanto a la pubescencia de la planta se observó un gradiente de norte a sur. Por lo tanto, para este carácter, no podemos confirmar la descripción de Müller (1940) de dos formas claramente diferenciadas *glabratrum*, con menor pubescencia en Ecuador, y *typicum* densamente pubescente en Perú.

No se encontró ninguna clina en el tamaño de la corola o la ejerción estigmática y tampoco aparece una correlación entre ellas y la heterocigosidad, en contraposición a lo afirmado por Rick *et al.* (1979) que observaron una corola más pequeña en las plantas autoincompatibles.

Al parecer, las formas encontradas por Müller serían las formas extremas de un continuo de variación.

Table 4.5.1. Datos morfológicos para cada región geográfica obtenidos de los datos de pasaporte

Zona Geográfica	Pubescencia de la planta	Tamaño floral	Exerción estigmática
A: Ecuador Central	baja	df	df
B: Noreste de Loja	media-baja	medio-grande	exerto
C: Oeste de Loja-Piura	media-baja	pequeño-medio-grande	ligeramente exerto
D: Huancabamba	media-alta	pequeño-medio-grande	muy exerto
E: Cajamarca-La Libertad	alta	pequeño-grande	ligeramente exerto
F:Ancash	alta	pequeño	ligeramente exerto
G:Sur de Perú	df	df	df

Nota:df: datos faltantes

4.6. Discusión

A la hora de realizar estudios de germoplasma, los análisis de la diversidad genética de plantas son muy importantes. Nos permiten definir la estructura poblacional de la especie estudiada y evaluar su nivel de erosión genética. Así mismo, ayuda a establecer sus prioridades de conservación. Permite una gestión, multiplicación y conservación más eficientes de las colecciones de germoplasma. Además facilita su utilización por parte de los mejoradores. En este sentido, *S. habrochaites* es una importante fuente de nuevos genes para la mejora genética del tomate (Sacks & St. Clair 1998). Presenta resistencia a muchas enfermedades y plagas, tolerancia al frío y caracteres de calidad del fruto (Rick 1982).

La caracterización molecular realizada en este estudio ha aportado importante información acerca de la estructura genética poblacional de *S. habrochaites* en todo su rango de distribución. Hemos incluido numerosas entradas de las zonas norte y central del rango de distribución, así como de la zona sur, menos estudiada, complementando de esta manera estudios realizados anteriormente (Rick *et al.* 1979, Ercolano *et al.* 2005).

Los análisis realizados, tanto el AMOVA, como los análisis multivariantes (ACoP) y los diferentes análisis cluster, incluyendo el realizado con STRUCTURE, para AFLP y SSR demuestran que la estructura genética de esta especie está principalmente correlacionada con una diferenciación geográfica. Esto ocurre también en otras especies de plantas (Bornet *et al.* 2002, Coulibaly *et al.* 2002, Phillips *et al.* 2008).

En base a criterios ecogeográficos, y confirmándolo con datos moleculares y climáticos, se definieron siete regiones diferenciadas: dos regiones en Ecuador, una en la frontera entre Ecuador y Perú y cuatro en Perú. El AMOVA mostró diferencias significativas entre los siete grupos ecogeográficos.

Para AFLPs, tanto el ACoP como el dendrograma agrupan por un lado las entradas de Ecuador y por otro las entradas de Perú, mientras que las entradas de la región de Huancabamba en Perú aparecen entre ambas agrupaciones. También para SSRs se agrupan las entradas de las diferentes áreas aunque no tan claramente como con AFLP. Esto es debido probablemente al gran número de loci analizados con los AFLP.

Otra indicación de la diferenciación geográfica es la exclusividad de algunos alelos microsatélite encontrados en algunas regiones: 9 alelos en el norte, 16 en la zona central y 17 en las poblaciones del sur. Las regiones con mayor diversidad fueron las de la parte central del área de distribución de la especie, desde el noreste de Loja hasta La Libertad. Rick *et al.* (1979) describió estas poblaciones centrales como los posibles ancestros de las poblaciones actuales.

