



UNIVERSIDAD POLITÉCNICA DE VALENCIA

TESIS DOCTORAL

*Análisis de la variabilidad en poblaciones naturales de Solanum,
secciones Lycopersicon y Basarthurum*

Presentada por:

ELENA ZURIAGA GARCÍA

**Para optar al título de Doctor Ingeniero Agrónomo
por la Universidad Politécnica de Valencia**

Dirigida por:

Dr. D. Jose Miguel Blanca Postigo

Dr. D. Fernando Nuez Viñals

Valencia, 2009

AGRADECIMIENTOS

Ha sido un largo viaje y durante todo el recorrido ha sido mucha la gente que me ha echado una mano de una forma u otra y a la que tengo mucho que agradecer.

En primer lugar quería agradecer a mis directores de tesis la dedicación que me han prestado y que me ha permitido llegar a buen puerto. Al Dr. Fernando Nuez quiero agradecerle la oportunidad que me brindó de trabajar en el COMAV y de desarrollar esta tesis. Al Dr. Jose Blanca agradecerle su gran apoyo y su entrega sin la que, sin dudarlo, no habría llegado hasta aquí.

Además, hay un montón de gente a la que me gustaría agradecer un montón de cosas. A todos mis amigos y compañeros del COMAV que han compartido conmigo buenos y malos momentos. Muchísimas gracias a mis amigos Laura P., Laura C., Ana P., Pascualet, Javi, Mariola...; a aquellos que me han dado un montón de buenos consejos como Ximo, Carmelo, Cris, Jaime C., Belén, Jaume P., Santi, Carmina, Miguel, María F...; a todos los de banco e invernaderos, en especial a Jose Vicente y Pascual que siempre han estado dispuestos a ayudarme... No quiero dejarme a nadie, así que, en general, un beso a todos.

También quiero agradecer a mis nuevos amigos y compañeros del departamento de Frutales del IVIA la acogida que me han brindado y su apoyo en la fase final de esta tesis. Un beso grande a todos.

A mis amigas Raquel y Sheila, un millón de gracias por estar ahí a cualquier hora.

Mil gracias a toda mi familia. Gracias por aguantar mis neuras, sobre todo en estos últimos tiempos. Gracias por todo vuestro apoyo y vuestra paciencia. En especial agradecer a mis padres el esfuerzo que han hecho a lo largo de su vida para que yo pueda haber llegado hasta aquí.

A mis padres

RESUMEN

En el Banco de Germoplasma del Instituto de Conservación y Mejora de la Agrodiversidad Valenciana de la Universidad Politécnica de Valencia (COMAV) se mantienen actualmente más de 7000 entradas de especies hortícolas, incluyendo cultivares locales españoles y especies silvestres. Entre estas entradas se incluyen 3500 de la sección *Lycopersicon* y 144 de la sección *Basarthurum* del género *Solanum*. Para que el empleo de estos recursos fitogenéticos sea eficiente es necesario conocer su estructura genética y su taxonomía. El conocimiento de la distribución de la variabilidad existente permite establecer criterios de conservación más racionales y un uso más eficiente de estos recursos para la mejora de las especies cultivadas.

Los objetivos de la presente tesis doctoral se centran en el estudio de las secciones del género *Solanum* mencionadas previamente. Ambas secciones incluyen especies cultivadas: tomate y pepino dulce. La clasificación y filogenia de las especies pertenecientes a cada una de estas secciones ha sido objeto de controversia durante mucho tiempo y, a pesar del esfuerzo realizado hasta el momento, todavía no se ha alcanzado un consenso definitivo. Esto puede ser debido a una separación reciente de estas especies, que permanecen estrechamente relacionadas y con capacidad de cruzamiento, o a la falta de trabajos en los que se incluya una representación adecuada de toda la variabilidad presente en estas secciones.

Para analizar la taxonomía de la sección *Lycopersicon* se ha empleado una amplia representación de materiales de cada una de las especies abarcando todo su rango de distribución. Los análisis realizados identifican 12 especies distintas. Esta clasificación confirma algunas de las nuevas especies propuestas en la sección recientemente, pero no en todos los casos. La nueva especie *S. huaylasense* aparece claramente separada de *S. peruvianum*. También se confirma la separación de *S. arcanum*, aunque en este caso los datos moleculares sugieren que *S. arcanum* podría estar compuesta por dos especies crípticas difíciles de distinguir morfológicamente. Por otro lado, *S. corneliomulleri*, otra de las nuevas especies, no es distinguible genéticamente de *S. peruvianum* según nuestros análisis.

Basándose en las relaciones filogenéticas se pueden establecer 3 grupos en la sección *Lycopersicon*: Arcanum (*S. arcanum*, *S. chmielewskii*, *S. neorickii*, *S. huaylasense*), *Lycopersicon* (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* y *S. galapagense*) y Eriopersicon (*S. peruvianum* s.str., y *S. chilense*). En una posición basal respecto a estos grupos aparecen las especies *S. habrochaites* y *S. pennellii*.

En la sección *Basarthrum* los datos presentados respaldan la agrupación de los materiales en 13 especies. Entre estas se incluyen dos nuevas especies, previamente no descritas. Los materiales pertenecientes a estas especies son morfológicamente distintos y están claramente agrupados en los análisis moleculares. En cuanto a las relaciones filogenéticas entre ellas, se confirma que *S. canense* es la especie más alejada del resto y *S. suaveolens* aparece en una posición intermedia entre ésta (a la que se parece morfológica y reproductivamente) y la serie *Caripensia* (de la que está más cerca genéticamente). Dentro de la serie *Caripensia* se observan claramente diferenciadas las especies *S. basendopogon*, *S. trachycarpum* y *S. tabanoense*. Mientras que el complejo *Caripense* queda constituido por las especies *S. caripense*, *S. filiforme*, *S. fraxinifolium* y *S. heiseri*.

Además del análisis general de la taxonomía se analiza también en detalle la estructura genética de una especie de cada una de las secciones. En el caso de la sección *Lycopersicon*, se analiza la distribución de la variación de *S. pimpinellifolium* en todo el rango de distribución de esta especie; incluyendo una abundante representación de los materiales procedentes de Ecuador que no se había incluido en estudios previos. El análisis de esta especie tiene gran interés puesto que se trata de la especie silvestre más cercana al tomate cultivado y se emplea habitualmente como fuente de caracteres de interés para la mejora de este cultivo.

Solanum pimpinellifolium muestra una gran diferenciación entre los materiales procedentes de Perú y Ecuador. En el caso de las entradas ecuatorianas se observa una fuerte estructuración geográfica, mientras que en el caso de las peruanas la diferenciación entre las distintas zonas geográficas es pequeña. Estas diferencias genéticas podrían estar relacionadas con las diferentes condiciones climáticas y ecológicas que hemos observado entre ambos países. La costa peruana muestra unas características climáticas muy uniformes, mientras que en Ecuador se observa una gran variación climática y ecológica. Por otro lado, los datos de variabilidad y heterocigosidad señalan el norte de Perú como la posible región el origen de esta especie. De ser así, las poblaciones de Ecuador y del sur de Perú serían fruto de migraciones posteriores.

En el caso de la sección *Basarthrum*, se ha analizado la variabilidad genética presente en *S. muricatum*, la única especie cultivada de la sección. A diferencia de lo que ocurre en el caso de otras especie cultivadas, el pepino dulce mantiene una gran diversidad. La comparación entre la variabilidad presente en la especie cultivada y las silvestres permite concluir que *S. muricatum* presenta variabilidad genética proveniente de varias especies silvestres del complejo *Caripense*.

SUMMARY

The Germplasm Bank of The Institute of Conservation and Improvement of the Valencian Agrodiversity of the Polytechnical University of Valencia (COMAV) holds and maintains more than 7000 horticultural species accessions, including local landraces and wild species. 3500 of these accessions belong to the genus *Solanum* section *Lycopersicon* and 144 to the *Basarthurum* section.

To be able to use this resources efficiently it is necessary to know their genetic structure and taxonomy. Knowing their variability distribution rational criteria for the conservation of these resources can be established.

The objectives of the present doctoral thesis are focused in the study of the sections of the *Solanum* genus previously mentioned. Both sections include cultivated species: tomato and pepino. The classification and phylogeny studies of the species belonging to these sections have been controversial and, in spite of the effort carried out, a definitive consensus has not been reached yet. This could be explained by a recent separation of these species, that remain closely related and with hybridization capacity, or by the lack of previous works including an appropriate representation of all the variability present in these sections.

With the purpose of studying section *Lycopersicon* taxonomy a broad representation of each species has been employed covering their entire range of distribution. The analyses identified 12 different species. These classification confirms some of the new species proposed recently, but not all. The new species *S. huaylasense* appears clearly separated of *S. peruvianum*. Also the separation of *S. arcanum* is confirmed, although in this case the molecular data suggests that *S. arcanum* could be composed by two cryptic species morphologically difficult to distinguish. Contrary, *S. corneliomulleri*, another of the new species, could not be genetically differentiated from *S. peruvianum* according to our analyses.

Based on the phylogenetic relationships 3 groups could be established in the section *Lycopersicon*: Arcanum (*S. arcanum*, *S. chmielewskii*, *S. neorickii*, *S. huaylasense*), *Lycopersicon* (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae*, *S. galapagense*) and *Eriopersicon* (*S. peruvianum* s.str., *S. chilense*). The other two species from this genus, *S. habrochaites* and *S. pennellii* appear in a basal position.

In the section *Basarthurum* the classification of the materials in 13 species is supported by the presented data. Two new species not previously described have been found in this study. Materials

from these species are morphologically and molecularly different from the rest of species. Regarding to the phylogenetic relationships, it has been confirmed that *S. canense* is the most distant species and that *S. suaveolens* appears in an intermediate position between the latter and the series *Caripensia*. Within the series *Caripensia*, *S. basendopogon*, *S. trachycarpum* and *S. tabanoense* are observed as distinct species. The Caripense complex is formed by *S. caripense*, *S. filiforme*, *S. fraxinifolium* and *S. heiseri*.

Besides the general analysis of the taxonomy the genetic structure of a species from each section is also analyzed in detail. In the case of the section *Lycopersicon*, the variation distribution of *S. pimpinellifolium* is analyzed in the whole range of distribution of this species; including the materials from Ecuador previously not included in any study. The analysis of this species has great interest because it is the closest wild species to the cultivated tomato and it is often employed as source of interesting tomato agronomical characters.

A great differentiation is found in *Solanum pimpinellifolium* between the materials originated from Peru and Ecuador. In the case of the Ecuadorian accessions a strong differentiation is found, whereas in case of the Peruvian materials only a light differentiation between north and south is appreciated. These genetics differences could be related with the different climatic and ecological conditions present in both countries. The climate along the Peruvian coast is quite uniform, whereas in Ecuador a greater climatic and ecological variation between north and south is observed. On the other hand the information about the variability and heterocigosity suggest that the region of origin of this species could be located close to the northern Peru border. The populations from Ecuador and the south of Peru could be the result of posterior migrations.

In the case of the section *Basarthurum*, the genetic variability present in *S. muricatum*, the only one cultivated species of the section, has been analyzed. Unlike other cultivated species, pepino keeps a great diversity. Comparison between variability present in cultivated and wild species allows us to conclude that *S. muricatum* include genetic variability coming from several wild species of the complex Caripense.

RESUM

En el Banc de Germoplasma de l'Institut de Conservació i Millora de l'Agrodiversitat Valenciana de la Universitat Politècnica de València (COMAV) es mantenen actualment més de 7000 entrades d'espècies hortícoles, incloent-hi cultivars locals espanyols i espècies silvestres. Entre aquestes entrades s'inclouen 3500 de la secció *Lycopersicon* i 144 de la secció *Basarthurum* del gènere *Solanum*. Perquè l'ús d'estos recursos fitogenètics siga eficient és necessari conèixer la seua estructura genètica i la seua taxonomia. El coneixement de la distribució de la variabilitat existent permet establir criteris de conservació més racionals i un ús més eficient d'aquestos recursos per a la millora de les espècies cultivades.

Els objectius de la present tesi doctoral se centren en l'estudi de les seccions del gènere *Solanum* mencionades prèviament. Ambdues seccions inclouen espècies cultivades: tomaca i cogombre dolç. La classificació i filogènia de les espècies pertanyents a cadascuna d'aquestes seccions ha sigut objecte de controvèrsia durant molt de temps i, malgrat l'esforç realitzat fins al moment, encara no s'ha arribat a un consens definitiu. Açò pot ser degut a una separació recent d'aquestes espècies, que romanen estretament relacionades i amb capacitat de creuament, o a la falta de treballs en què s'incloua una representació adequada de tota la variabilitat present en estes seccions.

Per a analitzar la taxonomia de la secció *Lycopersicon* s'ha emprat una àmplia representació de materials de cadascuna de les espècies comprenent tot el seu rang de distribució. Les anàlisis realitzades identifiquen 12 espècies distintes. Aquesta classificació confirma algunes de les noves espècies proposades en la secció recentment, però no en tots els casos. La nova espècie *S. huaylasense* apareix clarament separada de *S. peruvianum*. També es confirma la separació de *S. arcanum*, encara que en aquest cas les dades moleculars suggerixen que *S. arcanum* podria estar composta per dos espècies críptiques difícils de distingir morfològicament. D'altra banda, *S. corneliomulleri*, una altra de les noves espècies, no és distingible genèticament de *S. peruvianum* segons les nostres anàlisis.

Basant-se en les relacions filogenètiques es poden establir 3 grups en la secció *Lycopersicon*: Arcanum (*S. arcanum*, *S. chmielewskii*, *S. neorickii*, *S. huaylasense*), Lycopersicon (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* i *S. galapagense*) i Eriopersicon (*S. peruvianum* s.str., i *S. chilense*). En una posició basal respecte a aquestos grups apareixen les espècies *S. habrochaites* i *S. pennellii*.

En la secció *Basarthrum* les dades presentades justifiquen l'agrupació dels materials en 13 espècies. Entre aquestes s'inclouen dos noves espècies, prèviament no descrites. Els materials pertanyents a aquestes espècies són morfològicament diversos i estan clarament agrupats en les anàlisis moleculars. Quant a les relacions filogenètiques entre elles, es confirma que *S. canense* és l'espècie més allunyada de la resta i *S. suaveolens* apareix en una posició intermèdia entre esta (a la que s'assembla morfològica i reproductivament) i la sèrie *Caripensia* (de la que està més prop genèticament). Dins la sèrie *Caripensia* s'observen clarament diferenciades les espècies *S. basendopogon*, *S. trachycarpum* i *S. tabanoense*. Mentres que el complex *Caripense* queda constituït per les espècies *S. caripense*, *S. filiforme*, *S. fraxinifolium* i *S. heiseri*.

A més de l'anàlisi general de la taxonomia, s'analitza també amb detall l'estructura genètica d'una espècie de cadascuna de les seccions. En el cas de la secció *Lycopersicon*, s'analitza la distribució de la variació de *S. pimpinellifolium* en tot el rang de distribució d'esta espècie; incloent una abundant representació dels materials procedents d'Equador que no s'havia inclòs en estudis previs. L'anàlisi d'esta espècie té gran interès, ja que es tracta de l'espècie silvestre més pròxima a la tomaca cultivada i s'empra habitualment com a font de caràcters d'interès per a la millora d'aquest cultiu.

Solanum pimpinellifolium mostra una gran diferenciació entre els materials procedents de Perú i Equador. En el cas de les entrades equatorianes s'observa una forta estructuració geogràfica, mentres que en el cas de les peruanes la diferenciació entre les distintes zones geogràfiques és xicoteta. Aquestes diferències genètiques podrien estar relacionades amb les diferents condicions climàtiques i ecològiques que hem observat entre ambdós països. La costa peruana mostra unes característiques climàtiques molt uniformes, mentres que a Equador s'observa una gran variació climàtica i ecològica. D'altra banda, les dades de variabilitat i heterocigositat assenyalen el nord de Perú com la possible regió l'origen de l'espècie. De ser així, les poblacions d'Equador i del sud de Perú serien fruit de migracions posteriors.

En el cas de la secció *Basarthrum*, s'ha analitzat la variabilitat genètica present en *S. muricatum*, l'única espècie cultivada de la secció. A diferència del que ocorre en el cas d'altres espècies cultivades, el cogombre dolç manté una gran diversitat. La comparació entre la variabilitat present en l'espècie cultivada i les silvestres permet concloure que *S. muricatum* presenta variabilitat genètica provinent de diverses espècies silvestres del complex *Caripense*.

| | |
|--|------------|
| I. INTRODUCCIÓN | 1 |
| 1. Implicaciones de los estudios de variabilidad en la mejora vegetal | 3 |
| 2. Avances en el estudio de la variabilidad | 5 |
| 3. Familia Solanaceae. Importancia económica y características | 8 |
| 4. Género <i>Solanum</i> | 9 |
| 4.1. Sección <i>Lycopersicon</i> | 15 |
| - Taxonomía y filogenia | 15 |
| - Características distintivas de la sección | 23 |
| - Distribución geográfica | 26 |
| - Cruzabilidad | 26 |
| - Utilización de las especies silvestres en la mejora | 28 |
| 4.2. Sección <i>Basarthrum</i> | 30 |
| - Taxonomía | 30 |
| - Origen del pepino dulce | 39 |
| - Utilización de estas especies en la mejora | 42 |
| 5.- El eclecticismo de las nuevas concepciones taxonómicas | 43 |
| II. OBJETIVOS | 49 |
| III. RESULTADOS | 53 |
| 1. Classification and phylogenetic relations in the <i>Solanum</i> section <i>Lycopersicon</i> based in AFLP and two nuclear gene sequences. <i>Genetic Resources and Crop Evolution</i> | 55 |
| 2. Genetic and bioclimatic variation in <i>Solanum pimpinellifolium</i> . <i>Genetic Resources and Crop Evolution</i> Supplementary material table | 73 88 |
| 3. The implications of AFLP data for the systematics of the wild species of <i>Solanum</i> section <i>Basarthrum</i> . <i>Systematic Botany</i> | 93 |
| 4. AFLP and DNA sequence variation in an Andean domesticate, pepino (<i>Solanum muricatum</i> , Solanaceae): implications for evolution and domestication. <i>American Journal of Botany</i> | 105 |
| IV. DISCUSIÓN GENERAL | 119 |
| 1. Clasificaciones taxonómicas Nuevas especies silvestres de la sección <i>Basarthrum</i> | 121 124 |
| 2. Relaciones filogenéticas Origen del pepino dulce | 125 126 |
| 3. Variabilidad intraespecífica | 127 |
| V. CONCLUSIONES | 131 |
| VI. BIBLIOGRAFÍA | 135 |

FIGURAS

1. Detalle de las flores de: A: *Lycianthes spp* (dehiscencia poral), B: *Solanum muricatum* (dehiscencia poral y anteras libres), C: *Solanum pimpinellifolium* (dehiscencia longitudinal y anteras soldadas) 10
2. Visión esquemática de las relaciones entre los clades principales dentro del género *Solanum* (Bohs 2005) 12
3. Representantes del “clado Potato” 13
4. Relaciones filogenéticas entre las especies pertenecientes al clado Potato 14
5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* 16
6. Esquema de las relaciones de cruzabilidad estudiadas por Rick (1979).
A: Entre las especies de la sección *Lycopersicon*, reconociendo los complejos *Esculentum* y *Peruvianum*. B: Entre las especies de las secciones *Lycopersicon*, *Lycopersicoides* y *Juglandifolium* 27
7. Variabilidad en frutos de pepino dulce (*S. muricatum*) 31
8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthurum* 33
9. Ejemplo visual en el que se representan: A: dos especies distintas, B: dos especies con algunos híbridos, C: dos subespecies, D: una única especie 47

TABLAS

1. Los 7 géneros más importantes de la familia Solanaceae por número de especies. (Fuente: PBI Solanum Project 2009) 9
2. Claves para distinguir las especies de la sección *Lycopersicon* (Peralta et al 2008) 24
3. Claves de identificación de las especies de la sección *Basarthurum* serie *Caripensia* (Anderson y Bernardello 1991) 32

I. INTRODUCCIÓN

1. IMPLICACIONES DE LOS ESTUDIOS DE VARIABILIDAD EN LA MEJORA

Los “Recursos Fitogenéticos para la Alimentación y la Agricultura” se definen, según la FAO (1996), como la diversidad del material genético contenido en las variedades tradicionales y en los cultivares modernos utilizados por los agricultores, así como en las especies silvestres relacionadas con los cultivos y otras silvestres que pueden utilizarse como alimento humano o para animales domésticos, para obtención de fibras y tejidos, protección, madera, energía, etc.

Los recursos fitogenéticos son recursos naturales que proporcionan la materia prima que, debidamente utilizada, originará mejores variedades de plantas. Esta materia prima se halla dispersa en cultivares locales y poblaciones naturales de plantas que han sido seleccionados durante miles de años por los agricultores y por la naturaleza mejorándose así características tales como su adaptación a las condiciones locales, productividad o resistencia a estreses bióticos o abióticos. Los recursos fitogenéticos son limitados y perecederos, y por ello se hace imprescindible su conservación, caracterización y utilización (Esquinas-Alcázar 1993).

La aparición de la agricultura hace unos 10000 años conllevó la domesticación de algunas especies de interés para el hombre. Los cultivos primitivos se fueron extendiendo desde sus lugares de origen, transportados primero durante las migraciones y después mediante las rutas comerciales, encontrándose con condiciones diversas de clima, suelo y vegetación (Martín 2001). Todo ello, provocó la aparición de una gran diversidad de tipos y un enorme número de variedades locales adaptadas a las necesidades humanas y al medio ambiente. Sin embargo, en los últimos años, debido a varios motivos como la exigencia de uniformidad en los mercados agrarios y la desaparición de las pequeñas unidades de autoconsumo, entre otros, se ha producido una sustitución de un enorme mosaico de variedades locales, heterogéneas y primitivas, por variedades comerciales constituidas por genotipos uniformes. Todo ello está provocando una pérdida enorme de la variabilidad acumulada durante años (Nuez et al. 1997).

El estrechamiento de la base genética de las especies cultivadas es un proceso irreversible, la variabilidad perdida ya no se puede recuperar. Esta erosión genética aumenta la vulnerabilidad agrícola de los cultivos y puede poner en peligro la seguridad alimentaria mundial (Martín 2001). Cabe citar, como ejemplo, el caso de la hambruna irlandesa a mediados del siglo XIX por un

ataque de *Phytophthora infestans* (Mont.) de Bary frente al que resultaron susceptibles las variedades cultivadas de patata, todas ellas procedentes de material homogéneo, y que provocó la muerte de casi un millón de personas y el desplazamiento de otros dos millones. Para solucionar el problema fue necesario localizar fuentes de resistencia y su introducción en variedades comerciales. Dichas fuentes de resistencia se encontraron entre los cultivares primitivos heterogéneos y las plantas silvestres de la región andina, centro de diversidad de la especie (Song et al. 2003). Otro caso más reciente es el ataque, en 1970, de *Helminthosporium maydis* Nisikado & Miyake que destruyó en Estados Unidos más del 50 por ciento de los maizales existentes en el sur del país, debido a que todos ellos procedían de semillas híbridas obtenidas mediante androesterilidad citoplasmática a partir de una sola variedad donante de citoplasma que era susceptible a esta enfermedad (NAS 1972).

Para prevenir estos problemas es imprescindible aumentar la base genética de los cultivos. Y para ello, una buena fuente de variabilidad genética son las especies silvestres relacionadas con las cultivadas. Estas especies se han utilizado para la mejora de cultivos durante siglos. De hecho, los cultivares modernos de la mayoría de cultivos contienen genes que provienen de alguna especie silvestre relacionada (Hajjar y Hodgkin 2007).

Sin embargo, pese a su gran interés, numerosas poblaciones naturales de muchas especies silvestres relacionadas con las especies cultivadas están sufriendo una importante erosión genética, provocada entre otros factores, por la elevada tasa de deforestación y sobreexplotación de áreas naturales. Es por ello que su conservación, tanto in-situ como ex-situ en bancos de germoplasma, se hace imprescindible. En 1993 se estableció el Sistema Mundial de la FAO para la Conservación y Uso Sostenible de los Recursos Fitogenéticos para la Alimentación y la Agricultura cuyos objetivos son “asegurar la conservación y promover la disponibilidad y utilización sostenible de los recursos genéticos, para las generaciones presentes y futuras” (FAO 1996).

A principios de los años 80, en el Departamento de Genética de la Universidad Politécnica de Valencia, comenzaron las actividades de recolección de recursos fitogenéticos de especies hortícolas que dieron como resultado el establecimiento del actual Banco de Germoplasma del Instituto de Conservación y Mejora de la Agrodiversidad Valenciana de la Universidad

Politécnica de Valencia (COMAV). En él se mantienen actualmente más de 7000 entradas de especies hortícolas, incluyendo cultivares locales españoles y especies silvestres, entre las que se incluyen unas 3500 de la sección *Lycopersicon* y 144 de la sección *Basarthurum*.

La utilización de la información de estudios biosistemáticos básicos puede ser de gran importancia para maximizar el potencial de los programas de mejora (Allard 1964). La identificación y clasificación de los recursos genéticos, así como la determinación de sus relaciones con las plantas domesticadas ayuda a los mejoradores a elegir los materiales más relevantes y útiles para sus programas de mejora (Prohens et al. 2003). Es decir, los estudios sobre la variabilidad genética entre y dentro de estas especies, la distribución de la diversidad genética en sus regiones de origen y la estructura de las poblaciones naturales nos proporcionan información relevante para la conservación de estos recursos genéticos, así como para la mejora de las especies cultivadas (Nuez et al. 2008).

2. AVANCES EN EL ESTUDIO DE LA VARIABILIDAD

Los estudios de variabilidad en plantas han sufrido una evolución espectacular en relativamente poco tiempo. Desde los análisis iniciales en los que únicamente se evaluaban caracteres morfológicos, se ha pasado a utilizar marcadores moleculares basados primero en un pequeño número de marcadores bioquímicos y más tarde en secuencias de ADN, hasta el empleo de soportes o matrices en las que se encuentran representados todos los genes de la especie en estudio.

La utilización de las diferencias fenotípicas como marcadores en el análisis genético aplicado a la Mejora Vegetal se remonta a las primeras décadas de la Genética como ciencia. 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 económico (Vázquez et al. 2000). En los estudios de diversidad, la elección de los caracteres a evaluar juega un papel importante, puesto que algunos caracteres pueden ser más variables entre individuos que otros y, por tanto, más discriminantes. Por ello, un buen punto de partida es la utilización de los descriptores editados por el International Plant Genetic Resources

Institute (IPGRI) como base para la elección de dichos caracteres. Entre otros, podemos encontrar los descriptores de tomate (IPGRI 1996) y de pepino dulce (IPGRI y COMAV 2004), cultivos pertenecientes a las secciones estudiadas en la presente tesis.

Sin embargo, el empleo de marcadores morfológicos presenta numerosas desventajas, como la escasez de caracteres útiles y discriminantes, su dependencia ambiental y su bajo nivel de polimorfismo, entre otras. Por todo ello, su empleo fue ampliamente desbancado por los marcadores moleculares.

Dentro de estos últimos, las isoenzimas constituyeron la primera generación de marcadores moleculares empleados. Descubiertas por Hunter y Markert en 1957, las isoenzimas son diferentes formas moleculares de la misma enzima que catalizan una reacción bioquímica específica y que pueden separarse por electroforesis debido a sus diferencias en carga y peso molecular (Vázquez et al. 2000). Suelen presentar una herencia codominante, no están habitualmente influidos por el medio y tienen una interacción nula o identificable con otros marcadores (García-Mas et al. 2000).

Sin embargo, los marcadores de ADN representan esencialmente un avance cuantitativo al añadir a las cualidades de los isoenzimas la de poder encontrarse en un número prácticamente ilimitado. Además de esto, pueden acceder a un rango de variabilidad mucho mayor, que abarca tanto las regiones codificantes como las que no lo son. 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 sus aplicaciones (García-Mas et al. 2000).

Los primeros marcadores basados en la variabilidad del ADN, los RFLPs (“Restriction Fragment Length Polymorphism”), se desarrollaron a principios de la década de los 80 (Botstein et al. 1980; Wyman y White 1980). Ejemplo del empleo de este tipo de marcadores en la familia de las solanáceas son, entre otros, los estudios de diversidad en patata (Gebhardt et al. 1989), tomate (Miller y Tanksley 1990) y pepino dulce (Anderson et al. 1996; Anderson y Jansen 1998). A continuación, les siguió una larga lista de marcadores que fueron apareciendo a medida que evolucionaban las técnicas de análisis molecular.

Tras la puesta a punto de la Reacción en Cadena de la Polimerasa (PCR) por Mullis en 1983, la evolución de las distintas técnicas ha sido imparable. Inicialmente, se desarrollaron técnicas para las que no eran necesarios conocimientos previos de la secuencia de los organismos en estudio, siendo los RAPDs (“Random Amplified Polymorphic DNA”) los primeros marcadores utilizados (Williams et al. 1990; Welsh y McClelland 1990), seguidos de los AFLPs (“Amplified Fragment Length Polymorphisms”) (Vos et al. 1995). Entre los primeros, cabe destacar su empleo en tomate por Williams y Clair (1993) comparando cultivares modernos y tradicionales, o el trabajo de Villand et al. (1998) comparando la variabilidad entre los centros primarios y secundarios de diversidad. En el caso de los AFLPs, Kardolus et al. (1998) los utilizaron por primera vez para estudiar las relaciones entre algunas especies del género *Solanum* (secciones *Petota* y *Lycopersicon*). Además, este tipo de marcadores han demostrado su utilidad en estudios de variabilidad, tanto interespecífica (en el caso de la patata, *S. tuberosum* L., en Spooner et al. 2005a) como intraespecífica (en el caso de *S. cheesmaniae* (L. Riley) Fosberg en Nuez et al. 2004).

Posteriormente, la disponibilidad de información de secuencias permitió desarrollar otros tipos de marcadores más específicos como los CAPSs (“Cleaved Amplified Polymorphic Sequence”) o los SCARs (“Sequence Characterized Amplified Regions”), aunque los más empleados han sido los microsatélites o SSRs (“Simple Sequence Repeats”). El nivel de polimorfismo de estos marcadores es el más alto cuando se compara con el resto de marcadores (Russell et al. 1997), pero su principal inconveniente reside en la dificultad para obtenerlos puesto que es necesario conocer las secuencias flanqueantes para poder diseñar los cebadores para su amplificación. Estos marcadores han sido ampliamente utilizados en estudios de diversidad en gran cantidad de especies cultivadas. Entre los estudios en solanáceas más recientes cabe destacar el de Spooner et al. (2007), en el que analizan 742 variedades tradicionales de patata.

Asimismo los SNPs (“Single-Nucleotide Polymorphisms”) son un tipo de marcador molecular cuyos alelos se diferencian en la variación de un solo nucleótido y tienen una frecuencia igual o superior al 1%. Los SNPs han sido muy estudiados por su asociación a enfermedades genéticas en humanos, sin embargo ha sido recientemente cuando se están empezando a utilizar en la selección asistida por marcadores (MAS) de los programas de mejora genética de cultivos al hacerse más rentable su empleo (Batley y Edwards 2007).

Los SNPs y las inserciones/delecciones (InDels) son los tipos de polimorfismos de ADN más abundantes y, teóricamente, se pueden encontrar dentro de cualquier secuencia genómica. Además el desarrollo de métodos de genotipado de alta productividad han hecho que estos marcadores sean muy prometedores. En plantas, se han realizado análisis sistemáticos en especies tales como *Arabidopsis thaliana* (L.) Heynh. (Clark et al. 2007), soja (Zhu et al. 2003) y vid (Lijavetzky et al. 2007).

Finalmente, el abaratamiento de los costes de secuenciación, así como la puesta a punto de métodos ultrarrápidos, hacen cada vez más viable la obtención de la secuencia de genomas completos. En un futuro próximo puede ser rentable su utilización en estudios de diversidad o filogenias y, por tanto, permitir el conocimiento completo de la distribución de la variación genética de las especies en estudio.

3. FAMILIA SOLANACEAE. IMPORTANCIA ECONÓMICA Y CARACTERÍSTICAS.

Entre las Angiospermas, la familia de las solanáceas es una de las más importantes para los seres humanos (Olmstead y Palmer 1992). En esta familia se incluyen especies alimenticias tan importantes como la patata (*Solanum tuberosum* L.), el tomate (*Solanum lycopersicum* L.), la berenjena (*Solanum melongena* L.) y el pimiento (*Capsicum* spp.). Asimismo, muchas plantas ornamentales muy populares pertenecen a las solanáceas, como algunas especies de los géneros *Petunia*, *Schizanthus* y *Salpiglossis*, entre otras. Algunas especies son mundialmente conocidas por sus usos medicinales o sus efectos psicotrópicos, como la mandrágora (*Mandragora autumnalis* Bertol.), el estramonio (*Datura stramonium* L.) o la belladona (*Atropa belladonna* L.). Además, varias de estas especies han jugado papeles importantes como plantas modelo para la investigación de numerosas cuestiones de genética, tales como el desarrollo de fruto (tomate y pimiento), desarrollo de tubérculos (patata), métodos defensivos de la planta (tomate y tabaco, *Nicotiana tabacum* L.) o la biosíntesis de antocianinas (petunia) (Wang et al. 2008).

La familia es cosmopolita, distribuyéndose por todo el globo con la excepción de la Antártida. La mayor diversidad de especies se halla en América del Sur y América Central. Las especies de esta familia viven en una gran diversidad de hábitats que incluyen desde selva lluviosa al

desierto, o altitudes que van desde el nivel del mar hasta las altas elevaciones de las montañas andinas (D'Arcy 1991).

Clasificada en la clase *Magnoliopsida*, subclase *Asteridae*, orden *Solanales*, su clasificación a niveles inferiores, tanto de subfamilia como de tribu o género, ha sufrido cambios con el tiempo. Tradicionalmente se han reconocido tres subfamilias: *Solanoideae*, *Cestroideae* y *Nolanoideae* (D'Arcy 1991; Hunziker 1979). Sin embargo, Hunziker (2001) considera a la subfamilia *Nolanoideae* como una familia distinta, *Nolanaceae*, y aumenta el número de subfamilias a seis: *Solanoideae*, *Cestroideae*, *Juanulloideae*, *Salpiglossoideae*, *Schizanthoideae* y *Anthocercidoideae*.

Se trata de una familia de plantas que comprende entre 3000-4000 especies agrupadas en unos 90 géneros aproximadamente (Knapp et al. 2004). No obstante, esa inmensa riqueza de especies no está uniformemente distribuida entre todos los géneros. Así, los 6 géneros más importantes de la familia concentran más del 60% de las especies (PBI Solanum Project 2009) (Tabla 1). De hecho, casi el 50% de la totalidad de especies de solanáceas pertenecen al género *Solanum*, de la tribu *Solaneae*, una de las 11 tribus comprendidas en la subfamilia *Solanoideae*.

Tabla 1. Los 6 géneros más importantes de la familia *Solanaceae* por número de especies. (Fuente: PBI Solanum Project 2009)

| Subfamilia | Tribu | Género | Número aprox. de especies |
|--------------------|--------------------|-------------------|---------------------------|
| <i>Solanoideae</i> | <i>Solaneae</i> | <i>Solanum</i> | 1000-2000 |
| | <i>Solaneae</i> | <i>Lycianthes</i> | 250 |
| | <i>Solaneae</i> | <i>Physalis</i> | 75 |
| | <i>Lycieae</i> | <i>Lycium</i> | 75 |
| <i>Cestroideae</i> | <i>Cestreae</i> | <i>Cestrum</i> | 250 |
| | <i>Nicotianeae</i> | <i>Nicotiana</i> | 70 |

4. GÉNERO *Solanum*

El género *Solanum*, de importancia económica mundial, incluye especies cultivadas tan importantes como la patata, el tomate y la berenjena. Asimismo, incluye un importante número de especies cultivadas a menor escala, comestibles por sus frutos, tubérculos u hojas, y como

fuelle de alcaloides de utilidad medicinal. Algunos ejemplos de estos cultivos menores son el tomate de árbol (*Solanum betaceum* Cav.), la naranjilla (*Solanum quitoense* Lam.) y el pepino dulce (*Solanum muricatum* Aiton) originarios del “Nuevo Mundo”, y la berenjena esarlata y africana (*Solanum aethiopicum* L. y *Solanum macrocarpon* L., respectivamente) del “Viejo Mundo”.

Solanum, como se ha mencionado, pertenece a la subfamilia *Solanoideae* que se caracteriza por poseer una semilla discoidal comprimida y con embrión enrollado, de diámetro más o menos uniforme. Todos los miembros de esta subfamilia tienen el mismo número cromosómico básico ($x=12$) (Hunziker 1979).

Dentro de la subfamilia, *Solanum* se ha clasificado tradicionalmente en la grande y compleja tribu *Solaneae*. *Solanum* se distingue de la mayoría de géneros de esta tribu por sus anteras con dehiscencia poral, un carácter presente en casi todas las especies de este género y compartido únicamente por el cercano género *Lycianthes* (Bohs 2005), el cual se diferencia de *Solanum* en la estructura del cáliz (D'Arcy 1986) (Figura 1). La reciente inclusión de la sección *Lycopersicon* dentro del género, con dehiscencia longitudinal, obliga a una redefinición de *Solanum*.

Figura 1. Detalle de las flores de: A: *Lycianthes* spp. (dehiscencia poral), B: *Solanum muricatum* (dehiscencia poral y anteras libres), C: *Solanum pimpinellifolium* (dehiscencia longitudinal y anteras soldadas).



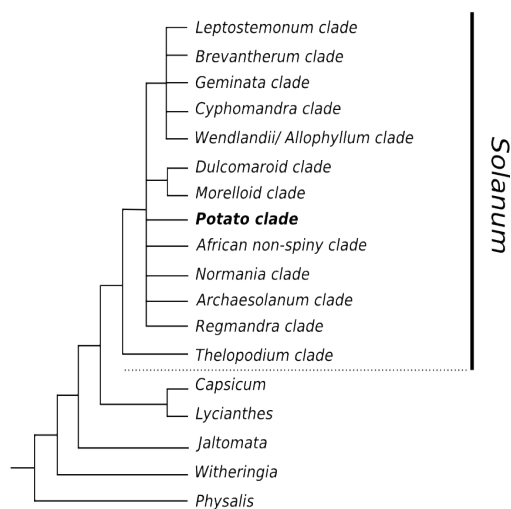
Sin embargo, las relaciones evolutivas entre los aproximadamente 34 géneros pertenecientes a esta tribu siguen sin estar completamente claras (Bohs y Olmstead 1997). Los primeros trabajos trataron de dividir el género en 2 grandes grupos basándose en la presencia o no de espinas (Linneo 1753), añadiendo a ello la morfología de las anteras (Dunal 1852), o teniendo en cuenta además el tipo de pilosidad (Seithe 1962). Posteriormente, Danert (1970) integró caracteres de patrones de ramificación y morfología de la raíz con los sistemas previos.

Estos trabajos sentaron las bases para el esquema de clasificación de D'Arcy (1972), el más ampliamente utilizado en nuestro días. En él se reconocen 7 subgéneros en *Solanum*, que varían en su tamaño desde el monotípico *Lyciosolanum* hasta los subgéneros *Leptostemonum*, *Potatoe*, y *Solanum*, que incluyen cientos de especies cada uno.

Posteriormente Nee (1999), Child y Lester (2001) y Hunziker (2001) realizaron clasificaciones de *Solanum* a nivel infragenérico basándose en datos morfológicos e ideas intuitivas sobre las relaciones de parentesco. Sin embargo, únicamente Nee (1999), centrándose en los taxones del Nuevo Mundo, aportó una lista explícita de las especies incluidas en cada subgénero, sección o serie.

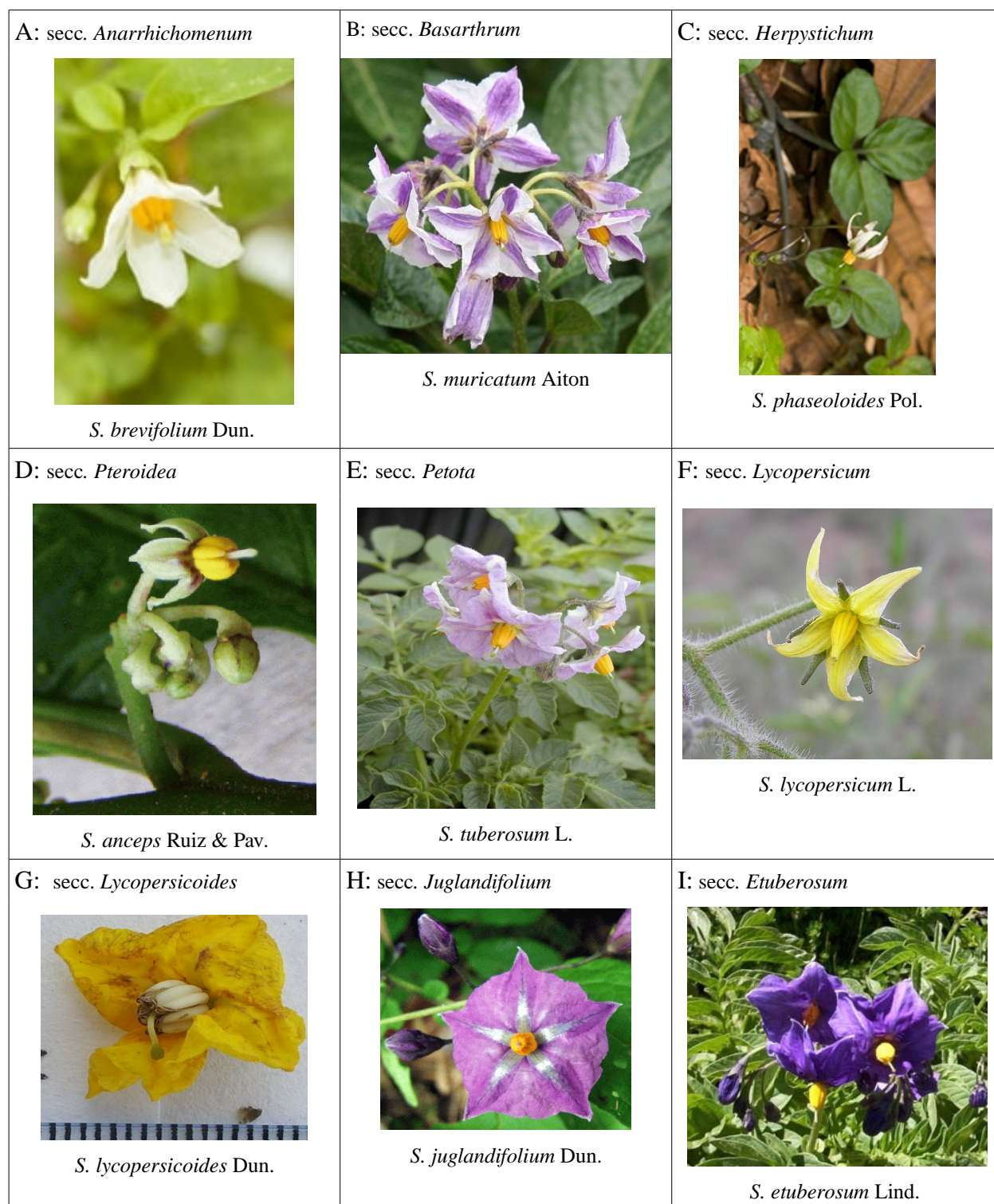
La monofilia de algunos grupos reconocidos por autores previos fue estudiada por Bohs (2005) realizando un análisis cladístico empleando datos moleculares de la secuencia del gen de cloroplasto *ndhF*. Posteriormente, este trabajo fue ampliado mediante la incorporación de secuencias de la región *trnT-F* de cloroplasto y del gen nuclear *waxy* (Weese y Bohs 2007). Según estos autores, la mayoría de los subgéneros tradicionalmente reconocidos no son monofiléticos, y proponen una clasificación alternativa del género *Solanum* en 13 linajes principales, sugiriendo nombres informales para los clados (Figura 2). Asimismo, sugieren esta clasificación como referencia a la hora de seleccionar los materiales para futuros estudios filogenéticos e identifican grupos naturales sobre los que focalizar nuevos trabajos de revisión.

Figura 2. Visión esquemática de las relaciones entre los clados principales dentro del género *Solanum* (Bohs 2005). Las especies estudiadas en esta tesis pertenecen al clado Potato.



Entre estos linajes descritos por Weese y Bohs (2007) cabe destacar el clado Potato por su gran importancia económica, al incluir la patata, el tomate y el pepino dulce. Este clado incluye la mayoría de las secciones del subgénero *Potatoe* de D'Arcy (1972), así como algunas del subgénero *Bassovia*. Consiste en 9 secciones (Figura 3): *Anarrhichomenum*, *Basarthrum* (que incluye el pepino dulce), *Herpystichum*, *Pterioidea*, *Petota* (que incluye la patata), *Lycopersicon* (que incluye el tomate), *Lycopersicoides*, *Juglandifolium* y *Etuberosum*.

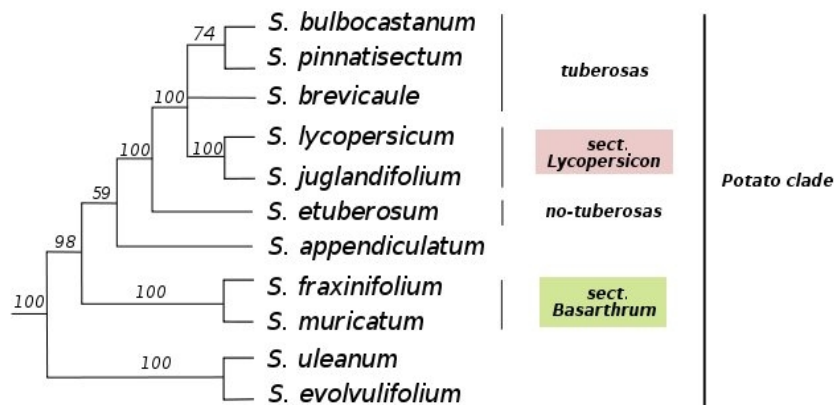
Figura 3. Representantes del “clado Potato”.



Este clado, principalmente suramericano, lo forman un grupo importante de plantas herbáceas a ligeramente leñosas, a menudo plantas trepadoras, la mayoría con hojas compuestas y algunas con rizomas o tubérculos (Bohs 2005). En la figura 4 se muestra un fragmento del dendrograma

mostrado por Weese y Bohs (2007) en el que se representan las relaciones filogenéticas entre las especies representativas de cada sección dentro de este clado. Las especies tuberosas, representadas por *S. bulbocastanum* Dun., *S. pinnatisectum* Dun., y *S. brevicaule* Bitter están estrechamente relacionadas con el tomate (*S. lycopersicum*) y la especie silvestre relacionada con éste, *S. juglandifolium* Dun.. Asimismo, se observa una estrecha relación entre *Solanum* secc. *Etuberosum*, representado por *S. etuberosum* Lind., y las especies tuberosas. Los autores confirman la inclusión de las especies de las secciones *Anarrhichomenum* y *Basarthrum* (representadas por *S. appendiculatum* Dun., *S. fraxinifolium* Dun., y *S. muricatum*) dentro del subgénero *Potatoe* (D’Arcy 1972, 1991; Child y Lester 2001). Y por último, *S. uleanum* Bitter (secc. *Pteroidea*; Knapp y Helgason 1997) aparece como hermana de *S. evolvulifolium* Greenm. (secc. *Herpystichum*; Nee 1999). Ambos taxa parecen hermanos del resto de especies del clado, pese a que estudios previos no sugerían una estrecha relación entre la sección *Pteroidea* y el clado Potato (Knapp y Helgason 1997).

Figura 4. Relaciones filogenéticas entre las especies pertenecientes al clado Potato. Modificado de Weese y Bohs (2007), donde analizan la filogenia completa del género *Solanum* empleando datos de secuencias de tres genes. Aparecen resaltadas las secciones estudiadas en la presente tesis.



Por último, cabe destacar que pese a la cantidad de información disponible y los esfuerzos realizados a lo largo de todos estos años, como los mismos autores Weese y Bohs (2007) aseguran, queda trabajo por hacer puesto que el conocimiento sobre la evolución dentro del género está lejos de estar completo.

4.1. Sección *Lycopersicon*





Taxonomía y filogenia

Al igual que en los niveles superiores, la taxonomía de esta sección ha sufrido numerosos cambios a lo largo del tiempo y todavía no se ha alcanzado un consenso definitivo. En la actualidad, se ha aceptado la inclusión de estas especies en el género *Solanum* sección *Lycopersicon*, tal y como sugirió Child (1990) a partir de datos morfológicos y como ha sido corroborado con datos moleculares (Spooner et al. 1993; Bohs y Olmstead 1997; Olmstead y Palmer 1997; Peralta y Spooner 2001). Una explicación detallada de todos estos cambios se puede ver en la introducción del primer artículo de la presente tesis.

La clasificación actual considera la existencia de 13 especies de tomates silvestres, incluyendo el tomate cultivado (Peralta et al. 2008) (Figura 5), divididas en 4 grupos:

- grupo *Lycopersicon*: incluye las 4 especies de frutos rojos o naranjas: *S. lycopersicum*, *S. pimpinellifolium* L., *S. cheesmaniae* (L. Riley) y Fosberg y *S. galapagense* S.C. Darwin & Peralta,
- grupo *Eriopersicon*: que incluye a las 5 especies *S. peruvianum* L., *S. corneliomulleri* J.F. Macbr., *S. huaylasense* Peralta, *S. habrochaites* S. Knapp & D.M. Spooner y *S. chilense* (Dunal) Reiche,
- grupo *Arcanum*: con las 3 especies: *S. arcanum* Peralta, *S. chmielewskii* (C.M. Rick et al.) D.M. Spooner et al. y *S. neorickii* D.M. Spooner et al.,
- grupo *Neolycopersicon*: que incluye únicamente a la especie *S. pennellii* Correll.