Además en estas áreas centrales la heterocigosidad es mayor, apoyando la idea de que estas poblaciones presentan una mayor alogamia. La heterocigosidad y diversidad genética alcanzan un máximo en la región de Huancabamba mostrando ocho alelos exclusivos donde Rick *et al.* (1979) ya habían descrito un gran polimorfismo con estudios con alozimas. Algunos autores han descrito esta área como un importante punto de biodiversidad de diferentes especies de plantas (Young y Reynel 1997, Weingend 2002). Esta alta diversidad genética puede ser debida a diferentes factores. Esta región podría ser la zona de origen de las especies, lo cual explicaría la gran diversidad encontrada. Además la localización de esta zona en el centro de diferentes regiones geoclimáticas y hábitats podría permitir la incorporación de diversidad de estas regiones colindantes, provocando un aumento de biodiversidad en esta zona, pero limitando el flujo genético entre el norte y el sur. Otro factor podría ser que este grupo tuviese una subestructura. Las entradas de Lambayeque y Cajamarca que aparecen separadas en los análisis de AFLP podrían ser indicadoras de este efecto, ya que podrían pertenecer a una población diferenciada. En un futuro se debería tener en cuenta esta zona, poco colectada, a la hora de plantearse nuevas expediciones de colecta.

Las poblaciones de las zonas ecogeográficas norte y sur muestran una menor heterocigosidad y una menor diversidad. Se debe remarcar la extremadamente baja heterocigosidad del sur de Perú considerando la gran cantidad de alelos por locus encontrados allí. Estos datos sugieren que en esta región, donde se han encontrado pequeñas poblaciones aisladas, las plantas tienden a la autopolinización. Estos resultados concuerdan con la autopolinización encontrada por Rick *et al.* (1979) en estas poblaciones marginales. También se han encontrado estos mismos patrones de variabilidad genética en *S. pimpinellifolium* (Rick *et al.* 1977, Zuriaga *et al.* 2009a).

En relación a la forma *glabratum* descrita por Muller (1940), nosotros no hemos encontrado dos formas claramente diferenciadas en relación a una reducción de la

pilosidad de tallos y hojas y una corola más pequeña. Únicamente encontramos un gradiente o clina para la pubescencia que aumenta de norte a sur, considerando los datos morfológicos extraídos de las hojas de pasaporte.

No encontramos ninguna clina en el tamaño de la corola, y no existe correlación entre el tamaño de las flores o la ejerción estigmática y la heterocigosidad. Estas observaciones se contradicen con las de Rick *et al.* (1979), que observaron una corola más pequeña en las plantas autocompatibles. Las principales diferencias en la morfología entre las poblaciones de las diferentes cuencas o valles podrían deberse al aislamiento geográfico de las mismas.

Al igual que ocurría en *S. pimpinellifolium* (Zuriaga *et al.* 2009a) se observa una alta correlación entre la diferenciación genética y el clima, aunque se encontró una mayor correlación entre las distancias genéticas y geográficas. Esto, podría ser debido a los diferentes hábitats de estas dos especies. *S. habrochaites*, vive en valles y montañas a grandes altitudes y zonas muy húmedas, mientras que *S. pimpinellifolium* habita sitios tan diferentes como el bosque tropical del norte de Ecuador o el desierto de la costa peruana. Diferentes poblaciones de *S. habrochaites* pueden vivir bajo las mismas condiciones climáticas a pesar de estar muy alejadas geográficamente, razón por la cual la diferenciación podría ser principalmente debida al aislamiento geográfico.

Nuestros resultados apoyan la hipótesis de evolución de *S. habrochaites* propuesta por Rick *et al.* (1979) y también descrita en otras especies del mismo género como *S. pimpinellifolium* (Rick *et al.* 1977, Zuriaga *et al.* 2009a). Ellos sugieren que las poblaciones centrales autoincompatibles podrían ser los biotipos más antiguos de *S. habrochaites* de los cuales habrían derivado los demás. A medida que las especies migraron hacia el norte y sur de su rango de distribución se irían convirtiendo en autocompatibles posiblemente por una selección favorecida por la baja densidad de población y por los genotipos menos frecuentes o por efecto fundador conforme a la ley de Baker (1955).

Este patrón de variación se asemeja al fenómeno de “emancipación de los recesivos”, observado por Vavilov (1927) cuando las poblaciones se alejan de su centro de diversidad. Este autor defiende que a medida que las poblaciones migran lejos de su

lugar de origen, la diversidad genética se ve disminuida y la frecuencia de los homocigotos alcanza máximos.

Conclusiones

Probé a saltar sin red ni hogar.
No sé volver. No se hacia donde. Ni con quién.
V. M.