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon*.

| | | | |
|---|---|---|--|
| <p>Nombre <i>S. lycopersicum</i> L.</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Frutos típicamente rojos, de diámetro superior a 2 cm. - Origen del tomate cultivado. | |
| <p>Nombre antiguo <i>L. esculentum</i> Miller</p> | | | |
| <p>Sistema reproductivo Autocompatible, autógama o alógama facultativa.</p> | | <p>Distribución geográfica</p>  | |
| <p>Ecosistema Aparece en regiones tropicales y su zona nativa está en Ecuador-Perú; mala hierba en cultivos recientes.</p> | | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  | |









| | | | |
|--|--|---|--|
| <p>Nombre <i>S. pimpinellifolium</i> L.</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto típicamente rojo, de diámetro menor de 2 cm. - Margen de la hoja más entero que <i>S. lycopersicum</i>. | |
| <p>Nombre antiguo <i>L. pimpinellifolium</i> (L.) Miller</p> | | | |
| <p>Sistema reproductivo Autocompatible; autógama o alógama facultativa.</p> | | <p>Distribución geográfica</p>  | |
| <p>Ecosistema Aparece como mala hierba en campos de cultivo en valles costeros y en aluviones interandinos secos.</p> | | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  | |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | | | |
|---|---|---|---|
| <p>Nombre <i>S. cheesmaniae</i> (L.Riley) Fosberg</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto de color amarillo o naranja. - Hoja dividida, pero menos que <i>S. galapagense</i>. | |
| <p>Nombre antiguo <i>L. cheesmaniae</i> L. Riley</p> | | | |
| <p>Sistema reproductivo Autocompatible; exclusivamente autógama.</p> | | | |
| <p>Ecosistema Endémica de las Islas Galápagos. Desde el nivel del mar hasta 500 m. de altitud.</p> | | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  | <p>Distribución geográfica</p>  |









| | | | |
|---|--|--|---|
| <p>Nombre <i>S. galapagense</i> S.C. Darwin & Peralta</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto de color amarillo o naranja. - Hoja muy dividida, con 3-4 lóbulos. | |
| <p>Nombre antiguo <i>L. cheesmaniae</i> L. Riley</p> | | | |
| <p>Sistema reproductivo Autocompatible; exclusivamente autógama.</p> | | | |
| <p>Ecosistema Endémica de las Islas Galápagos. Aparece en las rocas volcánicas a 1 m de la marca de la marea, y ocasionalmente tierra adentro. Entre 0-50 m.</p> | | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  | <p>Distribución geográfica</p>  |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | | |
|--|---|--|
| <p>Nombre <i>S. peruvianum</i> L.</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto verde a blanquecino, a veces con reflejos morados. - La inflorescencia no sobresale del extremo apical del tallo. - Brácteas en los pedicelos florales. - Cono de anteras curvado. - Pubescencia más corta que <i>S. corneliomulleri</i>. |
| <p>Nombre antiguo <i>L. peruvianum</i> (L.) Miller</p> | | |
| <p>Sistema reproductivo Típicamente autoincompatible, alógama.</p> | | |
| <p>Ecosistema En las "lomas" peruanas y ocasionalmente en desiertos costeros, desde el nivel del mar hasta los 2900m.</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | | |









| | | |
|--|--|--|
| <p>Nombre <i>S. corneliomulleri</i> J. F. Macbr.</p> | | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto verde con estrías verde oscuras o púrpuras. - La inflorescencia no sobresale del extremo apical del tallo. - Brácteas en los pedicelos florales. - Cono de anteras curvado. - Pubescencia más larga que <i>S. peruvianum</i>. |
| <p>Nombre antiguo <i>L. peruvianum</i> (L.) Miller; <i>L. glandulosum</i> C.F. Mull.</p> | | |
| <p>Sistema reproductivo Típicamente autoincompatible, alógama.</p> | | |
| <p>Ecosistema Elevaciones medias a altas de la laderas oeste de los Andes (1000-3000), en ocasiones en zonas más bajas tras corrimientos de tierra (huaycos).</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | | |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | | | |
|---|---|---|--|
| <p>Nombre <i>S. huaylasense</i> Peralta</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto típicamente verde con bandas verde oscuras. - Tallo y hojas escasamente pubescentes y de color verde vivo, la distinguen de <i>S. chilense</i>. | | |
| <p>Nombre antiguo <i>L. peruvianum</i> (L.) Miller</p> | | | |
| <p>Sistema reproductivo Típicamente autoincompatible, alógama.</p> | | | |
| <p>Ecosistema Localizada en el Callejón de Huaylas, y en el drenaje próximo del Río Fortaleza (Ancash, Perú); entre 1700 y 3000 m.</p> | | | |
| <p>Fruto</p>  <p>(fuente: Spooner et al 2008)</p> | <p>Flor</p>  <p>(fuente: Spooner et al 2008)</p> | <p>Hoja</p>  <p>(fuente: Spooner et al 2008)</p> | <p>Distribución geográfica</p>  |









| | | | |
|--|--|--|---|
| <p>Nombre <i>S. habrochaites</i> S. Knapp & D. M Spooner</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Elevada pilosidad. - Inflorescencia ramificada, 3 hojas por simpodio. - Ramificaciones que pueden alcanzar los 6 m de longitud. | | |
| <p>Nombre antiguo <i>L. hirsutum</i> Dunal</p> | | | |
| <p>Sistema reproductivo Típicamente autoincompatible, con poblaciones autocompatibles en los márgenes de la distribución.</p> | | | |
| <p>Ecosistema Se encuentra en variedad de tipos de bosque, desde 400 a 3600 m.</p> | | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  | <p>Distribución geográfica</p>  |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | | |
|--|---|--|
| <p>Nombre <i>S. chilense</i> (Dunal) Reiche</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto verdoso a blanquecino-verdoso, con bandas púrpuras. - Los pedúnculos de inflorescencia más largos y su porte erecto la distinguen de <i>S. peruvianum</i>. - Posee un buen sistema radicular. - Estigmas muy exertos. | |
| <p>Nombre antiguo <i>L. chilense</i> Dunal</p> | | |
| <p>Sistema reproductivo Autoincompatible, alógama.</p> | | |
| <p>Ecosistema Propia de hábitats secos: aparece en quebradas o bancos de rocas secas. Límite sur del rango de distribución de la sección <i>Lycopersicon</i>.</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| | | <p>Distribución geográfica</p>  |









| | | |
|--|---|---|
| <p>Nombre <i>S. arcanum</i> Peralta</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Inflorescencia no ramificada. - Cono de anteras recto. - Fruto verde con bandas verde oscuras. | |
| <p>Nombre antiguo <i>L. peruvianum</i> (L.) Miller</p> | | |
| <p>Sistema reproductivo Típica autoincompatible, alógama; aunque rara vez autocompatible, autógama o alógama facultativa.</p> | | |
| <p>Ecosistema Valles andinos costeros y del interior en el norte de Perú; en laderas rocosas, valles secos y lomas; 100-2500 m.</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| | | <p>Distribución geográfica</p>  |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | | |
|---|--|--|
| <p>Nombre <i>S. chmielewskii</i> (C. M. Rick et al.) D. M. Spooner et al.</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto típicamente verde con bandas verde oscuras. - Presenta estigmas exertos. | |
| <p>Nombre antiguo <i>L. chmielewskii</i> C. M. Rick et al.</p> | | |
| <p>Sistema reproductivo Autocompatible, alógama facultativa.</p> | | |
| <p>Ecosistema En valles interandinos; en medios rocosos y bien drenados; entre los 1500-3000 m.</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | | |









| | | |
|---|--|---|
| <p>Nombre <i>S. neorickii</i> D. M. Spooner et al.</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Fruto típicamente verde con bandas verde oscuras. - Presenta estigma inserto. - Flores más pequeñas que <i>S. chmielewskii</i>. | |
| <p>Nombre antiguo <i>L. parviflorum</i> C. M. Rick et al.</p> | | |
| <p>Sistema reproductivo Autocompatible, altamente autógena.</p> | | |
| <p>Ecosistema En valles interandinos, medios rocosos y bien drenados entre los 1500-2500 m de altitud</p> | | |
| <p>Fruto</p>  | <p>Flor</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | | |

Figura 5. Fichas informativas con los datos más relevantes de cada una de las especies silvestres pertenecientes a la sección *Lycopersicon* (cont.).

| | |
|--|---|
| <p>Nombre <i>S. pennellii</i> Correll</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Única con anteras no soldadas y dehiscencia poral. - Foliolos redondeados. - Fruto verde. - Regulación de los estomas única en el género que le confiere buena adaptación a ambientes xerofíticos, a pesar de no tener un buen sistema radicular. |
| <p>Nombre antiguo <i>L. pennellii</i> (Correll) D'Arcy</p> | |
| <p>Sistema reproductivo Autoincompatible, aunque algunas poblaciones son autocompatibles.</p> | |
| <p>Ecosistema Propia de laderas rocosas secas y áreas arenosas, desde el nivel del mar hasta los 3000 m de altitud.</p> | |
| <p>Fruto</p>  | <p>Flor</p>  |
| <p>Hoja</p>  | <p>Distribución geográfica</p>  |

Sin embargo, a pesar del esfuerzo realizado hasta la fecha por diversos investigadores (Palmer y Zamir 1982; McClean y Hanson 1986; Miller y Tanksley 1990; Alvarez et al. 2001; Marshall et al. 2001; Peralta y Spooner 2001, 2005; Spooner et al. 2005b; Peralta et al. 2008), las relaciones entre algunas de estas especies no quedan suficientemente claras, bien porque la significación estadística no es alta o por la existencia de resultados contradictorios entre los diferentes autores. Un ejemplo de este tipo se puede comprobar al observar la posición en los árboles filogenéticos de las poblaciones más septentrionales de *S. peruvianum* s.l., que en algunos casos aparecen más cercanas a *S. lycopersicum* (Peralta y Spooner 2001), en otros a las poblaciones meridionales de *S. peruvianum* s.l. (Marshall et al. 2001), y aún en otros a *S. chmielewskii* (Alvarez et al. 2001; Peralta et al. 2008).

Cabe destacar que en la mayoría de estos trabajos se ha utilizado un número muy reducido de entradas de cada una de las especies, por lo que es bastante probable que no se haya incluido una buena representación de la variabilidad existente. Esto podría llevarnos a confusión al estudiar las

relaciones entre las especies puesto que, al no tener en cuenta toda la variabilidad que contienen, estaríamos introduciendo un sesgo en dicha comparación. Además éste es un aspecto importante a tener en cuenta, puesto que permite observar la existencia de estructuración en algunas especies, como en el caso de *S. peruvianum* que ha sido dividida recientemente en otras cuatro especies (Peralta et al. 2005).

Por último, cabe mencionar que la mayoría de los trabajos citados son anteriores a la división de *S. peruvianum*, por lo que en ellos no se ha podido analizar este aspecto en profundidad.

Características distintivas de la sección

Las especies de esta sección se caracterizan por presentar los extremos de las anteras unidos formando un cuello, constituyendo el conjunto de las anteras un cono estaminal característico en forma de botella (Figura 1). De manera asociada, las anteras se abren longitudinalmente, a diferencia de la dehiscencia terminal o poral característica del resto de especies del género. El polen se libera en el interior del cono estaminal y emerge a través del cuello del cono. Durante mucho tiempo, el reconocimiento del valor taxonómico de este conjunto de caracteres permitió diferenciar los géneros *Lycopersicon* y *Solanum* (Miller 1754). Cabe destacar el caso de *S. pennellii*, que pese a carecer de dehiscencia poral, lo cual la clasificaría dentro del género *Solanum*, su facilidad de cruzamiento con *S. lycopersicum* así como la obtención de descendencia fértil, entre otros muchos caracteres, la colocan más cerca de las especies silvestres de tomate que de otras *Solanum* cercanas (Rick 1979).

Las anteras normalmente se desarrollan con la misma longitud y por tanto son rectas, sin embargo, en algunas especies, como *S. corneliomulleri*, *S. huaylasense*, *S. peruvianum* y *S. pennellii*, las anteras están ligeramente curvadas a causa de un crecimiento desigual del tercio superior que se desarrolla más que la parte inferior (Peralta et al. 2008) (Figura 5).

Los frutos de todas las especies de la sección son bayas globosas, biloculares, a excepción de algunas poblaciones silvestres y formas cultivadas de *S. lycopersicum*, que son multiloculares y de formas variadas (van der Knaap y Tanksley 2003). El color y la pubescencia de los frutos son caracteres importantes para diferenciar las especies y los grupos de especies. En la tabla 2 se

muestran las claves que permiten distinguir morfológicamente a las especies de esta sección (Peralta et al. 2008).

Tabla 2. Claves de las especies de la sección *Lycopersicon* (Peralta et al. 2008).

1a. Foliolos primarios redondeados; pedicelos florales articulados en la base; corola asimétrica; anteras separadas, sin estrechamiento estéril en el extremo apical, dehiscencia poral. >> **Grupo**

Neolycopersicon: *S. pennellii*

1b. Foliolos primarios elípticos o lanceolados; pedicelos florales articulados por la zona media; corola simétrica; anteras soldadas en forma de botella, con estrechamiento estéril en el extremo apical, dehiscencia longitudinal.

2a. Sin pseudoestípulas; inflorescencia sin brácteas y generalmente no ramificada (raramente bifurcada); frutos con pigmentos carotenoides >> **Grupo Lycopersicon**

3a. Interior del fruto maduro rojo; continentales de América del Sur (o cultivada)

4a. Plantas generalmente poco pubescentes (< 2 mm); margen foliar generalmente entero o ligeramente dentado o crenado; inflorescencia generalmente con más de 12 flores por ramificación; corola fuertemente estrellada; lóbulos divididos hasta casi la base; diámetro del fruto < 20 mm >> ***S. pimpinellifolium***

4b. Plantas generalmente poco pubescentes (>3 mm); margen foliar generalmente dentado, especialmente en la base, en ocasiones lobulado y con foliolos secundarios; inflorescencia generalmente con menos de 12 flores por ramificación; corola superficialmente estrellada; lóbulos divididos entre 1/3 y 1/2 de la base; diámetro del fruto > 20 mm >> ***S. lycopersicum***

3b. Interior del fruto maduro amarillo o naranja; endémicas de las Islas Galápagos

5a. Sin lóbulos foliares terciarios; sépalos más cortos que la longitud del fruto maduro >> ***S. cheesmaniae***

5b. Con lóbulos foliares terciarios y/o cuaternarios; sépalos más largos que la longitud del fruto maduro >> ***S. galapagense***

2b. Con pseudoestípulas; inflorescencia bracteada y con 1-2 ramificaciones (raramente 3); frutos verdes o con pigmentos antociánicos y con rayas verde oscura-púrpura.

6a. Inflorescencia generalmente no ramificada, raramente bifurcada; tubo de anteras recto; estigma inserto o ligeramente exerto (<1 mm) >> **Grupo Arcanum**

Tabla 2. Claves de las especies de la sección *Lycopersicon* (Peralta et al. 2008) (cont.)

7a. Corola >1,6-2 cm de diámetro; estigma ligeramente exerto.

8a. Peciolos, hojas e inflorescencias verdes, glabros a variablemente pubescente, con mezcla de tipos de tricomas; normalmente más de 7 flores/inflorescencia (5-20) >> *S. arcanum*

8b. Peciolos, hojas e inflorescencias verde pálido; pubescencia aterciopelada; normalmente menos de 7 flores/inflorescencia >> *S. chmielewskii*

7b. Corola <1-1,2 cm de diámetro; estigma inserto >> *S. neorickii*

6b. Inflorescencia con 2 o más ramificaciones; tubo de anteras recto o curvado; estigma exerto (>1 mm) >> **Grupo Eriopersicon**

9a. Tres hojas por simpodio; ramificaciones largas y rastreras, crecimiento desordenado, que llegan a los 6 m de extensión >> *S. habrochaites*

9b. Dos hojas por simpodio (raramente 2-3 en *S. chilense*); plantas erectas que posteriormente decaen; con menos de 3 m de longitud.

10a. Pedúnculo de la inflorescencia > 12 mm; la inflorescencia sobresale por encima del extremo apical del tallo; brácteas ausentes o, si están, aparecen en los primeros 2-4 pedicelos florales basales de la inflorescencia; cono de anteras recto o curvado.

11a. Tallo y hojas densamente pubescentes, de color blanquecino o típicamente verde-grisáceo; cono de anteras recto; costa sur de Perú y norte de Chile >> *S. chilense*

11b. Tallo y hojas escasamente pubescentes, de color verde vivo; cono de anteras recto (excepto las procedentes de la zona del Río Fortaleza); Ancash, Perú >> *S. huaylasense*

10b. Pedúnculo de la inflorescencia < 12 mm; la inflorescencia no sobresale por encima del extremo apical del tallo; brácteas en la mayoría de los pedicelos florales; cono de anteras siempre curvado.

12a. Pubescencia corta, aspecto grisáceo-verde pálido; margen de la hoja entero o ligeramente dentado o crenado >> *S. peruvianum*

12b. Pubescencia larga, aspecto verde; margen de la hoja dentado o crenado, a veces fuertemente dividido >> *S. corneliomulleri*

Distribución geográfica

Las especies silvestres de esta sección son nativas de la franja oeste de América del Sur (Figura 5). Se distribuyen desde el norte de Ecuador, a lo largo de Perú, hasta el norte de Chile, incluyendo las islas Galápagos (Warnock 1991). La única silvestre con una distribución mundial es *S. lycopersicum* var. *cerasiforme* que está más ampliamente distribuida y es probable que se haya introducido recientemente en México, Colombia, Bolivia y otros países de Sur América (Rick y Holle 1990), y desde allí al resto del mundo.

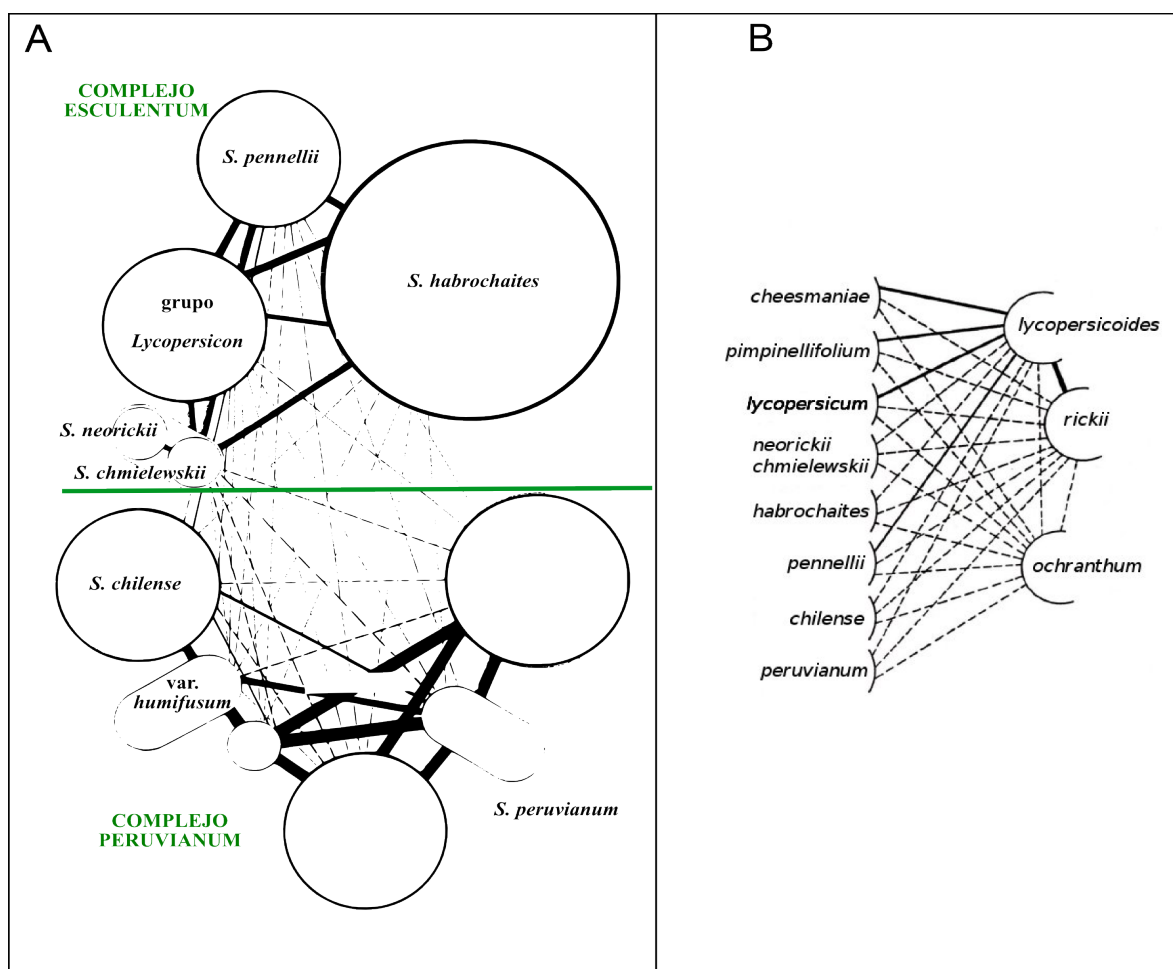
Estas especies silvestres crecen en una amplia variedad de hábitats, desde el nivel del mar hasta más de 3300 m de altitud (Rick 1973; Taylor 1986). Estos hábitats incluyen tanto las regiones bajas y áridas de la costa del Pacífico como los valles adyacentes y las regiones húmedas y altas en los Andes. Numerosos valles, formados por los drenajes de los ríos en el Pacífico, caracterizan la zona oeste de los Andes. Las poblaciones de tomates silvestres crecen a diferentes alturas en estos valles, estrechos y geográficamente aislados, y están adaptadas a condiciones particulares, tanto microclimáticas como edáficas. La geografía andina, la diversidad de hábitats ecológicos, y las diferencias climáticas contribuyen a la diversidad de estas especies (Warnock 1988).

Cruzabilidad

Un aspecto importante a tener en cuenta a la hora de la utilización directa de los recursos genéticos disponibles es la posibilidad de realizar cruzamientos entre la especie cultivada y las especies silvestres, fuente de caracteres de interés. Rick (1979) estudió las relaciones de cruzabilidad entre las especies silvestres de tomate, estableciendo dos grupos en función de la capacidad o no de cruzamiento con la especie cultivada: los complejos Esculentum (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae*, *S. chmielewskii*, *S. neorickii*, *S. habrochaites*, y *S. pennellii*) y Peruvianum (*S. peruvianum* y *S. chilense*), respectivamente (Figura 6a). Rick (1979) señala que este último grupo se podría utilizar como fuente de caracteres agronómicos si se realizasen cruzamientos mediante algunas “accesiones puente”, o mediante rescate de embriones, siempre que actúen como genitor masculino. Por otra parte, Rick también estudió la capacidad de cruzamiento de estas especies con otras más alejadas, como *S. lycopersicoides* Dun. y *S. rickii* Corr., de la sección *Lycopersicoides*, y *S. ochranthum* Dun., de la sección

Juglandifolium (Figura 6b). Observó que *S. lycopersoides* era compatible con *S. rickii*, y podía cruzar unilateralmente como genitor masculino con *S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* y *S. pennellii*. Por otro lado, *S. rickii* únicamente cruzaba con *S. lycopersoides*, y *S. ochranthum* falló en todos los cruces. Por ello, concluyó que la separación entre las especies pertenecientes a la actual sección *Lycopersicon* y el resto de *Solanum* era suficiente como para justificar su inclusión en otro género distinto (género *Lycopersicon*).

Figura 6. Esquema de las relaciones de cruzabilidad estudiadas por Rick (1979). A: Entre las especies de la sección *Lycopersicon*, reconociendo los complejos Esculentum y Peruvianum. B: Entre las especies de las secciones *Lycopersicon*, *Lycopersicoides* y *Juglandifolium*. El grosor de las líneas indica la cantidad de semilla producida en los cruces, mientras que las líneas discontinuas señalan cruces fallidos. Los círculos vacíos representan distintas poblaciones de *S. peruvianum*.



Utilización de las especies silvestres en la mejora

El tomate cultivado posee una diversidad genética muy pequeña (Williams y St. Clair 1993), mientras que las especies silvestres relacionadas son mucho más variables genéticamente y representan fuentes de variación para numerosos caracteres que han sido de utilidad en la mejora del tomate durante los últimos 70 años (Nuez 1995). Por ejemplo, la mayoría de genes de resistencia de los cultivares comerciales proceden de especies silvestres relacionadas. Asimismo, se han identificado, en accesiones de especies silvestres, caracteres de interés para la mejora de la calidad del fruto o tolerancia/resistencia tanto a estreses bióticos como abióticos (Foolad 2007). Recientemente, Labate et al. (2007) han realizado una interesante revisión al respecto.

En cuanto a la resistencia a enfermedades, tanto víricas como fúngicas o bacterianas, se han identificado fuentes de resistencia principalmente en *S. pimpinellifolium*, *S. peruvianum*, *S. habrochaites* y *S. chilense*. En algunos casos se ha visto que la resistencia a alguna enfermedad vírica era debida a la resistencia frente a su vector. Este es el caso de la resistencia a la enfermedad del rizado amarillo del tomate (Tomato yellow leaf curl disease, TYLCD) descrita en *S. habrochaites* y *S. pennellii*, y que se debe a que se impide la transmisión del virus por el vector natural *Bemisia tabaci* (Genn.). Sin embargo, su utilización práctica no ha tenido mucho éxito debido a la complejidad del control genético de la resistencia, así como a la dificultad para eliminar las características indeseables del parental donante ligadas a los genes de resistencia (Lawson et al. 1997; Momotaz et al. 2005).

En cuanto a otras resistencia a insectos, el tomate es susceptible a un amplio conjunto de plagas de artrópodos y los cultivares disponibles actualmente no muestran unos niveles de resistencia tan elevados como para que se puedan reducir los tratamientos con pesticidas. Conseguir un aumento de este tipo de resistencias, que permitan reducir estos tratamientos y por tanto reducir su impacto ambiental, es un objetivo interesante de mejora (Labate et al. 2007). Se han citado fuentes de resistencia a insectos en *S. habrochaites*, *S. peruvianum* y *S. pennellii* (Gentile y Stoner 1968; Gentile et al. 1969; Rick 1973; Kennedy y Yamamoto 1979; Williams et al. 1980; Ecole et al. 1999), aunque su introgresión en el tomate ha sido limitada en la práctica debido a las dificultades prácticas en el fenotipado de las resistencias (Stevens y Rick 1986).

En cuanto a estreses abióticos, cabe destacar la resistencia o tolerancia a la sequía y a la salinidad. En ambos casos, se han descritos fuentes de tolerancia o resistencia en prácticamente todas las especies silvestres. Además se ha observado que otras especies más alejadas, como es el caso de *S. lycopersicoides*, *S. rickii*, *S. juglandifolium* y *S. ochranthum*, pueden ser útiles en la mejora de la especie cultivada (Foolad 2007). Sin embargo, su utilidad en la práctica es más complicada debido al control multigénico de este tipo de caracteres (Labate et al. 2007), así como a la variación en la respuesta que se produce en función del estadio de desarrollo de la planta (Foolad y Lin 1997).

En el caso de la resistencia o tolerancia a la sequía, *S. pennellii* y *S. chilense* al ser propias de ambientes áridos o semiáridos, son buenas candidatas en la búsqueda de fuentes de resistencia. En el primer caso, *S. pennellii* muestra dos estrategias: su morfología “altamente eficiente en el uso del agua” (Martin y Thorstenson 1988) y la regulación de sus estomas (Kebede et al. 1994). En el caso de *S. chilense*, adaptada a las condiciones de extrema aridez del desierto de Atacama (norte de Chile), su estrategia se basa en el desarrollo de un potente sistema radicular (Sánchez Peña et al. 1995). Frente a salinidad, varias especies parecen útiles como fuentes de resistencia, como el caso de *S. pimpinellifolium*, *S. cheesmaniae*, *S. galapagense*, *S. pennellii* y *S. peruvianum* (Foolad 2004; Cuartero et al. 2006).

Otros caracteres de suma importancia son los referidos a la calidad del fruto. Recientemente, se ha publicado una revisión de este tema, distinguiendo entre la mejora de la calidad en tomate para industria (Gragera 2006) y para consumo en fresco (Roselló y Nuez 2006). En el primer caso, un aspecto muy importante es el color del fruto y, en relación con él, se han encontrado niveles importantes de variabilidad en *S. lycopersicum* y más aún en *S. pimpinellifolium* para contenido en licopeno y β -caroteno (Roselló y Nuez 2002).

Debido a la cantidad de tomate que consumimos se ha comprobado que su contribución a nuestra dieta en cuanto a vitamina A y C es importante (USDA 2002). Además, el licopeno, responsable del color rojo de los frutos, es un potente antioxidante. Por todo ello, un objetivo importante de mejora se centra en incrementar los valores de estos componentes. Sin embargo, la complejidad de estos caracteres, como el sabor del tomate, en el que hay numerosos genes implicados, hacen que su manejo en la práctica no sea sencillo, puesto que requieren el empleo de muchos QTLs

que explican parcialmente uno o varios de estos aspectos de calidad (Roselló y Nuez 2002). Se han caracterizado más de 20 genes en tomate que influyen en el tipo, cantidad y distribución de los carotenoides en el fruto (Labate et al. 2007). La mayoría de variantes de color se identificaron primero como mutantes en *S. lycopersicum*, pero también se han identificado en especies silvestres. Como ejemplo, el alelo *Beta* (*B*) se caracterizó en primer lugar a partir de descendientes transgresivos de fruto naranja, provenientes de un cruce entre *S. lycopersicum* y *S. habrochaites* (Lincoln et al. 1943; Kohler et al. 1947). Asimismo, también se han descrito numerosos QTLs introgresados de *S. pimpinellifolium*, *S. habrochaites*, *S. pennellii*, *S. peruvianum* y *S. neorickii*, que influyen sobre el color del fruto (Labate et al. 2007).

4.2. Sección *Basarthrum*

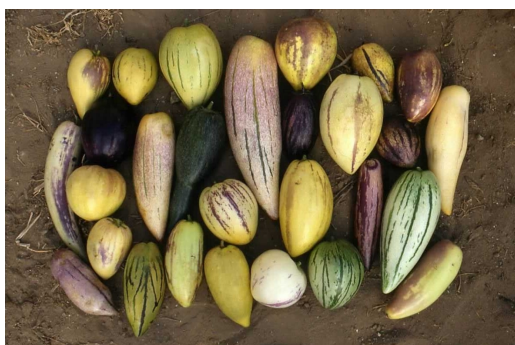
Taxonomía

La sección *Basarthrum* está formada por especies herbáceas, mayoritariamente perennes, nativas de la región andina de América del Sur, excepto algunas procedentes de las regiones montañosas de América Central (Correl 1962). Su distribución no es uniforme, algunas de estas especies se encuentran ampliamente distribuidas, mientras que otras aparecen restringidas a unas zonas muy concretas. Además, estas especies ocupan nichos ecológicos distintos, con rangos de altitudes que varían entre el nivel del mar y los 3300 m de altura (Prohens et al. 2003).

Esta sección se caracteriza por la presencia de una articulación basal en el pedicelo, es decir, las flores caen con el pedicelo unido, dejando una cicatriz en el eje de la inflorescencia (IPGRI y COMAV 2004).

En ella encontramos el pepino dulce (*Solanum muricatum*), cultivo andino aprovechado por sus frutos comestibles, los cuales son jugosos, aromáticos, dulces y pueden ser muy variables en forma y color (Figura 7).

Figura 7. Variabilidad en frutos de pepino dulce (*S. muricatum*).



Correll (1962) estructuró la sección *Basarthrum* en 6 series con un total de 22 especies. Sucesivos estudios de tipo biosistemático, en los que se analizaron datos cariotípicos, de morfología del polen, resultados de hibridaciones artificiales y/o materiales de herbarios (Anderson 1975; Anderson y Gensel 1976; Anderson 1979; Mione y Anderson 1992), y posteriormente de tipo molecular, empleando patrones de restricción de ADN cloroplástico (Spooner et al. 1993; Anderson et al. 1996; Anderson y Jansen 1998) han conducido a una reordenación de la sección. En la actualidad, la sección *Basarthrum* se ha separado en dos (sección *Anarrichomenum* y *Taeniotrichum*), quedando reducida a 4 series y 13 especies (Anderson y Jansen 1998) (Figura 8). Entre ellas se encuentran las dos nuevas especies descritas a partir de la caracterización molecular realizada en la presente tesis y que han sido nombradas como *S. catilliflorum* G.J. Anderson, Martine, Prohens & Nuez y *S. perlongistylum* G.J. Anderson, Martine, Prohens & Nuez (Anderson et al. 2006).

La distribución de las especies entre las 4 series es la siguiente:


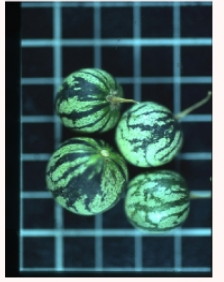

- *Canensa*: *S. canense* Rybd.,
- *Caripensia*: *S. basendopogon* Bitter, *S. caripense* Humb. & Bonpl. ex Dunal, *S. cochoae* G.J. Anderson & Bernardello, *S. filiforme* Ruiz López & Pavón, *S. fraxinifolium* Dunal in DC, *S. heiseri* G.J. Anderson, *S. tabanoense* Correll, *S. trachycarpum* Bitter & Sodiro, *S. catilliflorum* G.J. Anderson, Martine, Prohens & Nuez y *S. perlongistylum* G.J. Anderson, Martine, Prohens & Nuez,
- *Muricata*: *S. muricatum*,
- *Suaveolentia*: *S. suaveolens* Kunth & Boché.

Las 4 series están aisladas reproductivamente, a excepción de las series *Caripensia* y *Muricata*, puesto que *S. muricatum* puede cruzar con varias especies de *Caripensia* (Anderson et al. 1996). Del mismo modo, varias especies de *Caripensia* pueden cruzar produciendo híbridos fértiles. Todas las especies de la serie *Caripensia* son autoincompatibles, a excepción de *S. trachycarpum*, que es autocompatible y autógena (Mione y Anderson 1992). En la tabla 3 se muestran las claves que permiten distinguir las especies de esta serie (Anderson y Bernardello 1991). Dentro de la serie *Caripensia* se incluye el complejo Caripense que incluye la especie *S. caripense*, de amplia distribución, y 3 especies de distribución muy limitada: *S. filiforme*, *S. fraxinifolium* y *S. heiseri* (Anderson y Bernardello 1991).

Tabla 3. Claves de identificación de las especies de la sección *Basarthrum* serie *Caripensia* (Anderson y Bernardello 1991).

| |
|---|
| 1a. Hojas fundamentalmente simples |
| 2a. Inflorescencia ramificada >> <i>S. basendopogon</i> |
| 2b. Inflorescencia no ramificada |
| 3a. Corola rotada >> <i>S. trachycarpum</i> |
| 3b. Corola estrellada >> <i>S. tabanoense</i> |
| 1b. Hojas fundamentalmente pinnado-compuestas |
| 4a. Inflorescencia ramificada |
| 5a. Arbusto tipo parra, frutos con superficie lisa, normalmente 20-25 flores por inflorescencia >> <i>S. cochoae</i> |
| 5b. Porte erecto subarborescente, frutos verrugosos, normalmente menos de 20 flores por inflorescencia >> <i>S. basendopogon</i> |
| 4a. Inflorescencia simple >> complejo Caripense |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthrum*.

| | | |
|--|--|---|
| <p>Nombre <i>S. canense</i> Rybd.</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Inflorescencia bifurcada, con 17 flores de media. - Fruto globosos a ampliamente elipsoides, superficie lisa, verde con franjas verde pálido o verde oscuro, de 1,5-2,5 cm de longitud. - Herbácea anual. | |
| <p>Sistema reproductivo Autocompatible, autógama.</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 200 y 1400 m de altitud - Aparece en áreas modificadas | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |





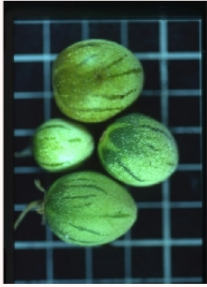
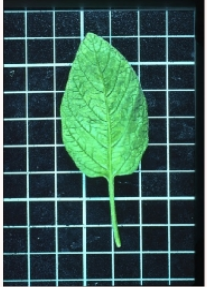
| | | |
|---|--|---|
| <p>Nombre <i>S. basendopogon</i> Bitter</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Pequeño arbusto - Frutos verrugosos, de 2-4 gr. - Como algunos <i>S. muricatum</i>, muestra inflorescencia multipara, hábito erecto, y a menudo hojas simples. - Cruza con <i>S. muricatum</i>, aunque con baja tasa de éxito. | |
| <p>Sistema reproductivo autoincompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 2400 y 3300 m de altitud - Especie no muy abundante | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthrum* (cont.).

| | | |
|---|--|---|
| <p>Nombre <i>S. trachycarpum</i> Bitter & Sodiro</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Arbusto pequeño de porte erecto. - Inflorescencias simples con 4-8 flores. - Superficie del fruto arrugada. | |
| <p>Sistema reproductivo autocompatible, parcialmente autógama.</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 1500 y 2500 m de altitud. - Propia de ambientes más xerofíticos que el resto de especies. | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |







| | | |
|---|--|---|
| <p>Nombre <i>S. tabanoense</i> Correll</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Arbusto tipo parra. - Fruto verde con bandas moradas, de 10-40 gr, son los más grandes entre las silvestres. - Hoja e inflorescencia simple, corola estrellada a diferencia de la mayoría de especies que la tienen rotada. | |
| <p>Sistema reproductivo Autoincompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 2600 - 3500 m de altitud. - Especie de distribución limitada. | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthrum* (cont.).

| | | |
|---|---|--|
| <p>Nombre <i>S. heiseri</i> G.J. Anderson</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Morfológicamente muy similar a <i>S. caripense</i>. - Se distingue por tener un menor número de flores por inflorescencia, menos foliolos y más pequeños. - Es más pubescente. | |
| <p>Sistema reproductivo Autoincompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Colectada en una única ocasión. - A 2800 m de altitud. - En una zona ligeramente húmeda y sombría. | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |


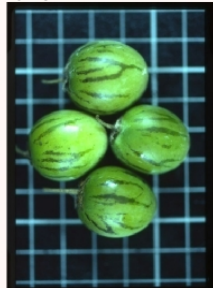




| | | |
|---|--|---|
| <p>Nombre <i>S. fraxinifolium</i> Dunal in DC</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Morfológicamente similar a <i>S. caripense</i>. | |
| <p>Sistema reproductivo Autoincompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 1100 y 2800 m | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthrum* (cont.).

| | |
|--|--|
| <p>Nombre <i>S. caripense</i> Humb & Bompl ex Dun</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Amplia variabilidad morfológica. - Forma típica con flores blancas a ligeramente violetas, y con 5 folíolos de media. - Fruto verde con bandas moradas, de 3-20 gr, comestible. - Inflorescencia simple. |
| <p>Sistema reproductivo Autoincompatible</p> | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 800 y 3800 m. - Aparece en gran variedad de hábitats, desde bastante húmedo hasta casi desierto. - Como mala hierba en márgenes de campos o cerca de casas. | |
| <p>Fruto</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | |







| | |
|---|--|
| <p>Nombre <i>S. filiforme</i> Ruiz López & Pavón</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Arbusto muy ramificado. - Densamente piloso. - Inflorescencia con pocas flores (<9) y pedúnculo menor de 2 cm. |
| <p>Sistema reproductivo Autoincompatible.</p> | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Única colecta realizada en las Lomas de Atiquipa, departamento de Arequipa. - Rango de altitud desconocido. | |
| <p>Fruto</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthurum* (cont.).

| | | |
|--|--|--|
| <p>Nombre <i>S. cochoae</i> GJ Anderson & Bernardello</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Única especie con hojas compuestas e inflorescencia ramificada, con más de 20 flores cada una. - Frutos de 2-6 gr, de color verde y superficie lisa. - Los híbridos con <i>S. muricatum</i> son fértiles. | |
| <p>Sistema reproductivo Autoincompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - La única colecta de material se realizó en la región peruana de Chachapoyas, en el departamento de Amazonas. - Entre 2600-2800 de altitud. | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |

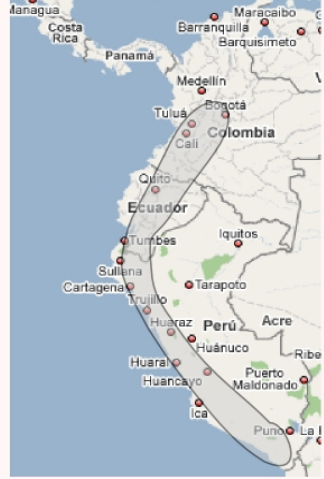





| | | |
|--|---|---|
| <p>Nombre <i>S. muricatum</i> Aiton</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Arbusto pequeño. - Fruto de coloración variable: verde, crema y morado, con la superficie lisa, entre 50-500 gr. - Hoja compuesta. | |
| <p>Sistema reproductivo Autocompatible</p> | | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre 0 y 2800 m. - Cultivada. | | <p>Distribución geográfica</p>  |
| <p>Fruto</p>  | <p>Hoja</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthrum* (cont.).

| | |
|--|---|
| <p>Nombre <i>S. suaveolens</i> Kunth & Bouché</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Herbácea anual. - Inflorescencia simple, con 9 flores de media, junto con <i>S. canense</i>, las que tienen las flores de menor tamaño. - Fruto verde con superficie lisa. |
| <p>Sistema reproductivo Autocompatible, autógama.</p> | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Entre el nivel del mar y 2300 m de altitud. - Aparece en hábitats modificados. | |
| <p>Fruto</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | |



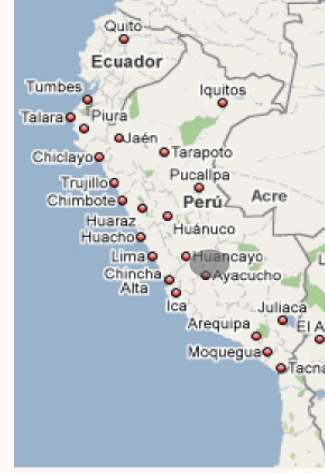



| | |
|---|---|
| <p>Nombre <i>S. catilliflorum</i> GJ Anderson, Martine, Prohens & Nuez</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Flores muy pequeñas. - Estilo corto. - Crecimiento con porte trepador. |
| <p>Sistema reproductivo Autocompatible, autógama.</p> | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Se encuentra a 2100 m de altitud. - En ambiente abierto y húmedo en Perú. | |
| <p>Fruto</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | |

Figura 8. Fichas informativas con los datos más relevantes de cada una de las especies pertenecientes a la sección *Basarthurum* (cont.).

| | |
|---|---|
| <p>Nombre <i>S. perlongistylum</i> GJ Anderson, Martine, Prohens & Nuez</p> | <p>Características morfológicas distintivas</p> <ul style="list-style-type: none"> - Estilo muy largo (8.5 mm); el más largo del grupo. - Flores más grandes. - Anteras de color amarillo limón, en lugar de amarillo-anaranjado. |
| <p>Sistema reproductivo Autoincompatible</p> | |
| <p>Ecosistema</p> <ul style="list-style-type: none"> - Se encuentra entre 2100 y 3200 m de altitud. - En bordes de caminos, taludes de carretera, quebradas y barrancos. | |
| <p>Fruto</p>  | <p>Hoja</p>  |
| <p>Distribución geográfica</p>  | |

Origen del pepino dulce

El pepino dulce sólo se encuentra en cultivo (Heiser 1969a; Anderson 1975; Heiser 1985), no conociéndose con exactitud su progenitor silvestre (Anderson et al. 1996). Sin embargo, se encuentra estrechamente relacionado con un grupo de especies silvestres incluidas en la serie *Caripensia* (Correll 1962; Anderson y Jansen 1998). Entre estas especies, los ancestros más probables del pepino dulce son: *S. caripense*, *S. tabanoense*, *S. basendopogon* y *S. cochoae* (Heiser 1969b; Anderson 1979; Anderson y Bernardello 1991; Anderson et al. 1996; Anderson y Jansen 1998). El pepino dulce y estas cuatro especies forman un grupo en que todas las especies pueden ser cruzadas entre sí (excepto *S. cochoae* con *S. tabanoense*), aunque el grado de éxito de los cruces y la fertilidad de los híbridos depende en gran medida de las especies consideradas y de la dirección del cruce (Heiser 1964; Anderson 1975; Anderson y Gensel 1976; Anderson 1977; Bernardello y Anderson 1990). De cada una de estas especies, sugeridas como posibles progenitores del pepino dulce, se tienen datos que apoyan esta hipótesis.

En cuanto a *S. caripense*, definida como “altamente compleja” por Correll (1962), su área de distribución es muy amplia y se solapa, en gran medida, con la de *S. muricatum* (Heiser 1969b). Algunas formas de *S. caripense*, compactas y con tendencia a producir ramas verticales, tienen una morfología parecida a la de *S. muricatum* (Prohens 1997). Aunque no pueda considerarse como una adventicia en sentido estricto, *S. caripense* se encuentra cerca de asentamientos humanos y en márgenes y vallas de campos de cultivo y sus frutos son recolectados y consumidos, especialmente por los niños, en Colombia y Ecuador (Correll 1962; Heiser 1964, 1969a, 1969b), e incluso se cultiva ocasionalmente (Fouqué 1973). Además, *S. caripense* es la especie silvestre más similar a *S. muricatum* en cuanto a cariotipo (Bernardello y Anderson 1990) y patrón de flavonoides (Anderson et al. 1987). Según Anderson y Jansen (1995) a partir de datos de patrones de restricción de ADN cloroplástico y ribosómico, *S. caripense* es uno de los ancestros más probables de *S. muricatum*. Además, los cruzamientos entre *S. caripense* y *S. muricatum* producen híbridos con meiosis regular y completamente fértiles, y han sido retrocruzados con éxito hacia cada uno de los parentales (Heiser 1964). Sin embargo, Brücher (1970) señala que las diferencias morfológicas entre *S. caripense* y *S. muricatum* excluyen al primero como ancestro de este último. Contrariamente a esta opinión, Heiser (1969b) y Anderson (1975) señalan que algunas plantas consideradas como *S. caripense* por Brücher, y en las cuales éste último se basa para descartar su posible contribución al origen del pepino dulce, no parecen ser *S. caripense* sino otra especie de la serie Caripensia (*S. fraxinifolium*).

En cuanto a *S. tabanoense*, fue propuesta por Brücher (1970) como el ancestro más probable del pepino dulce. El fruto de *S. tabanoense* es mayor que el de *S. caripense*, teniendo una mayor cantidad de pulpa y un aroma y sabor parecidos al de ciertos tipos de pepino dulce. Al igual que *S. caripense*, esta especie también se consume en las regiones donde se encuentra silvestre (Prohens 1997). Como en el caso anterior, los análisis de patrones de restricción de ADN cloroplástico y ribosómico apuntan a *S. tabanoense* como uno de los ancestros silvestres más probables del pepino dulce (Anderson y Jansen 1995; Anderson et al. 1996). Entre los argumentos en contra de *S. tabanoense* como ancestro del pepino dulce vemos que es una especie laxa, trepadora, mucho más pubescente que *S. caripense*, y con corola estrellada en lugar de rotada (Prohens 1997). Además, su distribución es más restringida, ya que sólo se encuentra en ciertas localidades de Colombia y Ecuador, en un rango de altitudes algo alto (2600 a 3500 m) si se compara con el del pepino dulce (Correll 1962; Heiser 1964, 1969b; Anderson 1975). Por

último, los híbridos entre *S. muricatum* y *S. tabanoense* tienen una baja fertilidad (Anderson 1979).

Otra especie que podría haber participado en el origen del pepino dulce es *S. basendopogon* (Correll 1962; Anderson 1979; Heiser 1964). Algunas plantas robustas de esta especie muestran ciertas similitudes con *S. muricatum* en el hábito de crecimiento, en la morfología de la hoja y en la presencia de inflorescencias ramificadas (Anderson 1979). Comparte con el pepino dulce una característica cariotípica poco usual en la sección *Basarthrum*: la presencia de un par de cromosomas telocéntricos (Bernardello y Anderson 1990). Sin embargo, *S. basendopogon* también es una planta poco frecuente y cuando se cruza con *S. muricatum* los frutos generalmente no presentan semillas (Anderson 1977).

Por último, *S. cochoae* presenta hábito erecto e inflorescencias ramificadas, al igual que *S. muricatum*, y da híbridos muy fértiles cuando se cruza con éste (Bernardello y Anderson 1990). También comparte con *S. muricatum* una característica cariotípica poco habitual en la sección *Basarthrum*: la presencia de un par de cromosomas subteloecéntricos (Bernardello y Anderson 1990). Sin embargo, esta especie es todavía menos frecuente que las anteriores; descrita por Bernardello y Anderson (1990) sólo ha sido citada en una localidad de Perú. Según Anderson et al. (1996), a partir de los datos de patrones de restricción de ADN cloroplástico y ribosómico, es poco probable que *S. cochoae* sea el ancestro de *S. muricatum*.

En resumen, de las cuatro especies consideradas, *S. caripense* y *S. tabanoense* son los ancestros más probables del pepino dulce, no descartándose que *S. muricatum* tenga un origen polifilético a partir de ambas especies (Anderson et al. 1996). Cladogramas obtenidos a partir de marcadores moleculares sitúan a algunos cultivares de *S. muricatum* muy cercanos a *S. caripense*, mientras que otros se encuentran muy próximos a *S. tabanoense* (Anderson et al. 1996; Anderson y Jansen 1998).

Por otra parte, se han señalado otras posibles hipótesis sobre el origen del pepino dulce como son: la posibilidad de que el ancestro silvestre de pepino dulce esté extinguido (Heiser 1969a, 1985; Anderson y Jansen 1995); que el pepino dulce pudiese ser el resultado de la combinación por hibridación de dos o más de las especies mencionadas anteriormente (Heiser 1985; Anderson

et al. 1987; Anderson et al. 1996); o por último, la posibilidad de que exista una especie silvestre de la serie *Muricata* que podría ser el ancestro silvestre del pepino dulce y que todavía no haya sido descubierta (Brücher 1970).

Utilización de estas especies en la mejora

El pepino dulce fue un cultivo importante en la región andina en tiempos precolombinos (Prohens et al. 1996). Sin embargo, fuera de su región de origen no ha alcanzado la importancia de otras solanáceas nativas del Nuevo Mundo, tales como el tomate, pimiento, patata o tabaco (Prohens et al. 2005). En la actualidad, debido a un incremento del interés por este cultivo en los mercados internacionales, han sido varios los intentos de introducirlo en varias regiones de clima mediterráneo (Cossio 1986; Péron et al. 1989; Pluda et al. 1993; Ercan y Akilli 1996). Sin embargo, en la mayoría de los casos los resultados no han sido satisfactorios puesto que, bajo las nuevas condiciones de cultivo, los materiales han mostrado niveles bajos de producción y/o calidad de fruto (Prohens et al. 1996).

Existe una gran diversidad de tipos y cultivares de pepino dulce y, al contrario que en otras plantas domesticadas, se ha visto que también existe una alta variabilidad a nivel molecular (Anderson et al. 1996; Rodríguez-Burruezo et al. 2003). A pesar de los logros conseguidos, uno de los principales objetivos de mejora sigue siendo el incremento de la calidad organoléptica del fruto, especialmente en referencia a su dulzura, lo que aumentaría su demanda en mercados potenciales de Europa y Norte América (Welles 1992). Sin embargo, el rango de variación para este carácter en la especie cultivada no es muy amplio (Rodríguez-Burruezo et al. 2002), por lo que las especies silvestres podrían ser fuentes de genes de interés para la mejora de este carácter.

Al igual que en otras solanáceas de interés (Kalloo y Bergh 1993), las especies silvestres pueden contribuir a la mejora genética del pepino dulce para otros caracteres como resistencias a estreses abióticos o bióticos, plagas y enfermedades, salinidad, sequía o frío. Sakata et al. (1989) encontraron una entrada de *S. caripense* resistente a *Verticillium dahliae* Kleb., y Pérez-Benlloch et al. (2001) encontraron algunos híbridos interespecíficos con *S. caripense* resistentes al virus del mosaico del tomate (ToMV).

Sin embargo, el aprovechamiento de estas especies silvestres como fuentes de mejora requiere un estudio preliminar de las mismas, de forma que podamos identificar los materiales que posean un mayor potencial. A continuación, se deberán introgresar los genes de interés desde la especie silvestre en el genoma del pepino dulce, lo que requiere la obtención de híbridos interespecíficos seguidos de retrocruces hacia la especie cultivada. Prohens et al. (2003) señalan que mediante programas de mejora sencillos se puede conseguir avances importantes en la mejora de calidad de esta especie.

Entre las especies silvestres de la sección *Basarthurum*, Prohens et al. (2003) señalan *S. caripense* y *S. tabanoense* como las más prometedoras para la mejora de la especie cultivada. Ambas cruzan fácilmente con el pepino dulce, tienen frutos de tamaño más o menos grande y alto contenido en sólidos solubles o ácido ascórbico. Otra especie interesante es *S. trachycarpum*, puesto que al estar adaptada a zonas muy secas, puede ser útil para la mejora de la tolerancia a la sequía (Prohens et al. 2003). Esta especie no cruza con el pepino dulce; sin embargo, se puede emplear *S. caripense* como puente genético para transferir los genes de interés desde *S. trachycarpum* a la cultivada.

Por último, la sección *Basarthurum* es muy cercana filogenéticamente a la patata y el tomate, dos cultivos muy importantes económicamente, por lo que esta sección podría ser un recurso útil para la mejora de estos cultivos (Trognitz y Trognitz 2004). En este sentido cabe destacar la obtención de híbridos somáticos interespecíficos entre tomate y pepino dulce (Sakamoto y Taguchi 1991).