5. Conclusiones

1. Al incluirse numerosas entradas de todo el rango de distribución de *S. habrochaites*, el estudio molecular realizado en este estudio ha aportado importante información acerca de la estructura genética poblacional de esta especie, complementando de esta manera los escasos estudios anteriores.

2. Los diferentes análisis realizados, tanto los análisis multivariantes como los análisis cluster, para AFLP y SSR demuestran que la estructura genética de esta especie está principalmente correlacionada con una diferenciación geográfica.

3. El Análisis Molecular de la Varianza mostró diferencias significativas entre los siete grupos ecogeográficos que se definieron *a priori* en base a criterios ecogeográficos y se confirmaron con datos moleculares y climáticos.

4. El Análisis de Coordinadas Principales obtenido para AFLPs mostró la existencia de una clina de norte a sur en la variación genética. Las entradas de los diferentes grupos geográficos se ordenaron secuencialmente de norte a sur mostrándose una clara separación entre ellos.

5. Para SSRs, también, se agruparon las entradas de las diferentes áreas aunque no tan claramente como en el caso de AFLPs, encontrándose alelos exclusivos del norte, sur y área central del área de distribución.

6. Los grupos de entradas de las regiones del centro del área de distribución mostraron una mayor variación genética y heterocigosidad, mientras que los grupos de los extremos mostraron menor variabilidad y una mayor homocigosis.

7. La heterocigosidad y diversidad genética alcanzan un máximo en la región de Huancabamba en Perú, que ya había sido descrita como un importante punto de biodiversidad para diferentes especies de plantas. Esto podría ser debido a que es parte del centro de origen de la especie; a la característica orografía del terreno que hace que sea un punto de entrada de material genético, pero no de salida; o a que se han considerado dos subpoblaciones diferentes dentro del mismo grupo geográfico.

8. En relación a los datos morfológicos de pasaporte, no encontramos los morfotipos *typicum* y *glabratum* considerados por Müller, que serían las formas

extremas de un continuo de variación. No existe correlación entre el tamaño de las flores o la ejerción estigmática y la heterocigosidad. Únicamente se encontró una clina en la pubescencia de norte a sur aunque que no asociada al grado de homocigosis.

9. Se observó una alta correlación entre la diferenciación genética y el clima, aunque la mayor correlación se dio entre las distancias genéticas y geográficas. Diferentes poblaciones de *S. habrochaites* pueden vivir bajo las mismas condiciones climáticas a pesar de estar muy alejadas geográficamente, por lo que la diferenciación estaría principalmente debida al aislamiento geográfico.

10. Los marcadores AFLPs han proporcionado información muy útil de cómo se distribuye la variación genética a lo largo del rango de distribución de la especie, mientras que los microsatélites han sido de gran utilidad para saber si ha habido intercambio de material genético entre las poblaciones y para conocer información acerca del grado de alogamia de las plantas en cada una de ellas.

11. Nuestros resultados apoyan la hipótesis de evolución de *S. habrochaites* propuesta por autores anteriores que sugieren que las poblaciones centrales autoincompatibles podrían ser los biotipos más antiguos de los cuales habrían derivado los demás. A medida que las especies migraron hacia el norte y sur de su rango de distribución se fueron convirtiendo en autocompatibles debido al efecto fundador, a la ventaja de fijar uno o pocos genotipos prósperos, a la dificultad para obtener el tamaño poblacional necesario o a la escasez de insectos polinizadores.

12. Los análisis de variabilidad genética realizados en este estudio son de gran ayuda para la gestión de los Bancos de Germoplasma. Podemos asegurar que las entradas utilizadas en este trabajo y que dicen ser de *S. habrochaites* sí lo son y si por cuestiones de espacio o económicas tuviésemos que escoger qué variedades se deberían conservar, esas serían las más variables, aunque rechazando las otras se perdieran alelos raros.

13. En cuanto al Mejoramiento de cultivos, gracias a los análisis de diversidad genética sabemos qué entradas son las más variables. En general son un importante recurso de caracteres agronómicos deseables.