5. EL ECLECTISMO DE LAS NUEVAS CONCEPCIONES TAXONÓMICAS

La especie es el nivel taxonómico básico en el estudio de la diversidad biológica, de la evolución y en la planificación de muchos programas de conservación, por lo tanto, es muy importante definir sus límites. Sin embargo, el concepto mismo de especie ha sido objeto de debate durante décadas. Se han propuesto numerosas definiciones de lo que es una especie y hoy en día se sigue debatiendo al respecto (Zimmer 2008).

La taxonomía moderna se fundó a partir del sistema de clasificación binomial creado por Linneo en el siglo XVIII. La nomenclatura binomial para identificar de manera unívoca a las especies solventaba los problemas de comunicación producidos por la variedad de nombres locales. Este autor fue uno de los primeros naturalistas en enfatizar el uso de similitudes entre organismos para construir un sistema de clasificación (Linneo 1735). De esa forma y sin saberlo, estaba clasificando a los organismos en virtud de sus similitudes genéticas y, por lo tanto, también evolutivas. Al mismo tiempo, Linneo propuso un esquema jerárquico de clasificación, donde las especies muy afines se agrupan en un mismo género, que a su vez se agrupan en órdenes mayores, estableciendo tres reinos (animal, vegetal y mineral) en el primer nivel, y subdividiendo los reinos en filos, los filos en clases, las clases en órdenes, los órdenes en familias, las familias en géneros y los géneros en especies. El nuevo orden creado por Linneo hizo más fácil la tarea de los taxónomos, pero tratar de delimitar las fronteras entre especies resulta a menudo frustrante.

Según Zimmer (2008), en opinión de Darwin estas discusiones eran consecuencia de querer “definir lo indefinible”. Las especies no estaban fijadas desde la creación sino que habían evolucionado. Cada grupo de organismos que denominamos especie se inicia como variedad de una especie más antigua. En el transcurso del tiempo, la selección natural la transforma conforme se adapta a su ambiente, mientras que otras variedades se extinguen. Una vieja variedad acaba siendo notablemente distinta de los demás organismos y es entonces cuando decimos que es una especie por derecho propio. Según Darwin (1859) el término “especie” es algo que se otorga arbitrariamente, por conveniencia, a un conjunto de individuos que se parecen mucho entre sí.

Hasta principios del siglo XX no se pudieron abordar las diferencias genéticas entre especies. Con la síntesis evolutiva moderna de los años 20 y 30, la investigación de estas diferencias genéticas llevó a una nueva forma de pensar. Las barreras que evitan la reproducción con otras especies hacen de una población una especie. Los genes fluyen entre los miembros de la especie cuando se aparean y permanecen dentro de la especie gracias a barreras reproductivas. El aislamiento geográfico constituye el origen mejor conocido para la aparición de tales barreras.

Esta forma de considerar la aparición evolutiva de nuevas especies condujo a una nueva noción de especie. Según Ernst Mayr, las especies no eran etiquetas de conveniencia, sino entidades reales, como las montañas o las personas. En 1942 propuso el **concepto biológico de especie**,

definida como el conjunto de poblaciones de organismos que pueden cruzarse entre sí (real o potencialmente) pero no con otras poblaciones.

Sin embargo, esta definición, aunque fue aplicada de forma general, tiene algunos problemas. En ella no se establece hasta qué punto debe estar aislada reproductivamente una especie para calificarla como tal. La definición se basa en el intercambio de información genética entre poblaciones (reproducción), no en el éxito reproductivo de los cruces: fertilidad y/o esterilidad. Una dificultad aún mayor hace referencia a las especies sin reproducción sexual, que pasarían a ser calificadas como “no-especies”. Además, este concepto no es claro con respecto a la relación entre poblaciones recientes y ancestrales, puesto que el concepto biológico se aplica a especies que coexisten en el tiempo.

En aquella época, Muller (1940a), en su estudio sobre la taxonomía y distribución del género *Lycopersicon*, apuntó que “si se aplica la regla arbitraria de que dos especies no son defendibles a menos que sean absolutamente distintas y no estén conectadas por tipos intermedios, se encuentra justificación a la opinión de que sólo existe una especie de tomate”. Sin embargo, el autor sugiere que la taxonomía no es un fin en sí misma, sino que tiene que ser útil, en particular respecto a la clasificación de los grupos económicamente importantes. Y que, por lo tanto, la variabilidad que se observa en los tomates no se puede obviar estableciendo un género monotípico y finalmente describe 6 especies, 3 variedades y 3 formas (Muller 1940b).

Se han propuesto otras definiciones de especie, siendo la más popular el **concepto filogenético de especie**, en el que se excluye el sexo de la ecuación y coloca en su lugar a la descendencia desde un antepasado común: “grupo de organismos que se pueden diagnosticar como diferente de otros grupos en el que hay un patrón de descendencia común”.

Los organismos emparentados comparten rasgos porque comparten antepasados. Se trata de ir obteniendo conjuntos más y más pequeños de organismos que comparten rasgos y cuando se alcanza un grupo que no puede dividirse ya más, ese grupo es una especie, según el concepto filogenético.

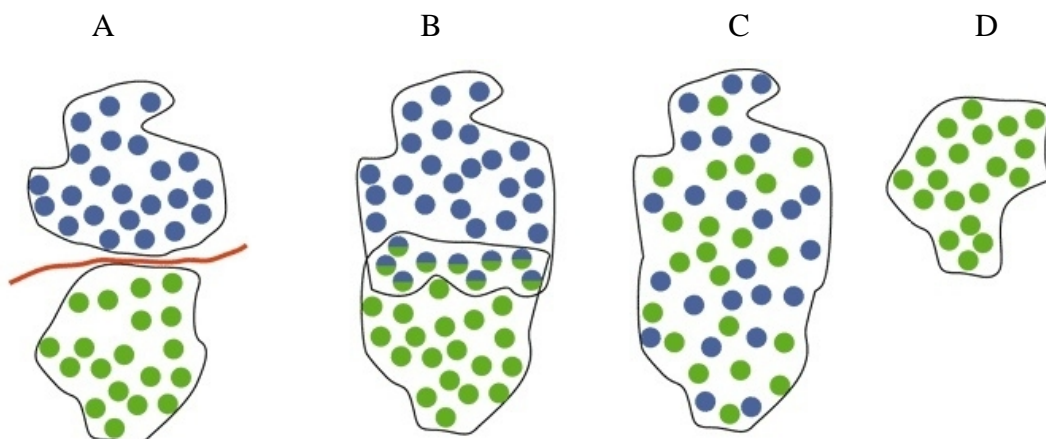
De Queiroz (1998) subrayó que la mayoría de conceptos de especies contrapuestos coinciden en aspectos básicos: aislamiento genético o ausencia de cohesión (barreras intrínsecas que limitan el intercambio de genes); falta de intercambio a nivel demográfico (aislamiento ecológico); presencia de caracteres diferenciales fijados; exclusividad; identidad y tendencias evolutivas independientes; agrupamientos genotípicos distinguibles. Todos se fundamentan en que una especie es un linaje evolutivo diferenciado. Para él, ésta es la definición fundamental de especie. De Queiroz piensa que métodos diferentes funcionan mejor en casos distintos.

Motivados por la confusión en torno a la idea de especie, algunos investigadores han empezado a crear clasificaciones filogenéticas que combinan las historias evolutivas con datos biológicos, conductuales, ecológicos y moleculares. Por ejemplo, Stockman y Bond (2008) estudiaron un género de araña, *Promyrmekiaphila*, hallada en California. Examinaron la historia evolutiva de la araña y su papel ecológico secuenciando varios genes de 222 ejemplares de 78 lugares. A partir de toda la información obtenida agruparon los animales en seis especies.

La clasificación de individuos en especies es de una gran utilidad en biología, pero hay que tener en cuenta que estamos ante procesos continuos y las dificultades aparecen cuando tratamos de establecer barreras discretas con significado biológico. Según una definición ideal, una especie estaría compuesta por grupos discretos de poblaciones, reproductivamente aisladas, formando grupos monofiléticos diferenciables genotípicamente y fenotípicamente. Sin embargo, esta situación está muy lejos de suceder en la práctica. En muchas ocasiones los límites no están claros y lo que se observa es un continuo, haciendo muy difícil la clasificación de los individuos.

En la figura 9 se muestra una representación gráfica en la que se simulan varias situaciones posibles: dos especies perfectamente definidas y separadas, dos especies distintas pero entre las que se producen algunos híbridos, dos subespecies entre las que el número de híbridos es importante, o una única especie. En la práctica, las diferencias entre estas situaciones pueden ser muy sutiles, puesto que se trata de un gradiente continuo entre cada una de ellas. Por este motivo, nos encontramos con tantos problemas en las definiciones cuando se quieren llevar al extremo. Por todo ello, la decisión de qué es una especie debería ser más flexible, con un enfoque integrador y apoyada por la mayor cantidad de datos posible (morfológicos, moleculares, ecológicos).

Figura 9. Ejemplo visual en el que se representan: A: dos especies distintas, B: dos especies con algunos híbridos, C: dos subespecies, D: una única especie.



II. OBJETIVOS

Para que el empleo de los recursos fitogenéticos sea eficiente, debemos conocer cual es la taxonomía, la estructura genética y las relaciones entre las distintas especies. El conocimiento de la distribución de la variabilidad existente permite establecer criterios de conservación más racionales y un uso más eficiente y más efectivo de estos recursos para la mejora de las especies cultivadas.

Los objetivos de la presente tesis doctoral se centran en el estudio de dos secciones del género *Solanum* que incluyen especies cultivadas, *Lycopersicon* y *Basarthurum*. Los objetivos son los siguientes:

- Estudiar la taxonomía y las relaciones filogenéticas en la sección *Lycopersicon*.
- Estudiar la distribución de la variación genética de *S. pimpinellifolium*, especie silvestre más cercana al tomate cultivado, en todo su rango de distribución.
- Estudiar la taxonomía y las relaciones filogenéticas en la sección *Basarthurum*.
- Estudiar la variabilidad intraespecífica existente en el pepino dulce, especie cultivada por sus frutos.

III. RESULTADOS

Capítulo 1.- Classification and phylogenetic relations in the Solanum section Lycopersicon based in AFLP and two nuclear gene sequences.

Genetic Resources and Crop Evolution

Classification and phylogenetic relationships in *Solanum* section *Lycopersicon* based on AFLP and two nuclear gene sequences

Elena Zuriaga · Jose Blanca · Fernando Nuez

Received: 24 September 2008 / Accepted: 17 November 2008
© Springer Science+Business Media B.V. 2008

Abstract The classification and phylogeny of the species belonging to *Solanum* section *Lycopersicon* is a complex issue that has not yet reached a widely accepted consensus. These species diverged recently, are still closely related and, in some cases, are still even capable of interspecific hybridization, thereby blurring the difference between intra- and interspecific variation. To help resolve these issues, in the present study, several accessions covering the natural range for each species were used. In addition, to avoid biases due to the molecular method employed, both AFLP markers and two nuclear-gene sequences, CT179 and CT66, were used to characterize the plant materials. The data obtained suggest a classification similar to those previously proposed by other authors, although with some significant changes. Twelve species were recognized as distinct based on this dataset. According to the data presented, the recently

proposed species, *S. corneliomulleri*, is indistinguishable from *S. peruvianum* s.str. In addition, both the sequence and the AFLP trees suggest that *S. arcanum* could represent a complex of populations composed of two cryptic species. With regard to phylogenetic relationships among these species, some clear groups were found: the Lycopersicon group formed by *S. pimpinellifolium*, *S. lycopersicum*, *S. cheesmaniae* and *S. galapagense*; the Arcanum group constituted by *S. chmielewskii*, *S. neorickii*, *S. arcanum* and *S. huaylasense*; and the Eriopersicon group made up of *S. peruvianum* and *S. chilense*. *Solanum pennellii* and *S. habrochaites* are not included in any group, but are the closest to the *S. lycopersicoides* outgroup.

Keywords Classification · *Lycopersicon* · Phylogeny · *Solanum* · Taxonomy

Jose Blanca and Fernando Nuez have contributed equally to this work and should be regarded as co-second authors.

E. Zuriaga · J. Blanca (✉) · F. Nuez
Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain
e-mail: jblanca@btc.upv.es

E. Zuriaga
e-mail: elzugar@btc.upv.es

F. Nuez
e-mail: fnuez@btc.upv.es

Introduction

Solanaceae contains between 3,000 and 4,000 species in about 90 genera (Knapp et al. 2004), the largest of which is *Solanum* L., with an estimated 1,500 species (Bohs 2007), almost half the diversity of the family. Within this genus, section *Lycopersicon* (Mill.) Wettst. includes tomatoes and their related wild species. The native distribution of this section ranges from the high Andes to coastal Ecuador, Peru and northern Chile, and it is also

found in the Galapagos Islands. The feral form of the cultivated *S. lycopersicum* L. has a worldwide distribution (Müller 1940).

The taxonomy of this section has been the subject of a long-lasting discussion. Even classification at the genus level has been controversial. In 1700, Tournefort, did not include the tomato in the *Solanum* genus, but instead in its own genus, *Lycopersicon*. In 1754, Linnaeus did not recognize *Lycopersicon*, but included the tomato in *Solanum*, although his contemporary, Miller (1754), formally described genus *Lycopersicon* based on Tournefort's classification (Esquinas-Alcazar and Nuez 2001). At present, tomatoes and their relatives are included within *Solanum* section *Lycopersicon*, as proposed by Child (1990) based on morphological characters. This classification has been supported by recent studies based on molecular data from both the chloroplast and nuclear genomes (Spooner et al. 1993; Bohs and Olmstead 1997; Olmstead and Palmer 1997; Peralta and Spooner 2001).

Different classification criteria have led to different numbers of species and subspecies in the section. In addition, conflicting models proposed to describe interspecific relationships have led to different complexes, series and groups (Peralta et al. 2008, p. 13). Müller (1940) and Luckwill (1943) divided the genus *Lycopersicon* into two subgenera (*Eulycopersicon* C. H. Müll. and *Eriopersicon* C. H. Müll.), although the two authors disagreed on the number of species or subspecies in each. Classification in these subgenera depended mainly on the color of the mature fruits, with species possessing glabrous and red to orange-colored fruits belonging to subgenus *Eulycopersicon* and species with pubescent and green fruits belonging to subgenus *Eriopersicon* (with the exception of *S. cheesmaniae*).

Later, Rick (1979) proposed an infrageneric classification based on crossing relationships into two complexes according to the ability (or inability) to cross with the cultivated tomato: “*esculentum*” and “*peruvianum*” complexes, respectively. *Solanum peruvianum* L. and *S. chilense* (Dunal) Reiche were classified in the “*peruvianum* complex” while the rest of the species were included in the “*esculentum* complex.”

Child (1990), treating tomatoes as being under genus *Solanum*, classified them into three series: *Lycopersicon* Child, *Eriopersicon* (C. H. Müll.) Child

and *Neolycopersicon* (Correll) Child. The first two series corresponded to the subgenera *Eulycopersicon* and *Eriopersicon* used by Müller (1940) and Luckwill (1943), while series *Neolycopersicon* only included *S. pennellii* Correll, as reviewed by Esquinas-Alcazar and Nuez (2001).

Recently, some new species have been considered as belonging to this section. Working with the tomatoes of the Galapagos Islands, Darwin et al. (2003) gave species rank to the taxon previously known ‘*L. cheesmanii* forma *minor*,’ describing *S. galapagense* S.C. Darwin et Peralta. They argued that this new species was clearly morphologically differentiated from the rest of the taxa present on the islands, and that it was more distinct from the others than *S. pimpinellifolium* L. and *S. lycopersicum* were from each other.

In addition, Peralta et al. (2005) recognized four species segregated from the highly variable *S. peruvianum* L. sensu lato: *S. peruvianum* sensu stricto, *S. corneliomulleri* J. F. Macbr., *S. arcanum* Peralta and *S. huaylasense* Peralta. Therefore, current classifications of the genus consider there to be 13 species of wild tomatoes, including the cultivated tomato (*S. lycopersicum*) and its weedy escaped forms (Peralta et al. 2008, p. 13).

The most recent classification was proposed by Peralta et al. (2008, p. 13). They divided the section into four groups: *Lycopersicon* group (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* (L. Riley) Fosberg and *S. galapagense*), *Neolycopersicon* group (*S. pennellii*), *Eriopersicon* group (*S. peruvianum* s.str., *S. corneliomulleri*, *S. huaylasense*, *S. habrochaites* S. Knapp et D.M. Spooner and *S. chilense*) and *Arcanum* group (*S. arcanum*, *S. chmielewskii* (C.M. Rick et al.) D.M. Spooner et al. and *S. neorickii* D.M. Spooner et al.).

Phylogenetic relationships among these species have been studied with various molecular approaches, including molecular markers based on restriction site data (cpDNA in Palmer and Zamir 1982; mtDNA in McClean and Hanson 1986; nuclear RFLPs in Miller and Tanksley 1990), on microsatellites (Alvarez et al. 2001), and on AFLPs (Spooner et al. 2005). Sequence data have also been employed, generated from the nuclear ITS rDNA region by Marshall et al. (2001) and the GBSSI gene sequence by Peralta and Spooner (2001). Recently, Peralta and Spooner (2005) also incorporated cladistic and

phenetic analyses of morphological data in order to clarify relationships generated from past studies.

Despite the great wealth of data collected, no consensus has been reached regarding phylogenetic relations, most likely a result of the complex structure of this group. Peralta et al. (2008, p. 53) noted that past studies left quite a few questions unanswered. Earlier attempts to develop phylogenetic relationships were discordant and the statistical support for them could be improved with more data. As an example of these conflicting results, one can examine the position of northern populations of *S. peruvianum* s.l. in recent phylogenetic trees. It has been found to be closely related to *S. lycopersicum* (Peralta and Spooner 2001), to southern populations of *S. peruvianum* s.l. (Marshall et al. 2001) and to *S. chmielewskii* (Alvarez et al. 2001; Peralta et al. 2008, p. 52).

In our opinion, one of the causes of these conflicting phylogenies could be the low intraspecific diversity employed in these studies. In general, very few accessions of each species were analyzed, and, thus, potential intraspecific variation was inadequately sampled. The role of intraspecific variation can be especially important in recently derived and closely related taxa. In fact, for some of these species, recent studies have found significant intraspecific structuring. Peralta et al. (2005) split *S. peruvianum* s.l. into four species, and Zuriaga et al. (2008) showed a clear subdivision between Ecuadorian and Peruvian populations of *S. pimpinellifolium*. Unfortunately, the studies carried out before 2005 did not take into account these specific substructures and, in general, used only one or a few accessions to represent each species.

The Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) has gathered extensive germplasm collections of these taxa through several expeditions organized with the collaboration of the Universidad Nacional Pedro Ruiz Gallo (UNPRG) (Peru), the Universidad Nacional Agraria La Molina (UNAM) (Peru), the Universidad Nacional de Piura (UNP) (Peru), the Universidad Nacional de Trujillo (UNT) (Peru), the Parque Nacional de Galápagos (PNG) (Ecuador) and the Universidad Nacional de Loja (UNL) (Ecuador) (Nuez et al. 2008). This germplasm collection, complemented by several accessions kindly provided by the Tomato Genetics Resource Center (TGRC), have allowed us to tackle classification and phylogenetic

relationships in section *Lycopersicon* by using many accessions that sample the entire range of distribution for each species (Taylor 1986; Warnock 1991; Darwin et al. 2003).

Moreover, to improve the quality of the data generated, we employed complementary molecular approaches, by analyzing ALFPs and DNA sequences from two genes. Several recent studies have successfully used both kinds of data to resolve phylogenetic relationships among closely related taxa at both interspecific and intraspecific levels (Blanca et al. 2007; Hodkinson et al. 2000; Xu and Sun 2001; van Ee et al. 2006). For DNA sequencing, we chose two markers successfully used in previous wild-tomato studies: CT066 and CT179 (Städler et al. 2005; Roselius et al. 2005; Städler et al. 2008).

Materials and methods

Plant materials

We analyzed 210 accessions including to all recognized species of *Solanum* section *Lycopersicon* and 3 accessions of *Solanum lycopersicoides* Dunal, which were included as a phylogenetic outgroup (Table 1). These accessions were selected to sample comprehensively the geographic ranges of each species, and also accounted for published information about intraspecific variation (Rick 1963; Rick et al. 1976, 1979; Nuez et al. 2004; Sifres et al. 2007, Peralta et al. 2008, p. 13; Zuriaga et al. 2008). The number of accessions per species ranged from 2 in *S. galapagense* to 46 in *S. chilense*.

Of these 213 accessions, 153 were collected by COMAV, UNPRG, UNAM, PNG, UNL, UNP and UNT. The other 60 accessions were supplied by TGRC. These reference accessions were especially useful in analyzing the division of *S. peruvianum* s.l. into several species, as well as in comparing our results with earlier previous works. The accessions collected in the expeditions were classified using the morphological data gathered in situ.

Molecular methods

For each accession, genomic DNA was isolated from one plant by using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA), following the

Table 1 List of the accessions used for the study of the phylogeny of *Solanum* section *Lycopersicon*, including current and original taxonomic classifications, accession identifiers, and geographic coordinates for the collection sites

| Accession code | Originally classified as | Latitude | Longitude |
|-----------------------|--------------------------|----------|-----------|
| <i>S. arcanum</i> | | | |
| Amazonas | | | |
| PE22 | <i>S. arcanum</i> | 0926–S | 07708–W |
| LA1395 | <i>S. arcanum</i> | 061300S | 0775100W |
| LA1626 | <i>S. arcanum</i> | 083030S | 0772200W |
| LA2185 | <i>S. arcanum</i> | 052900S | 0783100W |
| LA2917 | <i>S. neorickii</i> | 091800S | 0770100W |
| Cajamarca | | | |
| LA0441 | <i>S. arcanum</i> | 080000S | 0790700W |
| LA2157 | <i>S. arcanum</i> | 063000S | 0784912W |
| ECU0783 | <i>S. peruvianum</i> | 071947S | 0781102W |
| ECU0777 | <i>S. peruvianum</i> | 071817S | 0782839W |
| <i>S. cheesmaniae</i> | | | |
| GLP34 | <i>S. cheesmaniae</i> | 000933S | 0904924W |
| GLP39 | <i>S. cheesmaniae</i> | 001654S | 0903307W |
| LA1449 | <i>S. cheesmaniae</i> | 004429S | 0901849W |
| GLP44 | <i>S. cheesmaniae</i> | 001709S | 0903230W |
| GLP65 | <i>S. cheesmaniae</i> | 002452S | 0901724W |
| GLP27 | <i>S. cheesmaniae</i> | 003337S | 0901958W |
| GLP57 | <i>S. cheesmaniae</i> | 002445S | 0901704W |
| GLP35 | <i>S. cheesmaniae</i> | 001724S | 0903332W |
| GLP54 | <i>S. cheesmaniae</i> | 002201S | 0903505W |
| <i>S. chilense</i> | | | |
| LA2932 | <i>S. chilense</i> | 292900S | 0701000W |
| LA2930 | <i>S. chilense</i> | 252400S | 0702400W |
| T133 | <i>S. chilense</i> | 171808S | 0704254W |
| CH64 | <i>S. chilense</i> | 194025S | 0691015W |
| T091 | <i>S. chilense</i> | 174600S | 0695622W |
| LA1932 | <i>S. chilense</i> | 151700S | 0743700W |
| PER560 | <i>S. chilense</i> | 180615S | 0702023W |
| T079 | <i>S. chilense</i> | 175645S | 0701044W |
| T082 | <i>S. chilense</i> | 175225S | 0700539W |
| LA1969 | <i>S. chilense</i> | 173200S | 0700200W |
| T092 | <i>S. chilense</i> | 172633S | 0702655W |
| PER551 | <i>S. chilense</i> | 174305S | 0700822W |
| T077 | <i>S. chilense</i> | 173324S | 0700147W |
| CH58 | <i>S. chilense</i> | 195157S | 0692358W |
| CH25 | <i>S. chilense</i> | 184402S | 0694305W |
| T085 | <i>S. chilense</i> | 174827S | 0695945W |
| T078 | <i>S. chilense</i> | 173024S | 0700156W |
| CH34 | <i>S. chilense</i> | 184721S | 0694222W |

Table 1 continued

| Accession code | Originally classified as | Latitude | Longitude |
|---------------------------|---------------------------|----------|-----------|
| CH13 | <i>S. chilense</i> | 221514S | 0682848W |
| LA0460 | <i>S. chilense</i> | 174800S | 0700100W |
| CH46 | <i>S. chilense</i> | 182324S | 0693827W |
| CH18 | <i>S. chilense</i> | 182847S | 0685147W |
| T038 | <i>S. chilense</i> | 144934S | 0744113W |
| T074 | <i>S. chilense</i> | 174043S | 0700510W |
| CH30 | <i>S. chilense</i> | | |
| T076 | <i>S. chilense</i> | 173504S | 0700205W |
| LA1960 | <i>S. chilense</i> | 170500S | 0705200W |
| T136 | <i>S. chilense</i> | 172712S | 0704928W |
| T109 | <i>S. chilense</i> | 180428S | 0701852W |
| T075 | <i>S. chilense</i> | 173504S | 0700205W |
| CH15 | <i>S. chilense</i> | 221726S | 0682049W |
| CH57 | <i>S. chilense</i> | 191914S | 0692654W |
| CH59 | <i>S. chilense</i> | 194739S | 0691834W |
| LA2880 | <i>S. chilense</i> | 234700S | 0681500W |
| T128 | <i>S. chilense</i> | 171231S | 0704909W |
| T117 | <i>S. chilense</i> | 170627S | 0705055W |
| T123 | <i>S. chilense</i> | 170227S | 0704916W |
| T101 | <i>S. chilense</i> | 1721–S | 07019–W |
| CH55 | <i>S. chilense</i> | 192041S | 0693016W |
| CH60 | <i>S. chilense</i> | 194654S | 0691713W |
| LA0458 | <i>S. chilense</i> | 185700S | 0701100W |
| T041 | <i>S. chilense</i> | 144858S | 0743959W |
| T107 | <i>S. chilense</i> | 172250S | 0700829W |
| CH43 | <i>S. chilense</i> | | |
| T095 | <i>S. chilense</i> | 172318S | 0702020W |
| PER526 | <i>S. chilense</i> | 170923S | 0705134W |
| <i>S. chmielewskii</i> | | | |
| PER489 | <i>S. chmielewskii</i> | 134804S | 0725803W |
| LA1330 | <i>S. chmielewskii</i> | 132900S | 0724500W |
| PER451 | <i>S. chmielewskii</i> | 132403S | 0735344W |
| PER502 | <i>S. chmielewskii</i> | 135610S | 0730724W |
| PER512 | <i>S. chmielewskii</i> | 133401S | 0723536W |
| LA3643 | <i>S. chmielewskii</i> | 135100S | 0714900W |
| PER483 | <i>S. chmielewskii</i> | 134430S | 0725500W |
| LA3653 | <i>S. chmielewskii</i> | 134700S | 0725000W |
| PER499 | <i>S. chmielewskii</i> | 135155S | 0730128W |
| PER452 | <i>S. chmielewskii</i> | 132534S | 0735309W |
| LA1306 | <i>S. chmielewskii</i> | 125700S | 0740100W |
| <i>S. corneliomulleri</i> | | | |
| LA1473 | <i>S. corneliomulleri</i> | 114942S | 0763659W |
| LA1945 | <i>S. corneliomulleri</i> | 154630S | 0732400W |

Table 1 continued

| Accession code | Originally classified as | Latitude | Longitude |
|------------------------|---------------------------|----------|-----------|
| LA1379 | <i>S. corneliomulleri</i> | 104800S | 0770400W |
| T040 | <i>S. corneliomulleri</i> | 144916S | 0744035W |
| T118 | <i>S. corneliomulleri</i> | 170626S | 0705055W |
| T121 | <i>S. corneliomulleri</i> | 170315S | 0705153W |
| T024_1 | <i>S. corneliomulleri</i> | 145209S | 0744703W |
| T015 | <i>S. corneliomulleri</i> | 143310S | 0744858W |
| LA1305 | <i>S. corneliomulleri</i> | 132300S | 0752130W |
| LA1292 | <i>S. corneliomulleri</i> | 114534S | 0761802W |
| T036 | <i>S. corneliomulleri</i> | 145013S | 0744308W |
| T018 | <i>S. corneliomulleri</i> | 145106S | 0744754W |
| LA1283 | <i>S. corneliomulleri</i> | 120700S | 0762800W |
| T024_5 | <i>S. corneliomulleri</i> | 145209S | 0744703W |
| T047 | <i>S. corneliomulleri</i> | 162534S | 0730859W |
| T067 | <i>S. corneliomulleri</i> | 160924S | 0713841W |
| T024_10 | <i>S. corneliomulleri</i> | 145209S | 0744703W |
| T045 | <i>S. corneliomulleri</i> | 161614S | 0733006W |
| T053 | <i>S. corneliomulleri</i> | 155004S | 0723932W |
| T061 | <i>S. corneliomulleri</i> | 1609–S | 07211–W |
| <i>S. galapagense</i> | | | |
| LA0317 | <i>S. galapagense</i> | 0017–S | 09033–W |
| LA0483 | <i>S. galapagense</i> | 002314S | 0913211W |
| <i>S. habrochaites</i> | | | |
| ECU1498 | <i>S. habrochaites</i> | 010256S | 0800525W |
| LA2098 | <i>S. habrochaites</i> | 042200S | 0794800W |
| LA2861 | <i>S. habrochaites</i> | 034900S | 0791600W |
| ECU0531 | <i>S. habrochaites</i> | 040745S | 0795510W |
| LA1353 | <i>S. habrochaites</i> | 072200S | 0784900W |
| ECU0533 | <i>S. habrochaites</i> | 040212S | 0791120W |
| ECU0523 | <i>S. habrochaites</i> | 051253S | 0792610W |
| ECU0526 | <i>S. habrochaites</i> | 055942S | 0794303W |
| ECU0524 | <i>S. habrochaites</i> | 051648S | 0792816W |
| ECU0537 | <i>S. habrochaites</i> | 0358–S | 07928–W |
| LA1777 | <i>S. habrochaites</i> | 093300S | 0773524W |
| LA2314 | <i>S. habrochaites</i> | 062500S | 0775200W |
| ECU0652 | <i>S. habrochaites</i> | 052434S | 0793810W |
| ECU0544 | <i>S. habrochaites</i> | 052318S | 0793648W |
| ECU0527 | <i>S. habrochaites</i> | 0350–S | 08003–W |
| LA1753 | <i>S. habrochaites</i> | 115305S | 0762623W |
| LA1223 | <i>S. habrochaites</i> | 021200S | 0785000W |
| T039 | <i>S. habrochaites</i> | 144927S | 0744105W |
| <i>S. huaylasense</i> | | | |
| PE18 | <i>S. huaylasense</i> | 0849–S | 07752–W |
| PE20 | <i>S. huaylasense</i> | 0848–S | 07752–W |

Table 1 continued

| Accession code | Originally classified as | Latitude | Longitude |
|---------------------------|---------------------------|----------|-----------|
| LA2561 | <i>S. huaylasense</i> | 084900S | 0775200W |
| LA1983 | <i>S. huaylasense</i> | 084124S | 0775812W |
| PE19 | <i>S. huaylasense</i> | 0848–S | 07752–W |
| <i>S. lycopersicoides</i> | | | |
| T104 | <i>S. lycopersicoides</i> | 172003S | 0701537W |
| T090 | <i>S. lycopersicoides</i> | 174600S | 0695622W |
| T096 | <i>S. lycopersicoides</i> | 172232S | 0702002W |
| <i>S. lycopersicum</i> | | | |
| B046 | <i>S. lycopersicum</i> | | |
| ECU1474 | <i>S. lycopersicum</i> | 010350S | 0803529W |
| ECU1117 | <i>S. lycopersicum</i> | 020421S | 0775933W |
| ECU0096 | <i>S. lycopersicum</i> | | |
| ECU1068 | <i>S. lycopersicum</i> | 011906S | 0775320W |
| ECU1150 | <i>S. lycopersicum</i> | 024506S | 0781817W |
| LA1307 | <i>S. lycopersicum</i> | 123700S | 0794730W |
| LA2845 | <i>S. lycopersicum</i> | 060300S | 0765800W |
| B103 | <i>S. lycopersicum</i> | | |
| ECU1123 | <i>S. lycopersicum</i> | 021813S | 0780708W |
| <i>S. neorickii</i> | | | |
| ECU0301 | <i>S. neorickii</i> | 035401S | 0791853W |
| LA1626A | <i>S. neorickii</i> | 083030S | 0772200W |
| ECU1632 | <i>S. neorickii</i> | 032513S | 0790858W |
| ECU1626 | <i>S. neorickii</i> | 032455S | 0790933W |
| ECU1636 | <i>S. neorickii</i> | 041234S | 0791420W |
| PER493 | <i>S. neorickii</i> | 134923S | 0725803W |
| ECU1617 | <i>S. neorickii</i> | 032452S | 0791029W |
| LA1322 | <i>S. neorickii</i> | 132700S | 0722548W |
| LA2325 | <i>S. neorickii</i> | 065000S | 0775500W |
| ECU0450 | <i>S. neorickii</i> | 024747S | 0784539W |
| LA3655 | <i>S. neorickii</i> | 135600S | 0730100W |
| <i>S. pennellii</i> | | | |
| LA1272 | <i>S. pennellii</i> | 113112S | 0770000W |
| LA1940 | <i>S. pennellii</i> | 155830S | 0733700W |
| ECU0559 | <i>S. pennellii</i> | 050750S | 0810925W |
| T052 | <i>S. pennellii</i> | 155455S | 0723255W |
| T050 | <i>S. pennellii</i> | 155508S | 0723249W |
| ECU0729 | <i>S. pennellii</i> | 071718S | 0792852W |
| T021 | <i>S. pennellii</i> | 145141S | 0744721W |
| PE45 | <i>S. pennellii</i> | 0839–N | 07817–W |
| T139 | <i>S. pennellii</i> | 161642S | 0732835W |
| ECU0654 | <i>S. pennellii</i> | 050750S | 0810925W |
| LA2580 | <i>S. pennellii</i> | 093100S | 0780030W |
| ECU0831 | <i>S. pennellii</i> | 084515S | 0782905W |

Table 1 continued

| Accession code | Originally classified as | Latitude | Longitude |
|----------------------------|----------------------------|----------|-----------|
| T033 | <i>S. pennellii</i> | 145052S | 0744417W |
| LA1926 | <i>S. pennellii</i> | 143500S | 0745830W |
| <i>S. peruvianum</i> | | | |
| PER573 | <i>S. peruvianum</i> | 161709S | 0713852W |
| PER424 | <i>S. peruvianum</i> | 1154–S | 07630–W |
| ECU0766 | <i>S. peruvianum</i> | 071735S | 0783426W |
| PE27 | <i>S. peruvianum</i> | 0932–S | 07750–W |
| PER562 | <i>S. peruvianum</i> | 171541S | 0712952W |
| ECU0944 | <i>S. peruvianum</i> | 032431S | 0790946W |
| PER586 | <i>S. peruvianum</i> | 161017S | 0713918W |
| T140 | <i>S. peruvianum</i> | 130030S | 0762859W |
| PER429 | <i>S. peruvianum</i> | 114713S | 0761754W |
| ECU0106 | <i>S. peruvianum</i> | 0319–S | 07930–W |
| PER412 | <i>S. peruvianum</i> | 115555S | 0763747W |
| PER533 | <i>S. peruvianum</i> | 171620S | 0705050W |
| CH37 | <i>S. peruvianum</i> | 182404S | 0701555W |
| LA1913 | <i>S. peruvianum</i> | 142354S | 0751202W |
| T137 | <i>S. peruvianum</i> | 163725S | 0724233W |
| CH35 | <i>S. peruvianum</i> | | |
| LA2744 | <i>S. peruvianum</i> | 183300S | 0700900W |
| CH20 | <i>S. peruvianum</i> | 182415S | 0695843W |
| CH24 | <i>S. peruvianum</i> | 182705S | 0700333W |
| CH36 | <i>S. peruvianum</i> | 184853S | 0700915W |
| CH52 | <i>S. peruvianum</i> | 184703S | 0701207W |
| T113 | <i>S. peruvianum</i> | 175010S | 0710634W |
| T049 | <i>S. peruvianum</i> | 163435S | 0723743W |
| T012 | <i>S. peruvianum</i> | 143438S | 0745302W |
| T043 | <i>S. peruvianum</i> | 154854S | 0742259W |
| T110 | <i>S. peruvianum</i> | 180108S | 0705137W |
| CH45 | <i>S. peruvianum</i> | | |
| T051 | <i>S. peruvianum</i> | 155451S | 0723302W |
| <i>S. pimpinellifolium</i> | | | |
| Ecuador | | | |
| ECU1159 | <i>S. pimpinellifolium</i> | 032557S | 0783408W |
| LA1246 | <i>S. pimpinellifolium</i> | 035924S | 0792136W |
| ECU1349 | <i>S. pimpinellifolium</i> | 005922 N | 0793201W |
| ECU1516 | <i>S. pimpinellifolium</i> | 014153S | 0804649W |
| LA1429 | <i>S. lycopersicum</i> | 004903S | 0801256W |
| LA3124 | <i>S. cheesmaniae</i> | 004908S | 0900321W |
| ECU0602 | <i>S. pimpinellifolium</i> | 055942S | 0794303W |
| LA0411 | <i>S. pimpinellifolium</i> | 010600S | 0792900W |
| Peru | | | |
| ECU0733 | <i>S. pimpinellifolium</i> | 071821S | 0792645W |

Table 1 continued

| Accession code | Originally classified as | Latitude | Longitude |
|----------------|---|----------|-----------|
| PE15 | <i>S. pimpinellifolium</i> | 0931–N | 07755–W |
| LA1334 | <i>S. lycopersicum</i> | 162400S | 0731500W |
| ECU0574 | <i>S. pimpinellifolium</i> | 052559S | 0794432W |
| LA0369 | <i>S. pimpinellifolium</i> | 115700S | 0764300W |
| PE14 | <i>S. pimpinellifolium</i> | 0933–N | 07753–W |
| LA0373 | <i>S. pimpinellifolium</i> | 095624S | 0781348W |
| T008 | <i>S. pimpinellifolium</i> | 144053S | 0750805W |
| LA1689 | <i>S. pimpinellifolium</i> | 0513–S | 0803730W |
| LA1332 | <i>S. pimpinellifolium</i> | 145000S | 0745700W |
| LA2176 | <i>S. pimpinellifolium</i> | 050830S | 0790030W |
| T011 | <i>S. pimpinellifolium</i> | 143449S | 0745314W |
| T016 | <i>S. pimpinellifolium</i> | 143310S | 0744858W |
| ECU0719 | <i>S. pimpinellifolium</i> | 060936S | 0794202W |
| T093 | <i>S. pimpinellifolium</i> | 172543S | 0702453W |
| T004 | <i>S. pimpinellifolium</i> | 095555S | 0781315W |
| T046 | <i>S. pimpinellifolium</i> | 162219S | 0731605W |
| <i>Hybrid</i> | | | |
| LA1364 | <i>S. huaylasense</i> – <i>S. peruvianum</i> | 100900S | 0772830W |
| LA1984 | <i>S. arcanum</i> – <i>S. peruvianum</i> | 075400S | 0783300W |
| ECU1007 | <i>S. lycopersicum</i> – <i>S. pimpinellifolium</i> | 005726S | 0774857W |

manufacturer's protocol, and was stored at -20°C until used. DNA concentrations and qualities were quantified on agarose gels.

Visualization of AFLPs was carried out following procedures described by Nuez et al. (2004). The restriction enzymes *EcoRI* and *Tru9I* were used for digestion, and selective amplification employed three selective nucleotides on each primer. Six selective combinations of primers were used: *EcoRI* + ACA and *Mse* + CTA; *EcoRI* + AGC and *Mse* + CAA; *EcoRI* + AAC and *Mse* + CAA; *EcoRI* + ACC and *Mse* + CTA; *EcoRI* + AGG and *Mse* + CAT; and *EcoRI* + AAC and *Mse* + CTA. DNA fragments were separated in an ABI Prism 310 genetic analyzer (Applied Biosystems, Foster City, CA).

Two single-copy, nuclear markers, CT66 and CT179, were sequenced. CT66 corresponds to a putative arginine decarboxylase gene (Roselius et al. 2005), and its PCR amplification was done with the primers CT066_f (AAACGCTGTCCCTCTACCA)

and CT066_r (CTCACCAACACCATCAGCAA). CT179 is a putative delta-type tonoplast intrinsic protein gene (Roselius et al. 2005), and its PCR amplification was done with primers ct179_f (CGGCGTAGCTATTGGAAGTTT) and ct179_r (ATGGGTTGGAATAGCCTGAGA). All primers were designed to obtain sequences of approximately 700 bp. Our PCR conditions followed those of Roselius et al. (2005).

After amplification, PCR products were treated with Exonuclease I and precipitated with AcNa and ethanol, after adding 1 µg tRNA. DNA sequencing was performed in an ABI Prism 3100 genetic analyzer, following manufacturer instructions for the BigDye terminator v.3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA).

AFLP data analysis

For all 213 accessions the AFLP fragments were scored for presence or absence with Genographer 1.6 software (Benham et al. 1999), and a binary matrix was constructed. The reproducibility of the AFLP markers used was assessed studying two replicates for 30 randomly chosen accessions. The markers not reproducible in these replicates were not taken into account in the study. All analyses were made using homemade scripts based on R software v.2.6.0 (R Development Core Team 2007), available upon request. A neighbor-joining tree was built using Dice distances, and branch support was assessed by using 1000 bootstrapped data matrices.

DNA sequence data analysis

Sequences from 144 randomly selected accessions were assembled and edited with the Staden package v.1.4 (Bonfield 2004). Multiple alignments were made with ClustalW v.1.82 (Thompson et al. 1994) and edited in BioEdit v.7.0.9 (Hall 1999).

As these are single-copy genes, and all species analyzed are diploid, we considered individuals as homozygous or heterozygous depending on whether one or two haplotypes, respectively, were found in them. Heterozygotes were detected as double peaks in the chromatogram. Both haplotypes were inferred through haplotype subtraction (Clark 1990). The number of gaps and variable positions as well as

nucleotide and haplotype diversity were calculated with DnaSP v.4.50.3 (Rozas et al. 2003).

Phylogenetic analysis

Bayesian methods were employed to reconstruct phylogenetic trees with MrBayes software (Ronquist and Huelsenbeck 2003). Each Markov chain in the Bayesian search was run for 4×10^6 cycles, sampling every 100th cycle from the chain. Four chains were run simultaneously, and the initial 20% of trees from each run were discarded as burn-in samples. The remaining trees were used to construct majority-rule consensus trees.

Model comparison for single-gene data was performed in HyPhy (Kosakovsky Pond et al. 2005) by using the Akaike information criterion (AIC), based on Modeltest (Posada and Crandall 1998). The best-fitting evolutionary models (GTR + G + I for both genes, and non-absent sites for AFLP data) were implemented in the phylogenetic analyses for both single-gene and concatenated data.

MrBayes allows to combine different datasets to build phylogenies. One analysis was carried out combining the data from both DNA alignments and another one combining the AFLP and the sequence data. In this latter case to merge the dominant and codominant datasets an allele from each accession was randomly chosen.

Results

AFLP and gene sequence diversity

The six AFLP primer combinations produced 245 fragments, distributed between 60 and 380 bp, of which 221 (90.2%) were polymorphic. A marker was considered polymorphic at the 1% level. Multiple alignment of CT66 sequences resulted in a matrix of 665 bp, of which 96 (14.4%) were polymorphic. There was no variation in sequence length among taxa for this gene, and 66 different haplotypes were seen. CT179 sequence lengths ranged from 550 to 700 bp. Multiple alignment was problematic because the first half of the sequences was too variable, and several large indel events were present. Therefore, this region was eliminated from the alignment. Once this was done, the CT179 alignment resulted in a

Table 2 Nucleotide and haplotype diversity and level of AFLP polymorphism for the taxa studied. The sequence-based diversities are means calculated with data coming from both CT179 and CT66 genes. The AFLP polymorphism is the percentage of markers polymorphic at or above 1% level

| | Nucleotide diversity | Haplotype diversity | AFLP polymorphism % |
|----------------------------|----------------------|---------------------|---------------------|
| <i>S. peruvianum</i> | 0.013 | 0.92 | 64.4 |
| <i>S. chilense</i> | 0.009 | 0.89 | 54.6 |
| <i>S. arcanum</i> | 0.008 | 0.85 | 36.3 |
| <i>S. huaylasense</i> | 0.011 | 0.89 | 13.9 |
| <i>S. chmielewskii</i> | 0.002 | 0.45 | 17.1 |
| <i>S. neorickii</i> | 0.001 | 0.21 | 21.6 |
| <i>S. pimpinellifolium</i> | 0.001 | 0.30 | 20.0 |
| <i>S. cheesmaniae</i> | 0.000 | 0.00 | 6.9 |
| <i>S. lycopersicum</i> | 0.005 | 0.23 | 7.7 |
| <i>S. galapagense</i> | 0.000 | 0.00 | 1.2 |
| <i>S. habrochaites</i> | 0.006 | 0.72 | 42.0 |
| <i>S. pennellii</i> | 0.006 | 0.58 | 40.8 |

matrix of 339 characters, of which 64 (18.9%) were variable, and 78 different haplotypes were seen.

Levels of nucleotide and haplotype diversity and AFLP polymorphism were also calculated for each species (Table 2). The most diverse species was *S. peruvianum*. Members of the *Lycopersicon* group were considerably less variable, especially *S. cheesmaniae* and *S. galapagense*, which showed no nucleotide diversity and almost no AFLP polymorphism. Finally, *S. habrochaites* and *S. pennellii* showed intermediate variability. Correlations (r^2) between the different variability measures were 0.91 for nucleotide versus haplotype diversity, 0.75 for haplotype diversity versus AFLP polymorphisms, and 0.66 for and nucleotide diversity versus AFLP polymorphisms. In summary, all three diversity indicators were quite coherent. All three correlations were significant at the P 0.05 level.

Classification of the accessions

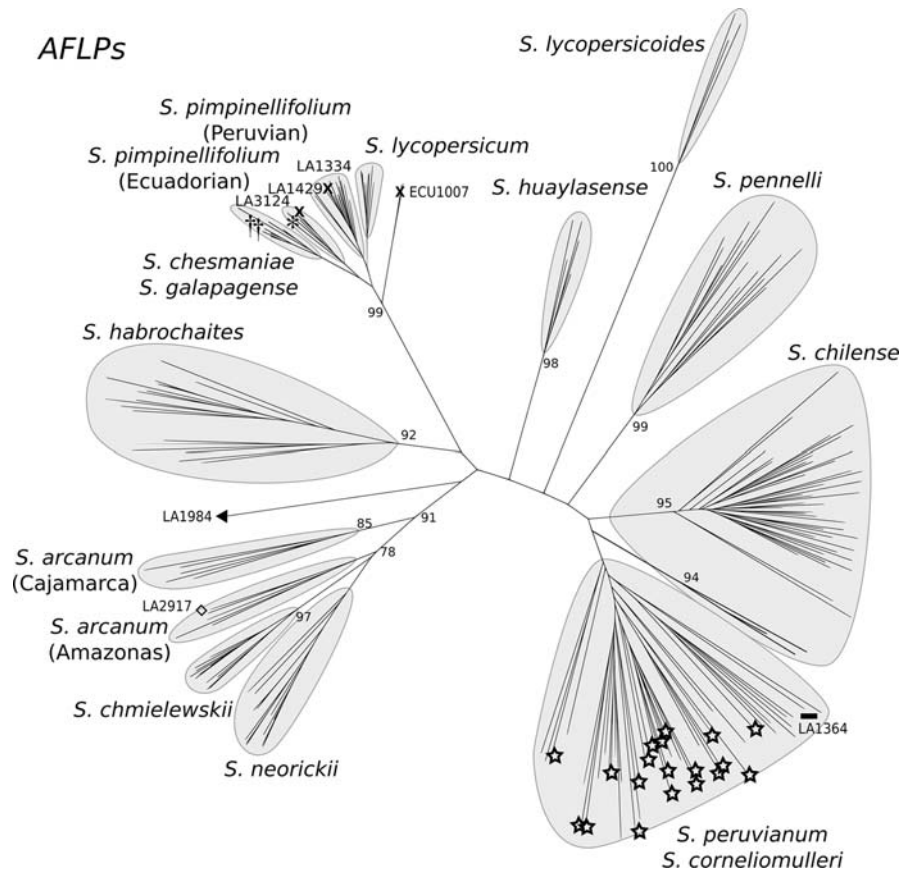
The accessions studied in the present work were either collected by the authors on different collection expeditions or obtained from genebanks. We have reclassified some of them, especially those belonging to *S. peruvianum* s.l., into new species (*S. peruvianum* s.str., *S. corneliomulleri*, *S. arcanum* and *S. huaylasense*). This reclassification was carried out by using the passport data, particularly the *in situ* characterization data and geographic origins. The coherence of this reclassification was tested by employing a molecular approximation.

Not many mismatches were detected between the reclassified data and the molecular AFLP and sequence data; the few misclassifications found are stated in Table 1. We used the molecular data provided by both the AFLP and the DNA sequences to build three trees: one with the accessions by using the AFLP data (Fig. 1) and two with the sequences CT179 and CT66, respectively (Figs. 2, 3). In these figures, putative misclassifications are designated with a symbol that corresponds to the species and the name of the accession.

The AFLP data provided the most straightforward clustering, producing an accession-based tree, and the different plant materials were neatly grouped into clusters that, in most cases, corresponded to the previous classifications. The sequence data provided haplotype phylogenies, not accession trees, which resulted in some differences from the AFLP tree. Despite this methodological difference, the AFLP and the CT179 datasets generated relatively congruent topologies, whereas CT66 differed in more details. We also carried out two extra, Bayesian-based clustering analyses: one joining the data from both genes, which resulted in a tree (data not shown) resembling that for CT179 alone, although with stronger posterior probability support, and another with all three datasets. In this latter case, the resulting topology (data not shown) resembled that from the AFLP data.

By using the AFLP data, it was possible to determine which *S. peruvianum* s.l. accessions belong to *S. huaylasense* because nearly all were in a single

Fig. 1 Accession-based tree based on Dice distances calculated from AFLP data and the Neighbor-joining clustering algorithm. Only when an accession did not cluster within the group defined by the rest of the accessions of its species is its name and specific symbol written in the tree. Also, to clarify relationships within the *S. peruvianum* cluster, all the *S. corneliomulleri* accessions have been labeled with their symbol. The symbols used are: ▲ *S. arcanum* (Amazonas), ◀ *S. arcanum* (Cajamarca), * *S. cheesmaniae*, ○ *S. chilense*, □ *S. chmielewskii*, ☆ *S. corneliomulleri*, † *S. galapagense*, ◇ *S. habrochaites*, = *S. huaylasense*, x *S. lycopersicum*, ◇ *S. neorickii*, ◆ *S. pennellii*, ● *S. peruvianum* s.str., and + *S. pimpinellifolium*. Numbers are percentage of bootstrap support



cluster supported by a 98% bootstrap value. A similar result was also obtained with the CT66 sequences, but not with those from CT179. In the CT179 tree, the *S. huaylasense* haplotypes were partially mingled with those from *S. arcanum*. The accession LA1364, also classified as *S. huaylasense* by Peralta et al. (2005), clustered at a basal position with *S. peruvianum* s.str. in the AFLP tree. LA1364 is heterozygous for both CT179 and CT66; in both cases one haplotype was clustered near *S. peruvianum* s.str. and the other near *S. arcanum*, so it might be a hybrid.

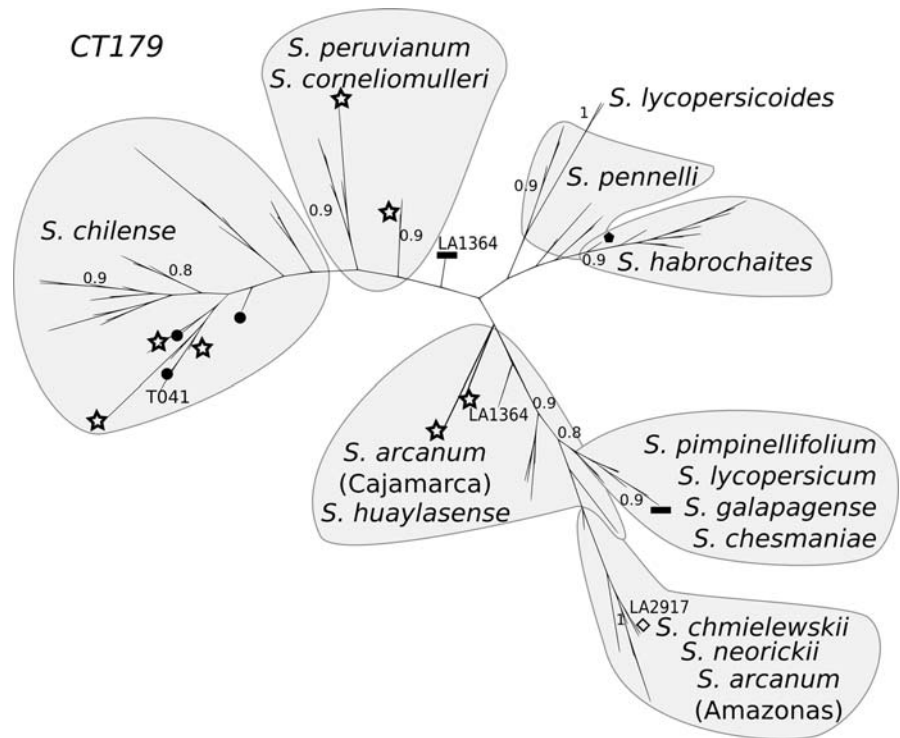
Other *S. peruvianum* s.l. accessions were reclassified as *S. arcanum*. They were grouped in two clusters with accessions of *S. arcanum* from the TGRC found in very well supported nodes and, in some cases, even sharing identical sequences. Also, we regarded LA2917 as *S. arcanum*, although it was provided to us as a *S. neorickii*. It was grouped with PE22 and LA1626, two *S. arcanum* accessions from the Ancash Department, and was clearly separated from the rest of *S. neorickii* (Fig. 1).

Some misclassification was also found in the *Lycopersicon* group, according to the molecular data presented. LA3124 was classified as *S. cheesmaniae* according to the passport data, but it grouped with Ecuadorian *S. pimpinellifolium* in the AFLP tree. LA1429 and LA1334 were previously considered *S. lycopersicum* but they clustered with Ecuadorian and Peruvian *S. pimpinellifolium*, respectively.

Intraspecific structure

The accessions clustered together in groups that corresponded to the different species. In some cases though there were single species divided in two groups or two species were merged together. *S. arcanum* accessions did not form a monophyletic cluster in any case. In the AFLP tree, they were distributed between two clusters that corresponded clearly to different geographic regions: (1) Amazonas and the eastern Ancash border and (2) Cajamarca and La Libertad (Fig. 4). Both groups were close to the

Fig. 2 CT179 sequence-based phylogram based on Bayesian methods. The naming convention and symbols are identical to the ones used in Fig. 1. Numbers are node posterior probabilities



clusters formed by the *S. neorickii* and *S. chmielewskii* accessions. This close relationship is even more striking in the sequence datasets. In the CT66 tree, *S. neorickii* and *S. arcanum* from Amazonas are clustered together; *S. chmielewskii* and *S. arcanum* from Ancash appeared close to them; and *S. arcanum* accessions from Cajamarca were grouped with some *S. peruvianum* s.str. In contrast, in the CT179 tree, *S. chmielewskii*, *S. neorickii* and *S. arcanum* from Ancash appeared grouped together, and *S. arcanum* from Amazonas and Cajamarca-La Libertad and *S. huaylasense* were located in a basal position.

The main natural habitats of the *Lycopersicon* group species are located to the north of those of those for *S. arcanum* (Fig. 4). The accessions from this group clustered together in all analyses. Four phenetic clusters were defined in the AFLP-based tree: *S. cheesmaniae*–*S. galapagense*, *S. lycopersicum*, and *S. pimpinellifolium* from Peru and Ecuador. Together, these groups clustered under a 99% bootstrap-supported node. The sequences obtained from these species showed little variability: only 3 and 4 haplotypes were found in CT66 and CT179, respectively.

Almost all accessions from *S. chilense* also clustered together under a 95% bootstrap-supported

node. Interestingly, some of the *S. chilense*-classified accessions were grouped in a cluster located as close to *S. chilense* as to *S. peruvianum*. All of these accessions originated in two valleys in Peru located where the Arequipa, Ayacucho and Ica regions meet.

Solanum corneliomulleri, one of the newly proposed species, inhabits central and southern Peru, sharing its habitat with *S. peruvianum* s.str. (Peralta et al. 2008, p. 37). The accessions provided by the TGRC to represent these species formed no clear clusters in analyses carried out with any of the three datasets. *Solanum corneliomulleri* was mixed in with the *S. peruvianum* s.str. cluster in the AFLP tree. The DNA sequences of *S. corneliomulleri* sequences were also clustered mainly with *S. peruvianum* s.str., although, some of them appeared in other regions of the trees. So to simplify our explanation of these results, we will not distinguish between *S. corneliomulleri* and *S. peruvianum* s.str., and will refer to both of them as *S. peruvianum*.

Although the *S. peruvianum* accessions were neatly clustered in the AFLP tree, their sequences were not monophyletic, as they formed two clusters in the CT179 tree and three in the CT66 tree. But this fact does not imply that there were two or three sets

Fig. 3 CT66 sequence-based phylogram based on Bayesian methods. The naming convention and the symbols are identical to the ones used in Fig. 1. Numbers are node posterior probabilities

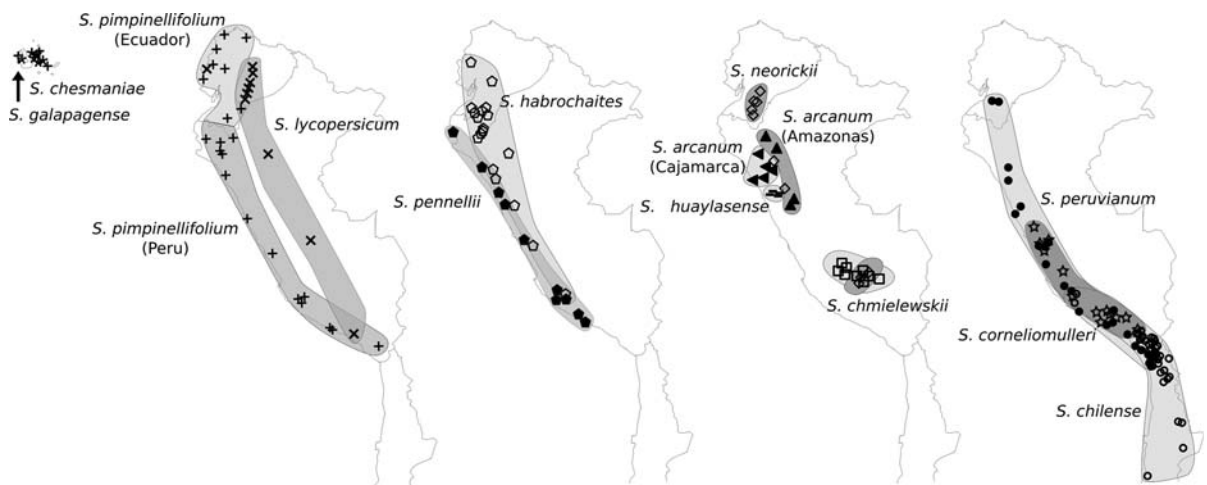
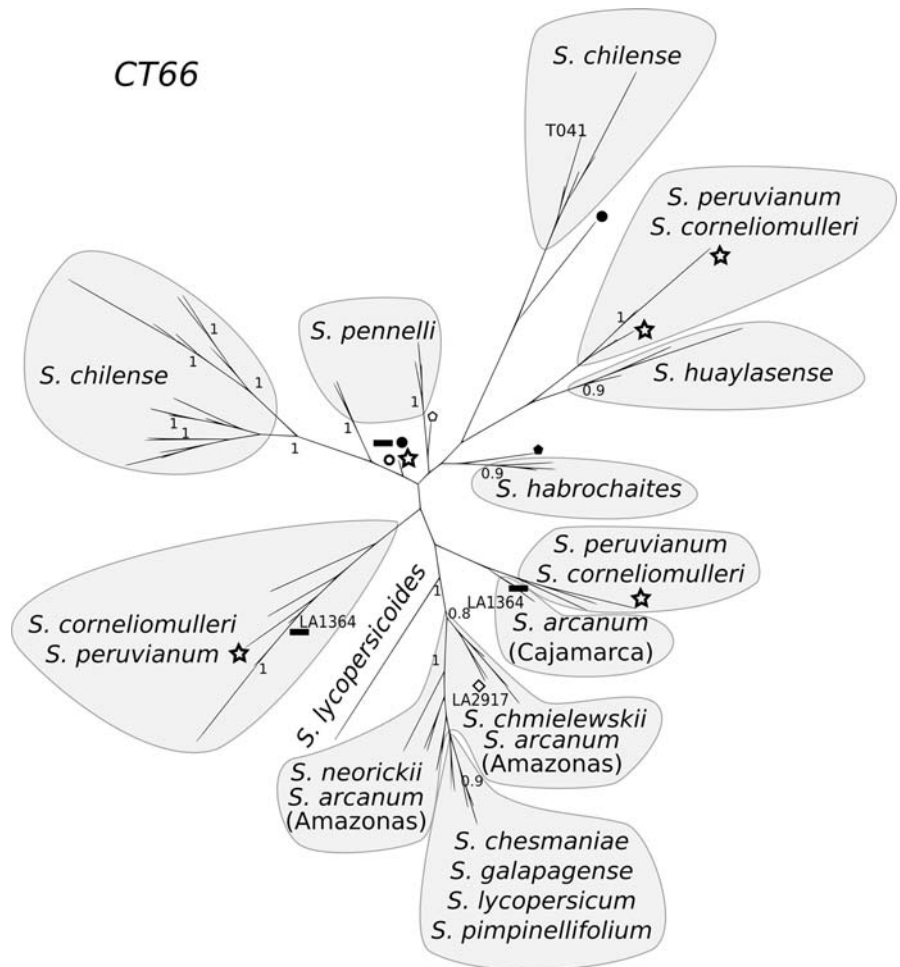


Fig. 4 Geographic locations of the accessions used. The symbols correspond to the species classification used in Fig. 1. The country borders represented are those of Ecuador, Peru and Chile

of different *S. peruvianum* accessions. There were numerous heterozygous plants, such as LA1292, T049 and T110, that held sequences from different clusters. Despite this fact, these species were grouped under well-supported phenetic clusters in the AFLP tree. This phenomenon was also observed in *S. chilense* and to a lesser degree in *S. pennellii*. In these cases, there were also heterozygous plants with alleles from both clusters, e.g., T085 and T091 (*S. chilense*) or LA1272 and LA2580 (*S. pennellii*).

The case of *S. habrochaites* was simpler, as both AFLP- and gene-based trees were in agreement; all accessions and haplotypes were grouped together in a unique cluster supported by a 92% bootstrap value. Also, all accessions of the outgroup *S. lycopersicoides* were perfectly grouped in all analyses.

Relationships among species

The AFLP tree and CT66 phylogram turned out to be rather uninformative regarding relationships among species, whereas the CT179 phylogram presented a more defined structure that suggested some intriguing phylogenetic hypotheses for future testing.

The data presented clearly confirmed the close relationship among *S. lycopersicum*, *S. cheesmaniae*, *S. galapagense* and *S. pimpinellifolium*. All of these species clustered together but could be distinguished using AFLPs, yet, according to the sequence data, they were almost indistinguishable. Another clearly supported group is formed by *S. arcanum*, *S. chmielewskii* and *S. neorickii*, all of whose accessions were clustered under a node supported by a 91% bootstrap value in the AFLP tree. They were also grouped together in the CT66 tree, as well as in the CT179 one, although, in this latter case, the *S. huaylasense* sequences were also clustered with *S. arcanum*. Regarding the relationship among *S. chmielewskii*, *S. neorickii* and *S. arcanum*, the AFLP data supported a clear geographic division of *S. arcanum*, although this is subtler in the sequence datasets. Neither of the sequence markers considered alone were capable of completely distinguishing this species from *S. chmielewskii* and *S. neorickii*. In both sequence-based phylograms, *S. chmielewskii* was clustered with *S. arcanum* from Ancash. In the case of CT179, even *S. neorickii* appeared close to the previously described cluster, while in the CT66 tree it is grouped with *S. arcanum* from the Amazonas

region. Another species closely related to this complex is *S. huaylasense*, which clustered with *S. arcanum* from Cajamarca in the CT179 tree. In the phylogram that combined the data from both markers, *S. huaylasense* was clearly segregated in from the rest of the group in a position basal to *S. arcanum* (data not shown).

Interestingly, in both sequence-based phylograms, the Lycopersicon group evolved from the Arcanum group. Moreover, this relationship is highly significant because it is supported by high posterior probabilities, 1 and 0.9 in the cases of CT66 and CT179, respectively. In the AFLP data, although both groups are close their relationship is not supported by a significant bootstrap value. This highly significant relationship between the two groups was due to the striking sequence similarity between them. For instance, the sequences found in the *S. arcanum* accession ECU0783 for CT66 and CT179 differed only by one and three nucleotides, respectively, from the sequences found in the Lycopersicon group. More support for this close relatedness comes from the tree combining all three datasets. In that case, these two groups were joined under a node supported by a 1 posterior probability (data not shown).

DNA sequences were incapable of distinguishing *S. galapagense* from *S. cheesmaniae*, and in the AFLP tree, both *S. galapagense* accessions also clustered with *S. cheesmaniae*. The study of more accessions of this species might improve the resolution inside this group. Two other species that appeared close both in the AFLP tree and in the CT179 phylogram were *S. peruvianum* and *S. chilense*. In the CT179 phylogram, *S. peruvianum* is basal to *S. chilense*, although statistical support for this relationship was not high. In the CT66 phylogram, the distribution of the sequences from these species was even more complex. In this case, some of the *S. peruvianum* sequences even clustered with, or close to, *S. huaylasense*, a relationship unsupported by any other data. Only when all three datasets were combined was strong statistical support obtained. In the combined tree, the node that joined *S. peruvianum* with *S. chilense* was supported by a 0.84 posterior probability.

The closest species to the outgroup, *S. lycopersicoides*, in the CT179 tree were *S. pennellii* and *S. habrochaites*. In addition, the CT179 and the combined phylograms also suggested that *S. pennellii* was

basal to *S. habrochaites*, but that result is neither supported by AFLP markers nor by CT66 sequences.

Discussion

Numerous relationships, groups, complexes, series and subgenera have been proposed to describe the structure of section *Lycopersicon*, which is likely a consequence of the complex scenario portrayed by the data gathered in this study. Accurate classification and description of species relationships within the section is not easy (Peralta et al. 2008, p. 53). Its species are evidently relatively young. According to Nesbitt and Tankley (2002), this section radiated around 7 million years ago, and, therefore, share close and blurred phylogenetic relationships. In addition, their intraspecific variability is, in some cases, quite important. All these factors have hampered past phylogenetic efforts.

Specific classification

Due to the extensive intraspecific variability used, a clear view of species-level classification in section *Lycopersicon* emerges. Some of recently proposed species are clearly supported by our data, whereas we recommend the reconsideration of the status of others. Here, we will focus our discussion especially on the species segregated from *S. peruvianum* s.l.

Solanum huaylasense accessions form a compact group clearly segregated from the rest of the species. This species is found in a narrowly circumscribed region, around Callejón de Huaylas along the Río Santa in the Department of Ancash and in the adjacent Río Fortaleza drainage (Rick 1986; Peralta et al. 2008, p. 35). The differentiation we observed could be a result of geographical isolation and ecological specialization.

In contrast to this monophyly, *S. arcanum* presents a more complex structure, as it formed two clusters. A clear geographical differentiation was noticed in the AFLP and in the sequence-based trees. These clusters corresponded to the accessions from the Amazonas and Ancash Departments, on one hand, and Cajamarca and La Libertad Departments, on the other. These two clusters conform, in part, to a classification proposed by Rick (1986). Rick studied crossability relations within *S. peruvianum*

populations from the Marañón region. He found three isolated groups: Marañón, Chotano and Chamaya. Our split is consistent with his proposal, but, unfortunately, none of the accessions that corresponded to the Chamaya group was available for the present study.

Rick decided not to name new species based on the groups he found in this complex because they were not easy to distinguish morphologically (Rick 1986). In fact, Peralta et al. (2008, p. 116) studied the morphology of these plants and proposed several assemblages: Marañón, Humifusum, Chotano and Lomas. Despite their morphological robustness, these latter groups are not supported by our genetic clustering, by the crossing relations shown by Rick, nor by geographic distribution. In summary, *S. arcanum* can be considered a complex by itself that might be composed of several cryptic species.

Another proposed split from *S. peruvianum* s.l. is *S. corneliomulleri*. Unfortunately, our results did not support the definition of *S. corneliomulleri* as a separate species. Its representative accessions were completely intermingled with those of *S. peruvianum* s.str. Städler et al. (2008) was of a similar opinion, justifying the treatment of all their samples as *S. peruvianum*, noting a lack of either molecular data or crossing results that would validate the proposed split of *S. corneliomulleri* from *S. peruvianum* s.str. In fact, Rick and Lamm (1955) and Rick (1963) showed that *Lycopersicon glandulosum* Müller (now classified as *S. corneliomulleri* by the TGRC) could cross with *S. peruvianum* and proposed that not to recognize it as a distinct species. Moreover, Rick (1963) pointed out that if *L. glandulosum* were recognized as distinct, then at least five other species should be defined.

If we were to consider only the sequence data, we might be tempted to propose new groups segregated from *S. peruvianum* and *S. chilense*, but we would be misled by incomplete data. Several *S. peruvianum* and *S. chilense* clusters are found in the CT179 tree and even more in the CT66 tree. The scenario portrayed by the AFLP data is quite different; according to these data, *S. chilense* and *S. peruvianum* s.str. form clear phenetic groups. This sequence structure does not imply a species substructure; numerous plants are found to bear haplotypes from the different clusters. This conflict could point to a large ancestral variability or, more probably, to the

existence of an ancestral structure that has recently evolved into the species that we now recognize. This observation complements the excess of low-frequency variants found by Arunyawat (2007) in *S. chilense* and, especially, in *S. peruvianum*. He also suggested that both species might have been influenced by the subdivision of their ancestral species. According to the AFLP tree, the *S. chilense*-classified accessions from the valley located between the Ica and the Arequipa regions could also form a new species closely related to *S. peruvianum* and *S. chilense*. But we believe that a more detailed study with more accessions from that region should be carried out before making such a conclusion.

Furthermore, a clear split of *S. pimpinellifolium* into two clusters, corresponding to the Peruvian and Ecuadorian accessions, respectively, was noticed in the AFLP tree. Interestingly, similar results were obtained recently in a *S. pimpinellifolium* study using microsatellite markers (Zuriaga et al. 2008). In that study, the authors correlated the split between these groups with sharp differences between Peruvian and Ecuadorian coastal climates. This split, however, is not supported by the sequence phylograms that cluster together *S. pimpinellifolium*, *S. lycopersicum*, *S. cheesmaniae* and *S. galapagense*. This lack of sequence diversity reflects the limited variation found within this complex and is also consistent with the hypothesis that these are all very recently split species known to easily hybridize in nature (Zuriaga et al. 2008).

Solanum cheesmaniae and *S. galapagense* appeared quite close in all of our datasets. This could be explained by the low genetic variability present in these species. *Solanum galapagense* can be differentiated by leaf morphology and a few other characters. But these same morphological characters were also found at a lower frequency in *S. cheesmaniae*, but only rarely in *S. lycopersicum* and *S. pimpinellifolium* (Peralta et al. 2008, p. 147). *Solanum galapagense* could be a restricted local endemic. In fact, Rick (1971) proposed both *S. cheesmaniae* and *S. galapagense* as forms of a single species.

Phylogeny

Recently, Peralta et al. (2008, p. 13) proposed a classification within section *Lycopersicon* by suggesting the following groups: Arcanum group (*S. arcanum*,

S. chmielewskii and *S. neorickii*), Lycopersicon group (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* and *S. galapagense*), Eriopersicon group (*S. peruvianum* s.str., *S. corneliomulleri*, *S. huaylasense*, *S. habrochaites* and *S. chilense*) and Neolycopersicon group (*S. pennellii*). Our results support some of these groups while questioning others.

Regarding the Arcanum group, our results support the classification proposed by Peralta et al. (2008). However, the CT179 data also imply a close relationship between *S. arcanum* and *S. huaylasense*, so we suggest that this latter species should be included in the Arcanum group and not in Eriopersicon. As mentioned above, *S. arcanum* showed strong geographical differentiation. Our AFLP and sequence data suggest a close relationship among *S. chmielewskii*, *S. neorickii* and the *S. arcanum* Amazonas groups, whereas the Cajamarca group would occupy a basal position. Rick (1963) proposed a differentiation of northern populations of *S. peruvianum* s.l. according to adaptation to different environments, and that seems to be the case for these species as well.

Interestingly, the CT179 phylogram and the AFLP tree suggest that the Arcanum group is close to both *S. peruvianum* and to the Esculentum group. This result is consistent with the crossability studies carried out by Rick (1986) on some of the Arcanum group accessions. He stated that races of *S. peruvianum* s.l. found in Río Marañón drainage, in Amazonas Department, could cross (with limitations) with *S. peruvianum*, *S. chilense* and also with *S. lycopersicum*. Rick (1986) also proposed that *S. arcanum* could be the origin of the Lycopersicon group. This latter relationship is also suggested by the CT179 phylogram, where the Lycopersicon group seems to evolve from the Arcanum group.

Neither the AFLP nor the sequence data support the grouping of *S. peruvianum* and *S. chilense* with *S. habrochaites*. *Solanum peruvianum* and *S. chilense* always appeared together in the trees, and they could be considered to form a group, but *S. habrochaites* occupies a basal position close to the outgroup, *S. lycopersicoides*. This basal position is in fact shared by *S. habrochaites* and *S. pennellii*.

Moreover, some broad trends are observed in the CT179 phylogram and AFLP tree. They sort the species along two axes, from self-incompatible to self-compatible and from south to north. *Solanum*

peruvianum and *S. chilense*, both self-incompatible and also the most southern species, were most distant from the *Lycopersicon* group composed of self-compatible, northern species. Located geographically and phylogenetically between them is the Arcanum group formed by a mixture of self-compatible and self-incompatible species. *Solanum pennellii* and *S. habrochaites* are found all over Peru, and they are the closest relative to the *S. lycopersicoides* outgroup, suggesting that could be the most ancient species of this section.

Acknowledgements We are deeply grateful to our colleagues from UNPRG, UNAM, PNG, UNL, UNP and UNT; thanks for the joint collection expeditions organized throughout the years, we enjoyed the privilege of their friendship and technical assistance. We also want to recognize the invaluable task carried out by the Tomato Genetics Resource Center, which provided important reference accessions. This work wouldn't have been possible without the support of the Fundación Banco Bilbao Vizcaya Argentaria.

References

- Alvarez AE, Van De Wiel CCM, Smulders MJM, Vosman B (2001) Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. *Theor Appl Genet* 103:1283–1292. doi:[10.1007/s001220100662](https://doi.org/10.1007/s001220100662)
- Arunyawat U (2007) Population structure and speciation history of two closely related wild tomato species. Ph.D. dissertation, Ludwig-Maximilians-Universität München, Germany
- 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
- Blanca JM, Prohens J, Anderson GJ, Zuriaga E, Cañizares J, Nuez F (2007) AFLP and DNA sequence variation in an Andean domesticated, pepino (*Solanum muricatum*, Solanaceae): implications for evolution and domestication. *Am J Bot* 94:1219–1229. doi:[10.3732/ajb.94.7.1219](https://doi.org/10.3732/ajb.94.7.1219)
- Bohs L (2007) Phylogeny of the cyphomandra clade of the genus *Solanum* (Solanaceae) based on ITS sequence data. *Taxon* 56:1012–1026
- Bohs L, Olmstead RG (1997) Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndhF* sequences. *Syst Bot* 22:5–17. doi:[10.2307/2419674](https://doi.org/10.2307/2419674)
- Bonfield J (2004) Staden package, version 1.4. Available at <http://staden.sourceforge.net>. Accessed 28 July 2008
- Child A (1990) A synopsis of *Solanum* subgenus *Potatoe* (G. Don) (D'Arcy) (*Tuberarium* (Dun.) Bitter (s.l.)). *Feddes Repert* 101:209–235
- Clark AG (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol Biol Evol* 7:111–122
- Darwin SC, Knapp S, Peralta IE (2003) Taxonomy of tomatoes in the Galapagos Islands: native and introduced species of *Solanum* section *Lycopersicon* (Solanaceae). *Syst Biodivers* 1:29–53. doi:[10.1017/S1477200003001026](https://doi.org/10.1017/S1477200003001026)
- Esquinas-Alcázar J, Nuez F (2001) Situación taxonómica, domesticación y difusión del tomate. In: Nuez F (ed) *El cultivo del tomate*. Mundi-Prensa, Madrid, pp 13–42
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hodkinson TR, Renvoize SA, Chonghaile GN, Stapleton CMA, Chase MW (2000) A comparison of ITS nuclear rDNA sequence data and AFLP markers for phylogenetic studies in *Phyllostachys* (Bambusoideae, Poaceae). *J Plant Res* 113:259–269. doi:[10.1007/PL00013936](https://doi.org/10.1007/PL00013936)
- Knapp SL, Bohs MN, Spooner DM (2004) Solanaceae: a model for linking genomics and biodiversity. *Comp Funct Genomics* 5:285–291. doi:[10.1002/cfg.393](https://doi.org/10.1002/cfg.393)
- Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679. doi:[10.1093/bioinformatics/bti079](https://doi.org/10.1093/bioinformatics/bti079)
- Luckwill LC (1943) The genus *Lycopersicon*: an historical, biological, and taxonomical survey of the wild and cultivated tomatoes. PhD dissertation, Aberdeen University Studies 120:1–44
- Marshall JA, Knapp S, Davey MR, Power JB, Cocking EC, Bennett MD, Cox AV (2001) Molecular systematics of *Solanum* section *Lycopersicum* (*Lycopersicon*) using the nuclear ITS rDNA region. *Theor Appl Genet* 103:1216–1222. doi:[10.1007/s001220100671](https://doi.org/10.1007/s001220100671)
- McClellan PE, Hanson MR (1986) Mitochondrial DNA sequence divergence among *Lycopersicon* and related *Solanum* species. *Genetics* 112:649–667
- Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor Appl Genet* 80:437–448
- Miller P (1754) *The gardeners dictionary abridged*, 4th edn. John and James Rivington, London
- Müller CH (1940) The taxonomy and distribution of the genus *Lycopersicon*. *Natl Hortic Mag* 19:157–160
- Nesbitt TC, Tanksley SD (2002) Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162:365–379
- Nuez F, Prohens J, Blanca JM (2004) Relationships, origin, and diversity of Galápagos tomatoes: implications for the conservation of natural populations. *Am J Bot* 91:86–99. doi:[10.3732/ajb.91.1.86](https://doi.org/10.3732/ajb.91.1.86)
- Nuez F, Díez MJ, Valcárcel JV, Cebolla-Cornejo J, Pérez A, Soler S, Roselló S, Adalid A, Galiana L, Sifres A, Picó B, Blanca JM, Frutos R (2008) Genetic resources of *Lycopersicon* at the Institute for the Conservation and Improvement of the Agrodiversity. *Acta Hort* 789:293–297
- Olmstead RG, Palmer JD (1997) Implications for phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Syst Bot* 22:19–29. doi:[10.2307/2419675](https://doi.org/10.2307/2419675)
- Palmer JD, Zamir D (1982) Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. *Proc Natl Acad Sci USA* 79:5006–5010. doi:[10.1073/pnas.79.16.5006](https://doi.org/10.1073/pnas.79.16.5006)

- Peralta IE, Spooner DM (2001) Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). *Am J Bot* 88:1888–1902. doi:[10.2307/3558365](https://doi.org/10.2307/3558365)
- Peralta IE, Spooner DM (2005) Morphological characterization and relationships of wild tomatoes (*Solanum* L. section *Lycopersicon*). In: Keating RC, Hollowell VC, Croat TB (eds) A festschrift for William G. D'Arcy: the legacy of a taxonomist. *Monogr Syst Bot Mo Bot Gard* 104:227–257
- Peralta IE, Knapp S, Spooner DM (2005) New species of wild tomatoes (*Solanum* section *Lycopersicon*: Solanaceae) from northern Peru. *Syst Bot* 30(2):424–434. doi:[10.1600/0363644054223657](https://doi.org/10.1600/0363644054223657)
- Peralta IE, Spooner DM, Knapp S (2008) Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicon*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). *Syst Bot Monogr* :84
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818. doi:[10.1093/bioinformatics/14.9.817](https://doi.org/10.1093/bioinformatics/14.9.817)
- R Development Core Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. Accessed 28 July 2008
- Rick CM (1963) Barriers to interbreeding in *Lycopersicon peruvianum*. *Evolution Int J Org Evolution* 17:216–232. doi:[10.2307/2406467](https://doi.org/10.2307/2406467)
- Rick CM (1971) *Lycopersicon*. In: Wiggins IL, Porter DM (eds) *Flora of the Galapagos Islands*. Stanford University Press, Stanford, pp 468–471
- Rick CM (1979) Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. In: Hawkes JG, Lester RN, Skelding AD (eds) *The biology and taxonomy of the Solanaceae*. Academic Press, New York, pp 667–678
- Rick CM (1986) Reproductive isolation in the *Lycopersicon peruvianum* complex. In: D'Arcy WG (ed) *Solanaceae: biology and systematics*. Columbia University Press, New York, pp 477–495
- Rick CM, Lamm R (1955) Biosystematic studies on the status of *Lycopersicon chilense*. *Am J Bot* 42:663–675. doi:[10.2307/2485327](https://doi.org/10.2307/2485327)
- Rick CM, Kesicki E, Fobes JF, Holle M (1976) Genetic and biosystematic studies on two new sibling species of *Lycopersicon* from interandean Peru. *Theor Appl Genet* 47:55–68. doi:[10.1007/BF00281917](https://doi.org/10.1007/BF00281917)
- 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 Syst Evol* 132:279–298. doi:[10.1007/BF00982390](https://doi.org/10.1007/BF00982390)
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574. doi:[10.1093/bioinformatics/btg180](https://doi.org/10.1093/bioinformatics/btg180)
- Roselius K, Stephan W, Städler T (2005) The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics* 171:753–763. doi:[10.1534/genetics.105.043877](https://doi.org/10.1534/genetics.105.043877)
- Rozas J, Sanchez-Del Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497. doi:[10.1093/bioinformatics/btg359](https://doi.org/10.1093/bioinformatics/btg359)
- Sifres A, Picó B, Blanca J, De Frutos R, Nuez F (2007) Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event 1997–1998. *Genet Resour Crop Evol* 54:359–377. doi:[10.1007/s10722-005-5725-4](https://doi.org/10.1007/s10722-005-5725-4)
- Spooner DM, Anderson GJ, Jansen RK (1993) Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (Solanaceae). *Am J Bot* 80:676–688. doi:[10.2307/2445438](https://doi.org/10.2307/2445438)
- Spooner DM, Peralta IE, Knapp S (2005) Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes. *Taxon* 54:43–61. *Solanum* L. section *Lycopersicon* (Mill.) Wettst.
- Städler T, Roselius K, Stephan W (2005) Genealogical footprints of speciation processes in wild tomatoes: demography and evidence for historical gene flow. *Evolution Int J Org Evolution* 59:1265–1270
- Städler T, Arunyawat U, Stephan W (2008) Population genetics of speciation in two closely related wild tomatoes (*Solanum* section *Lycopersicon*). *Genetics* 178:339–350. doi:[10.1534/genetics.107.081810](https://doi.org/10.1534/genetics.107.081810)
- Taylor IB (1986) Biosystematics of the tomato. In: Atherton JG, Rudich J (eds) *The tomato crop: a scientific basis for improvement*. Chapman and Hall, London, pp 1–34
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680. doi:[10.1093/nar/22.22.4673](https://doi.org/10.1093/nar/22.22.4673)
- van Ee BW, Jelinski N, Berry PE, Hipp AL (2006) Phylogeny and biogeography of *Croton alabamensis* (Euphorbiaceae), a rare shrub from Texas and Alabama, using DNA sequence and AFLP data. *Mol Ecol* 15:2735–2751
- Warnock SJ (1991) Natural habitats of *Lycopersicon* species. *HortScience* 26:466–471
- Xu F, Sun M (2001) Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. *Mol Phylogenet Evol* 21:372–387. doi:[10.1006/mpev.2001.1016](https://doi.org/10.1006/mpev.2001.1016)
- Zuriaga E, Blanca JM, Cordero L, Sifres A, Blas-Cerdán WG, Morales R, Nuez F (2008) Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genet Resour Crop Evol* . doi:[10.1007/s10722-008-9340-z](https://doi.org/10.1007/s10722-008-9340-z)

Capítulo 2.- Genetic and bioclimatic variation in

Solanum pimpinellifolium.

Genetic Resources and Crop Evolution

Genetic and bioclimatic variation in *Solanum pimpinellifolium*

Elena Zuriaga · José M. Blanca · Laura Cordero · Alicia Sifres ·
William G. Blas-Cerdán · Rafael Morales · Fernando Nuez

Received: 15 January 2008 / Accepted: 21 April 2008 / Published online: 16 May 2008
© Springer Science+Business Media B.V. 2008

Abstract *Solanum pimpinellifolium*, due to its close relationship to *S. lycopersicum*, has been a genetic source for many commercially important tomato traits. It is a wild species found in the coastal areas of Peru and Ecuador. In this study, the genetic variation of *S. pimpinellifolium* was studied using the diversity found in 10 microsatellites in 248 plants spread throughout its entire distribution area, including Ecuador, which has been underrepresented in previous studies. Peruvian and Ecuadorian accessions are genetically quite differentiated. A possible cause of these differences could be the non-uniform nature

of the coastal Ecuadorian and Peruvian climates, seeing as an important correlation between genetic differentiation and climate has been found. In addition, Ecuadorian and south Peruvian accessions have a lower genetic diversity and a higher homozygosity due to their higher autogamy, lower population size, and possible colonization bottlenecks. The Galápagos Islands population is an extreme case, with no diversity, likely caused by a recent colonization from the northern continental Ecuadorian region where genetically identical plants have been found.

Keywords Andean · Climate · Colonization · Genetic structure · Microsatellite · *Solanum pimpinellifolium*

Elena Zuriaga and José M. Blanca contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10722-008-9340-z) contains supplementary material, which is available to authorized users.

E. Zuriaga · J. M. Blanca · L. Cordero ·
A. Sifres · F. Nuez (✉)
Instituto de Conservación y Mejora de la Agrodiversidad
Valenciana (COMAV), Universidad Politécnica de
Valencia, Camino de Vera s/n 46022, Valencia, Spain
e-mail: fnuez@btc.upv.es

W. G. Blas-Cerdán
Universidad Nacional de Trujillo, Independencia
número 389, Trujillo, Peru

R. Morales
Facultad de Ciencias Agrícolas, Universidad Nacional de
Loja, casilla número 782, Loja, Ecuador

Introduction

Solanum pimpinellifolium L. is a small perennial bush common in the coastal areas of Peru and Ecuador. In central and southern Peru, *S. pimpinellifolium* is restricted to cultivated fields, roadsides and dumping grounds, and its distribution is sparse; but in Ecuador and northern Peru, it is found in dense populations located in undisturbed areas (Rick et al. 1977; Caicedo and Schall 2004). The former name of this species, *Lycopersicon pimpinellifolium* (L.) Mill., has been changed because recent studies, based on molecular data, have classified the genus *Lycopersicon* (Miller

1754; Luckwill 1943) as the genus *Solanum* L. sect. *Lycopersicon* (Mill.) Wettst. (Bohs and Olmstead 1997; Marshall et al. 2001; Peralta and Spooner 2000, 2001).

Although *S. pimpinellifolium* is seldom harvested for human consumption, it is widely used as a source of germplasm by tomato breeders (Rick and Chetelat 1995) because it can be easily hybridized with *S. lycopersicum* L. (formerly known as *L. esculentum* Mill.) (Rick et al. 1979) and because it has many economically important traits for commercial tomato varieties (Juvik et al. 1982; Mieslerova et al. 2000). The relationship between the cultivated and wild species is quite close: both are self-compatible and red-fruited. Both species differ in fruit size and leaf shape, but there are plants with intermediate traits that are difficult to classify (Rick and Fobes 1975).

The genetic variation structure of Peruvian *S. pimpinellifolium* has been studied with allozymes (Rick et al. 1977, 1978), nuclear DNA gene sequences (Caicedo and Schall 2004), and microsatellites (Sifres et al. 2007). Although the Peruvian region was analyzed extensively in these previous studies, Ecuador received little attention, mainly because there were not many accessions from Ecuador in the germplasm banks. These studies have shown that although *S. pimpinellifolium* is self-compatible, a wide variation in the rate of autogamy exists among populations, ranging from 100% to 16% (Rick et al. 1978). Differences in diversity have been reported as being correlated with this variation in the levels of outcrossing, with the highest rates in the northern Peruvian departments of Piura and Lambayeque (Rick et al. 1977).

The natural range of *S. pimpinellifolium* encompasses environments as different as the northern Ecuadorian tropical forest and the coastal Peruvian desert. The geographic distribution of different species has been shown to be greatly affected by climate (Woodward 1987), and at the same time genetic and climate correlations at both micro-geographic and regional scales have been shown in species as diverse as *Hordeum spontaneum* C. Koch (Owuor et al. 2003), *Triticum dicoccoides* (Körn. ex Asch. et Graebn.) Schweinf. (Li et al. 2000), and *Pinus edulis* Engelm. (Mitton and Duran 2004). Such environmental and genetic correlation has usually been interpreted as evidence of adaptive genetic differentiation. As *S. pimpinellifolium* thrives in such different climates, it could be a good model for

studying in some detail the relationship between genetic variation distribution and climate and ecosystem types.

The Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), with the collaboration of the Universidad Nacional de Loja (UNL) (Ecuador), the Universidad Nacional de Piura (UNP) (Peru), the Universidad Nacional Pedro Ruiz Gallo (UNPRG) (Lambayeque, Peru), and the Parque Nacional de Galápagos (PNG) (Ecuador), has organized several collection expeditions in the last few decades, covering the whole range of distribution of *S. pimpinellifolium*, from southern Peru to northern Ecuador, including the Galápagos Islands (Darwin et al. 2003), thus correcting Ecuador's underrepresentation. This wide germplasm collection has allowed the study of materials from throughout the region covered by this species. In order to compare the results with the previous studies on this subject done by other authors, accessions from international germplasm banks have been included. These reference materials also include some *S. lycopersicum* accessions.

Ten tomato-derived microsatellites (Smulders et al. 1997), previously reported to be polymorphic in *S. pimpinellifolium* (Sifres et al. 2007), were chosen as molecular markers representative of the genome variation. Microsatellites were used because they are codominant, highly polymorphic, and easily reproducible. This kind of marker is widely used in studies on variability in plants (Harter et al. 2004; Monrchen et al. 1996; Powell et al. 1996).

In this study, we examined the genetic variation of *S. pimpinellifolium* throughout its entire range of distribution, including samples collected by the COMAV Institute on its collection trips to previously under-collected regions as well as materials from other germplasm banks used as references. We also show the relationships between the genetic variation distribution of *S. pimpinellifolium* and the climate and ecosystems of the different areas that it inhabits.

Materials and methods

Plant materials

Solanum pimpinellifolium was collected during several collecting expeditions to Peru and Ecuador

organized by members of the COMAV Institute in collaboration with Peruvian and Ecuadorian Universities. These trips began in 1980 and have continued regularly up to the present day (Nuez and Cuartero 1984; Cuartero et al. 1984; Nuez et al. 1993, 1999; Nuez and Picó 1999). Passport data for each accession was taken as recommended by the International Plant Genetic Resources Institute (IPGRI 1996). The longitude, latitude and altitude of each accession were determined using a global positioning system (GPS; Magellan Meridian XL, San Dimas, California).

This study includes a total of 247 accessions. About 205 accessions were gathered by the COMAV Institute, the UNL, the UNPRG, and the PNG center (see Supplemental Data accompanying online version of this article), in addition to some materials from other institutes used as references: 35 from the Tomato Genetics Resource Center (TGRC), 3 from the World Vegetable Center (AVRDC), and 5 from the United States Department of Agriculture (USDA). These materials are divided into 213 *S. pimpinellifolium* and 33 *S. lycopersicum* plants, including 9 Andean, 17 non-Andean, and 7 *S. lycopersicum* var. *cerasiforme* (Dunal) Spooner, G.J. Anderson et R.K. Jansen accessions. Included in these materials are 67 Peruvian and 12 Galapagan accessions previously analyzed elsewhere (Sifres et al. 2007; Nuez et al. 2004). The UNL and UNPRG universities of Loja and Piura hold seed samples of the accessions. Longitude, latitude and accession codes are available for all materials used in the electronic supplementary material table.

DNA extraction and microsatellite analysis

Genomic DNA was isolated from young leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA) following the protocol of the manufacturer. All plants were screened for variation at ten polymorphic microsatellite loci (LE20592, LE21085, LELEUZIP, LEATPACAa, LEATPACAb, LEGAST1, LPHFS24, LEMDDNb, LEILV1B and LEGTOM5) (Smulders et al. 1997). Allele lengths were determined using an ABI Prism 310 Genetic Analyzer with the aid of Genescan and Genotyper software (Applied Biosystem). Only one plant from each accession was molecularly analyzed in order to avoid the possible effects caused by the

use of multiple close relatives on genetic diversity estimates.

Molecular data analysis

In order to classify the accessions into genetic clusters (Harter et al. 2004) several factorial correspondence analyses (FCA) were carried out using the Genetix program (Belkhir et al. 1996). Several broad clusters were defined using the data from an FCA carried out on all accessions. New FCAs were done for each cluster, and this information was used to subdivide the former groups. A genetic cluster was defined as a group that is found to be consistently isolated from the rest in the FCA. These clusters were compared with the ones obtained using the Bayesian model-based clustering proposed by Pritchard et al. (2000), implemented in the software STRUCTURE 2.1. (<http://pritch.bsd.uchicago.edu/software.html>). We used the basic admixture model with unlinked loci and correlated allele frequencies among groups, with the assumed number of populations (K) varying from 2 to 20, 5 replicate runs per K value, a burning period of 10^5 , and a post-burning simulation length of 1.5×10^6 . No a priori population information was used. The run showing the highest posterior probability of data was considered for each K value.

Both the FCA and the STRUCTURE-based clusterings showed a clear geographical correlation, and so the clustering of the accessions used in the rest of the work was based on geographic regions.

The genetic diversity of each geographic group was studied using the following standard genetic parameters calculated with the Genetix software: observed alleles (A), percentage of polymorphism (P), observed heterozygosity (H_o), and Nei's genetic diversity (D) (Nei 1977). All these parameters were calculated as the mean obtained for 100 bootstrapped samples for every geographic group.

These subsamples were prepared taking random individuals from all geographic groups and all of them had the same number of individuals (4). The variation inside the geographic groups was further investigated by an analysis of molecular variance (AMOVA) using ARLEQUIN 3.0 (Excoffier et al. 2005). The significance of the partitioning of genetic variance among geographic groups was also tested.

In order to visualize the genetic relationship among the geographic groups, a matrix of genetic Dc distances

(Cavalli-Sforza and Edwards 1967) was calculated, and a dendrogram was constructed by means of the Neighbor-Joining method using the Populations program, version 1.2.28 (Langella 2002). Bootstrap support for the phenogram nodes was constructed by bootstrapping the original data 1000 times with replacement over all of the loci. A complementary view of the same distance matrix data was built using a Principal Coordinates Analysis (PCoA) (Gower 1966) performed with the Ginkgo software (Bouxin 2005). In this view, geographic groups were represented by circles with radii proportional to diversity. The upper third of the lowest distances were represented by lines connecting the geographic groups.

Analysis of climate data

For the geographic analysis, the climate data set included in the DIVA-GIS software (Hijmans et al. 2001) was used, and the vegetation land cover map data was taken from the vegetation map of South America (Eva et al. 2002). Climate parameters were obtained for the geographical location of each accession and mean and variance values throughout the months were calculated. A principal component analysis (PCA) was done with the previously obtained climate matrix (Qian et al. 2005) using the Ginkgo software. We considered climate as being

defined by the first two principal components. These two values were computed for each accession and both were used to calculate a Euclidean distance matrix among the accessions.

The correlation between Dc genetic, geographic, and climate distance matrices were calculated using the ZT software (Bonnet and Van de Peer 2002).

Results

Genetic variation

In the studied accessions, between 3 and 16 alleles with a mean number of 8.7 were found in the 10 microsatellites analyzed. The *S. pimpinellifolium* accessions were grouped by their genetic similarity using several FCAs (Fig. 1) and the program STRUCTURE (Fig. 2).

An FCA multivariate analysis was carried out with the genetic data (Fig. 1). The first two dimensions accounted for 12.6% of the variation. In this representation, there are three clear genetic clusters, which correspond to Ecuadorian and Peruvian *S. pimpinellifolium*, and *S. lycopersicum*, respectively. Among these three groups there is a continuous range of variation and no clear limits. Inside Ecuador, the northern accessions tend to appear close to the

Fig. 1 FCA multivariate analysis done with the microsatellite data with the accessions represented according to their geographic origin

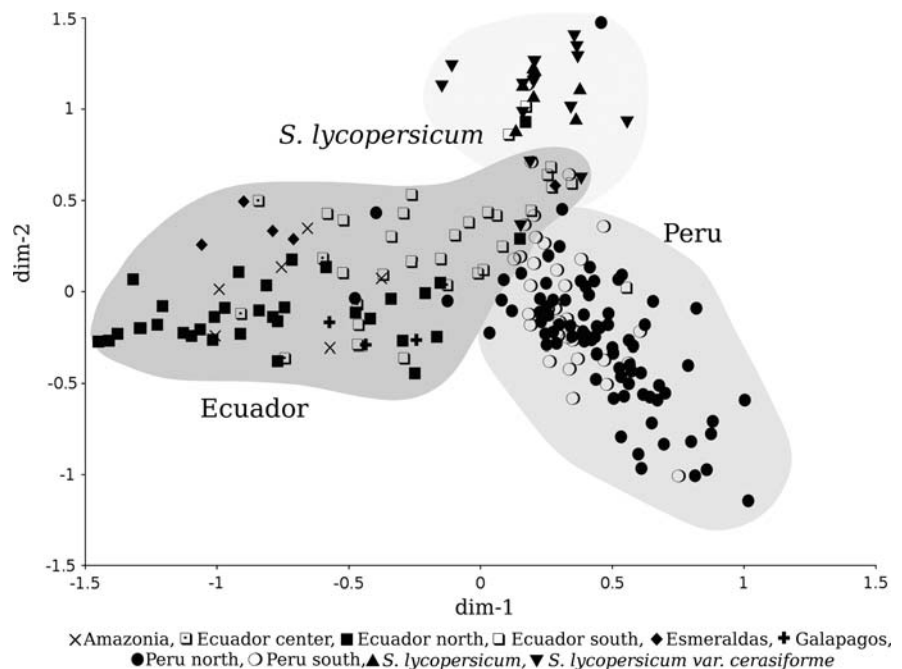
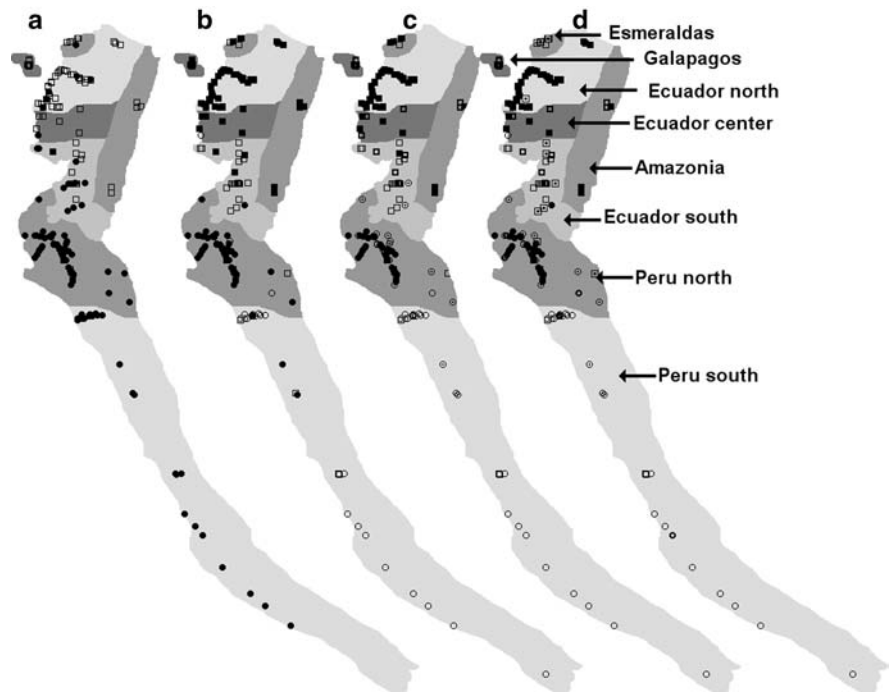


Fig. 2 Genetic divisions created by STRUCTURE. Accessions are represented in the Peru and Ecuador maps according to their point of collection. From **a** to **d**, different symbols represent different genetic clusters proposed by STRUCTURE in several runs carried out with the number of groups (K) equal to 2, 4, 6, and 8 (**a**, **b**, **c**, and **d** respectively). The regions painted in the map show the geographic clustering used in the rest of the study



bottom-left corner while the southern ones are closer to *S. lycopersicum*. Esmeraldas accessions form a group isolated from the rest when the third dimension is taken into consideration (data not shown).

An alternative method for dividing the genetic variability implemented in the STRUCTURE software has been widely used recently. We wanted to check if the geographic division obtained with the FCA analysis was also observed when this other method was used. STRUCTURE divides individuals into genetic clusters assuming a Hardy–Weinberg equilibrium within clusters, and, in some cases, is capable of calculating the most probable number of clusters (K). However, the authors warn that plant populations are usually not in equilibrium and thus K could be overestimated (Pritchard et al. 2000). This effect should be especially relevant in species with a high degree of autogamy, such as *S. pimpinellifolium*. For our genetic data, the Bayesian posterior probability reached no maximum and K could not be estimated, as occurred with a similar trend observed by other authors with sorghum data (Barnaud et al. 2007). Nonetheless, the STRUCTURE clusters obtained are also highly correlated with a geography-based genetic clustering. When STRUCTURE was asked to create 2 genetic clusters, a clear division between Ecuadorian and Peruvian accessions was

found (Fig. 2a). If K is raised up to 6, more splits in the north to south direction appear (Fig. 2b–c). No further clear geographic subdivisions appeared for K between 7 and 20: instead, genetic clusters with few plants appeared in the extra groups and the overall geographic structure was maintained. The clusters obtained for K equal to 8 are shown as an example (Fig. 2d).