Bibliografía

6. Bibliografía

- Anuario de Estadística Agroalimentaria (2008) [On line] En http://www.mapa.es/es/estadistica/pags/anuario/Anu_04/indice.asp
- Areshchenkova T, Ganal MW (1999) Long tomato microsatellites are predominantly associated with centromeric regions. *Genome* 42: 536- 544.
- Areshchenkova T, Ganal MW (2002) Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources. *Theoretical and applied genetics* 104: 229-235.
- Ayers T (1999) Biogeography of *Lysipomia* (Campanulaceae), a high elevation endemic: an illustration of species richness at the Huancabamba Depression, Peru. *Arnaldoa* 6: 13-28
- Baker HG (1955) Self-compatibility and establishment after long-distance dispersal. *Evolution* 9: 347-348
- Belkhir K, Borsa P, Chikhi L, Rafaste N, Bonhomme T (1996) Genetix 4.04 Logiciel sours Windows TM pour la genétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier <http://www.genetix.univ-montp2.fr/genetix/genetix.htm>
- Benham J, Jeung JU, Jasieniuk M, Kanazin V, Blake T (1999) Genographer: a graphical tool for automated fluorescent AFLP and microsatellite analysis. Department of Plant Science, Montana State University, Bozeman
- Bohs L (2007) Phylogeny of the cyphomandra clade of the genus *Solanum* (Solanaceae) based on ITS sequence data. *Taxon* 56: 1012-1026
- Bornet B, Goraguer F, Joly G, Brachard M (2002) Genetic diversity in European and Argentinian cultivated potatoes (*Solanum tuberosum* subsp. *tuberosum*) detected by inter-simple sequence repeats (ISSRs) *Genome* 45: 481-484
- Bouxin G (2005) Ginkgo, a multivariate analysis package. *Journal of Vegetables Science* 16: 355-359
- Bucheli, P (1999) Definition of biochemical and molecular markers (quality trait loci) for tomato flavour as tools in breeding. *Acta horticulturae* 301-306.

- Carter CD, Gianiagna TJ, Sacalis JN (1989) Sesquiterpenes in glandular trichomes of a wild tomato species and toxicity to the Colorado potato beetle. *Journal of Agriculture and Food Chemistry* 37: 1425
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution* 32: 550-570
- Child A (1990) A synopsis of *Solanum* subgenus *Potatoe* (G. Don) D'Arcy (*Tuberarium* (Dun.) Bitter s.l.). *Feddes Repert. Spec. Nov. Regni Veg.* 101: 209-235
- Correll DS (1958) A new species and some nomenclatural changes in *Solanum* section *Tuberarium*. *Madroño* 14: 232-236.
- Cortada L (2009) Response of tomato rootstocks carrying the Mi-resistance gene to populations of *Meloidogyne arenaria*, *M. incognita* and *M. javanica*. *European Journal of Plant Pathology* 124: 337-343.
- Cortada, L (2008) Variability in infection and reproduction of *Meloidogyne javanica* on tomato rootstocks with the Mi resistance gene. *Plant pathology* 57: 1125-1135.
- Coulibaly S, Pasquet R, Papa R, Gepts P (2002) AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types. *Theoretical and applied genetics* 104: 358-366
- D'Arcy WG (1991) The *Solanaceae* since 1976, with a review of its biogeography. En "Hawkes JG, Lester RN, Nee M, Estrada N (Eds) *Solanaceae* III: Taxonomy, Chemistry, Evolution, Royal Botanic Gardens, Kew": 75-137
- Darwin SC, Knapp S, Peralta I (2003) Taxonomy of tomatoes in the Galápagos Islands: native and introduced species of *Solanum* section *Lycopersicon* (Solanaceae). *Systematics and Biodiversity* 1 (1): 29-53
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun e, Lathrop, M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380: 152-154