Both the FCA and the STRUCTURE-based clustering showed a clear geographic division, and therefore, in the end, a geographic subdivision of the genetic diversity was proposed. The accessions were grouped according to their geographic region of origin: two regions were defined in Peru (Peru south and Peru north) and six in Ecuador (Ecuador south, Ecuador center, Ecuador north, Esmeraldas, Amazonia, and Galápagos) (Fig. 2). All the Amazonian accessions were clustered together because of the clear geographic division caused by the Andean ridge. This geography-based *S. pimpinellifolium* division was the one tested in the rest of the study.

Geographic group characterization

Several diversity indexes were calculated (Table 1). Overall, Peruvian *S. pimpinellifolium* is more variable than Ecuadorian *S. pimpinellifolium* according

to the mean number of alleles per locus (A) (7.6 to 5.2), although this is not the case if the Nei diversity index (D) is taken into account (0.40 to 0.41). The most diverse region is Peru north, which has even more alleles per locus than all of Ecuador considered together. Nevertheless, inside Ecuador there exists a lot of genetic variation and there are geographic groups with high Nei's D, such as Ecuador north (0.30) and south (0.43), and also those with very low D, such as Galápagos (0). *S. lycopersicum* possesses a lower diversity than *S. pimpinellifolium*, although it is worth mentioning that the Andean *S. lycopersicum* accessions have more variability than the non-Andean ones. In fact, these Andean accessions even have a Nei's D similar to some *S. pimpinellifolium* regions.

Another indication of the Peruvian diversity is that Peru possesses 33 exclusive alleles that do not appear in any other region, while in Ecuador and *S. lycopersicum* there are just 7 and 1 unique alleles, respectively. Peru north not only has the greatest diversity, but also possesses the highest heterozygosity (0.12). Peru south is characterized by low heterozygosity (0.04) and moderate diversity (0.32).

Table 1 Genetic diversity found in the studied geographic regions

| Population | <i>n</i> | Ho | D | P | A |
|----------------------------|----------|-------------|------|-----|-----|
| <i>S. pimpinellifolium</i> | 213 | 0.07 (0.05) | 0.45 | 0.8 | 8.5 |
| Peru | 119 | 0.10 (0.08) | 0.40 | 0.8 | 7.6 |
| Peru north | 83 | 0.12 (0.11) | 0.40 | 0.7 | 6.8 |
| Peru south | 36 | 0.04 (0.40) | 0.32 | 0.7 | 4.1 |
| Ecuador | 94 | 0.03 (0.03) | 0.41 | 0.7 | 5.2 |
| Ecuador north | 42 | 0.05 (0.06) | 0.30 | 0.6 | 3.3 |
| Ecuador center | 7 | 0.03 (0.06) | 0.30 | 0.5 | 2.3 |
| Ecuador south | 26 | 0.03 (0.05) | 0.43 | 0.8 | 4.2 |
| Esmeraldas | 5 | 0.02 (0.06) | 0.24 | 0.6 | 1.7 |
| Amazonia | 6 | 0 (0) | 0.22 | 0.4 | 1.7 |
| Galápagos | 8 | 0 (0) | 0 | 0 | 1.0 |
| <i>S. lycopersicum</i> | 33 | 0.01 (0.02) | 0.14 | 0.5 | 2.3 |
| Andean | 9 | 0.04 (0.06) | 0.21 | 0.5 | 1.8 |
| non-Andean | 17 | 0 (0) | 0.05 | 0.2 | 1.3 |
| var. <i>cerasiforme</i> | 7 | 0 (0) | 0.12 | 0.3 | 1.4 |

Notes: Number of individuals (*n*), observed heterozygosity (Ho) and its variance (in brackets), Nei diversity (D), proportion of polymorphic loci in which the most frequent allele is under 0.95 frequency (P), and mean number of alleles per locus (A)

All the Ecuadorian regions have, besides a lower variability, a lower heterozygosity that ranges from 0 in the Galápagos to 0.05 in Ecuador north.

Regional differentiation

The differentiation among all regions was confirmed by a permutation test on the Weir and Cockerham F for every pair of groups. According to this test, the differences between all pairs but one were significant. The only non-significant difference appeared in the pair formed by *S. lycopersicum* Andean and *S. lycopersicum* var. *cerasiforme*. AMOVA conducted on the geographic groups showed that all variance components were highly significant ($P < 0.001$). 21.1% of the variation was due to the Peruvian-Ecuadorian differentiation, 15.7% to the subdivisions inside each country, and 63.2% was found within the groups.

A 0.12 Gst between Peru and Ecuador also showed the differentiation of these two regions, as did the 0.28 Gst between both of these and *S. lycopersicum*.

To study the relationships among the geographic groups, a neighbor-joining dendrogram based on Cavalli Dc distances was constructed (Fig. 3b). In this tree, the Peruvian and Ecuadorian geographic groups are split and show no clear relationships between them. Bootstrap values for this dendrogram are low, with just four nodes with a value above 70%. To represent this distance matrix in a more informative way, a Principal Coordinate Analysis (PCoA) was carried out, and one third of the lowest distances were plotted in the graphic representation (Fig. 3a). The first two dimensions of this analysis account for 76.0 % of the total variation. In this view, it is clear that the Peruvian and Ecuadorian regions, although quite different, are linked by the Ecuador south geographic group. The central and northern Ecuadorian regions form a cluster with Amazonia and Galápagos, whereas Esmeraldas remains equidistant to *S. lycopersicum* and *S. pimpinellifolium*.

Climatic analysis

Geography and climate may influence the genetic variation of *S. pimpinellifolium*, and we have addressed the relative importance of these factors. In Fig. 4, the most relevant variations in the climate of the regions are presented. In Table 2, the main

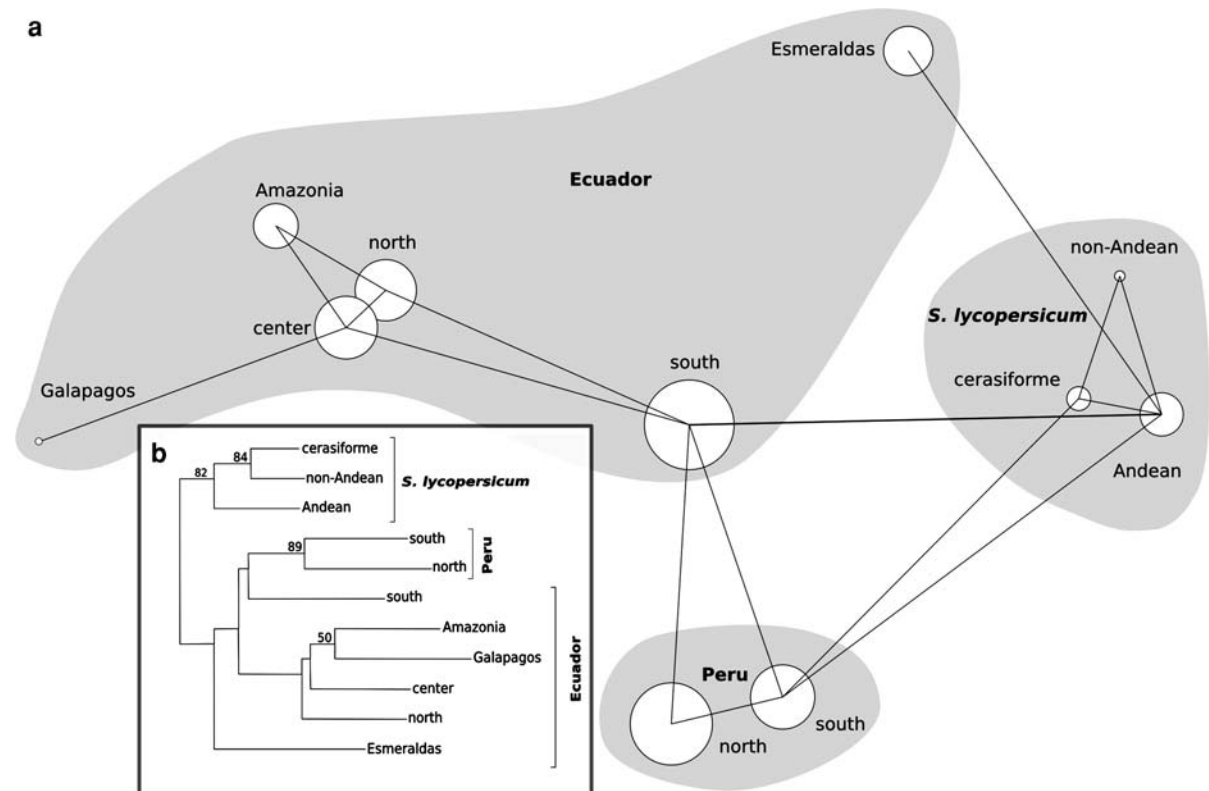


Fig. 3 (a) PCoA analysis of the Dc distances among the *S. pimpinellifolium* geographic regions and *S. lycopersicum*. Circle radii are proportional to diversity. Only a third of the

shortest distances are represented by lines; (b) NJ dendrogram based on the Dc distances—only bootstrap values above 50% have been printed

climatic parameters calculated for each of the geographic groups is shown. In every parameter measured there is a clear gradient in the north to south direction, precipitation ranges from less than 700 mm in coastal Peru to more than 2800 mm in some Ecuadorian regions (Fig. 4a), and temperature is also quite different between the colder Peru and the warmer Ecuador (Fig. 4c). In both parameters, temperature and rainfall, there is also a sharp difference in seasonality: whereas in Ecuador the climate is quite uniform throughout the year, in Peru there are substantial variations (Fig. 4b and d). This climate gradient is particularly abrupt between southern Ecuador and northern Peru, despite their being geographically close. The annual temperature ranges rise from 12.1 to 16.5°C, the precipitation drops from 938.5 to 203.5 mm, and the precipitation seasonality rises from 86.3 to 174.0 mm (Table 2).

All these differences are reflected in the coastal vegetation cover: northern Ecuador is covered by

tropical and degraded forests, central Ecuador by degraded vegetation, and southern Ecuador by grasslands and degraded forest. In Peru, the land cover is quite different; the north is covered by shrublands and the south by barren soil and desert (Fig. 4e).

To build a climate distance matrix for all the accessions, a Principal Component Analysis was done on the climate found in the geographic location of every accession (data not shown). The two principal components, which accounted for 66% of the climatic variation, were used to calculate the climate distances among the accessions. PCA1 depends mainly on the precipitation, relative humidity, and temperature variance throughout the year, and PCA2 varies with the diurnal temperature range and with mean temperature. To quantify the relationship among the genetic, climatic and geographic distances, several Mantel tests were carried out. The genetic and geographic correlation is 0.10, while the climatic and genetic correlation is 0.37, and the

Fig. 4 Climate (a–d) and land cover (e) in the Peruvian and Ecuadorian region

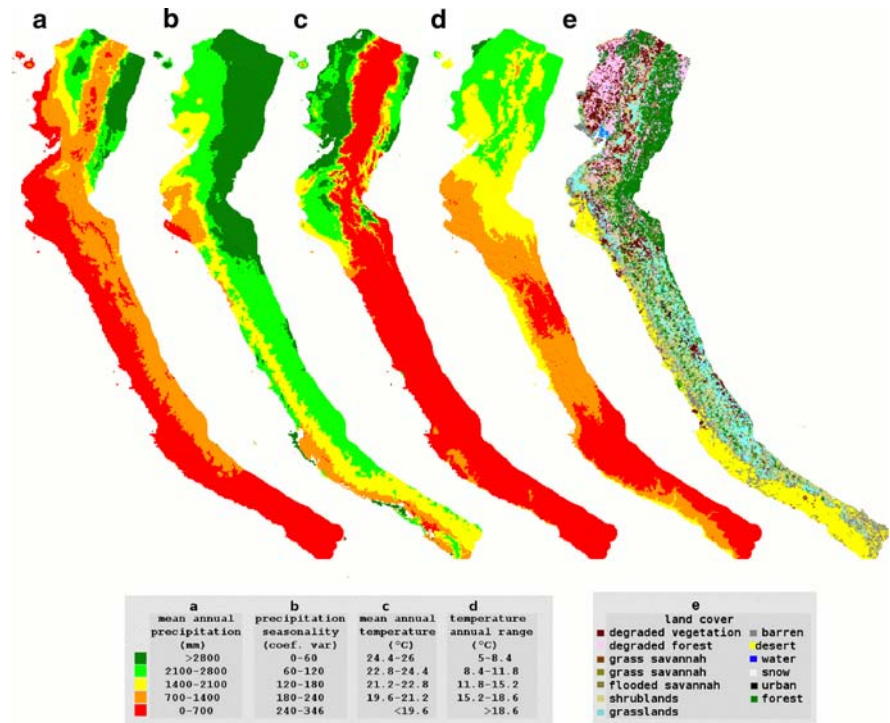


Table 2 Climatic mean values recorded in the accession collection points for each geographic region studied

| | Temperature (°C) | Temperature range (°C) | Precipitation (mm) | Precipitation seasonality (mm) |
|----------------|------------------|------------------------|--------------------|--------------------------------|
| Amazonia | 22.6 | 11.8 | 2982 | 48.3 |
| Galápagos | 22.6 | 12.1 | 621.1 | 184.3 |
| Ecuador north | 24.2 | 10.8 | 1403.6 | 71.4 |
| Ecuador center | 24.5 | 11.3 | 941.3 | 97.0 |
| Ecuador south | 23.9 | 12.1 | 938.5 | 86.3 |
| Peru north | 23.7 | 16.5 | 203.5 | 174.0 |
| Peru south | 20.1 | 17.1 | 155.6 | 194.8 |

Notes: Mean annual temperature, mean temperature range throughout the year, total annual precipitation and precipitation seasonality

geographic and climatic correlation is 0.34. All these correlations were significant according to their *P*-values.

Discussion

Genetic variation and climate correlation

The COMAV Institute, in collaboration with the UNL, UNPRG and PNG, has organized several

collection expeditions to gather *S. pimpinellifolium* samples from throughout its natural habitat (Rick and Fobes 1975; Warnock 1991), including the traditionally underrepresented Ecuador (Rick et al. 1977; Caicedo and Schall 2004; Sifres et al. 2007). *S. pimpinellifolium* from Ecuador and Peru is quite different as shown by the STRUCTURE clustering, the FCA, the geographic group dendrogram, and the AMOVA. This division is highly correlated with acute climatic differences that occur near the Peruvian-Ecuadorian border: for instance, the rainfall is 4 times greater in

the southern Ecuadorian accessions than in the geographically close northern Peruvian ones. These differences do not only affect *S. pimpinellifolium*, they also have an important effect on the overall ecosystems that characterize both countries; coastal Ecuador is covered by degraded forest and vegetation and Peru by shrublands, desert, and barren lands.

Climatic, geographic and genetic distance correlations also seem to indicate that climate and ecology could be more informative when explaining genetic differences than just the bare isolation by geographic distance. In fact, we have found a very low geographic and genetic distance correlation, as other authors previously reported in Andean wild potatoes (McGregor et al. 2002; Del Rio and Bamberg 2002). Ecological distance would be the most meaningful parameter, but it is very difficult to quantify. We have used climate distance, an easy parameter to estimate, as a fair approximation of ecological differences. Several authors have proposed that the local adaptation to different climatic conditions could account for isolation in a longer term. This climate-related adaptation has been shown recently in several species like wild barley (Cronin et al. 2007), *Brassica rapa* L. (Franks et al. 2007), and *Fagus sylvatica* L. (Jump et al. 2006). Another plausible mechanism that could account for the genetic isolation of geographically close regions with different climates and ecologies could be the adaptation of different pollinators and seed dispersal animals to these dissimilar bioclimatic environments, thus promoting the reproductive isolation of the plants living therein.

These bioclimatic and genetic correlations could also account for the Ecuadorian genetic subdivision. Whereas there is only a mild geographic differentiation within Peru, as has also been shown by other studies (Rick et al. 1977; Caicedo and Schall 2004; Sifres et al. 2007), Ecuador is divided into several geographic regions. Ecuadorian *S. pimpinellifolium* is found in isolated populations separated by great distances, which could contribute to the large inter-regional differences found in Ecuador. In northern Peru, however, *S. pimpinellifolium* distribution is almost continuous, and there are no important genetic differences in the plants collected there. Although this fact could account for part of the Ecuadorian distribution, it does not seem to be a sufficient cause. Southern Peru also shows a spotty plant distribution (Rick et al. 1977; Caicedo and Schall 2004), but this

geographic group is genetically very close to the northern Peruvian group and there is a lower genetic differentiation there. The climatic differences between Peru and Ecuador could be a possible explanation for this situation. While the coastal south Peru climate is quite uniform, regardless of the latitude (Fig. 4), the Ecuadorian climate and land cover varies heavily in the north to south direction. This bioclimatic variation could be associated with ecological differences and responsible for the important regional variations found in Ecuador.

Diversity, heterozygosity and history

There are other differences between Peruvian and Ecuadorian *S. pimpinellifolium*. Peru north is, according to the mean number of alleles (A), the most variable region. Peruvian *S. pimpinellifolium* is not just more diverse, it includes almost all alleles present in Ecuadorian *S. pimpinellifolium*. While there are 33 Peruvian alleles not present in Ecuador, only 7 exclusively Ecuadorian alleles have been found. These facts corroborate Rick et al.'s observations (Rick and Fobes 1975; Rick 1976; Rick et al. 1977). These authors suggest, as a possible explanation for this genetic richness, that the origin of the species could be located around the northern Peruvian departments of Piura and Lambayeque. The Ecuadorian genetic variation not present in Peru could be explained by new mutations which appeared after the migration, by introgressions from other species like *S. lycopersicum* var. *cerasiforme* (Rick and Holle 1990), or by losses in the original Peruvian *S. pimpinellifolium*.

Apart from the differences in diversity, there is also an important variation in the observed heterozygosity (Ho). The highest Ho is found in the north of Peru and the lowest in south Peru and Ecuador. This lack of heterozygosity in southern Peru corroborates previous observations by other authors (Rick et al. 1977; Caicedo and Schall 2004). A possible explanation for this fact could be based on the higher cross-pollination rate found in northern Peru (Rick et al. 1977, 1978). Since a clear correlation between genetic diversity and the degree of cross-pollination has also been found in the species closely related to *S. lycopersicum* (Roselius et al. 2005), the limited level of allogamy in the Ecuadorian regions could account for the lower heterozygosity and for the decreased diversity. The small effective population

size in the autogamous populations is likely to be the link between these two parameters.

The differences in the level of allogamy could also be due to differences in population density throughout *S. pimpinellifolium*'s natural range of distribution. In the southern departments of Peru and in Ecuador, as Rick observed (Rick et al. 1977), and as we have also noted in our expeditions, *S. pimpinellifolium* is characterized by a sparse distribution, a situation in contrast with the carpet of plants that completely cover northern Peru. In sparse distribution conditions, even outcrossing-capable plants could have problems finding neighbor candidates to cross-pollinate, and would choose the path of self-pollination. With these thin densities, the plants could be forced to have low heterozygosity, low diversity and strong selection towards autogamy.

A plausible model in agreement with all these data might be suggested. The original *S. pimpinellifolium*, located in northern Peru, migrated to Ecuador and southern Peru. This original colonization could have caused, in addition to a bottleneck and a loss of diversity, a selection towards autogamy in the new regions which reduced the effective population size, the heterozygosity, and the genetic diversity even more. The coastal Peruvian climate is quite homogeneous, and the similar ecological conditions could have facilitated the migration between southern and northern Peru, keeping these regions genetically close. The diversity of the Ecuadorian region could have hampered migration, keeping it more isolated with reduced real and effective population sizes and prone to further bottlenecks.

Nevertheless, caution is advised when the Peruvian origin hypothesis is evaluated. Even though our data are in agreement with that scenario, there are other possible explanations. For instance, it could be argued that the Ecuadorian genetic fragmentation and the reduced allogamy outside northern Peru, which implies a lower effective population size, has led to significant diversity losses responsible for the scenario that we have found, no matter where the original population was located.

Galapagan population origin

Some plant material from the Galápagos Islands, also included in a previous study (Nuez et al. 2004), was found, according to an AFLP-based dendrogram, to be

more related to the *S. lycopersicum* var. *cerasiforme* from the islands. This material was reanalyzed in this study. These Galapagan plants show no genetic diversity and are clustered with the Ecuadorian *S. pimpinellifolium* on the geographic region tree, although with low bootstrap values. These Galapagan plants are genetically identical among themselves, and are also indistinguishable from some *S. pimpinellifolium* plants from the Ecuadorian Los Ríos and Guayas regions. They also share the lowest Cavalli distance with the central and north Ecuadorian regions. Considering all this data, we propose to classify them as *S. pimpinellifolium*. The previous classification (Nuez et al. 2004) was probably due to a lack of typical *S. pimpinellifolium* from northern and central Ecuador in the AFLP-based dendrogram and to the relatedness of *S. pimpinellifolium* and *S. lycopersicum*. This population is likely to have been a colonization from some plants from Guayas or Los Ríos, the geographically and climatically nearest continental land, where genetically identical plants are found. A recent strong bottleneck, due to the colonization and low population size, would explain the extremely low diversity of this Galapagan region. Amazonian *S. pimpinellifolium* could also be a consequence of recent colonizations. These plants were found in cultivation lands outside their natural range of distribution and, according to the PCoA and the dendrogram, are genetically close to plants from northern Ecuador on the other side of the Andean range.

Interspecific gene flow

Solanum pimpinellifolium has a close relationship with *S. lycopersicum*. The morphological limits, based on fruit size and leaf shape, between these species are fuzzy, and there are plants with intermediate forms that cannot be clearly classified (Rick and Fobes 1975; Rick et al. 1978). Interspecific crosses are easily obtained and both species live together in the Andean region, so hybridization and gene flow could be occurring.

The most northern *S. pimpinellifolium*, located in Esmeraldas, could be related to its sister species. These plants were collected a hundred meters away from the coast, a not-so-common environment for *S. pimpinellifolium*. In the dendrogram, it is not grouped with the geographically closest northern Ecuadorian region or with any other Ecuadorian or

Peruvian *S. pimpinellifolium*, and in the PCoA, it is as close to *S. pimpinellifolium* as it is to *S. lycopersicum*. It could be the outcome of a natural, stabilized hybridization between both species.

Furthermore, Andean *S. lycopersicum* seems more diverse than the rest of the world's *S. lycopersicum*, although more accessions should be analyzed to confirm this observation. This possibly richer Andean genetic diversity might be associated with an inter-specific genetic flux occurring in this region, as was already proposed by other authors (Rick et al. 1974; Rick and Fobes 1975). This genetic exchange among cultivated and related wild species is a common phenomenon found when they are sympatric (Papa and Gepts 2003; Mason-Gamer et al. 1995; Rieseberg and Soltis 1991).

Conclusions

The intraspecific variation and the genetic diversity of *S. pimpinellifolium* were analyzed, and sharp differences were found between the Ecuadorian and Peruvian regions. The highest diversity and heterozygosity is found in northern Peru, and the molecular data is in agreement with the hypothesis of the Peruvian origin of the species. The Ecuadorian region is clearly subdivided. This genetic diversity is correlated with the different climate diversity found in the coastal areas of Ecuador and Peru. The origin of the Galapagan population has been traced to the region of central coastal Ecuador. The interspecific variation shows that *S. pimpinellifolium* and *S. lycopersicum* are very closely related and there is likely to be a genetic flow occurring between them.

Acknowledgements We are deeply grateful to Dr. Javier León, Dr. Roberto Mendoza, and Dr. Freddy Zuñiga of the Universidad Nacional de Piura; Dr. Ángel Díaz Celis, and Dr. Umberto Cardoso of the Universidad Nacional Pedro Ruiz Gallo; and to Dr. Rosa de Frutos Illán, Dr. Juan José Ruiz, María José Díez, and Dr. Jaume Prohens of the COMAV Institute. This study wouldn't have been possible without their assistance and kind affection during the collecting expeditions organized with them throughout the years. The BBVA contributed to the funding of this research.

References

- Barnaud A, Deu M, Garine E, McKey D, Joly HI (2007) Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theor Appl Genet* 114:237–248
- Belkhir K, Borsa P, Chikhi L, Rafaste N, Bonhomme T (1996) Genetix 4.04 Logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier. Website <http://www.genetix.univ-montp2.fr/genetix/genetix.htm> (Accessed 13 November 2007)
- Bohs L, Olmstead RG (1997) Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndhF* sequences. *Syst Bot* 22:5–17
- Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests. *J Stat Softw* 7:1–12
- Bouxin G (2005) Ginkgo, a multivariate analysis package. *J Veg Sci* 16:355–359
- Caicedo AL, Schall BA (2004) Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Mol Ecol* 13:1871–1882
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution* 32:550–570
- Cronin JK, Bundock PC, Henry RJ, Nevo E (2007) Adaptive climatic molecular evolution in wild barley at the *Isa* defense locus. *Proc Natl Acad Sci USA* 104:2773–2778
- Cuartero J, Nuez F, Díaz A (1984) Catalog of collections of *Lycopersicon* and *Solanum pennellii* from Northwest of Peru. *TGC Report* 34:43–46
- Darwin SC, Knapp S, Peralta IE (2003) Tomatoes in the Galápagos Islands: morphology of native and introduced species of *Solanum* section *Lycopersicon* (Solanaceae). *Syst Biodiv* 1:29–54
- Del Rio AH, Bamberg JB (2002) Lack of association between genetic and geographical origin characteristics for the wild potato *Solanum sucrense*. *Am J Potato Res* 79:335–338
- Eva HD, de Miranda EE, Di Bella CM, Gond V, Huber O, Sgrenzaroli M, Jones S, Coutinho A, Dorado A, Guimarães M, Elvidge C, Achard F, Belward AS, Bartholomé E, Baraldi A, De Grandi G, Vogt P, Fritz S, Hartley A (2002) A vegetation map of South America. Office for Official Publications of the European Communities
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50
- Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc Natl Acad Sci USA* 104:1278–1282
- Gower JC (1966) Some distance properties of latent roots and vector methods used in multivariate analysis. *Biometrika* 53:325–338
- Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA, Rieseberg LH (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature* 430:201–205
- Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: I. DIVA-GIS. *Plant Genet Resour Newsl* 127:15–19
- IPGRI (1996) Descriptor for Tomato (*Lycopersicon* spp.). IPGRI
- Jump AS, Hunt J, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Mol Ecol* 15:3469–3480

- Juvik JA, Berlinger MJ, Ben-David T, Rudich J (1982) Resistance among accessions of the genera *Lycopersicon* and *Solanum* to four of the main insect pest of tomato in Israel. *Phytoparasitica* 10:145–156
- Langella O (2002) Populations 1.2.28, Population genetic software. CNRS
- Li YC, Fahima T, Krugman T, Beiles A, Röder MS, Korol AB, Nevo E (2000) Parallel microgeographic patterns of genetic diversity and divergence revealed by allozyme, RAPD, and microsatellites in *Triticum dicoccoides* at Ammiad, Israel. *Conserv Genet* 1:191–207
- Luckwill LC (1943) The genus *Lycopersicon*: an historical, biological, and taxonomic survey of the wild and cultivated tomatoes. Aberdeen University Press, Aberdeen
- Marshall JA, Knapp S, Davey MR, Power JB, Cocking EC, Bennett MD, Cox AV (2001) Molecular systematics of *Solanum* section *Lycopersicum* (*Lycopersicon*) using the nuclear ITS rDNA region. *Theor Appl Genet* 103:1216–1222
- Mason-Gamer RJ, Holsinger KE, Jansen RK (1995) Chloroplast DNA haplotype variation within and among populations of *Coreopsis grandiflora*. *Mol Biol Evol* 12:371–381
- McGregor CE, van Treuren R, Hoekstra R, van Hintum ThJL (2002) Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for *ex situ* conservation. *Theor Appl Genet* 104:146–156
- Mieslerova B, Lebeda A, Chetelat RT (2000) Variation in response of wild *Lycopersicon* and *Solanum* sp. against tomato powdery mildew (*Oidium lycopersici*). *J Phytopathol* 148:303–311
- Miller P (1754) The gardener's dictionary. C. Rivington, London
- Mitton JB, Duran KL (2004) Genetic variation in piñon pine, *Pinus edulis*, associated with summer precipitation. *Mol Ecol* 13:1259–1264
- Monrchen M, Cuguen J, Michaelis G, Hanni C, Saumitou-Laprade P (1996) Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L. *Theor Appl Genet* 92:326–333
- Nei M (1977) F-statistics and analysis of gene diversity in subdivided populations. *Ann Hum Genet* 41:225–233
- Nuez F, Cuartero J (1984) Colectas de *Lycopersicon* y *Solanum pennellii* en el Noroeste de Perú. *Plant Genet Resour Newsl* 58:42–45
- Nuez F, Morales R, Ruíz JJ, Fernández de Córdova P, Soler S, Valdivieso E, Solórzano V (1993) Recolección de especies hortícolas en Ecuador. *Plant Genet Resour Newsl* 96:29–33
- Nuez F, Morales R, Prohens J, Fernández de Córdova P, Soler S, Valdivieso E, Solórzano V (1999) Germplasm of Solanaceae horticultural crops in the South of Ecuador. *Plant Genet Resour Newsl* 120:44–47
- Nuez F, Picó B (1999) Collections of vegetable crops and wild relatives in the Centre for Conservation and Breeding of the Agricultural Biodiversity (Spain). *Plant Genet Resour Newsl* 118:68
- Nuez F, Prohens J, Blanca JM (2004) Relationships, origin, and diversity of Galápagos tomatoes: implications for the conservation of natural populations. *Am J Bot* 91:86–99
- Owuor ED, Beharav A, Fahima T, Kirzhner VM, Korol AB, Nevo E (2003) Microscale ecological stress causes RAPD molecular selection in wild barley, Neve Yaar microsite, Israel. *Genet Resour Crop Evol* 50:213–223
- Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theor Appl Genet* 106:239–250
- Peralta IE, Spooner DM (2000) Classification of wild tomatoes: a review. *Kurtziana* 28:45–54
- Peralta IE, Spooner DM (2001) Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). *Am J Bot* 88:1888–1902
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Powell W, Morgante M, Doyle JJ, McNicol JW, Tingey SV, Rafalski AJ (1996) Genepool variation in genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. *Genetics* 144:793–803
- Qian H, Ricklefs RE, White PS (2005) Beta diversity of angiosperms in temperate floras of eastern Asia and eastern North America. *Ecol Lett* 8:15–22
- Rick CM (1976) Natural variability in wild species of *Lycopersicon* and its bearing on tomato breeding. *Genet Agr* 30:249–259
- Rick CM, Chetelat RT (1995) Utilization of related wild species for tomato improvement. *Acta Hort* 412:21–38
- Rick CM, Fobes JF (1975) Allozyme variation in the cultivated tomato and closely related species. *Bull Torrey Bot Club* 102:376–384
- Rick CM, Holle M (1990) Andean *Lycopersicon esculentum* var. *cerasiforme*: genetic variation and its evolutionary significance. *Econ Bot* 44:69–78
- Rick CM, Zobel RW, Fobes JF (1974) Four peroxidase loci in red-fruited tomato species: genetics and geographic distribution. *Proc Natl Acad Sci USA* 71:835–839
- Rick CM, Fobes JF, Holle M (1977) Genetic variation in *Lycopersicon pimpinellifolium*: evidence of evolutionary change in mating systems. *Plant Syst Evol* 127:139–170
- Rick CM, Holle M, Thorp RW (1978) Rates of cross-pollination in *Lycopersicon pimpinellifolium*: impact of genetic variation in floral characters. *Plant Syst Evol* 129:31–44
- 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 Syst Evol* 132:279–298
- Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic flow in plants. *Evol Trends Plants* 5:65–84
- Roselius K, Stephan W, Stadler T (2005) The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics* 171:753–763
- Sifres A, Picó B, Blanca JM, De Frutos R, Nuez F (2007) Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event of 1997–98. *Genet Resour Crop Evol* 54:359–377

- Smulders MJM, Bredemeijer G, Rus-kortekass 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. *Theor Appl Genet* 97:264–272
- Warnock SJ (1991) Natural habitats of *Lycopersicon* species. *HortScience* 26:466–471
- Woodward FI (1987) *Climate and plant distribution*. Cambridge University Press, Cambridge, UK

Supplementary material table.

| Accession code | GeneBank | Species | Latitude | Longitude |
|-----------------------|-----------------|--|-----------------|------------------|
| L00712 | AVRDC | <i>S. pimpinellifolium</i> | 125000S | 0724200W |
| L06065 | AVRDC | <i>S. pimpinellifolium</i> | 010600S | 0792900W |
| UPV9156 | CIAPAN | <i>S. lycopersicum</i> | | |
| CATIE20245 | COMAV | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 095900N | 0842000W |
| GLP-24 | COMAV | <i>S. pimpinellifolium</i> | 003140S | 0901921W |
| GLP-26 | COMAV | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 003140S | 0901921W |
| GLP-30 | COMAV | <i>S. pimpinellifolium</i> | 003341S | 0902001W |
| PL508 | COMAV | <i>S. lycopersicum</i> | 040223S | 0794719W |
| PL509 | COMAV | <i>S. lycopersicum</i> | 040223S | 0794719W |
| PL510 | COMAV | <i>S. lycopersicum</i> | 040223S | 0794719W |
| PL511 | COMAV | <i>S. lycopersicum</i> | 040223S | 0794719W |
| PL515 | COMAV | <i>S. lycopersicum</i> | 051648S | 0792816W |
| PL516 | COMAV | <i>S. pimpinellifolium</i> | 055942S | 0794303W |
| PL517 | COMAV | <i>S. pimpinellifolium</i> | 055942S | 0794303W |
| PL518 | COMAV | <i>S. lycopersicum</i> | 055942S | 0794303W |
| PL519 | COMAV | <i>S. lycopersicum</i> | 055942S | 0794303W |
| PL520 | COMAV | <i>S. lycopersicum</i> | 055942S | 0794303W |
| PL560 | COMAV | <i>S. pimpinellifolium</i> | 045500S | 0802000W |
| PL561 | COMAV | <i>S. pimpinellifolium</i> | 045300S | 0802200W |
| PL562 | COMAV | <i>S. pimpinellifolium</i> | 045600S | 0803200W |
| PL563 | COMAV | <i>S. pimpinellifolium</i> | 050900S | 0801000W |
| PL564 | COMAV | <i>S. pimpinellifolium</i> | 051400S | 0800600W |
| PL565 | COMAV | <i>S. pimpinellifolium</i> | 051700S | 0795700W |
| PL566 | COMAV | <i>S. pimpinellifolium</i> | 051700S | 0795700W |
| PL567 | COMAV | <i>S. pimpinellifolium</i> | 051100S | 0803700W |
| PL568 | COMAV | <i>S. pimpinellifolium</i> | 051100S | 0803700W |
| PL569 | COMAV | <i>S. pimpinellifolium</i> | 051700S | 0795600W |
| PL570 | COMAV | <i>S. pimpinellifolium</i> | 051900S | 0795500W |
| PL571 | COMAV | <i>S. pimpinellifolium</i> | 052000S | 0795000W |
| PL572 | COMAV | <i>S. pimpinellifolium</i> | 052600S | 0794400W |
| PL573 | COMAV | <i>S. pimpinellifolium</i> | 052600S | 0794400W |
| PL574 | COMAV | <i>S. pimpinellifolium</i> | 052500S | 0794400W |
| PL575 | COMAV | <i>S. pimpinellifolium</i> | 051600S | 0804100W |
| PL577 | COMAV | <i>S. pimpinellifolium</i> | 051900S | 0804200W |
| PL578 | COMAV | <i>S. pimpinellifolium</i> | 052600S | 0804500W |
| PL579 | COMAV | <i>S. pimpinellifolium</i> | 053100S | 0804900W |
| PL580 | COMAV | <i>S. pimpinellifolium</i> | 053100S | 0804900W |
| PL581 | COMAV | <i>S. pimpinellifolium</i> | 051600S | 0800600W |
| PL582 | COMAV | <i>S. pimpinellifolium</i> | 052300S | 0800300W |
| PL583 | COMAV | <i>S. pimpinellifolium</i> | 052600S | 0800100W |
| PL585 | COMAV | <i>S. pimpinellifolium</i> | 053300S | 0795800W |
| PL586 | COMAV | <i>S. pimpinellifolium</i> | 053300S | 0795800W |
| PL587 | COMAV | <i>S. pimpinellifolium</i> | 053500S | 0795800W |
| PL588 | COMAV | <i>S. pimpinellifolium</i> | 053500S | 0795700W |
| PL589 | COMAV | <i>S. pimpinellifolium</i> | 053500S | 0795700W |
| PL591 | COMAV | <i>S. pimpinellifolium</i> | 054900S | 0794900W |
| PL592 | COMAV | <i>S. pimpinellifolium</i> | 054905S | 0794959W |
| PL593 | COMAV | <i>S. pimpinellifolium</i> | 054900S | 0794900W |
| PL594 | COMAV | <i>S. pimpinellifolium</i> | 055500S | 0794600W |
| PL595 | COMAV | <i>S. pimpinellifolium</i> | 055500S | 0794600W |
| PL596 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL597 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL598 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL599 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL601 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL602 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |

| Accession code | GeneBank | Species | Latitude | Longitude |
|-----------------------|-----------------|--|-----------------|------------------|
| PL603 | COMAV | <i>S. pimpinellifolium</i> | 045100S | 0804700W |
| PL604 | COMAV | <i>S. pimpinellifolium</i> | 045100S | 0805000W |
| PL647 | COMAV | <i>S. pimpinellifolium</i> | 051700S | 0804114W |
| PL648 | COMAV | <i>S. lycopersicum</i> | 052734S | 0804557W |
| PL650 | COMAV | <i>S. pimpinellifolium</i> | 045300S | 0802200W |
| PL655 | COMAV | <i>S. pimpinellifolium</i> | 055900S | 0794300W |
| PL658 | COMAV | <i>S. pimpinellifolium</i> | 052900S | 0800000W |
| PL660 | COMAV | <i>S. pimpinellifolium</i> | 044800S | 0801700W |
| QL119 | COMAV | <i>S. pimpinellifolium</i> | 004248S | 0802400W |
| QL137 | COMAV | <i>S. pimpinellifolium</i> | 010219S | 0801907W |
| QL156 | COMAV | <i>S. pimpinellifolium</i> | 021741S | 0804227W |
| QL159 | COMAV | <i>S. pimpinellifolium</i> | 022154S | 0801802W |
| QL59 | COMAV | <i>S. pimpinellifolium</i> | 004657N | 0793543W |
| QL75 | COMAV | <i>S. pimpinellifolium</i> | 002000S | 0792000W |
| UPV13126 | COMAV | <i>S. pimpinellifolium</i> | 063400S | 0783800W |
| UPV13127 | COMAV | <i>S. pimpinellifolium</i> | 071500S | 0790800W |
| UPV13518 | COMAV | <i>S. pimpinellifolium</i> | 093300N | 0775300W |
| UPV13569 | COMAV | <i>S. lycopersicum</i> | | |
| UPV13922 | COMAV | <i>S. lycopersicum</i> | | |
| UPV13926 | COMAV | <i>S. lycopersicum</i> | | |
| UPV13950 | COMAV | <i>S. pimpinellifolium</i> | 093100N | 0775500W |
| UPV14277 | COMAV | <i>S. pimpinellifolium</i> | 035707S | 0792608W |
| UPV14297 | COMAV | <i>S. pimpinellifolium</i> | 034749S | 0793679W |
| UPV14335 | COMAV | <i>S. pimpinellifolium</i> | 032005S | 0794149W |
| UPV14336 | COMAV | <i>S. pimpinellifolium</i> | 031844S | 0793743W |
| UPV14337 | COMAV | <i>S. pimpinellifolium</i> | 031905S | 0793517W |
| UPV14341 | COMAV | <i>S. pimpinellifolium</i> | 023548S | 0792826W |
| UPV14344 | COMAV | <i>S. pimpinellifolium</i> | 022806S | 0792716W |
| UPV14345 | COMAV | <i>S. pimpinellifolium</i> | 022806S | 0792716W |
| UPV14361 | COMAV | <i>S. pimpinellifolium</i> | 031829S | 0792053W |
| UPV15820 | COMAV | <i>S. lycopersicum</i> | 37----N | 002----W |
| UPV16459 | COMAV | <i>S. lycopersicum</i> | 28----N | 015----W |
| UPV17029 | COMAV | <i>S. pimpinellifolium</i> | 053521S | 0795813W |
| UPV19003 | COMAV | <i>S. pimpinellifolium</i> | 004318S | 0901935W |
| UPV19004 | COMAV | <i>S. pimpinellifolium</i> | 004315S | 0901937W |
| UPV19005 | COMAV | <i>S. pimpinellifolium</i> | 004141S | 0901926w |
| UPV19006 | COMAV | <i>S. pimpinellifolium</i> | 004044S | 0901841W |
| UPV19007 | COMAV | <i>S. pimpinellifolium</i> | 004131S | 0902024W |
| UPV19009 | COMAV | <i>S. pimpinellifolium</i> | 004103S | 0902205W |
| UPV19859 | COMAV | <i>S. lycopersicum</i> | 40----N | 002----W |
| UPV19883 | COMAV | <i>S. pimpinellifolium</i> | 072000S | 0793400W |
| UPV20199 | COMAV | <i>S. pimpinellifolium</i> | 115448S | 0763010W |
| UPV20199 | COMAV | <i>S. pimpinellifolium</i> | 115448S | 0763010W |
| UPV20380 | COMAV | <i>S. pimpinellifolium</i> | 115458S | 0763747W |
| UPV20380 | COMAV | <i>S. pimpinellifolium</i> | 115458S | 0763747W |
| UPV21104 | COMAV | <i>S. lycopersicum</i> | 39----N | 000----W |
| UPV22073 | COMAV | <i>S. lycopersicum</i> | | |
| UPV22220 | COMAV | <i>S. pimpinellifolium</i> | 071000S | 0790100W |
| UPV22339 | COMAV | <i>S. pimpinellifolium</i> | 005426S | 0774813W |
| UPV22359 | COMAV | <i>S. pimpinellifolium</i> | 010210S | 0774128W |
| UPV22362 | COMAV | <i>S. pimpinellifolium</i> | 010248S | 0774743W |
| UPV22504 | COMAV | <i>S. pimpinellifolium</i> | 032557S | 0783408W |
| UPV22505 | COMAV | <i>S. pimpinellifolium</i> | 032557S | 0783408W |
| UPV22511 | COMAV | <i>S. pimpinellifolium</i> | 033527S | 0783406W |
| UPV22674 | COMAV | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 21----N | 089----W |
| UPV22675 | COMAV | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 20----N | 098----W |
| UPV23574 | COMAV | <i>S. pimpinellifolium</i> | 004448N | 0781437W |
| UPV23590 | COMAV | <i>S. pimpinellifolium</i> | 004719N | 0781732W |

| Accession code | GeneBank | Species | Latitude | Longitude |
|-----------------------|-----------------|----------------------------|-----------------|------------------|
| UPV23601 | COMAV | <i>S. pimpinellifolium</i> | 005102N | 0782531W |
| UPV23607 | COMAV | <i>S. pimpinellifolium</i> | 005259N | 0783000W |
| UPV23616 | COMAV | <i>S. pimpinellifolium</i> | 005231N | 0794631W |
| UPV23621 | COMAV | <i>S. pimpinellifolium</i> | 005922N | 0793201W |
| UPV23628 | COMAV | <i>S. pimpinellifolium</i> | 005936N | 0793320W |
| UPV23639 | COMAV | <i>S. pimpinellifolium</i> | 005130N | 0795232W |
| UPV23647 | COMAV | <i>S. pimpinellifolium</i> | 001450S | 0790919W |
| UPV23649 | COMAV | <i>S. pimpinellifolium</i> | 001439S | 0791640W |
| UPV23657 | COMAV | <i>S. pimpinellifolium</i> | 001319S | 0792920W |
| UPV23660 | COMAV | <i>S. pimpinellifolium</i> | 001237S | 0792958W |
| UPV23661 | COMAV | <i>S. pimpinellifolium</i> | 001102S | 0793115W |
| UPV23662 | COMAV | <i>S. pimpinellifolium</i> | 001042S | 0793155W |
| UPV23663 | COMAV | <i>S. pimpinellifolium</i> | 000733S | 0793507W |
| UPV23664 | COMAV | <i>S. pimpinellifolium</i> | 000504S | 0794059W |
| UPV23668 | COMAV | <i>S. pimpinellifolium</i> | 000528S | 0794529W |
| UPV23677 | COMAV | <i>S. pimpinellifolium</i> | 000019N | 0795300W |
| UPV23680 | COMAV | <i>S. pimpinellifolium</i> | 000209N | 0795702W |
| UPV23693 | COMAV | <i>S. pimpinellifolium</i> | 000416N | 0800325W |
| UPV23696 | COMAV | <i>S. pimpinellifolium</i> | 000044N | 0800523W |
| UPV23703 | COMAV | <i>S. pimpinellifolium</i> | 000253S | 0800811W |
| UPV23706 | COMAV | <i>S. pimpinellifolium</i> | 000732S | 0801301W |
| UPV23710 | COMAV | <i>S. pimpinellifolium</i> | 001804S | 0801945W |
| UPV23712 | COMAV | <i>S. pimpinellifolium</i> | 001947S | 0802045W |
| UPV23713 | COMAV | <i>S. pimpinellifolium</i> | 002523S | 0802653W |
| UPV23714 | COMAV | <i>S. pimpinellifolium</i> | 002557S | 0802711W |
| UPV23715 | COMAV | <i>S. pimpinellifolium</i> | 002658S | 0802704W |
| UPV23718 | COMAV | <i>S. pimpinellifolium</i> | 003828S | 0802443W |
| UPV23734 | COMAV | <i>S. pimpinellifolium</i> | 004944S | 0802940W |
| UPV23737 | COMAV | <i>S. pimpinellifolium</i> | 005136S | 0802810W |
| UPV23745 | COMAV | <i>S. pimpinellifolium</i> | 005828S | 0803818W |
| UPV23746 | COMAV | <i>S. pimpinellifolium</i> | 010350S | 0803529W |
| UPV23747 | COMAV | <i>S. pimpinellifolium</i> | 010350S | 0803529W |
| UPV23757 | COMAV | <i>S. pimpinellifolium</i> | 010419S | 0801052W |
| UPV23762 | COMAV | <i>S. pimpinellifolium</i> | 010235S | 0800517W |
| UPV23768 | COMAV | <i>S. pimpinellifolium</i> | 010256S | 0800525W |
| UPV23776 | COMAV | <i>S. pimpinellifolium</i> | 012018S | 0804201W |
| UPV23788 | COMAV | <i>S. pimpinellifolium</i> | 014153S | 0804649W |
| UPV23789 | COMAV | <i>S. pimpinellifolium</i> | 015350S | 0804353W |
| UPV23791 | COMAV | <i>S. pimpinellifolium</i> | 021607S | 0804519W |
| UPV23798 | COMAV | <i>S. pimpinellifolium</i> | 022157S | 0801809W |
| UPV23799 | COMAV | <i>S. pimpinellifolium</i> | 020700S | 0793560W |
| UPV23808 | COMAV | <i>S. pimpinellifolium</i> | 022458S | 0793727W |
| UPV23810 | COMAV | <i>S. pimpinellifolium</i> | 024000S | 0793700W |
| UPV23824 | COMAV | <i>S. pimpinellifolium</i> | 025850S | 0794337W |
| UPV23851 | COMAV | <i>S. pimpinellifolium</i> | 031922S | 0794956W |
| UPV23858 | COMAV | <i>S. pimpinellifolium</i> | 031941S | 0794720W |
| UPV23860 | COMAV | <i>S. pimpinellifolium</i> | 031917S | 0794259W |
| UPV23871 | COMAV | <i>S. pimpinellifolium</i> | 031927S | 0793946W |
| UPV2388 | COMAV | <i>S. pimpinellifolium</i> | 031942S | 0793629W |
| UPV3378 | COMAV | <i>S. lycopersicum</i> | 41----N | 002----E |
| UPV3598 | COMAV | <i>S. lycopersicum</i> | 39----N | 006----W |
| UPV3644 | COMAV | <i>S. lycopersicum</i> | 38----N | 001----W |
| UPV5745 | COMAV | <i>S. lycopersicum</i> | 40----N | 002----W |
| UPV6441 | COMAV | <i>S. pimpinellifolium</i> | 050551S | 0801613W |
| UPV6445 | COMAV | <i>S. pimpinellifolium</i> | 051019S | 0800819W |
| UPV6446 | COMAV | <i>S. pimpinellifolium</i> | 051602S | 0800612W |
| UPV6447 | COMAV | <i>S. pimpinellifolium</i> | 052220S | 0800326W |
| UPV6453 | COMAV | <i>S. pimpinellifolium</i> | 053459S | 0795751W |

| Accession code | GeneBank | Species | Latitude | Longitude |
|----------------|----------|--|----------|-----------|
| UPV6454 | COMAV | <i>S. pimpinellifolium</i> | 053526S | 0795742W |
| UPV6455 | COMAV | <i>S. pimpinellifolium</i> | 055522S | 0794553W |
| UPV6464 | COMAV | <i>S. pimpinellifolium</i> | 060050S | 0794041W |
| UPV6465 | COMAV | <i>S. pimpinellifolium</i> | 060050S | 0794041W |
| UPV6472 | COMAV | <i>S. pimpinellifolium</i> | 060936S | 0794202W |
| UPV6473 | COMAV | <i>S. pimpinellifolium</i> | 060936S | 0794202W |
| UPV6474 | COMAV | <i>S. pimpinellifolium</i> | 061325S | 0794230W |
| UPV6480 | COMAV | <i>S. pimpinellifolium</i> | 071148S | 0792542W |
| UPV6485 | COMAV | <i>S. pimpinellifolium</i> | 071821S | 0792645W |
| UPV6486 | COMAV | <i>S. pimpinellifolium</i> | 071821S | 0792645W |
| UPV6489 | COMAV | <i>S. pimpinellifolium</i> | 071825S | 0792645W |
| UPV6490 | COMAV | <i>S. pimpinellifolium</i> | 071825S | 0792645W |
| UPV6491 | COMAV | <i>S. pimpinellifolium</i> | 071825S | 0792645W |
| UPV6492 | COMAV | <i>S. pimpinellifolium</i> | 071911S | 0792213W |
| UPV6495 | COMAV | <i>S. pimpinellifolium</i> | 071500S | 0791400W |
| UPV6496 | COMAV | <i>S. pimpinellifolium</i> | 071320S | 0791251W |
| UPV6497 | COMAV | <i>S. pimpinellifolium</i> | 071320S | 0791251W |
| UPV6498 | COMAV | <i>S. pimpinellifolium</i> | 071320S | 0791251W |
| UPV6499 | COMAV | <i>S. pimpinellifolium</i> | 071414S | 0791126W |
| UPV6501 | COMAV | <i>S. pimpinellifolium</i> | 071244S | 0791126W |
| UPV6502 | COMAV | <i>S. pimpinellifolium</i> | 071300S | 0785800W |
| UPV6503 | COMAV | <i>S. pimpinellifolium</i> | 071300S | 0785800W |
| UPV6505 | COMAV | <i>S. pimpinellifolium</i> | 071343S | 0784958W |
| UPV6596 | COMAV | <i>S. pimpinellifolium</i> | 050600S | 0795400W |
| UPV6597 | COMAV | <i>S. pimpinellifolium</i> | 050100S | 0795300W |
| UPV6620 | COMAV | <i>S. pimpinellifolium</i> | 045200S | 0795700W |
| UPV6626 | COMAV | <i>S. pimpinellifolium</i> | 050800S | 0801600W |
| UPV6678 | COMAV | <i>S. pimpinellifolium</i> | 040837S | 0795050W |
| UPV6679 | COMAV | <i>S. pimpinellifolium</i> | 040237S | 0794119W |
| UPV6701 | COMAV | <i>S. pimpinellifolium</i> | 031819S | 0792107W |
| UPV6717 | COMAV | <i>S. pimpinellifolium</i> | 084005S | 0781848W |
| UPV6920 | COMAV | <i>S. lycopersicum</i> | 39----N | 006----W |
| UPV8895 | COMAV | <i>S. lycopersicum</i> | 43----N | 003----W |
| UPV9151 | COMAV | <i>S. lycopersicum</i> | | |
| UPV9172 | COMAV | <i>S. lycopersicum</i> | | |
| LA0400 | TGRC | <i>S. pimpinellifolium</i> | 051500S | 0795700W |
| LA1236 | TGRC | <i>S. pimpinellifolium</i> | 001400S | 0791000W |
| LA1245 | TGRC | <i>S. pimpinellifolium</i> | 032730S | 0795800W |
| LA1258 | TGRC | <i>S. pimpinellifolium</i> | 012900S | 0800600W |
| LA1261 | TGRC | <i>S. pimpinellifolium</i> | 014900S | 0793100W |
| LA1307 | TGRC | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 130700S | 0741300W |
| LA1312 | TGRC | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 133700S | 0734700W |
| LA1380 | TGRC | <i>S. pimpinellifolium</i> | 051609S | 0800302W |
| LA1381 | TGRC | <i>S. pimpinellifolium</i> | 053544S | 0795359W |
| LA1388 | TGRC | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | 110800S | 0752000W |
| LA1429 | TGRC | <i>S. pimpinellifolium</i> | 004800S | 0801400W |
| LA1469 | TGRC | <i>S. pimpinellifolium</i> | 055136S | 0794724W |
| LA1470 | TGRC | <i>S. pimpinellifolium</i> | 060112S | 0794048W |
| LA1471 | TGRC | <i>S. pimpinellifolium</i> | 061836S | 0794500W |
| LA1607 | TGRC | <i>S. pimpinellifolium</i> | 130524S | 0762324W |
| LA1636 | TGRC | <i>S. pimpinellifolium</i> | 132700S | 0760500W |
| LA1670 | TGRC | <i>S. pimpinellifolium</i> | 175000S | 0703100W |
| LA1683 | TGRC | <i>S. pimpinellifolium</i> | 045212S | 0810636W |
| LA1684 | TGRC | <i>S. pimpinellifolium</i> | 050712S | 0801012W |
| LA1688 | TGRC | <i>S. pimpinellifolium</i> | 045300S | 0802230W |
| LA1689 | TGRC | <i>S. pimpinellifolium</i> | 051300S | 0803730W |
| LA1781 | TGRC | <i>S. pimpinellifolium</i> | 003600S | 0802500W |
| LA1923 | TGRC | <i>S. pimpinellifolium</i> | 144000S | 0751700W |

| Accession code | GeneBank | Species | Latitude | Longitude |
|-----------------------|-----------------|----------------------------|-----------------|------------------|
| LA1933 | TGRC | <i>S. pimpinellifolium</i> | 152700S | 0742700W |
| LA1936 | TGRC | <i>S. pimpinellifolium</i> | 154936S | 0740115W |
| LA2182 | TGRC | <i>S. pimpinellifolium</i> | 055551S | 0783946W |
| LA2188 | TGRC | <i>S. pimpinellifolium</i> | 055900S | 0781100W |
| LA2652 | TGRC | <i>S. pimpinellifolium</i> | 045300S | 0804100W |
| LA2653 | TGRC | <i>S. pimpinellifolium</i> | 044500S | 0803500W |
| LA2656 | TGRC | <i>S. pimpinellifolium</i> | 034800S | 0804200W |
| LA2725 | TGRC | <i>S. pimpinellifolium</i> | 134300S | 0755200W |
| LA2850 | TGRC | <i>S. pimpinellifolium</i> | 005600S | 0804300W |
| LA2852 | TGRC | <i>S. pimpinellifolium</i> | 004900S | 0802900W |
| LA2854 | TGRC | <i>S. pimpinellifolium</i> | 012000S | 0803500W |
| LA2915 | TGRC | <i>S. pimpinellifolium</i> | 055905S | 0794443W |
| PI127807 | USDA | <i>S. pimpinellifolium</i> | 125000S | 0724200W |
| PI365927 | USDA | <i>S. pimpinellifolium</i> | 115403S | 0763948W |
| PI365959 | USDA | <i>S. pimpinellifolium</i> | 162400S | 0731500W |
| PI379007 | USDA | <i>S. pimpinellifolium</i> | 065000S | 0780100W |
| PI379055 | USDA | <i>S. pimpinellifolium</i> | 001500S | 0793000W |

***Capítulo 3.- The implications of AFLP data for the
systematics of the wild species of
Solanum section Basarthrum.***

Systematic Botany

The Implications of AFLP Data for the Systematics of the Wild Species of *Solanum* Section *Basarthrum*

J. PROHENS,¹ G. J. ANDERSON,² J. M. BLANCA,¹ J. CAÑIZARES,¹ E. ZURIAGA,¹ and F. NUEZ¹

¹Centro de Conservación y Mejora de la Agrodiversidad Valenciana, Universidad Politécnica de Valencia, Camino de Vera 14, 46022 Valencia, Spain

²Department of Ecology and Evolutionary Biology, 75 North Eagleville Road, Storrs, Connecticut 06269, U.S.A.

Communicating Editor: Thomas G. Lammers

ABSTRACT. The wild species associated with the Andean domesticate, the pepino, and closely associated with the wild and cultivated potatoes, are morphologically and ecologically variable. We studied 10 of the 11 known species, represented by 35 accessions, of *Solanum* section *Basarthrum*, plus material of two putative new species. Given the morphological variability, and cryptic species, molecular studies were appropriate. Amplified Fragment Length Polymorphisms (AFLP) were utilized because they are highly polymorphic, and cover most of the genome. Some 98% of the 292 fragments recovered proved informative. A neighbour joining cluster analysis and principal coordinates analysis largely supported previous taxonomic distinctions based on decades of morphological and biosystematic study. However, two new distinct molecular elements were identified, one autogamous, the other self incompatible, that will be described as new species, and that mimic a species pair of wild tomatoes native to the same region of Peru. The most diverse taxonomic group, the series *Caripensia*, also proved to be the most diverse genetically (85% of loci polymorphic), allowing morphologically similar species to be distinguished. The AFLP data and species distributions in this series support an hypothesis of rapid evolution and peripheral isolation as evolutionary mechanisms in the geographically and ecologically diverse series *Caripensia*.

Solanum section *Basarthrum* has proven to be fertile for studies of systematics, evolution, reproductive biology, domestication, and plant breeding (Correll 1962; Anderson 1979; Mione and Anderson 1992; Anderson et al. 1996; Prohens et al. 2003). Several characteristics make this group a model for these types of studies. A few species are autogamous, while most are self-incompatible and allogamous (Mione and Anderson 1992); a range of crossing relationships from full intercrossability and fertility of hybrids to impenetrable crossing barriers exist among the species (e.g., Anderson 1979), and the section is morphologically, ecologically, and geographically diverse. *Solanum* sect. *Basarthrum* includes widely distributed species and very rare species known from only one location; the species occupy different ecological niches, and grow at elevations from 0 to 3300 m (Correll 1962; Prohens et al. 2003). Some intercompatible species are sympatric while others are allopatric (Correll 1962; Anderson 1975, 1977, 1979). The section includes one cultigen, the pepino (*Solanum muricatum* Aiton), and thus studies on the relationships of the wild species to the domesticate, the process of domestication in general, and the use of wild species in the genetic improvement of the pepino have been important (Anderson et al. 1996; Prohens et al. 2003, 2004; Rodríguez-Burruezo et al. 2003). *Solanum* sect. *Basarthrum* is phylogenetically very close to potatoes (*Solanum* sect. *Petota*) and tomatoes (*Solanum* sect. *Lycopersicon*; Lester 1991; Spooner et al. 1993), two groups with great economic importance, and for which *Solanum* sect. *Basarthrum* represents a tertiary gene pool for breeding (Sakamoto and Taguchi 1991; Trognitz and Trognitz 2004). Finally, because of its relationship to potatoes and tomatoes, *Solanum* sect. *Bas-*

arthrum is useful for comparative studies of evolution between tuberous and non-tuberous groups of *Solanum*. Over three decades ago, Donovan Correll, the last monographer of section (which he included as the closest relatives in his treatment of the tuberous and non-tuberous relatives of the potatoes), considered many of the species in this section more variable and difficult than the notoriously variable potatoes (pg. 43, 1962).

Wild species of sect. *Basarthrum* are herbaceous, mostly perennial, and native to the Andean region of South America, with some also in the mountains of Central America (Correll 1962). The section includes 12 wild species distributed among four series: *Canensa* (*S. canense* Rybd.), *Caripensia* (*S. basendopogon* Bitter, *S. caripense* Humb. & Bonpl. ex Dunal, *S. cochoae* G.J. Anderson & Bernardello, *S. filiforme* Ruiz López & Pavón, *S. fraxinifolium* Dunal in DC, *S. heiseri* G.J. Anderson, *S. tabanoense* Correll, *S. taeniotrichum* Correll, *S. trachycarpum* Bitter & Sodirol), *Articulata* (*S. santae-marthae* Bitter), and *Suaevolentia* (*S. suaevolens* Kunth & Bouché) (Anderson and Jansen 1998). These four series are reproductively isolated from each other, but within series *Caripensia* several species can be intercrossed and may give fertile hybrids (Anderson 1975, 1977; Anderson and Bernardello 1991; Prohens and Nuez 2001). In addition, several species of *Caripensia* can be crossed with the cultivated *S. muricatum* (Heiser 1964; Anderson 1975, 1977; Bernardello and Anderson 1990; Anderson and Bernardello 1991). All currently known species of series *Caripensia* are self-incompatible with the exception of *S. trachycarpum*, which is self-compatible and autogamous (Mione and Anderson 1992). Within series *Caripensia*, the 'Caripense complex' includes the widespread *S. caripense*, and three species

TABLE 1. Species and collections studied. Each accession is listed with accession number, collector, and locality. ^a Accession from the G. J. Anderson collection; ^b Accession from the J. Prohens collection. ^c Vouchers in CONN.

| |
|--|
| <i>S. basendopogon</i> Bitter. 38 ^{a,c} , J. G. Hawkes, Peru (Dep. La Libertad); 43 44 ^{b,c} , J. G. Hawkes, Peru (Dep. La Libertad); BIRM/S 0032 ^b , J. G. Hawkes, Peru (Dep. La Libertad) |
| <i>S. canense</i> Rybd. 46 ^{a,c} , D'Arcy, Panama (Canita); 498 ^{a,c} , G. J. Anderson, Costa Rica (Monte Verde, Prov. Alajuela); 509 ^{a,c} , G. J. Anderson, Costa Rica (near Tilarán, Prov. Guanacoste) |
| <i>S. caripense</i> Humb. & Bonpl. ex Dun. 220 239 ^{a,c} , G. J. Anderson, 220: Ecuador (near Tixan), 239: Ecuador (near Urdaneta); BIRM/S 1034 ^b , G. J. Anderson, Ecuador; BIRM/S 1251 ^b , Alan Child, Ecuador (Prov. Loja); E-7 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Pichincha); EC-36 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Loja); EC-40 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Loja); PI-243342 ^b , Unknown, Costa Rica |
| <i>S. cochoae</i> G. J. Anderson & Bernardello. 934 ^{a,c} , Carlos Ochoa, Peru (Prov. Amazonas) |
| <i>S. filiforme</i> Ruiz López & Pavón. 1452 ^{a,c} , Carlos Ochoa, Peru (Prov. Tarma) |
| <i>S. fraxinifolium</i> Dunal in DC. 500 ^{a,c} , G. J. Anderson, Costa Rica (near Monte Verde); BIRM/S 1809 ^b , Barbara Pickersgill, Peru (Selva de San Francisco, Prov. Ayacucho) |
| <i>S. heiseri</i> G. J. Anderson. 6800 ^{a,c} , Charles B. Heiser, Colombia, Cerro de Monserrate (Bogotá); BIRM/S 1979 ^b , Charles B. Heiser, Colombia |
| <i>S. suaveolens</i> Kunth & Bouché. BIRM/S 1573 ^b , Unknown, Colombia (Caguan) |
| <i>S. tabanoense</i> Correll. EC-22 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Loja); EC-26 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Loja); E-257 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Loja) |
| <i>S. tabanoense</i> <i>S. caripense</i> hybrid. 769 239 ^{a,c} , G. J. Anderson, 769: Colombia (near El Encanto), 239: Ecuador (near Urdaneta) |
| <i>S. trachycarpum</i> Bitter & Sodiro. 7087 ^{a,c} , Charles B. Heiser, Ecuador; 169 ^{a,c} , G. J. Anderson, Ecuador (Mitad del Mundo; Prov. Pichincha); E-34 ^b , F. Nuez & J. Prohens, Ecuador (Prov. Cotopaxi) |
| PNS1. P-71 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Abancay); P-74 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Abancay); P-80 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Abancay); P-82 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Abancay) |
| PNS2. P-51 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Ayacucho); P-61 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Ayacucho); P-62 ^{b,c} , F. Nuez & J. Prohens, Peru (Prov. Ayacucho); BIRM/S 1122 ^b , Unknown, Peru (Dep. Cusco) |

with limited distributions: *S. filiforme*, *S. fraxinifolium*, and *S. heiseri* (Anderson and Bernardello 1991). Species of this complex are distinguished from the rest of species of series *Caripensia* by having pinnately compound leaves, rotate corollas, and unbranched inflorescences.

Previous authors have employed a wide range of methods to understand the systematics of this group, including morphological studies (Correll 1962; Anderson 1975, 1977; Anderson and Gensel 1976), crossing studies (Anderson 1975, 1977, 1979; Prohens and Nuez 2001), chemotaxonomy (Simpson 1979; Anderson et al. 1987), cytology (Bernardello and Anderson 1990; Stiefkens et al. 1999), and chloroplast and ribosomal DNA studies (Anderson et al. 1996; Anderson and Jansen 1998). These studies have led to a reorganization of the section, including the reestablishment of a second section (sect. *Anarrhichomenum*) (Anderson and Jansen 1998), the suggestion that the monotypic series *Articulata* (*S. sanctae-marthae*) also be treated as a separate section (Anderson 1977), and the recognition of two new species, (*S. heiseri* [Anderson 1975] and *S. cochoae* [Anderson and Bernardello 1991]).

The study of nuclear DNA offers a powerful tool to elucidate the systematic relationships of plants (Brower et al. 1996; Henry 1997). Despite its utility, an extensive study of the nuclear genome has not been applied to *Solanum* sect. *Basarthrum*. Here we use amplified fragment length polymorphisms (AFLPs) to study the systematics of wild species of this section. AFLPs cover most of the genome and exhibit a high level of polymorphism in the Solanaceae (Milbourne et al. 1997; Furini and Wunder 2004; Nuez et al. 2004, Spooner et

al. 2005), allow a large number of loci to be scored in a single reaction, and have a much better repeatability among laboratories than other markers such as random amplified polymorphisms of DNA (RAPDs; Jones et al. 1997). Furthermore, previous work (Kardolus et al. 1998) indicates that AFLPs are of great utility in the study of the biosystematics of the related *Solanum* sect. *Petota*. Thus, we expected that AFLPs would be useful to address systematic questions among wild species of *Solanum* sect. *Basarthrum*.

MATERIALS AND METHODS

Plant Material. We examined 35 accessions from *Solanum* sect. *Basarthrum* (Table 1) corresponding to 10 of the known wild species, plus materials corresponding to two putative new species, and to an artificially obtained interspecific hybrid between *S. tabanoense* and *S. caripense*. Seeds and living material are no longer available for the very rare (one locality) *S. sanctae-marthae* and *S. taeniotrichum*. The number of accessions for each species studied varied between seven, for the abundant and widespread *S. caripense*, and one for *S. suaveolens*, the rare *S. cochoae*, and *S. filiforme*, for which only one collection is known. Selection of accessions was based on the availability of materials and on the geographical range of distribution of each species.

Materials corresponding to two very distinctive collections (Anderson et al. in preparation) collected in Central Peru in 2001 were labeled as "PNS1" and "PNS2" ("putative new species"). PNS1, unlike all known *S. caripense* but similarly to *S. canense*, *S. suaveolens* and *S. trachycarpum*, is self-compatible, likely autogamous, bears smaller flowers, and though self compatible, has a style that protrudes ca. 1 mm beyond the staminal column. The other—PNS2—is self-incompatible, possesses a strongly exerted style (several mm beyond the terminus of the staminal column), generally has larger flowers and leaves, and has significantly longer hairs on leaves, stems, etc. (0.8 mm vs. 0.5 mm). Accession BIRM/S 1122, which originally was considered as *S. caripense*, instead is gen-

TABLE 2. Oligonucleotide adaptors and primers used for the AFLP analysis.

| | Restriction enzyme | Sequence |
|-------------|--------------------|-----------------------------|
| E-0 Adaptor | Eco RI | 5'-CTCGTAGACTGCGTACC-3' |
| | | 3'-CTGACGCATGGTTAA-5' |
| | | 5'-GACGATGAGTCTGAG-3' |
| M-0 Adaptor | Mse I | 3'-TACTCAGGACTCAT-5' |
| E-A | Eco RI | 5'-AGACTGCGTACCAATTCA-3' |
| M-C | Mse I | 5'-GATGAGTCTGAGTAAC-3' |
| E-ACA | Eco RI | 5'-AGACTGCGTACCAATTCCACA-3' |
| E-AAC | Eco RI | 5'-AGACTGCGTACCAATTCAAC-3' |
| M-CAA | Mse I | 5'-GATGAGTCTGAGTAACAA-3' |
| M-CTA | Mse I | 5'-GATGAGTCTGAGTAACAA-3' |

omically very similar to PNS2, and, as a consequence, was included in the PNS2 group.

DNA Isolation and AFLPs. Total genomic DNA was isolated from young leaves of single plants using the DNeasy® Plant Mini Kit (Qiagen Inc.), following the protocol of the manufacturer. DNA concentration was quantified on agarose and stored at -20°C until used. A sample of 0.25 µg DNA sample was digested by the enzyme combination EcoRI and MseI at 37°C for 2 h. Ligation was performed with the AFLP® Core Reagent Kit (Invitrogen Corp.) following the instructions of the manufacturer. After ligation, the reaction mixture was diluted 1:10 in TE buffer.

For the pre-selective amplification, a 5 µl aliquot from the DNA dilution was added to a 25 µl solution containing 2.5 µl of 10× buffer, 0.5 µl of primer EcoA (10 µM), 0.5 µl of primer MseC (10 µM), 1.0 µl of dNTPs (10 mM), and 0.8 units of *Taq* polymerase (Roche). After pre-amplification, DNA was diluted again 1:10 in TE buffer. The selective amplification was performed on 2 µl aliquots using four combinations of primers (Table 2). DNA fragments were separated in an ABI Prism 310 genetic analyzer (Applied Biosystems). Resulting fragments were scored as binary traits (1 = present, 0 = absent) using Genographer 1.6 software (Benham 2001).

Analysis of AFLP Data. Pairwise genetic similarities were estimated with the Dice (Sorensen) similarity coefficient $S_{ij} = 2a / (2a + b + c)$, where a is the number of bands shared by i and j , b is the number of bands present in i and absent in j , and c is the number of bands present in j and absent in i . The resulting genetic similarity matrices were used to generate a principal coordinates analysis (PCoA). Calculations were performed using the NTSYSpc2.0 software package (Applied Biostatistics Inc.). Bootstrapped distance matrices were calculated using the Phyltools 1.32 software (Buntjer 2001) and used to test the stability of the neighbour joining tree constructed with the Phylip 3.62 package (Felsenstein 2004b).

Genetic diversity for AFLP data was estimated with the Popgene software (Yeh and Boyle 1997). The proportion of polymorphic fragments (P) and the Nei's gene diversity index were calculated (Nei 1987). For the species represented by more than two accessions, the Nei minimum distance (Nei 1978) and the non-biased G_{st} (Nei and Chesser 1983), computed using the Genetix 4.05.2 software (Belkhir et al. 2004) and assuming complete homozygosity (Mohammadi and Prasanna, 2003), were estimated.

RESULTS

A total of 292 AFLP fragments were recognized, of which 266 (98%) were polymorphic and informative (Table 3). Of these, 18 were specific to series *Canensa*, 82 to series *Caripensia*, 0 to series *Suaevolentia*, and seven to the group formed by PNS1 and PNS2. When fragments specific to a series, but universal to all members of that series were sought, we found 13 in series

TABLE 3. Gene diversity statistics estimated from AFLP data for *Solanum* section *Basarthrum*, series *Caripensia* and PNS1 + PNS2. The hybrid between *S. caripense* and *S. tabanoense* (769 239) was excluded from the analyses. H_T = total gene diversity.

| Groups | Number of accessions | Polymorphic loci (%) | H_T |
|---|----------------------|----------------------|--------------|
| Total | 34 | 97.6 | 0.230 |
| Series <i>Canensa</i> (<i>S. canense</i>) | 3 | 15.0 | 0.062 |
| Series <i>Caripensia</i> (7 species) | 22 | 85.3 | 0.201 |
| Series <i>Suaevolentia</i> (<i>S. suaevolens</i>) | 1 | — | — |
| "PNS1 + PNS2" (2 putative new species) | 8 | 29.5 | 0.092 |
| Series <i>Caripensia</i> | 22 | 85.3 | 0.201 |
| <i>S. basendopogon</i> | 3 | 16.4 | 0.067 |
| <i>S. cochoae</i> | 1 | — | — |
| <i>S. trachycarpum</i> | 3 | 20.2 | 0.085 |
| <i>S. tabanoense</i> | 3 | 29.5 | 0.112 |
| " <i>S. caripense</i> " complex | 12 | 56.5 | 0.159 |
| <i>S. caripense</i> | 7 | 38.7 | 0.114 |
| <i>S. heiseri</i> | 2 | 5.5 | 0.023 |
| <i>S. fraxinifolium</i> | 2 | 10.3 | 0.043 |
| <i>S. filiforme</i> | 1 | — | — |
| Putative new species | 8 | 29.5 | 0.092 |
| PNS1 | 4 | 13.7 | 0.049 |
| PNS2 | 4 | 20.2 | 0.071 |

Canensa, none in series *Caripensia*, none in series *Suaevolentia* and one in the group formed by PNS1 and PNS2.

The total gene diversity (H_T) for wild species of *Solanum* sect. *Basarthrum* was 0.23 (Table 3). The greatest gene diversity was found in series *Caripensia* (more than 85% polymorphic loci), presumably the result of this being a series including several species. Not surprisingly, the level of polymorphism within each species is much lower, with polymorphism ranging from 5.5% in *S. heiseri* (2 accessions) to 38.7% in *S. caripense* (7 accessions) (Table 3). The great difference in the percentage of polymorphism of series *Caripensia* vs. the individual species in the series results from the fact that polymorphic fragments within a species are monomorphic or absent in the remainder of species. Because of this, the relative magnitude of AFLP differentiation among species (G_{st}) is very high ($G_{st}=0.51$). AFLP diversity of *S. trachycarpum*, the only autogamous species of series *Caripensia*, is similar to or greater than that of the rest of species of the series, which are self-incompatible (Table 3). In contrast, in the group composed of PNS1+PNS2, the diversity in PNS2 (self-incompatible) is much greater than in PNS1 (autogamous). However, we included three geographically distinct collections of *S. trachycarpum*, and though there are four collections of PNS1, all are from the same locality.

The usefulness of AFLPs as phenetic markers was assessed with a Neighbour-Joining cluster analysis. *So-*

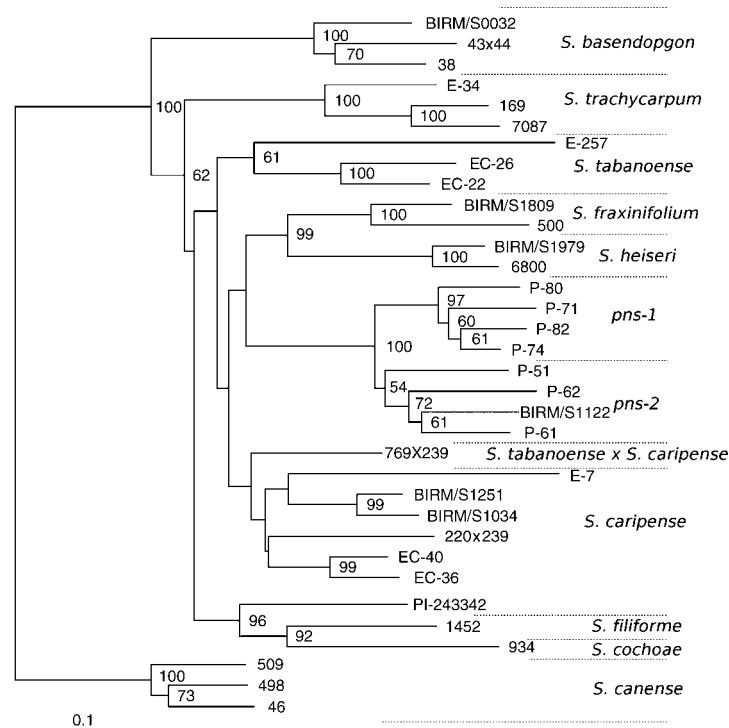


FIG. 1. UPGMA phenogram of 34 accessions of *Solanum* section *Basarthrum* plus an interspecific hybrid between *S. caripense* and *S. tabanoense* based on AFLPs. *Solanum suaveolens* was excluded from the phenogram because it caused a long branch attraction problem (Felsenstein 2004a). Phenetic relationships were derived from Dice AFLP-based pairwise genetic distances. The numbers adjacent to some nodes indicate bootstrap confidence values (percentages; 1000 replications). Nodes without numbers had bootstrap values of less than 50.

lanum suaveolens showed a very long branch, indicating a rate of genetic divergence different from that of the other species, and was removed from the cluster analysis in order to avoid a long-branch attraction problem. The inclusion of a species with such a different rate of evolution could lead to misleading phenograms (Felsenstein 2004a). The phenogram produced using genetic distance data among accessions (Fig. 1) clearly separates three consistent main clusters consisting of: 1) *S. canense*, 2) *S. basendopogon*, and 3) the rest of accessions, including the two putative new species. Within the large latter plexus, there are four strongly supported groups (bootstrap values greater than 90%): a) *S. trachycarpum*, b) a unit with *S. cochoae*, *S. filiforme*, and one errant *S. caripense* (PI 243342), c) *S. fraxinifolium* and *S. heiseri*, and d) a branch uniting the two putative new species. A cluster with a significant, but lower bootstrap value (61%) unites the three *S. tabanoense* collections as well. Similarly, all accessions but one of *S. caripense* are included in one subcluster (with 42% bootstrap support). As would be expected, the hybrid between *S. tabanoense* and *S. caripense* showed a basal position in the *S. caripense* subcluster (Fig. 1).

In order to obtain more information on the relationships among species we also performed a principal co-

ordinates analysis (PCoA). In the PCoA with all the accessions, the first and second coordinate explain 14.3% and 11.2% of the variation and mostly account for the separation of three groups formed by *S. canense*, PNS1+PNS2, and the rest of the accessions. The third coordinate (8.0% of the variation explained) fundamentally separates *S. basendopogon* from the rest of the accessions. The PCoA generally confirm the cluster analysis, and verifies that *S. basendopogon* is genetically differentiated from the remaining species of series *Caripensia*. However, it also shows that PNS1 and PNS2 are genetically differentiated from the rest of the members of series *Caripensia*, a result that could not be inferred from the cluster analysis. In addition, the PCoA shows that *S. suaveolens* plots closer to species of series *Caripensia* than to *S. canense*.

Because the three first principal coordinates only explain a bit more than one third (34.9%) of the variation detected, and mainly separate the series or species that are morphologically very different, a second PCoA was performed in which the four most distinct species (*S. basendopogon*, *S. canense*, PNS1, and PNS2) were excluded. Thus, this PCoA included the species of series *Caripensia* (excluding the highly distinct *S. basendopogon*) and *S. suaveolens*. In this restricted analysis,

TABLE 4. Nei minimum distance (above the diagonal) and relative magnitude of gene differentiation (Gst; below the diagonal) between species of section *Basarthrum*. Only species with two or more accessions have been included.

| | <i>S. basendopogon</i> | <i>S. canense</i> | <i>S. caripense</i> | <i>S. fraxinifolium</i> | <i>S. heiseri</i> | PNS1 | PNS2 | <i>S. tabanoense</i> | <i>S. trachycarpum</i> |
|-------------------------|------------------------|-------------------|---------------------|-------------------------|-------------------|------|------|----------------------|------------------------|
| <i>S. basendopogon</i> | | 0.27 | 0.16 | 0.21 | 0.25 | 0.25 | 0.22 | 0.21 | 0.26 |
| <i>S. canense</i> | 0.47 | | 0.23 | 0.30 | 0.30 | 0.30 | 0.29 | 0.25 | 0.30 |
| <i>S. caripense</i> | 0.26 | 0.38 | | 0.11 | 0.13 | 0.14 | 0.13 | 0.10 | 0.16 |
| <i>S. fraxinifolium</i> | 0.37 | 0.49 | 0.13 | | 0.15 | 0.20 | 0.17 | 0.17 | 0.22 |
| <i>S. heiseri</i> | 0.48 | 0.55 | 0.21 | 0.30 | | 0.21 | 0.20 | 0.18 | 0.25 |
| PNS1 | 0.49 | 0.55 | 0.22 | 0.41 | 0.49 | | 0.07 | 0.19 | 0.25 |
| PNS2 | 0.40 | 0.49 | 0.20 | 0.41 | 0.40 | 0.09 | | 0.17 | 0.21 |
| <i>S. tabanoense</i> | 0.28 | 0.35 | 0.09 | 0.19 | 0.24 | 0.30 | 0.22 | | 0.21 |
| <i>S. trachycarpum</i> | 0.41 | 0.47 | 0.24 | 0.35 | 0.44 | 0.45 | 0.36 | 0.26 | |

the first coordinate explains 15.3% of the variation, and clearly separates *S. trachycarpum* from the rest of the species. The second and third coordinates (11.5% and 10.4% of the variation explained) segregate *S. suaveolens* from the rest of species and also clearly separate *S. tabanoense* and the different species of the Caripense complex, which were intermingled in the first PCoA with all the materials. The total variation explained by the first three coordinates in the second PCoA analysis is 37.2%.

The lowest values of genetic distances and Gst are found among the members of the Caripense complex, as well as between *S. caripense* and *S. tabanoense* (Table 4). Also, the genetic differentiation of *S. canense* is confirmed, given that it displays the highest genetic distance and Gst values of all the species. The putative new species, PNS1 and PNS2, although reproductively distinct from each other, are genetically very similar, with a genetic distance between them of 0.07 and a Gst of 0.09.

DISCUSSION

Though no revolutionary changes were suggested, these new molecular data proved very useful in clarifying some of the most difficult systematic distinctions in this morphologically complex group. Gratifyingly, our nuclear genome data support most of the previous taxonomic classification and reorganizations performed on the basis of morphological, cytological, flavonoid, and organelle DNA data (e.g., Anderson and Jansen 1998). It is notable that although within section *Basarthrum* we have found great genetic variation, which supports a number of series and species, as well as treatment of this taxon as a distinct section, there are also a number of species for which, although molecular genetic distinction is possible, the genetic distinctions are not great. Based on these results, diversification in series *Caripensia*, and, in particular in the Caripense complex, seems to have taken place recently, with several species perhaps arising nearly simultaneously.

As crossing data in the past have done (Anderson

1975), the AFLP data highlighted unusual material (PNS1, PNS2) that will be recognized as new species (Anderson, Martine, Prohens and Nuez in review). This material, from southern Peru, displays very distinct AFLP patterns (Fig. 1). In addition, there are morphological features (plant habit, flower size, style length, pubescence features) that distinguish these new species. However, these putative new species are morphologically clearly part of *Solanum* sect. *Basarthrum*. Interestingly, it is reproductive features (compatibility and autogamy in this case) that first suggested a distinction between these species, analogous to the taxonomic outcome (i.e., recognition of a dioecious species) of reproductive studies of related solanums more than 20 years ago (Anderson and Levine 1982), and exemplified in other angiosperm groups in general (Anderson et al. 2002).

Though morphologically different from each other, PNS1 and PNS2 are genetically very similar in AFLP profile. It is of considerable interest that a parallel pattern to that found for PNS1 and PNS2 occurs in two related species of *Solanum* sect. *Lycopersicon* from this same region (Rick et al. 1976). The two species, *S. chmielewskii* (C.M. Rick, Kesicki, Fobes & M. Holle) D.M. Spooner, G.J. Anderson & R.K. Jansen (syn. *L. chmielewskii* C.M. Rick, Kesicki, Fobes & M. Holle) and *S. neorickii* D.M. Spooner, G.J. Anderson & R.K. Jansen (syn. *L. parviflorum* C.M. Rick, Kesicki, Fobes & M. Holle) are relatively closely related to *Solanum* sect. *Basarthrum* (Spooner et al. 1993). PNS1 has a rather diminutive appearance and smaller plant parts, including less showy flowers, and inserted or slightly exerted stigma, features that are also present in the nearly completely autogamous *S. neorickii* (Rick et al. 1976). Like *S. chmielewskii*, PNS2 has larger plant parts and more robust stems, exerted stigmas, and is allogamous. In addition, paralleling the syndrome in *S. chmielewskii* and *S. neorickii* (Bretó et al. 1993; Marshall et al. 2001; Spooner et al. 2005), PNS1 and PNS2 are genetically very similar. Rick et al. (1976) suggested that *S. neorickii* evolved from *S. chmielewskii* and became genetically isolated from the parent by virtue of its

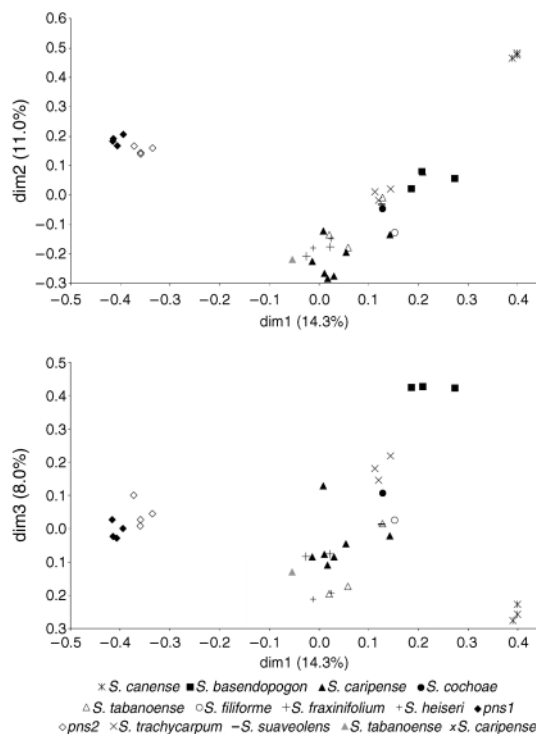


FIG. 2. Principal coordinates analysis (PCoA) plot of 35 accessions of *Solanum* section *Basarthrum* for the three principal coordinates, estimated from the AFLP similarity matrix. A. First and second coordinates. B. First and third coordinates.

tendency towards inbreeding. A similar scenario might apply to these new species, with PNS1 being the autogamous derivative of the allogamous PNS2.

Similarly distinct is the morphologically unmistakable *S. canense*, which is also the most genetically different species; it occupies a distinct position on the PCoA diagram (Fig. 2), and in the phenogram (Fig. 1). *Solanum suaveolens* is the species morphologically most similar to *S. canense*, but was originally placed by Correll (1962) in a separate series because he considered the morphological resemblance superficial. At the molecular level, the AFLP data (Fig. 1), supported by earlier molecular studies (Simpson 1979; Anderson et al. 1987, 1996), argue that the many points of morphological similarity are misleading, and are likely derived from a convergence on self compatibility and autogamy (Anderson and Jansen 1998). Thus, genetically, *S. suaveolens* occupies a position relatively much closer to the series *Caripensia* than to *S. canense* (Fig. 2). Nonetheless, molecularly *S. suaveolens* is also different from species of the series *Caripensia* as shown in the second and third coordinate of the PCoA (Fig. 3). Although only one collection of *S. suaveolens* was included, because it is molecularly distinct from *S. canense* and morphologically and reproductively different from the

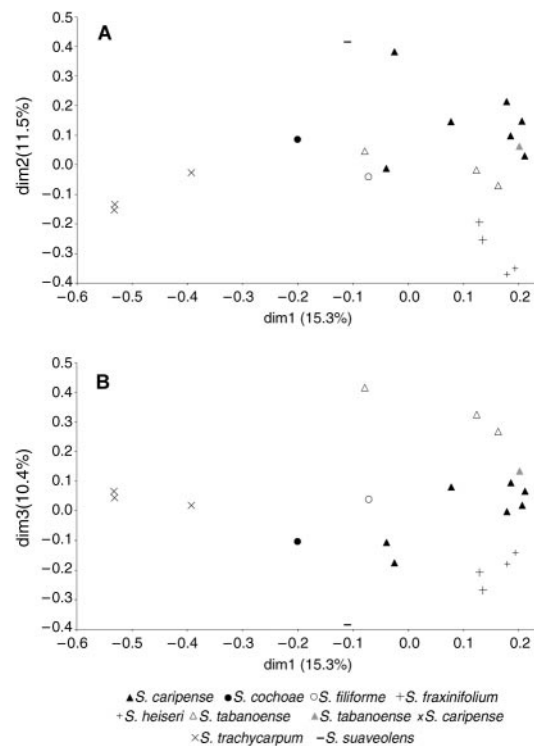


FIG. 3. Principal coordinates analysis (PCoA) plot of 19 accessions of series *Caripensia* (except *S. basendopogon*) and *Suaveolentia* from *Solanum* section *Basarthrum* for the three principal coordinates, estimated from the AFLP similarity matrix. A. First and second coordinates. B. First and third coordinates.

Caripensia complex, its treatment in a separate series (as the sole member) continues to be justified (as in Correll 1962). The remaining species comprising series *Caripensia* form a genetically similar group in which many members are at least partially interfertile (Heiser 1964; Anderson 1975, 1977, 1979; Anderson and Bernardello 1991; Prohens and Nuez 2001). Within this group, *S. basendopogon* forms a molecularly separate entity, a treatment supported by the habit of the plants as well as a number of morphological features (e.g., the unique warty fruits; Anderson and Bernardello 1991). *Solanum trachycarpum* is also molecularly separated from the remaining species, as is shown in the phenogram and in the PCoA (particularly on the first coordinate). Presently it is the only species of series *Caripensia* that is known to be self-compatible (Mione and Anderson 1992), and that is characterized by growing in dry areas, a habitat quite distinct from other members of *Solanum* sect. *Basarthrum*.

Additional support for the transfer by Anderson and Jansen (1998) of *S. tabanoense* from Correll's (1962) series *Appendiculata* to the series *Caripensia* is provided by the AFLP data. Although *S. tabanoense* is located in the AFLP plots with the other species of the *Caripensia*

complex (*S. caripense*, *S. filiforme*, *S. fraxinifolium*, and *S. heiseri*), on the basis of a quite distinct morphology (e.g., stellate corollas, often simple leaves, and fruits at least 4× the size of all other wild species) Correll 1962; Anderson 1977) and biochemical differences (Simpson 1979; Anderson et al. 1987), we believe it is best treated outside the Caripense complex.

The two species morphologically most similar to *S. caripense* and to each other, *S. fraxinifolium* and *S. heiseri*, are unequivocally distinct in the phenogram (Fig. 1) and in the PCoA analyses (Figs. 3a, 3b), so their status as different species is supported.

Solanum cochoae and *S. filiforme* are rare species indeed, each essentially represented by single collections (from Peru) (Correll 1962; Anderson and Bernardello 1991). Nonetheless, they are genetically distinct from the remainder of species. The reason for their rarity is not known; perhaps they too, like PNS1 and PNS2, are established peripheral isolates, or old established hybrids. The series *Caripensia* and the Caripense complex within it continue to be of interest: new species are discovered regularly when the group is studied intensively in the field or laboratory, such as *S. heiseri* (Anderson 1975), *S. cochoae* (Anderson and Bernardello 1991), the new taxa (PNS1 and PNS2) alluded to herein, or material sent to one of us (Anderson) by Lynn Bohs and Mike Nee from field work in the Andes.

Artificial hybrids in the greenhouse and experimental garden can be made between some of the species pairs within the Caripense complex, and some are at least partially fertile (Anderson 1975, 1977; Prohens and Nuez 2001), and on some occasions in the field, species have been observed growing in contiguous areas. This suggests that gene flow among various of these species could occur. Such gene flow would, of course, help explain the great morphological and molecular diversity of the most common species, *S. caripense*, and also why some elements (such as *S. caripense* PI243342) fall out associated with other groups (see phenogram, Fig. 1). Therefore, one might consider treating the whole Caripense complex as a single species as has been done in wild potatoes (van den Berg et al. 2002). However, each of the species possesses morphological features sufficient to allow them to be recognized as distinct entities. In addition, the molecular analyses make it clear that the four known species included in the Caripense complex (plus the new species, PSN1 and PSN2) can be separated from each other. In addition, no obvious hybrids have been identified in field settings, though we expect they occur. Furthermore, the topographically complex habitat that these Andean species occupy facilitates geographic isolation and poses barriers to pollen flow for these buzz-pollinated species. Given all these factors, and additionally that it is easier for plant breeders to identify potentially useful breeding material when natural

variation is distinguished by recognition of species (Prohens et al. 2003; Rodríguez-Burruezo et al. 2003), we maintain these entities as distinct species.

Species of the Caripense complex are possibly quite recent in origin, an hypothesis supported by the relatively large number of variant species and by the fact that intensive field work continues to reveal new variants and species. Intensive studies of morphology and intercrossability of *S. caripense* through much of its extensive range have failed to indicate any clear clines of variation (Anderson 1975, 1977, 1979). We also speculate that, based on its wide morphological and genetic variation and the patchiness of its rugged mountainous habitat, that the broadly distributed (Costa Rica through Peru; Correll 1962) *S. caripense* might be the ancestral plexis from which the new entities of this complex radiated, perhaps non-dichotomously, but as more-or-less simultaneous peripheral isolates via dispersal into remote areas. This parapatric model of speciation, perhaps involving metapopulations, is one of the chief ways in which Levin (1993, 2000) envisions new species being formed.

The data on AFLP diversity present in each species (with the exception of *S. cochoae*, *S. filiforme*, and *S. suaveolens*, for which only one accession each was studied) indicate that among the wild species there are no large differences in the diversity within self-compatible versus self-incompatible species. The greatest diversity is found in *S. caripense*, a self-incompatible species, but, as indicated, it is a species with a very wide distribution. The lowest diversity is found in *S. heiseri*, a species with a very limited distribution.

In conclusion, AFLP data have proved to be a powerful tool for examining the relationships among the wild species of *Solanum* sect. *Basarthrum*. Rare or uncommon species (e.g., *S. cochoae*, *S. filiforme*, *S. fraxinifolium*, *S. heiseri*, PNS1, and PNS2) that are difficult to distinguish morphologically have been supported as distinct elements of the same complex series. In others, shared morphological features (e.g., between the self-compatible *S. canense* and *S. suaveolens*) are revealed as products of convergence rather than common origin.

Finally, unfortunately, but perhaps revealingly, there is no clear hierarchical structure in the Caripense complex. This was also the result in previous molecular work (Anderson and Jansen 1998). This argues strongly for relatively rapid evolution in this group and/or potential hybridization among its members. The combination of both factors seems likely.

ACKNOWLEDGEMENTS. We thank the Dean of the College of Liberal Arts and Sciences at the Univ. of Connecticut, and the Univ. Politécnica de Valencia for support. Mona Anderson, Maryke Schlehofer, Robynn Shannon, Christopher Martine, Adrián Rodríguez, Mariola Plazas, Toya Cuenca, Lynn Bohs, Carlos Ochoa, Richard Lester, Matt Opel, and Paul Neal provided help with the

manuscript, greenhouse cultures and seed selection. The comments of two anonymous reviewers were helpful as well.

LITERATURE CITED

- ANDERSON, G. J. 1975. The variation and evolution of selected species of *Solanum* section *Basarthrum*. *Brittonia* 27: 209–222.
- . 1977. The variation and evolution of selected species of *Solanum* section *Basarthrum* (Solanaceae) II. *Brittonia* 29: 116–128.
- . 1979. Systematic and evolutionary consideration of *Solanum* section *Basarthrum*. Pp. 549–562 in *The biology and taxonomy of the Solanaceae*, eds. J. G. Hawkes, R. N. Lester, A. D. Skelding. Linnean Society Symposium Series No. 7. London: Academic Press.
- and L. M. BERNARDELLO. 1991. The relationships of *Solanum cochoae* (Solanaceae), a new species from Peru. *Novon* 1: 127–133.
- and P. G. GENSEL. 1976. Pollen morphology and the systematics of *Solanum* section *Basarthrum*. *Pollen et Spores* 13: 533–552.
- and R. K. JANSEN. 1998. Biosystematic and molecular systematic studies of *Solanum* section *Basarthrum* and the origin and relationships of the pepino (*S. muricatum*). *Monographs in Systematic Botany from the Missouri Botanical Garden* 66: 17–32.
- and D. A. LEVINE. 1982. Three taxa constitute the sexes of a single dioecious species of *Solanum*. *Taxon* 31: 667–672.
- , R. K. JANSEN, and Y. KIM. 1996. The origin and relationships of the pepino, *Solanum muricatum* (Solanaceae): DNA restriction evidence. *Economic Botany* 50: 369–380.
- , S. D. JOHNSON, P. R. NEAL, and G. BERNARDELLO. 2002. Reproductive biology and plant systematics: the growth of a symbiotic association. *Taxon* 51: 637–653.
- , T. P. STEINHARTER, and G. COOPER-DRIVER. 1987. Foliar flavonoids and the systematics of *Solanum* section *Basarthrum*. *Systematic Botany* 12: 534–540.
- BELKHIR, K., P. BORSA, L. CHIKHI, N. RAUFASTE, and F. BONHOMME. 2004. GENETIX, logiciel sous Windows™ pour la génétique des populations. Montpellier, France: Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II.
- BENHAM, J. 2001. Genographer 1.6. <http://hordeum.oscs.montana.edu/genographer>. Distributed by the author.
- BERNARDELLO, L. M. and G. J. ANDERSON. 1990. Karyotypic studies in *Solanum* section *Basarthrum* (Solanaceae). *American Journal of Botany* 77: 420–431.
- BRETÓ M. P., M. J. ASINS, and E. A. CARBONELL. 1993. Genetic variability in *Lycopersicon* species and their genetic relationships. *Theoretical and Applied Genetics* 86: 113–120.
- BROWER, A. V. Z., R. DE SALLE, and A. VOGLER. 1996. Gene trees, species trees, and systematics: a cladistic perspective. *Annual Review of Ecology and Systematics* 27: 423–450.
- BUNTJER, J. B. 2001. Phylogenetic computer tools, v. 1.3. Wageningen: Wageningen University.
- CORRELL, D. S. 1962. *The potato and its wild relatives*. Renner: Texas Research Foundation.
- FELSENSTEIN, J. 2004a. Inferring phylogenies. Sunderland: Sinauer Associates, Inc.
- . 2004b. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, Washington.
- FURINI, A. and J. WUNDER. 2004. Analysis of eggplant (*Solanum melongena*)-related germplasm: morphological and AFLP data contribute to phylogenetic interpretations and germplasm utilization. *Theoretical and Applied Genetics* 108: 197–208.
- HEISER, C. B. 1964. Origin and variability of the pepino (*Solanum muricatum*): a preliminary report. *Baileya* 12: 151–158.
- HENRY, R. J. 1997. Practical applications of plant molecular biology. London: Chapman & Hall.
- JONES, C. J., K. J. EDWARDS, S. CASTAGLIONE, M. O. WINFIELD, F. SALA, C. VAN DE WIEL, G. BREDEMEIJER, B. VOSMAN, M. MATTHES, A. DALY, R. BRETTSCHEIDER, P. BETTINI, M. BUIATTI, B. MAESTRI, A. MALCEVSCI, N. MARMIROLI, R. AERT, G. VOLCKAERT, J. RUEDA, R. LINACERO, A. VÁZQUEZ, and A. KARP. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381–390.
- KARDOLUS, J. P., H. J. VAN ECK, and R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87–103.
- LESTER, R. N. 1991. Evolutionary relationships of tomato, potato, pepino and wild species of *Lycopersicon* and *Solanum*. Pp. 283–301 in *Solanaceae III: taxonomy, chemistry, evolution*, eds. J. G. Hawkes, R. N. Lester, M. Nee, and N. Estrada. London: Royal Botanic Gardens Kew and Linnean Society of London.
- LEVIN, D. A. 1993. Local speciation in plants: the rule, not the exception. *Systematic Botany* 18: 197–208.
- LEVIN, D. A. 2000. The origin, expansion, and demise of plant species. New York: Oxford University Press.
- MARSHALL, J. A., S. KNAPP, M. R. DAVEY, J. B. POWER, E. C. COCKING, M. D. BENNETT, and A. V. COX. 2001. Molecular systematics of *Solanum* section *Lycopersicum* (*Lycopersicon*) using the nuclear ITS rDNA region. *Theoretical and Applied Genetics* 103: 1216–1222.
- MILBOURNE, D., R. MEYER, J. E. BRADSHAW, E. BAIRD, N. BONAR, J. PROVAN, W. POWELL, and R. WAUGH. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Molecular Breeding* 3: 127–136.
- MIONE, T. and G. J. ANDERSON. 1992. Pollen-ovule ratios and breeding system evolution in *Solanum* section *Basarthrum* (Solanaceae). *American Journal of Botany* 79: 279–287.
- MOHAMMADI, S. A. and B. M. PRASANNA. 2003. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Science* 43: 1235–1248.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.
- . 1987. *Molecular Evolutionary Genetics*. Pp. 176–187. New York: Columbia University Press.
- and R. H. CHESSE. 1983. Estimation of fixation indices and gene diversities. *Annals of Human Genetics* 47: 253–259.
- NUEZ, F., J. PROHENS, and J. M. BLANCA. 2004. Relationships, origin, and diversity of Galápagos tomatoes: implications for the conservation of natural populations. *American Journal of Botany* 91: 86–99.
- PROHENS, J. and F. NUEZ. 2001. The use of wild species in the improvement of pepino (*Solanum muricatum*), Pp. 297–306 in *Solanaceae V: advances in taxonomy and utilization*, eds. G. Barendse, R. van den Berg, T. Mariani, G. van der Weerden. Nijmegen: Nijmegen University Press.
- , G. J. ANDERSON, A. RODRIGUEZ-BURRUEZO, C. B. HEISER, and F. NUEZ. 2004. *Descriptors for pepino (Solanum muricatum)*. Rome: International Plant Genetic Resources Institute.
- , ———, ———, and F. NUEZ. 2003. Exploiting wild species for the genetic improvement of the pepino (*Solanum muricatum*). *Journal of Applied Botany* 77: 21–27.
- RICK, C. M., E. KESICKI, J. F. FÖBER, and M. HOLLE. 1976. Genetic and biosystematic studies on two new sibling species of *Lycopersicon* from Interandean Peru. *Theoretical and Applied Genetics* 47: 55–68.
- RODRÍGUEZ-BURRUEZO, A., J. PROHENS, and F. NUEZ. 2003. Wild relatives can contribute to the improvement of fruit quality in pepino. *Euphytica* 129: 311–318.
- SAKAMOTO, K. and T. TAGUCHI. 1991. Regeneration of intergeneric

- somatic hybrid plants between *Lycopersicon esculentum* and *Solanum muricatum*. *Theoretical and Applied Genetics* 81: 509–513.
- SIMPSON, M. J. A. 1979. A chemotaxonomic study of *Solanum* section *Basarthurum*. M.Sc. thesis. Birmingham: University of Birmingham.
- SPOONER, D. M., G. J. ANDERSON, and R. K. JANSEN. 1993. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (Solanaceae). *American Journal of Botany* 80: 676–688.
- , I. E. PERALTA, and S. KNAPP. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum* L. Section *Lycopersicon* (Mill.) Wettst.). *Taxon* 54: 43–61.
- STIEFKENS, L., G. BERNARDELLO, and G. J. ANDERSON. 1999. Karyotypic studies in artificial hybrids of *Solanum* sections *Anarrhichomenum* and *Basarthurum* (Solanaceae). *Australian Journal of Botany* 47: 147–155.
- TROGNITZ, F. and B. TROGNITZ. 2004. Mapping genes of *Solanum caripense* involved in resistance to *Phytophthora infestans*, the causal agent of potato late blight. Pp. 249–253 in *Genetic variation for plant breeding*, eds. J. Vollmann, H. Grausgruber, and P. Ruckebauer. Vienna: EUCARPIA & BOKU—University of Natural Resources and Applied Life Sciences.
- VAN DEN BERG, R. G., G. J. BRYAN, A. DEL RÍO, and D. M. SPOONER. 2002. Reduction of species in the wild potato *Solanum* section *Petota* series *Longipedicellata*: AFLP, RAPD and chloroplast SSR data. *Theoretical and Applied Genetics* 105: 1109–1114.
- YEH, F. C. and T. J. B. BOYLE. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129: 157.