- Dimock, MB (1983) The role of glandular trichomes in the resistance of *Lycopersicon hirsutum f. glabratum* to *Heliothis zea*. *Entomologia experimentalis et applicata* 33: 263-268.
- Ercolano MR, Sebastiano A, Monti L, Frusciante L, Barone A (2005) Molecular characterization of *Solanum habrochaites* accessions. *Journal of Genetics and Breeding* 59: 15-20
- Esquinas-Alcázar J, Nuez F (1995) Situación Taxonómica, Domesticación y Difusión del Tomate. En “Nuez F (Ed) El cultivo del tomate. Ediciones Mundi-Prensa. Madrid, Barcelona, México”: 14-42
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50
- FAOSTAT 2009. Agriculture data [On line] Disponible en <http://faostat.fao.org/site/612/default.aspx#ancor> 03/12/2009
- Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) v3.69. Department of Genome Sciences and Department of Biology. University of Washington. Washington. EEUU. Disponible en <http://evolution.genetics.washington.edu/phylip.html>
- Foolad MR, Merk HL, Ashrafi H (2008) Genetics, genomics and breeding of late blight and early blight resistance in tomato. *Critical Review of Plant Science* 27: 75-107
- García-Mas J, Graciano E, Aranzana MJ, Monforte A, Oliver M, Ballester J, Viruel MA, Arús P (2000) Marcadores de ADN: concepto, tipos, protocolos. En “Nuez F, Carrillo JM (Eds) Los marcadores genéticos en la mejora vegetal. Universidad Politécnica de Valencia”: 91-151
- Gisbert C, Rus AM, Bolarin MC, Lopez-Coronado JM, Arrillaga I, Montesinos C, Caro M, Serrano R, Moreno V (2000) The yeast HAL1 gene improves salt tolerance of transgenic tomato. *Plant Physiology* 123: 393-402.
- Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of heredity* 88: 335-342

- Guo ZH, Weston PA, Snyder JC (1993) Repellency to 2-spotted spider-mite, *tetranychus-urticae* koch, as related to leaf surface-chemistry of *Lycopersicon hirsutum* accessions. *Journal of Chemical Ecology* 19: 2965-2979
- Hamada H, Petrino MG, Kakunaga T (1982) A novel repeated element with Z-DNA-forming potencial is widely found in evolutionary diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the USA* 79: 6465-6469
- Harlan JR (1992) *Crops and man*. American Society of Agronomy, Inc. p.217-246
- Juvik JA (1982) Resistance among accessions of the genera *Lycopersicon* and *Solanum* to four of the main insect pests of tomato in Israel. *Phytoparasitica* 10: 145-156.
- Kaloo G (1991) Breeding for environmental stress resistance in tomato. En “Kaloo G (ED) Genetic improvement of tomato. Springer- Verlag, Berlín y Heidelberg”: 153-165
- Kaloo G (1993) Tomato. En “Kaloo G, Bergh BO. Genetic improvement of vegetable crops. Pergamon Press”: 645-666
- Kenedy GC, Yamamoto RT, (1979) Resistance in a wild tomato to the tobacco hornworm: the presence of a toxic factor. *Entomologia experimentalis et applicata* 26: 121-126
- Knapp SL, Bohs MN, Spooner DM (2004) Solanaceae: a model for linking genomics and biodiversity. *Comparative and Functional Genomics* 5: 285-291
- Langella O (2002) Populations 1.2.28, Population genetic software. CNRS <http://bioinformatics.org/~tryphon/populations/>
- Leite G, Picanco M, Guedes R, Zanuncio J (2001) Role of plant age in the resistance of *Lycopersicon hirsutum* f. *glabratum* to the tomato leafminer *Tuta absoluta* (*Lepidoptera:Gelechiidae*). *Scientia horticultrae* 89: 103-113
- Lin SYH (1987) Activity of volatile compounds in glandular trichomes of *Lycopersicon* species against two insect herbivores. *Journal of chemical ecology* 13: 837-850.
- Luckwill LC (1943) The genus *Lycopersicon*. A historical, biological and taxonomic survey of the wild and cultivated tomatoes. *Aberdeen University Studies* 120: 1-44

- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-202.
- Mapa Bioclimático de Ecuador (1978) Ministerio de Agricultura y Ganadería. Programa Nacional de Regionalización Agraria, PRONAREG, Ecuador.
- Mapa Hidrográfico del Perú. República del Perú, Instituto Geográfico Nacional, Sistemas de Información Geográfica.
- Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics* 80: 437-448.
- Monforte A, Tanksley S (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L-esculentum* genetic background: A tool for gene mapping and gene discovery. *Genome* 43: 803-813
- Muller CH (1940) A revision of the genus *Lycopersicon*. United States Department of Agriculture. Miscellaneous Publications No 382
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70: 3321-3323
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* 76: 5269-5273
- Page RDM (2001) TreeView 1.6.6 University of Glasgow. Glasgow. Escocia. Disponible en <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html> 12/06/2003
- Palmer JD, Zamir D (1982) Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. *Proceedings of the National Academy of Sciences of the United States of America* 79: 5006-5010.
- Paterson BD, Mutton L, Paull RE, Nguyen VQ (1987) Tomato pollen development: stages sensitive to chilling and a natural environment for the selection of resistant genotypes. *Plant Cell and Environment* 10: (5) 363-368
- Peralta I, Knapp S, Spooner DM (2005) New species of wild tomatoes (*Solanum* section *Lycopersicon*: *Solanaceae*) from Northern Peru. *Systematic Botany* 30: 424-434