***Capítulo 4.- AFLP and DNA sequence variation
in an Andean domesticate, pepino
(Solanum muricatum, Solanaceae):
implications for evolution and domestication.***

American Journal of Botany

AFLP AND DNA SEQUENCE VARIATION IN AN ANDEAN DOMESTICATE, PEPINO (*SOLANUM MURICATUM*, SOLANACEAE): IMPLICATIONS FOR EVOLUTION AND DOMESTICATION¹

JOSÉ M. BLANCA,² JAIME PROHENS,^{2,4} GREGORY J. ANDERSON,³ ELENA ZURIAGA,² JOAQUÍN CAÑIZARES,² AND FERNANDO NUEZ²

²Instituto para la Conservación y Mejora de la Agrodiversidad Valenciana, Universidad Politécnica de Valencia, Camino de Vera 14, 46022 Valencia, Spain; and ³Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs, Connecticut 06269 USA

The pepino (*Solanum muricatum*) is a vegetatively propagated, domesticated native of the Andes, where it grows with wild relatives. We used AFLPs and a 1-kb sequence of the 3-methylcrotonyl-CoA carboxylase gene to study variation of 27 accessions of *S. muricatum* and 35 collections of 10 species of wild relatives (*Solanum* section *Basarthrum*). A total of 298 AFLP fragments and 29 DNA sequence haplotypes were detected. Cluster and principal coordinate analyses and other genetic parameters estimated from both types of markers, show that *S. muricatum* is closely related to the species from one of the series (*Caripensia*) of section *Basarthrum* and that >90% of the variation of the cultigen is also represented in that series. Pepino is highly diverse, either because it is not monophyletic or it has been subjected to regular introgression with wild species, or both. Although a continuous distribution of the genetic variation occurred within the cultivated species, three genetic clusters were recognized. Cluster 1 is mostly centered in Ecuador, cluster 2 in Ecuador and Peru, and cluster 3 in Colombia and Ecuador. Cluster 3 also includes all modern cultivars studied. These results and other evidence suggest that northern Ecuador/southern Colombia is the main center of pepino diversity and the center of origin. The high genetic variation of this cultigen indicates that domestication does not always produce a genetic bottleneck.

Key words: AFLP; Andean region; crop evolution; DNA sequence; pepino; *Solanum muricatum*; *Solanum* section *Basarthrum*.

The pepino (*Solanum muricatum* Aiton) is an herbaceous Andean domesticate grown for its juicy and aromatic fruits. Although it was a very important crop in the Andean region in pre-Columbian times (Prohens et al., 1996), its 20th century prominence has not equaled that of its close relatives the tomato (*Solanum lycopersicum* L.) and potato (*Solanum tuberosum* L.). However, in the last three decades, there has been growing interest in the pepino from exotic fruit markets, and its cultivation has spread from its ancestral home in the Andes of South America to other countries such as New Zealand, Spain, and the Netherlands (Nuez and Ruiz, 1996).

The study of the molecular variation of the pepino is of interest for several reasons. Although the seeds of pepino plants are fertile and produce vigorous offspring, this crop is primarily propagated vegetatively by cuttings (Heiser, 1964; Anderson, 1979; Morley-Bunker, 1983), and as a consequence, its genetic structure could be different from that of seed-propagated crops. Also, the pepino is fairly easy to intercross with several wild species, and the cultigen is grown in the areas of natural distribution of some of the wild species (Heiser,

1964; Anderson, 1975, 1977); therefore, natural hybridization and introgression with the closely related wild species is likely (e.g., Anderson et al., 1996). Pepino materials grown in Andean South America correspond to local landraces that have not been subjected to modern breeding (Nuez and Ruiz, 1996), but there are also many cultivars that have been developed elsewhere (e.g., Dawes and Pringle, 1983; Prohens et al., 2002). Because of this, the pepino is an interesting model for the study of (a) the intraspecific variation structure and (b) relationships with wild relatives, in a vegetatively propagated crop with a wide “natural” distribution where the crop still grows in close proximity to its closest wild relatives. Studying intraspecific variation and relationships with wild relatives will provide information for understanding the evolution and domestication of the pepino specifically and of the process of domestication generally.

Previous studies of the molecular variation of pepino have included analyses of patterns of flavonoids (Anderson et al., 1987) and cp-DNA restriction fragment length polymorphisms (RFLPs) (Spooner et al., 1993; Anderson et al., 1996; Anderson and Jansen, 1998). These first molecular studies determined that the wild species most closely related to the pepino were those of the series *Caripensia* (*S. basendopogon* Bitter, *S. caripense* Humb. and Bonpl. ex Dun., *S. cochoae* G. J. Anderson and Bernardello, *S. filiforme* Ruiz López and Pavón, *S. fraxinifolium* Dunal in DC, *S. heiseri* G. J. Anderson, and *S. tabanoense* Correll) and that intraspecific diversity existed. However, the molecular variation assessed with these markers was limited by the amount of the genome covered and was not appropriate for a detailed study of intraspecific variation of the cultigen, *S. muricatum*.

Polymorphisms in the DNA sequence of an individual gene

¹ Manuscript received 12 June 2006; revision accepted 14 May 2007.

The authors thank the following for support: the National Science Foundation, the Dean of the University of Connecticut College of Liberal Arts and Sciences, the Ministerio de Ciencia y Tecnología (RF2004-00002-00-00), the European Union (RESGEN PL98-113), and the Universidad Politécnica de Valencia. M. Anderson assisted with field collections, C. Morse and C. Martine helped with greenhouse work, M. Plazas and A. Rodríguez assisted with greenhouse and molecular work, and M. Opel provided critical comments on an earlier draft of the manuscript. The authors are grateful to these contributors and also to C. Heiser, whose work originally sparked their interest in pepino.

⁴ Author for correspondence (e-mail: jprohens@btc.upv.es)

can provide information on the evolutionary history of a species (Kimura, 1983), and they have been used to study domestication in cultivated plants (Schaal and Olsen, 2000; Olsen and Purugganan, 2002; Clark et al., 2004). These data, however, provide information on only a limited portion of the genome; therefore, multivariate techniques applied to sets of multilocus markers, such as the amplified fragment length polymorphisms (AFLP; Vos et al., 1995), may be more useful in identifying groups of genetically similar individuals. AFLPs show a high level of polymorphism in *Solanaceae* (Kardolus et al., 1998; Nuez et al., 2004; Spooner et al., 2006), and their use allows a large number of loci to be scored in a single reaction with much better repeatability among laboratories than other markers such as random amplified polymorphisms of DNA (RAPDs) (Jones et al., 1997). Furthermore, AFLPs have proved to be of great interest for the study of domestication processes in potato (Spooner et al., 2006), a crop closely related to pepino. Other types of markers that could be useful in such studies, such as simple sequence repeats (SSRs), are not available in the pepino. In this work, we use an approach that combines the variation in AFLPs and in the sequence of a nuclear gene to study the genetic relationships of the cultivated pepino to its closely related wild species, as well as the structure of its genetic variation over its natural range, i.e., that associated with its presumed region of origin and pre-Columbian cultivation in Colombia, Ecuador, and Peru (Heiser, 1964; Anderson, 1975; Anderson et al., 1996; Prohens et al., 1996). Because two complementary molecular techniques are used, results obtained are more robust than those based on a single technique. Although our study focuses on pepino, our data are relevant to a general understanding of evolution and domestication of crops.

MATERIALS AND METHODS

Plant materials—Materials studied included 22 accessions of *Solanum muricatum* that we have classified as the typical cultivated *S. muricatum* (our *S. muricatum* sensu stricto). In addition, there were five accessions that we consider to be putative introgressant forms, or natural hybrids between *S. muricatum* and wild species of the series *Caripensia*: we have called these *S. muricatum* sensu lato (Table 1).

To assess the relationships of the pepino with its wild relatives, we included 35 accessions representing the 12 closely related species from *Solanum* section *Basarthurum* (Table 2). The number of accessions included for each of the wild species depended on abundance (e.g., one accession for rare species) and availability of seeds. Among these materials there is an interspecific hybrid (*S. tabanoense* × *S. caripense*) as well as an odd accession of *S. caripense* (PI-243342) that, based on morphological characteristics, seems to have been introgressed with genes from *S. muricatum* (presumably during its repeated propagation in a seed bank). These accessions were also part of a separate AFLP study on the systematics of wild *Basarthurum* species (Prohens et al., 2006). Included among the wild species in the latter study were two newly described species (*S. catilliflorum* G. J. Anderson, Martine, Prohens and Nuez and *S. perlongistylum* G. J. Anderson, Martine, Prohens and Nuez) (Anderson et al., 2006).

To understand the distribution of genetic variation, we compared genetic clusters (derived from the AFLP analyses described in the “Results” section, subsection Structure of variation in the cultigen). This approach has been successfully applied to studies of other domesticates (Harter et al., 2004; Song et al., 2006).

Generation of molecular data—For each accession, genomic DNA was extracted from one plant using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA) following the protocol of the manufacturer. The AFLP analyses (Vos et al., 1995) were carried out as described elsewhere (Nuez et al., 2004), using three selective combinations of primers (EcoNed: AGACTGCG

TACCAATTCACC-Mse: GATGAGTCCTGAGTAAc, EcoFam: AGACTGCGTACCAATTCCTT-Mse: GATGAGTCCTGAGTAACTT, Eco-Hex: AGACTGCGTACCAATTCCTC-Mse GATGAGTCCTGAGTAA). DNA fragments were separated in an ABI Prism 310 genetic analyzer (Applied Biosystems, Foster City, California, USA).

A preliminary survey carried out by sequencing six nuclear regions (CD15, CT81, CT165, CT267, CD23, CT137) from the tomato map (Tanksley et al., 1992) on several accessions of pepino and five other species of the closely related *Solanum* section *Lycopersicon* subsection *Lycopersicon* (i.e., *S. chilense*, *S. habrochaites*, *S. lycopersicum*, *S. peruvianum* and *S. pimpinellifolium*) (Spooner et al., 2005) showed that the most variable region was CT137, which corresponds to a 3-methylcrotonyl-CoA carboxylase gene (MCC) (McKean et al., 2000). Therefore, this gene was chosen for further analysis in our materials.

The PCR amplification of the MCC gene was done with primers 006d (GTCCAAATCTGAGGCAAGTGG) and 006r (GGTAGCAGAGATG CAAAGGCTC). After amplification, the PCR products were treated with Exonuclease I and precipitated with NaCl and ethanol. DNA sequencing was performed using an ABI Prism 3100 genetic analyzer following manufacturer instructions for the BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems). Four sequencing reactions were carried out to obtain a high-quality sequence of each PCR product. Apart from the primers 006d and 006r, internal primers 006d2 (GCTAAAGTTGTAGAGCGGTATGAC) and 006r2 (AAAAGCAGCCACGGAACCACA) were used for amplification.

Because all species analyzed are diploid (Bernardello and Anderson, 1990; Anderson et al., 2006) and we did not find any evidence of more than two haplotypes per individual, we assumed that the materials studied have a single copy of the MCC gene. Consequently, we considered the individuals as homozygous or heterozygous depending on whether one or two haplotypes, respectively, were found. Heterozygotes were detected as double peaks in the chromatogram. In most cases, both haplotypes were inferred through haplotype subtraction (Clark, 1990). In the cases with deletions in heterozygosis or with poor chromatogram quality, PCR products were cloned using the pGEM-T vector System II (Promega, Madison, Wisconsin, USA), and each haplotype was sequenced separately.

AFLP data analysis—The AFLP fragments were scored as binary traits (1 = present, 0 = absent) using Genographer 1.6 software (Benham et al., 1999). Because AFLPs are dominant and anonymous markers, phenetic methods are appropriate (Koopman et al., 2001). Principal coordinate analyses (PCoA) based on Dice distance matrices were performed using the NTSys 2.02g software package (Rohlf, 1996). Nei distances (Nei and Li, 1979) were calculated for 1000 bootstrapped data matrices using Phyltools 1.32 software (Buntjer, 1997), and a neighbor-joining phenogram was built with the Phylip 3.62 package (Felsenstein, 1989). Based on this analysis, the genetic clusters were defined. Genetic diversity (h) (Nei, 1973) was estimated with the Popgene 1.32 software (Yeh and Boyle, 1987). The gene differentiation among genetic clusters (G_{st}) was calculated with the Genetix Software 4.05.2 (Belkhir et al., 2004) using the unbiased estimator (Nei and Chesser, 1983). One hundred permuted data sets were also prepared with the Genetix software.

DNA sequence data analysis—The Staden package v1.4 (Bonfield, 2004) was used to obtain complete sequences from the partial readings, and ClustalW v1.82 (Thompson et al., 1994) was used to align them. The cluster analysis was done using the Kimura 2 parameter as an evolution model and neighbor-joining as a clustering method using the Phylip package (Felsenstein, 1989). DNAsp software 4.0.6 (Rozas et al., 2003) was used to calculate the recombination parameter R (Hudson, 1987); the minimum number of recombination events R_m (Hudson and Kaplan, 1985); the genetic differentiation between groups, K_{st} (Hudson et al., 1992, equation 11); and the neutrality tests (Tajima, 1989; Fu and Li, 1993). Nucleotide diversity (π [Nei, 1987, equation 10.5] and θ [per site] [Nei, 1987, equation 10.3]) and haplotype diversity (Nei, 1987, equation 8.4) are means of the values calculated by DNAsp for two groups. To form these two groups, a randomly chosen allele from each accession was included in group 1 and the other allele in group 2.

RESULTS

AFLP and DNA sequence diversity—Scoring of the AFLP fragments in the range of 60–380 base pairs resulted in 298

TABLE 1. *Solanum muricatum* collections studied.

| Taxon | Accession | Collector | Origin ^a |
|--------------------------------|----------------------------|---------------------------------|---|
| <i>S. muricatum</i> Aiton s.s. | 32 ^b | B. Pickersgill | Peru (Ayacucho) |
| | 35 ^b | B. Pickersgill | Peru (Ayacucho) |
| | 98 ^b | B. Pickersgill | Peru (Chimbote) |
| | 642 ^b | G.J. Anderson | Ecuador (Quito) |
| | 734 ^b | G.J. Anderson | Colombia (San Antonio) |
| | 844 ^b | G.J. Anderson | Ecuador (Patate) |
| | 910 ^b | G.J. Anderson | Peru (Chiclayo) |
| | 913 ^b | G.J. Anderson | Peru (Lima) |
| | 1450 ^b | L. Bohs | Ecuador (Baños) |
| | 1451 ^b | L. Bohs | Peru (Huancayo) |
| | 6318 ^b | C.B. Heiser | Colombia (Titaitata) |
| | 7038 ^b | C.B. Heiser | Ecuador |
| | 37-A ^c | F. Nuez | Ecuador (Gualaceo) |
| | E-37 ^c | F. Nuez and J. Prohens | Ecuador (Riobamba) |
| | EC-46 ^c | F. Nuez and J. Prohens | Ecuador (Loja) |
| | El Camino ^{c,d} | Commercial cv. | New Zealand |
| | JP-261 ^{c,d} | J. Prohens | Unknown (supermarket in Spain) |
| | Kawj ^{c,d} | Commercial cv. | New Zealand |
| | Puzol ^{c,d} | Commercial cv. | Spain (hybrid New Zealand × Chilean material) |
| | Sweet Round ^{c,d} | Commercial cv. | Spain (from Chilean material) |
| Viru ^c | C. Gisbert | Peru (Trujillo) | |
| Vetas Verdes ^c | F. Nuez and J. Prohens | Ecuador (Ibarra) | |
| <i>S. muricatum</i> Aiton s.l. | 751 ^b | G.J. Anderson | Colombia (Popayán) ^e |
| | 868 ^b | G.J. Anderson | Ecuador (Imbabura) |
| | EC-12 ^c | F. Nuez and J. Prohens | Ecuador (Loja) ^f |
| Popayan ^c | F. Nuez | Colombia (Popayán) ^g | |
| SL ^c | Breeding line | Chile ^h | |

^a Indicates the country of origin. When known, the city or village closest to the collection site is given in parentheses for the local Andean materials. For the other materials, we give all the collection information we have (also in the parentheses).

^b Accession from the G.J. Anderson collection. Vouchers in CONN.

^c Accession from the J. Prohens collection. Seeds deposited in the germplasm bank of the Instituto de Conservación y Mejora de la Agrodiversidad Valenciana.

^d “Modern” cultivars.

^e Fertile plant growing near wild pepino relatives.

^f Derived from seed from a typical pepino fruit, but resultant plants morphologically reminiscent of hybrids or backcrosses with a wild relative.

^g Sterile plant, as is common for interspecific pepino hybrids.

^h Cultivar of uncertain Chilean origin, with high-sugar, low-yield fruits typical of some interspecific hybrids and wild pepino relatives.

bands, all of which were polymorphic. For *S. muricatum*, 204 of the 298 bands were present in at least one of the accessions, and all were polymorphic as well.

One thousand and sixty-six nucleotide sites of the MCC gene were examined. Of these, 676 correspond to noncoding regions (three introns and the 3' untranslated region) and 390 to coding fragments (three exons) (Fig. 1). Nucleotide substitutions amount to 9.5% in the noncoding sites and 3.5% in the coding regions. Four insertions/deletions (indels) were present, two corresponding to single nucleotides, one to 2 bp, and the other to 55 bp.

A total of 29 haplotypes were considered (Fig. 1). Of these, *S. muricatum* presents seven, *S. caripense* seven, *S. basendopogon* four, and *S. tabanoense* two; the remaining species were characterized by a single haplotype each. The complete nucleotide sequences were deposited at GenBank (accession numbers EF200591–EF200696). The heterozygosity for the MCC locus in *S. muricatum* is 0.5. The haplotypes muricatum 3a and 3b, which represent 9.6% of the haplotypes found in *S. muricatum*, are identical to the *S. caripense* haplotypes caripense 1a and caripense 1b. *Solanum heiseri* has a haplotype identical to *S. caripense* haplotype caripense 1a and to *S.*

muricatum haplotype muricatum 3a, and the *S. filiforme* haplotype has the pepino haplotype muricatum 1b.

Tajima and Fu and Li's tests were performed to test for neutrality on the *S. muricatum* and *S. caripense* sequence sets. Nonsignificant results were obtained with both tests (data not shown).

Relationships between cultivated and wild relatives—The AFLP data were subjected to a series of PCoA analyses. In the first PCoA, all the wild and cultivated accessions were included (Fig. 2A). In this case, the two main axes account for 27.7% of the variation. Despite the distance between the main *S. muricatum* group and the species of series *Caripensia*, a continuous dispersion of accessions is found between them. The *S. muricatum* accessions located closer to series *Caripensia* accessions are: SL, Popayan, 751, EC-12, and 868, and as indicated in Table 1, these are considered putative hybrids or introgressed forms (the *S. muricatum* sensu lato group). The species furthest from *S. muricatum* (*S. canense*, *S. basendopogon*, *S. catilliflorum*, and *S. perlongistylum*) were removed to more carefully examine the relationships among *S. muricatum* and the most closely allied wild relatives, and the PCoA was rerun. In this new analysis (data not shown), *S.*

TABLE 2. Wild *Solanum* section *Basarthurum* collections studied.

| Taxon | Accession | Collector | Origin ^a |
|--|--------------------------|------------------------|--|
| <i>S. basendopogon</i> Bitter | 38 ^b | J.G. Hawkes | Peru (Dep. La Libertad) |
| | 43 × 44 ^b | J.G. Hawkes | Peru (Dep. La Libertad) |
| | BIRM/S 0032 ^c | J.G. Hawkes | Peru (Dep. La Libertad) |
| <i>S. canense</i> Rybd. | 46 ^b | D'Arcy | Panama (Canita) |
| | 498 ^b | G.J. Anderson | Costa Rica (Monte Verde, Prov. Alajuela) |
| | 509 ^b | G.J. Anderson | Costa Rica (near Tilarán, Prov. Guanacoste) |
| <i>S. caripense</i> Humb. and Bonpl. ex Dun. | 220 × 239 ^b | G.J. Anderson | 220: Ecuador (near Tixan, Prov. Chimborazo) 239: Ecuador (near Urdaneta, Prov. Loja) |
| | BIRM/S 1034 ^c | G.J. Anderson | Ecuador |
| | BIRM/S 1251 ^c | A. Child | Ecuador (Prov. Loja) |
| | E-7 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Pichincha) |
| | EC-36 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Loja) |
| | EC-40 ^b | F. Nuez and J. Prohens | Ecuador (Prov. Loja) |
| | PI-243342 ^c | Unknown | Costa Rica |
| <i>S. catilliflorum</i> G. J. Anderson, Martine, Prohens and Nuez | P-71 ^c | F. Nuez and J. Prohens | Peru (Prov. Abancay, Dep. Apurímac) |
| | P-74 ^c | F. Nuez and J. Prohens | Peru (Prov. Abancay, Dep. Apurímac) |
| | P-80 ^c | F. Nuez and J. Prohens | Peru (Prov. Abancay, Dep. Apurímac) |
| | P-82 ^c | F. Nuez and J. Prohens | Peru (Prov. Abancay, Dep. Apurímac) |
| <i>S. cochoae</i> G. J. Anderson and Bernardello | 934 ^b | C. Ochoa | Peru (Prov. Amazonas; Dep. Loreto) |
| <i>S. filiforme</i> Ruiz López and Pavón | 1452 ^b | C. Ochoa | Peru (Prov. Tarma, Dep. Junín) |
| <i>S. fraxinifolium</i> Dunal in DC | 500 ^b | G.J. Anderson | Costa Rica (near Monte Verde, Prov. Puntarenas) |
| | BIRM/S 1809 ^c | B. Pickersgill | Peru (Selva de San Francisco, Dep. Ayacucho) |
| <i>S. heiseri</i> G. J. Anderson | 6800 ^b | C.B. Heiser | Colombia (Cerro de Monserrate, Dep. Bogotá) |
| | BIRM/S 1979 ^c | C.B. Heiser | Colombia |
| <i>S. perlongistylum</i> G. J. Anderson, Martine, Prohens and Nuez | P-51 ^c | F. Nuez and J. Prohens | Peru (Prov. Huanta, Dep. Ayacucho) |
| | P-61 ^c | F. Nuez and J. Prohens | Peru (Prov. La Mar, Dep. Ayacucho) |
| | P-62 ^c | F. Nuez and J. Prohens | Peru (Prov. La Mar, Dep. Ayacucho) |
| | BIRM/S 1122 ^c | Unknown | Peru (Dep. Cusco) |
| | BIRM/S 1573 ^c | Unknown | Colombia (Caguan, Dep. Huila) |
| <i>S. tabanoense</i> Correll | EC-22 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Loja) |
| | EC-26 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Loja) |
| | E-257 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Loja) |
| <i>S. tabanoense</i> × <i>S. caripense</i> hybrid | 769 × 239 ^b | G.J. Anderson | 769: Colombia (near El Encanto, Dep. Amazonas) 239: Ecuador (near Urdaneta, Prov. Loja) |
| <i>S. trachycarpum</i> Bitter and Sodiro | 7087 ^b | C.B. Heiser | Ecuador |
| | 169 ^b | G.J. Anderson | Ecuador (Mitad del Mundo, Prov. Pichincha) |
| | E-34 ^c | F. Nuez and J. Prohens | Ecuador (Prov. Cotopaxi) |

^a Indicates the country of origin. When known, the place of collection, province, or department is given within parentheses.

^b Accession from the G.J. Anderson collection. Vouchers in CONN.

^c Accession from the J. Prohens collection. Seeds deposited in the germplasm bank of the Instituto de Conservación y Mejora de la Agrodiversidad Valenciana.

trachycarpum lies far away in the second axis from the group formed by *S. muricatum* and the remaining species of the series *Caripensia*, so the PCoA analysis was run again without *S. trachycarpum*. This new analysis (Fig. 2B) shows that the group formed by *S. muricatum* sensu stricto plus *S. filiforme* and *S. cochoae* are separated in the first axis from the wild species. The *S. muricatum* sensu lato accessions and the odd *S. caripense* PI-243342 are located in the middle of these two groups.

A phenogram was constructed with the AFLP data (Fig. 3A). All the *S. muricatum* accessions are grouped under a node

supported by an 89% bootstrap value. This group also includes *S. cochoae*, *S. filiforme*, and *S. caripense* PI-243342. All *S. muricatum* sensu lato accessions (except 'Popayan' and SL) are grouped together in one part of the diagram in a basal position in relation to the bulk of the wild species. The lack of bootstrap support in the structure of the lower nodes makes it difficult to interpret which of the rest of the wild species is most closely associated with the pepino.

Structure of variation in the cultigen—The AFLP data were subjected to a third PCoA that focused only on the *S.*

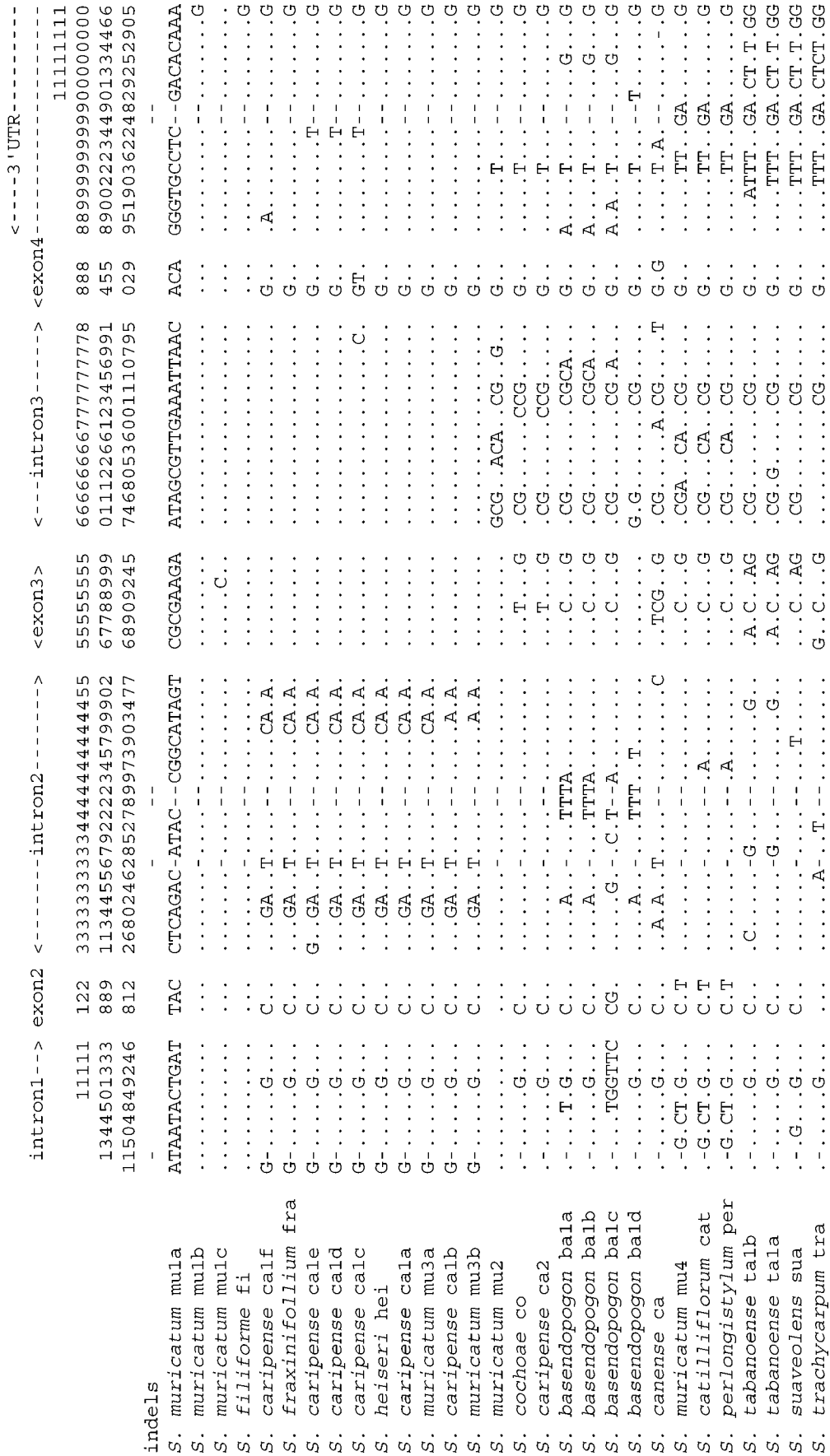


Fig. 1. Multiple sequence alignment of variable positions in *Solanum* section *Basaristrum* accessions. Positions conserved relative to haplotype 1a are shown as dots and deletions as hyphens. Numbering is indicated above the alignment for each position. Indels between positions 944 and 988 are 55 nucleotides long. Location of exon-intron boundaries are shown by vertical empty lines.

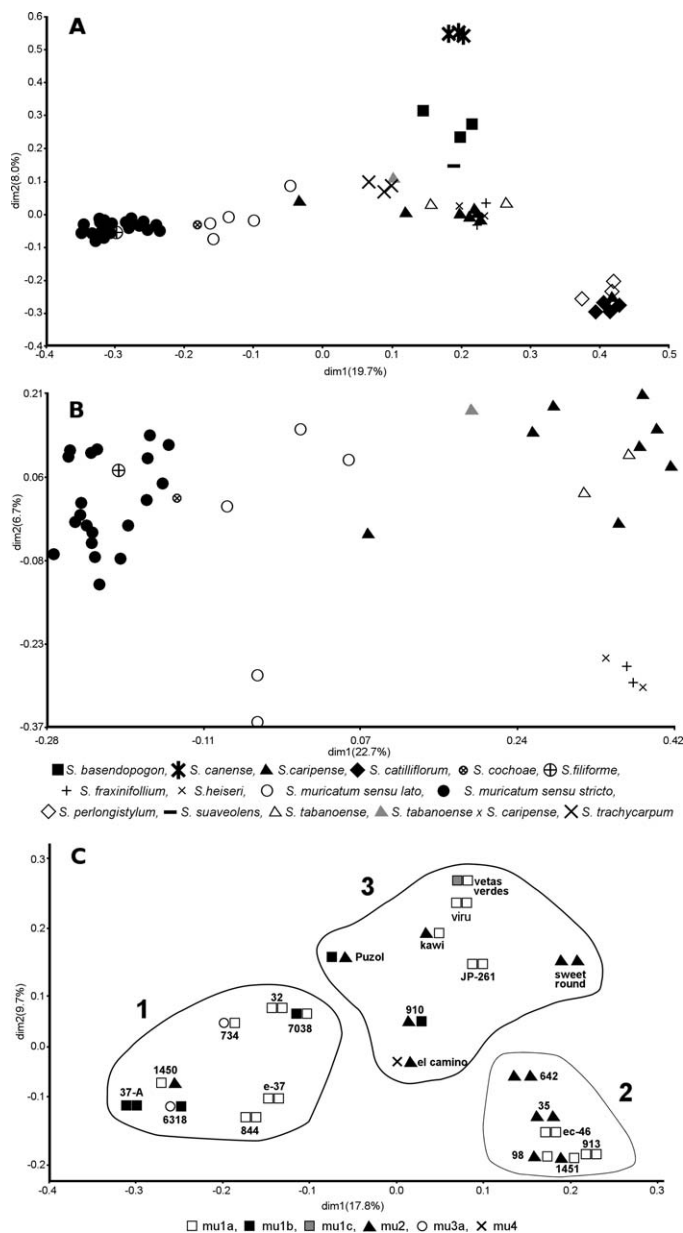


Fig. 2. Principal coordinate analysis (PCoA) plot of (A) *Solanum muricatum* and all studied wild species of *Solanum* section *Basarthurum*, (B) *S. muricatum* with species only from the *S. caripense* complex, and (C) accessions of *S. muricatum* s.s. In (C), each accession is represented by its haplotype composition. Axis labels indicate the percentage explained by the first (dim1) and second (dim2) dimensions of the PCoA.

muricatum sensu stricto accessions (Fig. 2C). Even though a continuous distribution is evident, the *S. muricatum sensu stricto* accessions can be assembled into three genetic clusters that appear after the construction of a neighbor-joining tree built with the AFLP data from the *S. muricatum sensu stricto* accessions (data not shown). Two clusters (labeled 1 and 2) are clearly differentiated (Fig. 2C) and were supported by bootstrap values of 40 and 55, respectively, in the AFLP-based phenogram (following a similar approach to that of Harter et al., 2006). Cluster 3 seems to be intermediate between clusters 1 and 2.

Genetic cluster 1 is centered in Ecuador, with an outlier in southern Peru (Fig. 4). Genetic cluster 2 is located in Peru and Ecuador. Both of these clusters include only traditional landraces from the Andean region. Cluster 3 is a mixture of traditional varieties from Peru and Ecuador as well as modern cultivars developed all around the world.

The genetic diversity within the genetic clusters of *S. muricatum* is relatively high and quite variable (Table 3). The genetic cluster 2 is the least diverse based on AFLP estimates and nucleotide diversity. Genetic cluster 3, which, apart from local landraces, contains the modern cultivars, had the greatest diversity. *Solanum muricatum* is as variable as the very widespread wild species *S. caripense*; perhaps not surprisingly, genetic cluster 3 by itself is almost as diverse as *S. caripense*.

The haplotype diversity data also confirm that *S. muricatum* is a highly diverse species (Table 3). The *S. muricatum* haplotypes are widely dispersed in the neighbor-joining tree constructed using the MCC gene data (Fig. 3B). For instance, two of the haplotypes of *S. muricatum* are grouped with those of *S. heiseri*, *S. fraxinifolium*, and the *S. caripense* haplotypes (all except *caripense* 2). Three others (*muricatum* 1a, 1b, 1c) are linked to *S. filiforme* and, with an 86% bootstrap value, to the former group. Haplotype *muricatum* 4 is quite different from the other *S. muricatum* haplotypes and is nearly identical to the shared haplotype found in *S. catilliflorum* and *S. perlongistylum*.

The positions of three of the *S. muricatum* haplotypes mu1a, mu1b, and mu1c in the tree (Fig. 3B) could indicate that other unstudied materials were the source of some of the variation in the pepino. However, careful study of the data itself, i.e., the sequences given in Fig. 1, shows that these haplotypes could result from recombination within the gene studied. For example, haplotypes *muricatum* 1b and 2 accumulate their 11 nucleotide differences in a fragment of just 316 bp while the rest of the sequence is identical. And haplotype *muricatum* 1b could result from a recombination between two *S. caripense* haplotypes (*caripense* 2 and *caripense* 1b) plus four additional mutations. The statistics testing genetic differentiation among species also support the contention that the cultivated pepino and the wild relatives share much of the molecular diversity described. In support of that hypothesis, note that only 9% of the markers in *S. muricatum sensu stricto* are unique (i.e., not found in the wild species of series *Caripensia*). Adding further support to this argument is that two thirds of the markers in the wild species are found within *S. muricatum sensu stricto*.

DISCUSSION

The origin of most domesticates turns out to be more complicated than supposed at first analysis and may involve different scenarios, such as polyphyletic origins and the existence of complex interrelationships among the wild relatives and the domesticate (Olsen and Schaal, 1999; Nesbitt and Tanksley, 2002; Clark et al., 2004; Sukhotu et al., 2004). The evolutionary history of the pepino is no exception. Four decades of previous study have yielded many hypotheses on the origin of this Andean cultigen (e.g., Correll, 1962; Heiser, 1964, 1969; Brücher, 1966). The earliest hypotheses suggested different single species as possible progenitors (e.g., Heiser, 1964; Brücher, 1966). On the basis of a wide range of analyses utilizing all the tools in systematic botany/evolutionary research, indicate that, other species have probably been

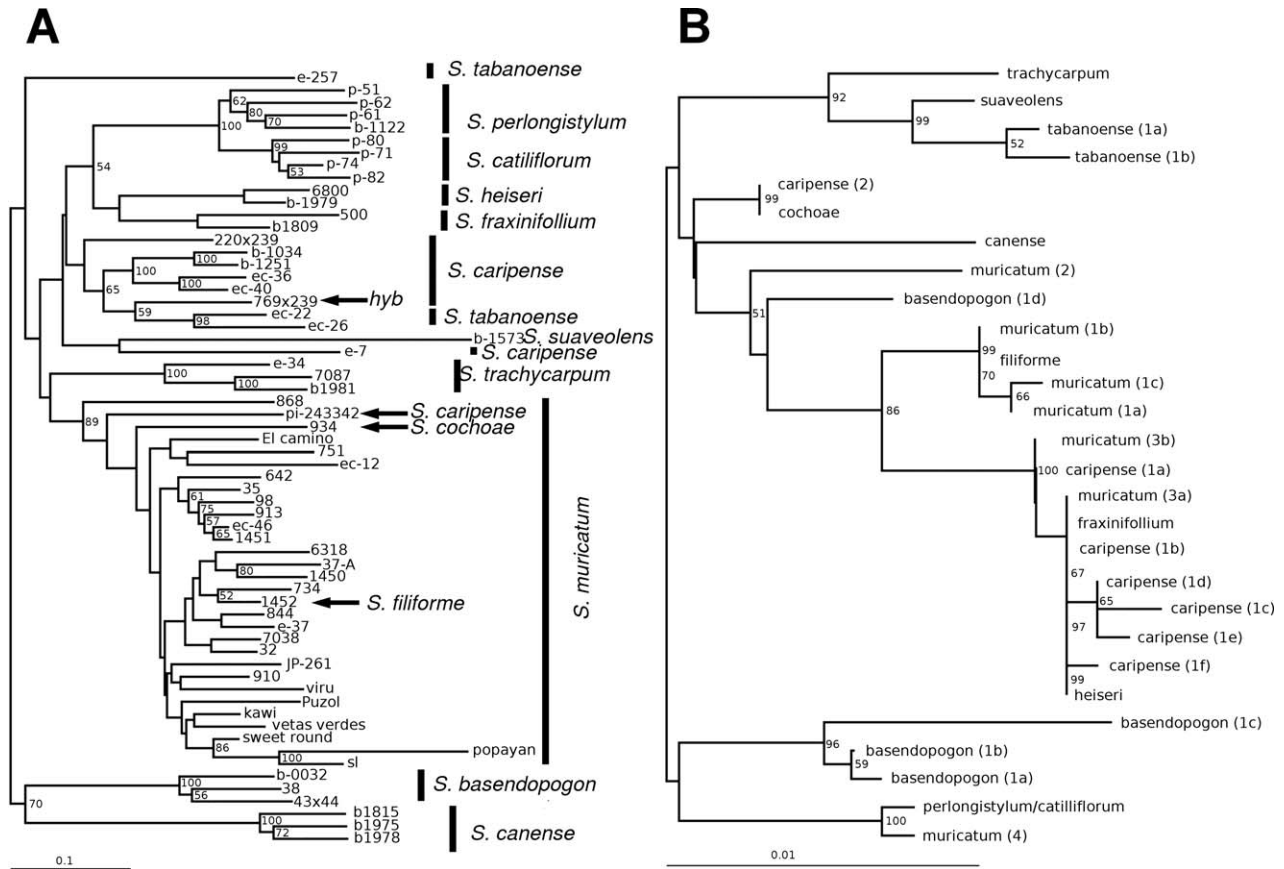


Fig. 3. Trees based on neighbor-joining analysis of (A) AFLP data of all accessions included in the study and (B) haplotypes derived from the MCC gene sequence. Haplotype codes are indicated in Fig. 1. In both trees, branch lengths are proportional to distance. Bootstrap values greater than 50% are shown in both trees. Bars represent the genetic distances.

involved in the diversification of the pepino through post-origin hybridization, even if the pepino originated from a single species (e.g., Anderson and Jansen, 1998; Prohens et al., 2003). The molecular tools available over the last 15 years have made our hypotheses about the origin of domesticates, the species, locations, and especially the underlying processes (e.g., interspecific hybridization) much more precise. The genetic data we present here, based on two complementary types of molecular data, strengthen our understanding of the origins of this complicated cultigen, and we hope they also demonstrate that the origin and evolution of domesticates—a process involving human manipulation—is not the simple one-directional process often envisioned.

Relationships between the cultigen and the wild relatives—The AFLP data indicate that *S. muricatum* is genetically close to *S. filiforme*, *S. cochoae*, and *S. caripense* PI-243342, relationships that could lead to the interpretation that these accessions are the wild ancestors of pepino. However, we prefer to be cautious and to take into account other information from other studies that indicate that these accessions represent special cases. *Solanum caripense* accession PI-243342 is quite different from the rest of the *S. caripense* accessions and became part of our study via stock from a seed bank. Because of its reputed origin from an area near the northern edge of the range of *S. caripense* (Costa Rica) and considerably outside of

the ancient range for *S. muricatum* and because it resembles materials derived from artificial hybrids between *S. caripense* and *S. muricatum*, we suggest that this unusual accession may be derived from hybridization between the original Costa Rican *S. caripense* accession and some material of *S. muricatum*, presumably during its repeated multiplication ex situ in germplasm collections. *Solanum filiforme* and *S. cochoae* are rare species, and only one accession of each is available. The haplotype found in *S. filiforme* is very similar to those of *S. muricatum*, while that from *S. cochoae* is identical to one of the *S. caripense* haplotypes. Thus, these species could be true wild species closely related to the origin of *S. muricatum* or, more likely, stabilized hybrids (and a fertile, sexually reproducing entity in the case of *S. cochoae*; Anderson and Bernardello, 1991) between *S. muricatum* and a wild species, probably the widespread and highly polymorphic *S. caripense*.

Apart from the three exceptions mentioned, *S. muricatum* is not linked with a single wild species either in the AFLP or in the haplotype trees. Furthermore, in the haplotype tree, *S. muricatum* sequences are dispersed through the tree. This suggests that several wild species have contributed to the origin and evolution of *S. muricatum*. A similar lack of association of the domesticate with a single wild species and the dispersion of the cultigen across parts of the tree were described previously with cpDNA data in *S. muricatum* (Anderson et al., 1996). Analogous scenarios have been described for other cultivated

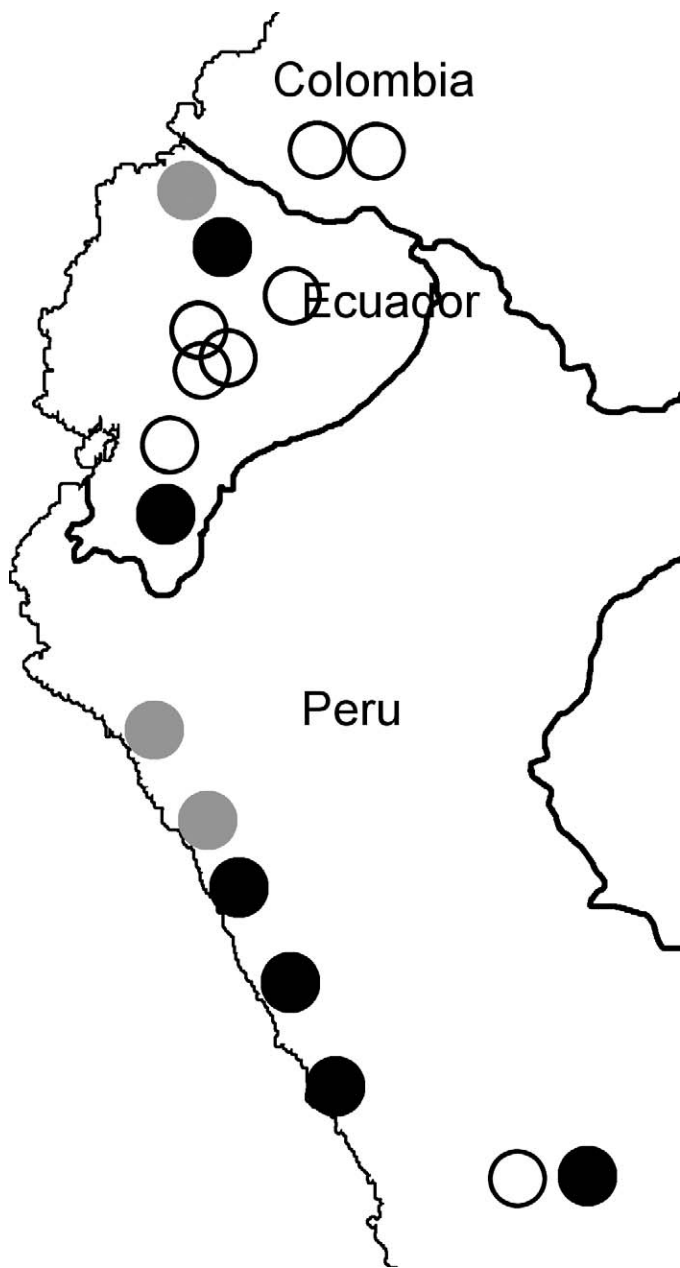


Fig. 4. Geographic locations of the Andean *Solanum muricatum* accessions. Different symbols represent accessions from different *S. muricatum* genetic clusters (from Fig. 2C): open circles for genetic cluster 1, solid circles for cluster 2, gray circles for cluster 3.

species (e.g., Olsen and Schaal, 1999; Clark et al., 2004). It seems clear from the haplotype data that several wild species have either contributed to the *S. muricatum* genome or have become associated with it through hybridization. For instance, haplotypes muricatum 3a and 3b are identical to caripense haplotypes 1a and 1b from *S. caripense*, and haplotype muricatum 4 is only two nucleotides different from the *S. catilliflorum* and *S. perlongistylum* haplotypes. This hypothesis of hybridization is strengthened by the fact that the haplotypes are not only found in the putative pepino \times wild species hybrids (*S. muricatum* sensu lato) but also in some of the *S. muricatum* sensu stricto accessions. However, some of the *S.*

TABLE 3. Estimates of diversity based on AFLP (h ; Nei, 1973) and sequence-based indexes (θ [per site] [Nei, 1987, equation 10.3], nucleotide diversity π [Nei, 1987, equation 10.5], and haplotype diversity H_d [Nei, 1987, equation 8.4]) for *Solanum muricatum* and the wild species with a larger number of accessions.

| Groups | AFLP | | Sequence | | | | |
|---|------------|------|-----------|------------|----------|--------|-------|
| | Accessions | h | Sequences | Haplotypes | θ | π | H_d |
| <i>S. muricatum</i> | | | | | | | |
| (all accessions) | 27 | 0.19 | 52 | 7 | 0.0058 | 0.0068 | 0.71 |
| <i>S. muricatum</i> s.s. | 22 | 0.15 | 44 | 6 | 0.0057 | 0.0061 | 0.69 |
| Cluster 1 | 8 | 0.15 | 16 | 4 | 0.0040 | 0.0036 | 0.64 |
| Cluster 2 | 6 | 0.09 | 12 | 2 | 0.0052 | 0.0063 | 0.53 |
| Cluster 3 | 8 | 0.18 | 16 | 5 | 0.0067 | 0.0072 | 0.75 |
| <i>S. caripense</i> | 9 | 0.13 | 12 | 6 | 0.0042 | 0.0037 | 0.86 |
| <i>S. catilliflorum</i> + <i>S. perlongistylum</i> | 8 | 0.11 | 10 | 2 | 0.0004 | 0.0004 | 0.40 |

muricatum haplotypes (i.e., muricatum 1a, 1b and 2) are found only in the cultigen (as well as in *S. filiforme*, whose haplotype is identical to muricatum 1b).

The haplotypes muricatum 3a, 3b, and 4 found in some accessions of *S. muricatum* are likely the result of post-origin hybridization, perhaps as a result of "hybrid capture" (Rieseberg and Soltis, 1991; Mason-Gamer et al., 1995; Papa and Gepts, 2003). The genetic flow between *S. muricatum* and related wild species is to be expected, given that this cultigen is grown in the natural area of distribution of its wild relatives (Heiser, 1964; Anderson, 1975, 1977). In fact, several putative natural hybrids have been identified (via morphology and/or molecular variation) in this study. Furthermore, hand-pollinated hybrids between the pepino and various wild species are relatively easy to make in greenhouse and garden cultures (Anderson and Jansen, 1998; Prohens and Nuez, 2001). Part of the haplotype richness in the pepino could be the consequence of hybridizations at several places and at several times, and the lack of genetic or geographic barriers today, and presumably in the past, between the cultivated and the wild species could explain the high diversity in *S. muricatum* (Anderson et al., 1996; Prohens et al., 2003). This is not the case in other cultivated relatives like tomato (Nesbitt and Tanksley, 2002), which was domesticated far from the natural area of distribution of most of its wild relatives and has a limited diversity.

We believe that it is very likely that recombinations are present and that their existence will confuse the relationships shown by the haplotype tree. We consider the existence of the few nodes with bootstrap values greater than 70% to be one indication of recombination. When only nodes with more than 70% support are considered, a tree with very low resolution and with no significant relationships results. For instance, the haplotype distinguishing *S. canense*, the most distant species based on morphological, ecological (Anderson et al., 1996), and genetic data (Prohens et al., 2006), does not place this species at a significant distance from the others. Alternative trees were constructed using maximum likelihood, Bayesian, and parsimony approaches with similar results (data not shown). The likely reason for the failure of these methods is that all assume no recombination. Network approaches also provided no resolution. A similar lack of concordance between the haplotype tree and the species tree has been described for *Manihot* spp. (Olsen and Schaal, 1999). This situation could be

explained by a recent origin of these wild species. A similar complex situation has been found in the wild relatives of other domesticated species in the Solanaceae (Sukhotu et al., 2004). Given all of this, we can conclude that it is the species in series *Caripensia* that are the closest wild relatives (given the *S. muricatum* haplotypes muricatum 1a, 1b, 1c, 3a, and 3b and typical haplotypes of the series *Caripensia* all grouped together, supported by an 86% bootstrap value). However, it is not possible to infer exactly which wild species may have contributed the unique pepino haplotypes (muricatum 1a, 1b, 1c, and 2).

The complex origin of *S. muricatum* could help explain why the AFLP-based tree does not show clear-cut relationships among the cultivated and the wild species (except for *S. filiforme*, *S. cochoae*, and the odd *S. caripense* PI-243342). Previous proposals suggested *S. tabanoense* and *S. caripense* as possible progenitors of *S. muricatum* (Anderson et al., 1996). Although the participation of *S. caripense* is supported by the haplotype data herein, the involvement of *S. tabanoense* is not as clear from these data. Chloroplast RFLP DNA data, however, strongly implicate *S. tabanoense* in the origin of the pepino as well (Anderson et al., 1996). Furthermore, hybrids between the cultigen and *S. tabanoense* are easy to make (Heiser, 1964, 1969; Anderson, 1979), and there is by far the greatest morphological similarity between the fruits of *S. tabanoense* and some cultivars of pepino (Heiser, 1964, 1969). In addition, the simple leaves of typical *S. tabanoense* mirror those of some pepino cultivars. Thus, we would not dismiss the role of this species in the origin and evolution of the pepino.