- Peralta I, Knapp S, Spooner DM (2008) Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). En “Anderson C (Ed) Systematic Botany Monographs vol 84” The American Society of Plant Taxonomists
- Peralta I, Spooner DM (2000) Classification of wild tomatoes: a review. *Kurtziana*. Tomo 28 (1): 45-54
- Phillips N, Larson S, Drost D (2008) Detection of genetic variation in wild populations of three *Allium* species using amplified fragment length polymorphisms. *HortScience* 43: 637-643
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959 Disponible en <http://pritch.bsd.uchicago.edu/structure.html> 03/12/2009
- Qian H, Ricklefs RE, White PS (2005) Beta diversity of angiosperms in temperate floras of eastern Asia and eastern North America. *Ecology Letters* 8: 15-22
- Quiros CF, Stevens MA, Rick CM, Kok-Yokomi ML (1977) Resistance in the tomato to the pink form of the potato aphid (*Macrosiphum euphorbiae* Thomas): the role of anatomy, epidermal hairs, and foliage composition. *Journal of the American Society of Horticulture and Science* 102: 166
- Rick CM, Chetelat RT (1991) The breakdown of self-incompatibility in *Lycopersicon hirsutum*. En “Hawkes JG, Lester RN, Nee M, Estrada N (Eds) Solanaceae III: Taxonomy, Chemistry, Evolution. Royal Botanic Gardens, Kew”: 253-256
- Rick CM (1960) Hybridization between *Lycopersicon Esculentum* and *Solanum Pennellii*: Phylogenetic and Cytogenetic Significance. *Proceedings of the National Academy of Sciences of the United States of America* 46: 78-82.
- Rick CM (1963) Barriers to interbreeding in *Lycopersicon peruvianum*. *Evolution* 17: 216-232
- Rick CM (1971) The Tomato Ge Locus: Linkage Relations and Geographic Distribution of Alleles. *Genetics* 67: 75-85.
- Rick CM (1973) Potential genetic resources in tomato species: clues from observations in native habitats. En “Hollaender A, Srb AM (Eds) Genes, Enzymes and Populations. Plenum Press. Nueva York. EEUU”

- Rick CM (1978) El Tomate. *Investigación y Ciencia* 25: 45-55
- Rick CM (1982) The potential of exotic germplasm for tomato improvement. En: "Vasil IK, Scowcroft WR, Frey KJ (Eds) *Plant Improvement and Somatic Cell Genetics*. Academic Press, Inc." pp 1-28
- Rick CM (1991) Tomato paste: a concentrated review of genetic highlights from the beginnings to the advent of molecular genetics. *Genetics* 128: 1-5.
- Rick CM, Fobes JF, Holle M (1977) Genetic variation in *Lycopersicon pimpinellifolium*: evidence of evolutionary change in mating systems. *Plant Systematics and Evolution* 127: 139-170
- Rick CM, Fobes JF, Tanksley SD (1979) Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. *Plant Systematics and Evolution* 132: 279-298
- Rick CM, Yoder JI (1988) Classical and molecular genetics of tomato: highlights and perspectives. *Annual Review of Genetics* 22: 281-300.
- Ronen G (1999) Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. *Plant journal: for cell and molecular biology* 17: 341-351.
- Roselló S, Nuez F (2006) Mejora de la calidad del tomate para fresco. En "Llácer G, Díez MJ, Carrillo JM, Badenes ML (Eds) *Mejora genética de la calidad en plantas*. SECH, SEG, UPV": 333-359
- Sacks E, St. Clair D (1998) Variation among seven genotypes of *Lycopersicon esculentum* and 36 accessions of *L. hirsutum* for interspecific crossability. *Euphytica* 101: 185-191
- Sifres A (2003) Variabilidad de *Lycopersicon* spp. en su área de distribución natural influenciada por el fenómeno climatológico de "El Niño". PhD Thesis, Universidad Politécnica de Valencia, Valencia, Spain
- Smulders MJM, Bredemeijer G, Rus-Kortekaas W, Arens P, Vosman B (1997) Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics* 97: 264-272

-
- Solanaceae Genomics Network (2006) Maps and Markers [On line] En <http://www.sgn.cornell.edu/> 03/12/2009
- Soler S, Picó B, Sifres A, Frutos R de, Nuez F (2000) Multiple virus resistance in a collection of *Lycopersicon spp.* 5th EFPP Congress (18-22 Sep 2000, Taormina, Italia): 17
- Spooner DM, Anderson GJ, Jansen RK (1993) Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes and pepinos (*Solanaceae*). American Journal of Botany 80: 676-688
- Spooner DM, Peralta I (2001) Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum L.* section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). American Journal of Botany 88: 1888
- Stevens MA, Rick CM (1986) Genetics and breeding. En “The tomato crop: a scientific basis for improvement. Atherton JG, Rudich J (Eds) Chapman & Hall, London” p. 35
- Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063-1066.
- Taylor IB (1986) Biosystematics of the tomato. En: “Atherton JG, Rudich J (Eds) The Tomato Crop. A scientific basis for improvement. Chapman and Hall, London & New York”: 1-34
- Tomato Genetics Resource Center (2009) Seed Catalogue [On line] <http://tgrc.ucdavis.edu/>
- Vavilov NI (1927) Geographical regularities in the distribution of the genes of cultivated plants. Bulletin of Applied Botanical and Genetic Plant Breeding 17: 411-428
- Vidavsky F, Czosnek H (1998) Tomato breeding lines resistant and tolerant to tomato yellow leaf curl virus issued from *Lycopersicon hirsutum*. Phytopathology 88: 910-914
- Vos P, Hogers R, Bleeker M, Reijans M, Van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic acids research 23: 4407-4414.

- Warnock SJ (1988) A review of taxonomy and phylogeny of the genus *Lycopersicon*. *HorScience* 23: 669-673
- Warnock SJ (1991) Natural habitats of *Lycopersicon* species. *Horscience* 26 (5): 446-441
- Webb RE, Stoner AK, Gentile AG (1971) Resistance to leafminers in *Lycopersicon* accessions. *Journal of the American Society for Horticultural Science* 95: 65-67
- Weber JL (1990) Informativeness of human (dC-dA)_n-(dG-dT)_n polymorphism. *Genomics* 7: 524-530
- Weigend M (2002) Observations on the biogeography of the Amotape-Huancabamba Zone in Northern Peru. *The botanical review* 68: 38-54
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370
- Williams EG, Kenedy GG, Yamamoto RT, Thacker JD, Bordner J (1980) 2-tridecadone: a naturally occurring insecticide from the wild tomato *Lycopersicon hirsutum* f. *glabratum*. *Science* 207: p.888
- Witting-Bissinger B (2008) Novel arthropod repellent, BioUD, is an efficacious alternative to Deet. *Journal of medical entomology* 45: 891-898
- Yeh FC, Yang RC, Boyle TBJ, Ye ZH and Mao JX (1997) POPGENE, the user-friendly shareware for population genetic analysis. *Molecular Biology and Biotechnology Centre, University of Alberta, Alberta*
<http://www.ualberta.ca/~fyeh/index.htm>
- Young KR, Reynel C (1997) Huancabamba Region, Peru and Ecuador. In: Davis SD, Heywood VH, Herrera-MacBryde O, Villa-Lobos J, Hamilton AC (Eds) *Centres of plant Diversity: A guide and strategy for their conservation*, 3: North America, Middle America, South America, Caribbean Island. IUCN, Cambridge, England
- Zamir D (2001) Improving plant breeding with exotic genetic libraries. *Natural Review of Genetics* 2: 983-989

- Zhang, Peng-Ying (2008) Effects of crop development on the emission of volatiles in leaves of *Lycopersicon esculentum* and its inhibitory activity to *Botrytis cinerea* and *Fusarium oxysporum*. *Journal of integrative plant biology* 50: 84-91.
- Zuriaga E, Blanca J, Cordero L, Sifres A, Blas-Cerdan W, Morales R (2009a) Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genetic Resources and Crop Evolution* 56: 39-51
- Zuriaga E, Blanca J, Nuez, F (2009b) Classification and phylogenetic relationships in *Solanum* section *Lycopersicon* based on AFLP and two nuclear gene sequences. *Genetic Resources and Crop Evolution* 56: 663-678

Anexo

Anexo:

Representación de las frecuencias alélicas encontradas para los diferentes loci microsatélite en *Solanum habrochaites*.