Variation structure within the cultigen—*Solanum muricatum* varies considerably in both the AFLP and sequence data. Only the materials studied corresponding to the widespread wild species *S. caripense* have a diversity comparable to that of the cultivated species. A similar result for these species was reported in the cpDNA RFLPs study (Anderson et al., 1996), although the variation in *S. muricatum* was not as high as that described here, probably because of the nature of the markers used. Together, these results suggest that there have been no bottlenecks associated with the domestication of the pepino. The pepino also showed overall high heterozygosity, likely due to its high diversity per se, as well as to its reproductive behavior in cultivation, which is based primarily on vegetative propagation, a mode that favors heterotic hybrids (Rodríguez-Burruero et al., 2003). The variation in this domesticate, the pepino, is notable in comparison to the much more important congeneric domesticate (Spooner et al., 1993), the tomato. That is, the tomato shows much less genetic diversity compared with its closest wild relative *S. pimpinellifolium* L. (Bretó et al., 1993; Rus-Kortekaas et al., 1994) than does the pepino with its closest relative. Perhaps this again indicates that the tomato was domesticated in Mexico, far from most of its related wild species in South America.

The AFLP and sequence data indicate that there is some intraspecific differentiation in the pepino, although there are no strongly isolated groups. The clusters 1 and 3, which are centered in the northern Andean area, are more diverse than cluster 2, which is centered in Peru. The genetic diversity of the pepino in the northern areas is considerably increased when the *S. muricatum* sensu lato accessions are added. However, adding these accessions may not make much sense, because they clearly involve selection, from many unknown stocks around the world and include material judged to be hybrids or

backcross products. Pepino diversity, however, may also be substantial in these northerly areas. That is, there may be more hybridization with wild species in the north, because in the north pepinos have been and are regularly being cultivated in kitchen gardens or small fields adjacent to populations of the wild species (J. Prohens, personal observation). Alternatively or additionally, field work conducted some decades ago by Charles Heiser and Richard Schultes suggested that southern Colombia/northern Ecuador was a locus of possible origin of the pepino (Schultes and Romero-Castañeda, 1962; Heiser, 1964). That hypothesis was based on morphological diversity of pepinos and the fact that some of the fruits in that region bore seeds. The former would be reasonably strong evidence, but the latter is neither restricted to the Colombia/Ecuador region nor indicative of “primitiveness” because many pepino fruits, throughout the natural Andean range of this cultigen, bear seeds (Anderson, 1979). It has been proposed that the geographic area with the greatest diversity is likely to be the center of the species domestication (Vavilov, 1951; Harlan, 1992). In that case, northern Ecuador and southern Colombia could be the center of pepino domestication, with Peru constituting a secondary center and a more recent area of crop expansion.

We suggest that geographic differentiation of genetic diversity of cultivated pepinos could have been considerably weakened by the much greater human interchange among communities and countries in the past 50 years or so. The authors' experience, supplemented by personal communications from C. Heiser (a specialist in Andean-cultivated Solanaceae at the University of Indiana), indicates that there has been a dramatic change in markets, kitchen gardens, and local cultivation in the last half of the 20th century (Heiser, 1969, 1985; Nuez et al., 1999). First, travel for local people in the Andes is now much easier. An indication of the greater interchange, and of the curiosity of gardeners, is that many crops from outside the region appear more prominently in markets (markets that, as a consequence, have lost much of their local or regional “flavor”). In addition, household kitchen gardens are less common today, and when they exist, the crops in them are often not local. And, finally, there are now even commercial field-scale production systems for crops such as the pepino that previously came to market primarily from small, local kitchen gardens. Such mass production systems not only introduce new commercially developed varieties into areas, but also may swamp out the local forms in markets. In addition, of course, botanists and horticulturalists collected pepinos and other crops and developed new cultivars, leading to confusing patterns of local landraces and eventually to the homogenization of local agriculture. Thus, it would actually be more surprising if the geographic patterns of variation of cultivated species such as the pepino were not heterogeneous and ambiguous. One would thus expect a relatively weak signal in the genetic structure of geographic groups, as we have found in this study.

The high diversity in the breeding materials from outside the Andean region (included in genetic cluster 3) is notable; this species seems to have avoided the typical bottleneck associated with the spread of the cultigen from its region of origin to the rest of the world. One possible explanation for this unusually high diversity could be related to the heterosis previously described in this cultivated species. Rodríguez-Burruero et al. (2003) demonstrated that the greater the genetic distance among parents in the cultivated pepino, the greater the hybrid

vigor of the offspring—the result expected. Thus, modern plant breeders would have been drawn to the greater vigor and diversity displayed by intercultural hybrids. Of course, botanists and plant breeders (including the present authors) are also drawn to “new” diversity, novelties, and forms, and thus, new hybrid combinations also attract attention. This hypothesis is further supported by the known intercultural hybrid ancestry of some of the accessions included (El Camino, Kawi, Sweet Round, and Puzol; Prohens et al., 2002).

Conclusions—The AFLP and haplotype data link *S. muricatum* to a group of wild species of the series *Caripensia*, although it is not possible to ascertain whether there is a single species genetically closest to the pepino. In this respect, there is genetic material from several wild species in the cultigen. However, not all haplotypes present in *S. muricatum* have been found in the wild accessions studied, indicating that other wild materials have contributed to the origin and evolution of *S. muricatum*. The highly diverse pepino retains a considerable proportion of the diversity present in its wild relatives. During domestication and subsequent selection, the pepino seems not to have passed through the same stringent bottleneck as some other related domesticates (e.g., the tomato). That it is likely a polyphyletic species and that it is incompatible with several sympatric wild species from series *Caripensia* may account for this high genetic diversity. Several accessions of *S. muricatum* sensu lato that are intermediate between *S. muricatum* sensu stricto and the wild relatives make a strong argument for the (regular) occurrence of gene flow. These results indicate that domestication does not always result in a reduced diversity in the cultigen.

LITERATURE CITED

- ANDERSON, G. J. 1975. The variation and evolution of selected species of *Solanum*, section *Basarthurum*. *Brittonia* 27: 209–222.
- ANDERSON, G. J. 1977. The variation and evolution of selected species of *Solanum*, section *Basarthurum* (Solanaceae). II. *Brittonia* 29: 116–128.
- ANDERSON, G. J. 1979. Systematic and evolutionary consideration of species of *Solanum* section *Basarthurum*. In J. G. Hawkes, R. N. Lester, and A. D. Skelding [eds.], *The biology and taxonomy of the Solanaceae*, 549–562. Linnean Society Symposium Series No. 7. Academic Press, London, UK.
- ANDERSON, G. J., AND L. M. BERNARDELLO. 1991. The relationships of *Solanum chochoae*, a new species from Peru. *Novon* 1: 127–133.
- ANDERSON, G. J., AND R. K. JANSEN. 1998. Biosystematic and molecular systematic studies of *Solanum* section *Basarthurum* and the origin and relationships of the pepino dulce (*S. muricatum*). *Monographs in Systematic Botany from the Missouri Botanical Garden* 68: 17–32.
- ANDERSON, G. J., R. K. JANSEN, AND Y. KIM. 1996. The origin and relationships of the pepino, *Solanum muricatum* (Solanaceae): DNA restriction fragment evidence. *Economic Botany* 50: 369–380.
- ANDERSON, G. J., C. T. MARTINE, J. PROHENS, AND F. NUEZ. 2006. *Solanum perlongistylum* and *S. catilliflorum*, new endemic Peruvian species of *Solanum*, section *Basarthurum*, are close relatives of the domesticated pepino, *S. muricatum*. *Novon* 16: 161–167.
- ANDERSON, G. J., T. P. STEINHARTER, AND G. COOPER-DRIVER. 1987. Foliar flavonoids and the systematics of *Solanum* section *Basarthurum*. *Systematic Botany* 12: 534–540.
- BELKHIR, K., P. BORSA, L. CHIKHI, N. RAUFASTE, AND F. BONHOMME. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.
- BENHAM, J., J. U. JEUNG, M. JASIENIUK, V. KANAZIN, AND T. BLAKE. 1999. Genographer: a graphical tool for automated fluorescent AFLP and microsatellite analysis. Department of Plant Science, Montana State University, Bozeman, Montana, USA.
- BERNARDELLO, L. M., AND G. J. ANDERSON. 1990. Karyotypic studies in *Solanum* section *Basarthurum* (Solanaceae). *American Journal of Botany* 77: 420–431.
- BONFIELD, J. 2004. Staden package, version 1.4. Available at <http://staden.sourceforge.net>.
- BRETÓ, M. P., M. J. ASINS, AND E. A. CARBONELL. 1993. Genetic variability in *Lycopersicon* species and their genetic relationships. *Theoretical and Applied Genetics* 86: 113–120.
- BRÜCHER, H. 1966. *Solanum caripense* HBK (subsect. *Basarthurum*) in Venezuela. *Feddes Repertorium* 73: 216–221.
- BUNTIJER, J. B. 1997. Phylogenetic computer tools, version 1.3. Wageningen University, Wageningen, Netherlands.
- CLARK, A. G. 1990. Inference of haplotypes from PCR-amplified samples of diploid populations. *Molecular Biology and Evolution* 7: 111–122.
- CLARK, R. M., E. LINTON, J. MESSING, AND J. F. DOEBLEY. 2004. Pattern of diversity in the genomic region near the maize domestication gene tb1. *Proceedings of the National Academy of Sciences, USA* 101: 700–707.
- CORRELL, D. S. 1962. The potato and its wild relatives. Texas Research Foundation, Renner, Texas, USA.
- DAWES, S. N., AND G. J. PRINGLE. 1983. Subtropical fruits from South and Central America. In G. Wratt and H. C. Smith [eds.], *Plant breeding in New Zealand*, 33–35. Butterworths, Wellington, New Zealand.
- FELSENSTEIN, J. 1989. PHYLIP—phylogeny inference package, version 3.2. *Cladistics* 5: 164–166.
- FU, Y. X., AND W. H. LI. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
- HARLAN, J. R. 1992. Crops and man. American Society of Agronomy, Madison, Wisconsin, USA.
- HARTER, A. V., K. A. GRADNER, D. FALUSH, D. L. LENTZ, R. A. BYE, AND L. H. RIESEBERG. 2004. Origin of extant domesticated sunflowers in eastern North America. *Nature* 430: 201–205.
- HEISER, C. B. 1964. Origin and variability of the pepino (*Solanum muricatum*): a preliminary report. *Baileya* 12: 151–158.
- HEISER, C. B. 1969. Nightshades: the paradoxical plants. Freeman, San Francisco, California, USA.
- HEISER, C. B. 1985. Of plants and people. University of Oklahoma Press, Norman, Oklahoma, USA.
- HUDSON, R. R. 1987. Estimating the recombination parameter of a finite population model without selection. *Genetical Research* 50: 245–250.
- HUDSON, R. R., D. D. BOOS, AND N. L. KAPLAN. 1992. A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution* 9: 138–151.
- HUDSON, R. R., AND N. L. KAPLAN. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111: 147–164.
- JONES, C. J., K. J. EDWARDS, S. CASTAGLIONE, M. O. WINFIELD, F. SALA, C. VAN DE WIEL, G. BREDEMEIJER, B. VOSMAN, M. MATTHES, A. DALY, R. BREITSCHEIDER, P. BETTINI, M. BUIATTI, B. MAESTRI, A. MALCEVSCI, N. MARMIROLI, R. AERT, G. VOLCKAERT, J. RUEDA, R. LINACERO, A. VÁZQUEZ, AND A. KARP. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381–390.
- KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87–103.
- KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK.
- KOOPMAN, W. J. M., M. J. ZEVENBERGEN, AND R. G. VAN DEN BERG. 2001. Species relationships in *Lactuca* s.l. (Lactucaceae, Asteraceae) inferred from AFLP fingerprints. *American Journal of Botany* 88: 1881–1887.
- MASON-GAMER, R. J., K. E. HOLSINGER, AND R. K. JANSEN. 1995. Chloroplast DNA haplotype variation within and among populations of *Coreopsis grandiflora*. *Molecular Biology and Evolution* 12: 371–381.

- McKEAN, A. L., J. KE, J. SONG, P. CHE, S. ACHENBACH, B. J. NIKOLAU, AND E. S. WURTELE. 2000. Molecular characterization of the non-biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase. *Journal of Biological Chemistry* 275: 5582–5590.
- MORLEY-BUNKER, M. J. S. 1983. A new commercial crop, the pepino (*Solanum muricatum*, Ait.) and suggestions for further development. *Annual Report of the Royal New Zealand Institute of Horticulture* 11: 8–19.
- NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* 70: 3321–3323.
- NEI, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York, New York, USA.
- NEI, M., AND R. K. CHESSER. 1983. Estimation of fixation indices and gene diversities. *Annals of Human Genetics* 47: 253–259.
- NEI, M., AND W. H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269–5273.
- NESBITT, T. C., AND S. D. TANKSLEY. 2002. Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162: 365–379.
- NUEZ, F., R. MORALES, J. PROHENS, P. FERNÁNDEZ DE CÓRDOVA, S. SOLER, E. VALDIVIESO, AND V. SOLÓRZANO. 1999. Germplasm of Solanaceae horticultural crops in the south of Ecuador. *Plant Genetic Resources Newsletter* 120: 44–47.
- NUEZ, F., J. PROHENS, AND J. M. BLANCA. 2004. Relationships, origin, and diversity of Galapagos tomatoes: implications for the conservation of natural populations. *American Journal of Botany* 91: 86–99.
- NUEZ, F., AND J. J. RUIZ. 1996. El pepino dulce y su cultivo. Food and Agricultural Organization of the United Nations, Rome, Italy.
- OLSEN, K. M., AND M. D. PURUGGANAN. 2002. Molecular evidence on the origin and evolution of glutinous rice. *Genetics* 162: 941–950.
- OLSEN, K. M., AND B. A. SCHAAL. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences, USA* 96: 5586–5591.
- PAPA, R., AND P. GEPTS. 2003. Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theoretical and Applied Genetics* 106: 239–250.
- PROHENS, J., G. J. ANDERSON, J. M. BLANCA, J. CAÑIZARES, E. ZURIAGA, AND F. NUEZ. 2006. The implications of AFLP data for the systematics of the wild species of *Solanum* section *Basarthurum*. *Systematic Botany* 31: 206–214.
- PROHENS, J., G. J. ANDERSON, A. RODRÍGUEZ-BURRUEZO, AND F. NUEZ. 2003. Exploiting wild species for the genetic improvement of the pepino (*Solanum muricatum*). *Journal of Applied Botany* 77: 21–27.
- PROHENS, J., M. LEIVA-BRONDO, A. RODRÍGUEZ-BURRUEZO, AND F. NUEZ. 2002. 'Puzol': a facultatively parthenocarpic hybrid of pepino. *HortScience* 37: 418–419.
- PROHENS, J., AND F. NUEZ. 2001. The use of wild species in the improvement of pepino (*Solanum muricatum*). In G. Barendse, G. van den Berg, T. Mariani, and G. van der Weerden [eds.], *Solanaceae V: advances in taxonomy and utilization*, 297–306. Nijmegen University Press, Nijmegen, Netherlands.
- PROHENS, J., J. J. RUIZ, AND F. NUEZ. 1996. The pepino (*Solanum muricatum*, Solanaceae): a "new" crop with a history. *Economic Botany* 50: 255–268.
- RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic flow in plants. *Evolutionary Trends in Plants* 5: 65–84.
- RODRÍGUEZ-BURRUEZO, A., J. PROHENS, AND F. NUEZ. 2003. Performance of hybrid segregating populations of pepino (*Solanum muricatum*) and its relation to genetic distance among parents. *Journal of Horticultural Science and Biotechnology* 78: 911–918.
- ROHLF, F. 1996. NTsys-pc: numerical taxonomy and multivariate system, version 2.0. Exeter Software, Soetauher, New York, USA.
- ROZAS, J., J. C. SANCHEZ-DELBARRIO, X. MESSEGUER, AND R. ROZAS. 2003. DnaSP, DNA polymorphisms analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
- RUS-KORTEKAAS, W., M. J. M. SMULDERS, P. ARENS, AND B. VOSMAN. 1994. Direct comparison of levels of genetic variation in tomato detected by a GACA-containing microsatellite probe and by random amplified polymorphic DNA. *Genome* 37: 375–381.
- SCHAAL, B. A., AND K. M. OLSEN. 2000. Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences, USA* 97: 7024–7029.
- SCHULTES, R. E., AND R. ROMERO-CASTAÑEDA. 1962. Edible fruits of *Solanum* in Colombia. *Harvard University Botanical Museum Leaflets* 19: 235–286.
- SONG, B. H. A., M. J. CLAUS, A. PEPPER, AND T. MITCHELL-OLDS. 2006. Geographic patterns of microsatellite variation in *Boechera stricta*, a close relative of *Arabidopsis*. *Molecular Ecology* 15: 357–369.
- SPOONER, D. M., G. J. ANDERSON, AND R. K. JANSSEN. 1993. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes and pepinos (*Solanaceae*). *American Journal of Botany* 80: 676–688.
- SPOONER, D. M., K. MCLEAN, G. RAMSAY, R. WAUGH, AND G. J. BRYAN. 2006. A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences, USA* 102: 14694–14699.
- SPOONER, D. M., I. E. PERALTA, AND S. KNAPP. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum* L. section *Lycopersicon* (Mill.) Wettst.). *Taxon* 54: 43–62.
- SUKHOTU, T., O. KAMIJIMA, AND K. HOSAKA. 2004. Nuclear and chloroplast DNA differentiation in Andean potatoes. *Genome* 47: 46–56.
- TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- TANKSLEY, S. D., M. W. GANAL, J. P. PRINCE, M. C. DE-VICENTE, M. W. BONIERBALE, P. BROUN, T. M. FULTON, J. J. GIOVANNONI, S. GRANDILLO, G. B. MARTIN, R. MESSEGUER, J. C. MILLER, L. MILLER, A. H. PATERSON, O. PINEDA, M. S. RÖDER, R. A. WING, W. WU, AND N. D. YOUNG. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132: 1141–1160.
- THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.
- VAVILOV, N. I. 1951. *The origin, variation, immunity, and breeding of cultivated plants*. Ronald Press, New York, New York, USA.
- VOS, P., R. HOGERS, M. BLEEKER, M. REIJNS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, AND M. KUIPER. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.
- YEH, F. C., AND T. J. B. BOYLE. 1987. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129: 157.

IV. DISCUSIÓN GENERAL

Los resultados obtenidos en esta tesis nos han permitido aportar nueva información sobre la taxonomía y la filogenia de las especies pertenecientes a las secciones *Lycopersicon* y *Basarthurum* del género *Solanum*. En ambas secciones hemos revisado las clasificaciones propuestas en estudios anteriores aportando nuevos datos moleculares que en algunos casos confirman dichas clasificaciones y en otros están en desacuerdo con ellas.

Hemos estudiado en detalle la variabilidad de una especie de cada sección. En el primer caso, nos centramos en *S. pimpinellifolium* al tratarse de la especie silvestre más cercana a la especie cultivada. En el segundo caso, estudiamos la variación en la especie cultivada porque en los estudios anteriores se había utilizado un bajo número de entradas y no se había hecho un uso extensivo de los marcadores moleculares más resolutivos.

Una vez estudiada la taxonomía de estas especies, hemos pasado a analizar las relaciones entre las mismas. En todos los casos los distintos abordajes moleculares empleados, marcadores microsatélite, AFLPs y secuencias de genes nucleares, han demostrado su utilidad para estos estudios.

I. Clasificaciones taxonómicas

Tal y como se ha comentado en la introducción y como se ha comprobado en los trabajos realizados, el establecimiento de los límites que definen y separan una especie de otra no es una tarea fácil. Estos límites no siempre están claros y la posibilidad de cruzamiento entre los individuos dibuja un escenario continuo en el que establecer la línea divisoria no es sencillo.

La taxonomía de las dos secciones estudiadas, así como las relaciones filogenéticas entre sus especies, ha sido objeto de debate durante largo tiempo por parte de numerosos investigadores. A pesar de todos los esfuerzos realizados hasta el momento no se ha alcanzado un consenso definitivo. Los principales motivos de estas discrepancias pueden deberse a la divergencia reciente de las especies, que permanecen estrechamente relacionadas e incluso mantienen la capacidad de cruzamiento y de producción de híbridos fértiles en algunos casos, y a la escasez de estudios en los que se incluya una amplia representación de la variabilidad natural existente.

En ambas secciones, los resultados obtenidos muestran la existencia de una gran diversidad de situaciones. En la mayoría de los casos se confirma la entidad de especies reconocidas por el resto de los autores, pero en algunos otros casos esos límites no están tan claros.

En el caso de la sección *Lycopersicon*, recientemente se ha propuesto la consideración de nuevas especies, como es el caso de *S. galapagense* (Darwin et al. 2003) o la segregación de *S. peruvianum* en cuatro: *S. arcanum*, *S. huaylasense*, *S. corneliomulleri* y *S. peruvianum* s.str. (Peralta et al. 2008). Sin embargo, según nuestros resultados no todas las especies propuestas parecen formar grupos genéticos diferenciados y por lo tanto creemos que deberían ser reconsideradas.

Solanum huaylasense, que parece ser el resultado de una diferenciación por aislamiento geográfico, puesto que se encuentra en una región muy concreta (en el Callejón de Huaylas, en el curso del río Santa, en el departamento peruano de Ancash (Rick 1986; Peralta et al. 2008)), sí parece formar una entidad genética distinta al resto. Asimismo, *S. arcanum* aparece diferenciada del resto de entradas de *S. peruvianum*, aunque mostrando una estructura más compleja que *S. huaylasense*.

Por el contrario, nuestros datos no apoyan la definición de *S. corneliomulleri* como una especie distinta. Las entradas empleadas como referencias aparecen completamente mezcladas con el resto de *S. peruvianum* s. str., especie con la que comparte el hábitat, en el centro y sur de Perú. De hecho, Rick y Lamm (1955) y Rick (1963) vieron que *Lycopersicon glandulosum* (clasificada actualmente como *S. corneliomulleri* por el TGRC) podía cruzar con *S. peruvianum* y, por tanto, sugirieron que no se consideraran especies distintas. En cuanto a la diferenciación morfológica entre ambas especies, según las claves propuestas por Peralta et al. (2008) para distinguirlas, *S. corneliomulleri* es más pubescente y presenta el margen de la hoja más dentado o dividido que *S. peruvianum*. Sin embargo, a nivel práctico, estas diferencias pueden ser muy sutiles y la variabilidad presente en estas especies hace que la distinción entre ambas sea difícil.

A diferencia del caso anterior, en la sección *Basarthurum* encontramos un ejemplo en el que su posible utilidad práctica justifica el mantenimiento de la separación entre especies. El complejo Caripense está formado por 4 especies: *S. caripense*, *S. filiforme*, *S. fraxinifolium* y *S. heiseri*.

Entre varias de ellas es posible conseguir híbridos, en algunos casos fértiles (Anderson 1975, 1977; Prohens y Nuez 2001) y además estas especies coexisten en ciertas áreas. La existencia de flujo genético podría explicar la gran variabilidad existente en *S. caripense* y la existencia de algunas entradas intermedias entre algunas especies que no pueden asignarse claramente a ninguna de ellas. Por ello, se podría plantear la reducción de todas las especies del complejo Caripense a una sola, tal y como se realizó en el caso de la patata (van der Berg et al. 2002). Sin embargo, morfológica y genéticamente pueden distinguirse las 4 especies del complejo; además mantenerlas como especies independientes puede ser de utilidad para la identificación de material que pueda ser empleado en programas de mejora. Por ello, en este caso sugerimos el mantenimiento de las mismas como especies distintas, a diferencia de lo que ocurre con *S. peruvianum* y *S. corneliomulleri* en la sección *Lycopersicon*.

Otra de las especies propuestas recientemente en la sección *Lycopersicon* es *S. galapagense*. En este caso podría tratarse de un endemismo local muy restringido. Aparece en todos los casos estrechamente relacionada con *S. cheesmaniae*. Su diferenciación morfológica no es muy fuerte, distinguiéndose básicamente en algunos caracteres de la morfología de la hoja, que además se encuentran en baja frecuencia en *S. cheesmaniae*. Rick (1971) ya sugirió que se trataba de dos formas de la misma especie y no de dos especies distintas, tal y como han sugerido recientemente los autores Darwin et al. (2003).

Hemos mencionado la dificultad existente para establecer los límites entre las especies cuando aparecen individuos con características intermedias. Tal es el caso que encontramos en el norte de Ecuador, donde el grupo de entradas de *S. pimpinellifolium* de la región de Esmeraldas podrían estar relacionadas tanto con *S. lycopersicum* como con *S. pimpinellifolium*. Ambas especies están estrechamente relacionadas, los límites en sus características morfológicas no están claros, apareciendo formas intermedias que dificultan su clasificación (Rick y Fobes 1975; Rick et al. 1978). Este grupo de entradas aparece tan cercano al resto de *S. pimpinellifolium* como a *S. lycopersicum*, por lo que podría tratarse del resultado de una hibridación natural que se ha estabilizado.

Nuevas especies silvestres de la sección Basarthrum

En esta sección regularmente se describen nuevas especies, tales como el caso de *S. heiseri* (Anderson 1975) o *S. cochoae* (Anderson y Bernardello 1991). Entre los materiales analizados en esta tesis, se incluyeron dos colectas recientes, procedentes del centro de Perú, denominadas PNS (“Putative New Species”). La primera, PNS1, colectada en la provincia de Abancay, es autocompatible y probablemente autógama, puesto que la ejerción estigmática no es muy acusada. En cuanto a PNS2, colectada en la provincia de Ayacucho, es autoincompatible y presenta un estilo fuertemente exerto. Los resultados moleculares obtenidos muestran la clara distinción entre ambas especies y el resto, apareciendo muy próximas entre sí. Esta situación recuerda al caso de las especies de la sección *Lycopersicon*, *S. neorickii* y *S. chmielewskii*. PNS1, con una apariencia más menuda, muestra partes más pequeñas, incluyendo flores menos vistosas, y una ejerción estigmática escasa o nula, tal y como se observa en *S. neorickii*. En el caso de PNS2, las partes de la planta son más grandes y la ejerción estigmática es fuerte, tal y como ocurre en *S. chmielewskii*. Rick et al. (1976) sugirieron la posibilidad de que *S. neorickii* hubiera evolucionado a partir de *S. chmielewskii* y que se hubieran mantenido separadas por la autogamia de la primera. En el caso de estas dos nuevas especies podría tratarse del mismo caso, siendo PNS1 la derivada autógama de la alógama PNS2. Tras estos resultados que indicaban su entidad como especies distintas, ambas fueron caracterizadas morfológicamente y descritas por Anderson et al. (2006). La especie PNS1 se denominó *S. catilliflorum* G.J. Anderson, Martine, Prohens & Nuez, mientras que la especie PNS2 se denominó *S. perlongistylum* G.J. Anderson, Martine, Prohens & Nuez.

Un caso similar lo constituyen las especies *S. cochoae* y *S. filiforme*, que únicamente han sido colectadas en una ocasión en Perú. Se trata de especies muy infrecuentes y molecularmente son distintas del resto. Podría ser un caso parecido al de PNS1 y PNS2, en el que se trate de aislados periféricos o híbridos antiguos que se han mantenido.

II. Relaciones filogenéticas

Al igual que ocurre con la taxonomía de las especies, las relaciones filogenéticas entre éstas también han sido objeto de discusión y controversia en los últimos años y no se ha alcanzado un consenso definitivo. En este sentido, al estudiar las relaciones filogenéticas en la sección *Lycopersicon* observamos algunas diferencias con respecto a la clasificación en cuatro grupos sugerida por Peralta et al. (2008).

En primer lugar, en el caso del grupo Arcanum (formado por *S. arcanum*, *S. chmielewskii* y *S. neorickii*), nuestros resultados sugieren la inclusión de *S. huaylasense* en este grupo y no en Eriopersicon, puesto que en todos los análisis aparece más estrechamente relacionado con *S. arcanum* que con el resto de especies.

En base a nuestras observaciones, el grupo Eriopersicon estaría compuesto únicamente por las especies *S. peruvianum* y *S. chilense*, incluyéndose dentro de la primera la recientemente descrita *S. corneliomulleri*. Al contrario de lo que sugiere Peralta et al. (2008), *S. habrochaites* no aparece agrupada con estas especies, sino ocupando una posición basal compartida con *S. pennelli*. Esta posición de ambas podría señalarlas como las especies de esta sección que se separaron en primer lugar, al ser las más cercanas a *S. lycopersicoides*, especie de una sección distinta y empleada en el análisis como control fuera de grupo.

Cabe destacar que en los árboles filogenéticos se observa una tendencia en la distribución de las especies a lo largo de dos ejes según su reproducción, desde autoincompatibles a autocompatibles, y su distribución geográfica, de sur a norte. *Solanum peruvianum* y *S. chilense* (autoincompatibles y del sur) aparecen en el extremo opuesto a las especies del grupo *Lycopersicon* (autocompatibles y del norte). Situadas geográfica y filogenéticamente en una posición intermedia aparece el grupo Arcanum, formado por una mezcla de especies autocompatibles y autoincompatibles. Esto concuerda con los resultados de cruzabilidad de Rick (1986), el cual sugirió que las poblaciones de *S. peruvianum* del norte de la distribución, que se corresponden con la actual *S. arcanum*, podrían ser el origen del grupo *Lycopersicon*, hipótesis que se ve refrendada por nuestros resultados.

Respecto a la sección *Basarthrum*, *S. canense* aparece alejada del resto de especies, lo que concuerda con su diferenciación morfológica y su clasificación en una serie distinta. *Solanum suaveolens*, la especie más parecida morfológicamente a *S. canense*, fue clasificada en una serie distinta por Correll (1962), al considerar ese parecido como superficial. De acuerdo con este autor, *S. suaveolens* aparece más cerca genéticamente de la serie *Caripensia* que de *S. canense*, pero a su vez es algo distinta de las especies pertenecientes a esa serie. Por lo tanto, parece estar justificado el mantenimiento de su clasificación en la serie *Suaveolentia*, tal y como propuso Correll (1962), puesto que es molecularmente distinta de *S. canense* y morfológica y reproductivamente diferente del complejo *Caripense*.

En cuanto a las especies pertenecientes a la serie *Caripensia*, *S. basendopogon* forma una entidad claramente diferenciada, que se ve refrendada por caracteres morfológicos tales como el hábito de crecimiento (Anderson y Bernardello 1991). Asimismo, *S. trachycarpum* también es molecularmente distinta del resto de especies, es la única especie autocompatible de la serie (Mione y Anderson 1992) y se caracteriza por crecer en hábitats secos, a diferencia del resto de especies de la sección. Por otro lado, nuestros resultados apoyan la transferencia de *S. tabanoense* desde la serie *Appendiculata* a la serie *Caripensia* realizada por Anderson y Jansen (1998), aunque creemos que debe permanecer fuera del complejo *Caripense* por las diferencias morfológicas (Correll 1962) y bioquímicas (Simpson 1979; Anderson et al. 1987) con el resto de especies de este complejo. Respecto a *S. fraxinifolium* y *S. heiseri*, muy similares entre sí y respecto a *S. caripense*, parece sin embargo que son suficientemente distintas como para justificar su consideración como especies distintas.

Las especies del complejo tienen probablemente un origen reciente. Los datos observados sugieren que *S. caripense* puede ser el ancestro a partir del cual surgieron el resto de especies del complejo. Posiblemente, de forma más o menos simultánea, se produjeron aislamientos periféricos debido a la dispersión en áreas remotas.

Origen del pepino dulce:

La especie cultivada, *S. muricatum*, no aparece ligada a una única especie en los resultados de ninguno de los dos tipos de marcadores empleados, aunque las especies del complejo *Caripense*

son las más cercanas a ella. Esto podría sugerir que varias especies silvestres han contribuido al origen y evolución de la cultivada, coincidiendo con los resultados obtenidos empleando cpDNA (Anderson et al. 1996). Escenarios análogos se han descrito en otras especies como cassava (Olsen y Schaal 1999) o maíz (Clark et al. 2004).

Estudios previos señalaban a *S. tabanoense* y *S. caripense* como posibles progenitores del pepino dulce (Anderson et al. 1996). Sin embargo, así como nuestros resultados apoyan la participación de *S. caripense*, por el contrario no permiten clarificar el papel de *S. tabanoense*. Aunque, por su facilidad de hibridación y su gran similitud morfológica, no podemos descartar su implicación en el origen y evolución de *S. muricatum*. Además, la presencia de alelos exclusivos en *S. muricatum*, que no aparecen en ninguna otra especie en estudio, sugiere que otra especie silvestre desconocida haya intervenido también.

III. Variabilidad intraespecífica

En ambas secciones estudiadas se ha observado la existencia de una gran diversidad genética, pero ésta no está uniformemente distribuida. En ambos casos, hay especies muy variables y otras con una variabilidad muy reducida. Además en ambas secciones algunas especies muestran una fuerte estructuración interna.

Entre las especies de la sección *Lycopersicon*, la especie que muestra una mayor variabilidad genética, tanto empleando AFLPs como secuencias génicas, es *S. peruvianum*. Algunos miembros del grupo *Lycopersicon* son mucho menos variables, especialmente *S. cheesmaniae* y *S. galapagense*, que no mostraron variabilidad haplotípica y apenas polimorfismo en el caso de los AFLPs. En una posición intermedia encontramos a *S. pennelli* y *S. habrochaites*.

En cuanto a la existencia de cierta subestructuración, merecen mención especial los casos de *S. arcanum* y *S. pimpinellifolium*. En el primer caso, *Solanum arcanum* muestra una estructura compleja, en la que se observan dos grupos pertenecientes a los departamentos de Amazonas y Ancash por un lado, y los de Cajamarca y La Libertad, por otro, ambos separados por una importante barrera geográfica como son los Andes. Estos resultados concuerdan con la división

propuesta por Rick (1986), que estudiando la cruzabilidad de las entradas de *S. peruvianum* (tal y como estaban clasificadas antes de la división propuesta por Peralta et al. 2008) describió tres grupos, dos de los cuales se corresponden con los dos grupos que vemos en nuestros análisis. Es posible que el tercer grupo no lo hayamos podido identificar porque lamentablemente no había representantes de esa región geográfica disponibles en el momento de iniciar el presente trabajo. Rick (1986) sugirió la no división de estos grupos en distintas especies, puesto que su diferenciación morfológica no era sencilla. Por el contrario, nuestros datos no concuerdan con los resultados aportados recientemente por Peralta et al. (2008). Estos autores sugieren, basándose en datos morfológicos, la existencia de cuatro grupos distintos en esta especie, distintos a los que vemos con nuestros datos y que tampoco coinciden con los obtenidos a partir de datos de cruzabilidad de Rick (1986) ni con la distribución geográfica. Por todo ello, consideramos *S. arcanum* como un complejo, que debe estar compuesto por varias especies crípticas, y sugerimos la elaboración de un estudio en detalle para clarificar esta situación.

Al analizar la distribución de la variabilidad genética de *S. pimpinellifolium* en todo el rango de distribución de la misma, la primera observación clara es que esta especie también muestra un par de grupos, en este caso compuestos por las entradas peruanas y ecuatorianas respectivamente. Además, dentro de estos grupos se observa un comportamiento distinto. Mientras que en el caso de las entradas procedentes de Ecuador se observa una fuerte estructuración, en el caso de Perú únicamente se ve una suave diferenciación entre norte y sur. Se observa que la región del norte de Perú es la que muestra una mayor diversidad, con un número importante de alelos exclusivos de esa región, y heterocigosidad, lo que podría explicarse por la mayor tasa de alogamia descrita en esta región (Rick et al. 1977, 1978).

Según nuestros resultados, el clima y la ecología parecen ser más informativos a la hora de explicar las diferencias genéticas en *S. pimpinellifolium* que únicamente el aislamiento por distancia geográfica en el caso de esta especie. Esto podría ser debido a un aislamiento de las poblaciones por adaptaciones a distintas condiciones ambientales o ecológicas. Por ejemplo, por la adaptación de distintos polinizadores o animales responsables de dispersar las semillas en cada región. Si entre dos regiones donde se encuentran poblaciones de esta especie existe otra región en la que los polinizadores no pueden sobrevivir, no podrían atravesarlo y el aislamiento entre ambas poblaciones sería aún mayor.

Estas diferencias bioclimáticas podrían explicar también la diferenciación observada dentro de Ecuador en *S. pimpinellifolium*. Mientras que la costa peruana, región donde se encuentra esta especie, muestra unas características climáticas muy uniformes, en Ecuador se observa una gran variación de norte a sur, que esta acompañada de variación en la vegetación.

A partir de estos datos, podemos establecer una hipótesis sobre el origen y posterior distribución de la especie. Al parecer, *S. pimpinellifolium* migró hacia Ecuador y al sur de Perú desde el norte de Perú, de donde es originaria. Esto provocó una pérdida de diversidad debida a cuellos de botella, y una selección hacia la autogamia en las nuevas regiones, lo que a su vez provocó una reducción del tamaño poblacional efectivo, la heterocigosidad y la diversidad genética. Sin embargo, las regiones del norte y sur de Perú aparecen genéticamente próximas, lo que puede ser debido a que las similitudes climáticas y ecológicas a lo largo de la costa peruana hayan favorecido las migraciones entre individuos de ambas regiones.

Por último, cabe destacar que entre las entradas de *S. lycopersicum* utilizadas como referencias externas, las procedentes de la región andina muestran mayor variabilidad genética que las del resto del mundo y que, pese a que sería necesario un análisis más profundo, ésto podría indicar la existencia de flujo genético interespecífico en la región, tal y como proponen otros autores (Rick et al. 1974; Rick y Fobes 1975).

En cuanto a la sección *Basarthurum*, la diversidad presente en cada una de las especies silvestres no muestra grandes diferencias entre las especies autoincompatibles y autocompatibles. La especie con mayor diversidad genética es *S. caripense*, que también se trata de la especie más ampliamente distribuida. Por el contrario, la especie con menos variabilidad es *S. heiseri*, que tiene una distribución muy restringida.

Respecto a la variación dentro de la especie cultivada, al contrario de lo que cabría esperar, *S. muricatum* ha mostrado una gran diversidad, únicamente comparable a la observada en *S. caripense*. Resultados similares se obtuvieron trabajando con cpDNA (Anderson et al. 1996). Esto sugiere que no se produjeron cuellos de botella importantes tras la domesticación del pepino dulce. Además este cultivo muestra una gran heterocigosidad, que puede explicarse por su reproducción vegetativa. Esta situación es muy distinta a la que encontramos en el tomate

cultivado, en el que la variabilidad genética es muy pequeña, y en el que parece que los cuellos de botella han sido muy importantes (Nesbitt y Tanksley 2002).

El origen polifilético de la especie cultivada puede ser la causa principal de su variabilidad. Además, se han identificado posibles híbridos naturales en este trabajo, lo que es de esperar puesto que las especies silvestres y la cultivada comparten el área de distribución. Parte de la amplia variabilidad haplotípica presente en *S. muricatum* puede ser consecuencia de múltiples hibridaciones en distintos lugares y momentos, y la falta de barreras tanto genéticas como geográficas entre las silvestres y la cultivada puede explicar la gran diversidad observada en esta última (Anderson et al. 1996; Prohens et al. 2003).

En cuanto a la diferenciación intraespecífica en *S. muricatum*, pese a que no se observan grupos fuertemente aislados, sí que observamos cierta subestructuración. Los materiales procedentes del norte de la región andina muestran más variabilidad que los de la región central de Perú, hecho que concuerda con la opinión de varios autores (Schultes y Romero-Castañeda 1962; Heiser 1964). Según estos autores, la región de origen del pepino dulce se encuentra en el sur de Colombia y norte de Ecuador, basándose en la amplia variabilidad morfológica presente en la región. Además en esta región se cultiva frecuentemente el pepino dulce en pequeños campos o en los jardines de las casas, donde es frecuente que aparezcan estas especies silvestres (Prohens, observación personal). Por todo ello, el sur de Colombia y el norte de Ecuador podría ser el centro de origen del pepino dulce, mientras que Perú podría ser un centro secundario y una región de expansión del cultivo más reciente.

El hecho de que la diferenciación geográfica en *S. muricatum* no sea muy fuerte podría explicarse por un incremento reciente en el intercambio de material entre comunidades. Este fenómeno parece haber aumentado en los últimos 50 años, puesto que se han observado cambios en los mercados, donde antes se veían únicamente cultivos caseros y locales, y en los que, por una mayor facilidad en los desplazamientos, se han introducido cultivos de otras regiones. Además, la aparición de sistemas de producción a gran escala están desplazando los cultivares locales.

V. CONCLUSIONES

1. Las especies identificadas en la sección *Lycopersicon* son las siguientes: *S. peruvianum*, *S. chilense*, *S. arcanum*, *S. huaylasense*, *S. chmielewskii*, *S. neorickii*, *S. pimpinellifolium*, *S. cheesmaniae*, *S. lycopersicum*, *S. galapagense*, *S. habrochaites* y *S. pennellii*.
2. El establecimiento de *S. corneliomulleri* como especie distinta, tal y como proponen algunos autores, no se ve corroborado por nuestros datos, no siendo posible su distinción de *S. peruvianum* s. str.
3. Entre las especies propuestas recientemente, se confirma que *S. huaylasense* forma un grupo coherente y distinto del resto de *S. peruvianum*.
4. *Solanum arcanum* es distinta del resto de *S. peruvianum*. En este caso se observa una fuerte estructuración, que aconseja la realización de un estudio en profundidad sobre la misma.
5. Entre las especies de la sección *Lycopersicon* se podrían establecer los siguientes grupos: Arcanum (*S. arcanum*, *S. chmielewskii*, *S. neorickii*, *S. huaylasense*), Lycopersicon (*S. lycopersicum*, *S. pimpinellifolium*, *S. cheesmaniae* y *S. galapagense*), Eriopersicon (*S. peruvianum* s.str., y *S. chilense*) y en una posición basal las especies *S. habrochaites* y *S. pennellii*.
6. El grupo Arcanum aparece en una posición intermedia entre los grupos Lycopersicon (autocompatibles y del norte de la distribución) y Eriopersicon (autoincompatibles y del sur).
7. *Solanum pimpinellifolium* muestra una gran diferenciación entre los materiales procedentes de Perú y Ecuador. Este hecho se ve refrendado empleando distintos tipos de marcadores moleculares, tanto microsatélites, como AFLPs y secuencias de genes nucleares.
8. Dentro de cada uno de estos grupos geográficos se observa un comportamiento distinto. En el caso de las entradas procedentes de Ecuador se observa una fuerte estructuración, mientras que en el caso de Perú únicamente se ve una ligera diferenciación entre norte y sur.
9. Las condiciones climáticas y ecológicas parecen ser más informativas para explicar las diferencias genéticas entre poblaciones de *S. pimpinellifolium* que únicamente el aislamiento por distancia geográfica. La costa peruana muestra unas características climáticas muy uniformes, mientras que en Ecuador se observa una gran variación de norte a sur, que esta acompañada de variación en la vegetación.

10. La existencia de una mayor variabilidad y heterocigosidad en el norte de Perú, parece situar en esta región el origen de *S. pimpinellifolium*, desde donde migró hacia Ecuador y al sur de Perú. Las regiones del norte y sur de Perú han permanecido genéticamente próximas, probablemente gracias a migraciones de individuos entre ambas regiones favorecidas por las similitudes climáticas y ecológicas a lo largo de la costa peruana.
11. En la sección *Basarthrum* se ha observado la existencia de una gran diversidad, confirmándose además que los materiales procedentes de dos colectas recientes realizadas en Perú, denominadas provisionalmente PNS1 y PNS2, son distintos entre sí y respecto del resto de especies descritas hasta el momento.
12. En cuanto a las relaciones filogenéticas entre las especies de esta sección, se confirma que *S. canense* es la especie más alejada del resto. Además se justifica que *S. suaveolens* se clasifique en una serie propia, *Suaveolentia*, intermedia entre *S. canense* (a la que se parece morfológica y reproductivamente) y la serie *Caripensia* (de la que está más cerca genéticamente).
13. Dentro de la serie *Caripensia* se observan claramente diferenciadas las especies *S. basendopogon*, *S. trachycarpum* y *S. tabanoense*. El complejo *Caripense* queda constituido por las especies *S. caripense*, *S. filiforme*, *S. fraxinifolium* y *S. heiseri*. Estas especies se pueden distinguir morfológica y molecularmente por lo que a pesar de coexistir en ciertas áreas y de ser capaces de hibridar, la posible utilidad práctica justifica su mantenimiento como especies distintas.
14. *Solanum caripense*, especie ampliamente distribuida, parece ser el ancestro a partir del cual surgieron recientemente el resto de especies de este complejo, quizá a partir de aislamientos en áreas remotas.
15. *Solanum muricatum* no aparece ligada a una única especie silvestre, lo que sugiere que fueron varias las que han contribuido a su origen y evolución. A partir de los resultados podemos concluir que las especies del complejo Caripense son las más cercanas al pepino dulce.

VI. BIBLIOGRAFÍA

Allard RW. 1964. Principles of Plant Breeding. John Wiley & Sons, New York.

Alvarez AE, Van De Wiel CCM, Smulders MJM, Vosman B. 2001. Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. Theoretical and Applied Genetics 103: 1283-1292.

Anderson GJ. 1975. The variation and evolution of selected species of *Solanum*, section *Basarthurum*. Brittonia 27: 209-222.

Anderson GJ. 1977. The variation and evolution of selected species of *Solanum* section *Basarthurum* (Solanaceae) II. Brittonia 29:116-128.

Anderson GJ. 1979. Systematic and evolutionary consideration of *Solanum* section *Basarthurum*. En: Hawkes, J.G.; Lester, R.N.; Skelding, A.D. (eds.), The biology and taxonomy of the Solanaceae. Linnean Society Symposium Series N° 7. Londres, Reino Unido, 549-562.

Anderson, GJ, Bernardello LM. 1991. The relationships of *Solanum cochoae* (Solanaceae), a new species from Peru. Novon 1:127-133.

Anderson GJ, Gensel PG. 1976. Pollen morphology and the systematics of *Solanum* section *Basarthurum*. Pollen et Spores 13:533-552.

Anderson GJ, Jansen RK. 1995. Contributions to the study of the origin and relationships of the “Pepino dulce”, *Solanum muricatum*. Abstract of the Annual Meeting of the Botanical Society of America 316. San Diego, CA, USA.

Anderson GJ, Jansen RK. 1998. Biosystematic and molecular systematic studies of *Solanum* section *Basarthurum* and the origin and relationships of the pepino (*S. muricatum*). Monographs in Systematic Botany from the Missouri Botanical Garden 66: 17-32.

Anderson GJ, Jansen RK, Kim Y. 1996. The origin and relationships of the pepino, *Solanum muricatum* (Solanaceae): DNA restriction fragment evidence. Economic Botany 50: 369-380.

Anderson GJ, Martine CT, Prohens J, Nuez F. 2006. *Solanum perlongistylum* and *S. catilliflorum*, new endemic Peruvian species of *Solanum*, section *Basarthrum*, are close relatives of the domesticated pepino, *S. muricatum*. *Novon* 16: 161-167.

Anderson GJ, Steinharter TP, Cooper-Driver G. 1987. Foliar flavonoids and the systematics of *Solanum* section *Basarthrum*. *Systematic Botany* 12:534-540.

Batley J, Edwards D. 2007. SNP applications in plants. En: Oraguzie NC, Rikkerink EHA, Gardiner SE, y De Silva HN. (eds.), *Association Mapping in Plants*. Springer, New York, 95-102.

Bernardello LM, Anderson GJ. 1990. Karyotypic studies in *Solanum* section *Basarthrum* (Solanaceae). *American Journal of Botany* 77:420-431.

Bohs L. 2005. Major clades in *Solanum* based on ndhF sequence data. En: Keating RC, Hollowell VC, y Croat TB (eds.), *A festschrift for William G. D'Arcy: the legacy of a taxonomist*, St. Louis: Missouri Botanical Garden Press, 27-49.

Bohs L, Olsmted RG. 1997. Phylogenetic relationships in *Solanum* (Solanaceae) based on ndhF sequences. *Systematic Botany* 22(1): 5-17.

Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314-331.

Brücher, H. 1970. Chromosomenzahlen argentinischer, chilenischer and venezolanischer Wildkartoffeln (*Solanum* sect. *Tuberarium*). *Cytologia* 35:153-170.

Child A. 1990. A synopsis of *Solanum* subgenus *Potaoe* (G. Don) (D'Arcy) (*Tuberarium* (Dun.) Bitter (s.l.)). *Feddes Repertorium* 101: 209-235.

Child A, Lester RN. 2001. Synopsis of the genus *Solanum* L. and its infrageneric taxa. En: van den Berg RG, Barendse GWM, van der Weerden GM, y Mariani C (eds.), *Solanaceae V: advances in taxonomy and utilization*. Nijmegen, The Netherlands: Nijmegen University Press, 39-52.

Clark RM, Linton E, Messing J, Doebley JF. 2004. Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proceedings of the National Academy of Sciences, USA 101: 700-707.

Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA, Chen H, Frazer KA, Huson DH, Schölkopf B, Nordborg M, Rättsch G, Ecker JR, Weigel D. 2007. Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science 317:338-342.

Correll DS. 1962. The potato and its wild relatives. Renner: Texas Research Foundation.

Cossio F. 1986. Il pepino o pera-melone, frutto esotico da sperimentare in orticoltura. L'Informatore Agrario-Verona 42 (29):49-51.

Cuartero J, Bolarin MC, Assins MJ, Moreno V. 2006. Increasing salt tolerance in the tomato. Journal of Experimental Botany 57:1045-1056.

D'Arcy WG. 1972. Solanaceae studies II: typification of subdivisions of *Solanum*. Annals of the Missouri Botanical Garden 59: 262-278.

D'Arcy WG. 1986. The calyx in *Lycianthes* and some other genera. Annals of the Missouri Botanical Garden 73: 117-127.

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 & Linnean Society: London, 75-137.

Danert S. 1970. Infragenerische Taxa der Gattung *Solanum* L. Kulturpflanze 18: 253-297.

Darwin C. 1859. El origen de las especies.

Darwin SC, Knapp S, Peralta IE. 2003. Taxonomy of tomatoes in the Galapagos Islands: native and introduced species of *Solanum* section *Lycopersicon* (Solanaceae). Systematics and Biodiversity 1: 29-53.

De Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. En: Howard DJ, y Berlocher SH (eds.), *Endless Forms: Species and Speciation*. Oxford University Press, Oxford, England, 57-75.

Dunal MF. 1852. Solanaceae. En: De Candolle AP (Ed.). *Prodromus Systematis Naturalis Regni Vegetalis*. Victoris Masson, París, 13:1-690.

Ecole CC, Picaço M, Jham GN, Guedes RNC. 1999. Variability in *Lycopersicon hirsutum* f. *typicum* and possible compounds involved in its resistance to *Tuta absoluta*. *Agricultural and Forest Entomology* 1:249-254.

Ercan N, Akilli M. 1996. Reasons for parthenocarpy and the effects of various hormone treatments on fruit set in the pepino (*Solanum muricatum* Ait.). *Scientia Horticulturae*. 66:141-147.

Esquinas-Alcázar JT. 1993. La diversidad genética como material básico para el desarrollo agrícola. En: Cubero JI, Moreno MT (eds.), *La Agricultura del Siglo XXI*. Mundi-Prensa, Madrid, 79-102.

FAO. 1996. Informe sobre el Estado de los Recursos Fitogenéticos en el Mundo. Organización de las Naciones Unidas para la Agricultura y la Alimentación. Roma, Italia.

Foolad MR. 2004. Recent advances in genetics of salt tolerance in tomato. *Plant Cell, Tissue and Organ Culture* 76:101-119.

Foolad MR. 2007. Genome Mapping and Molecular Breeding of Tomato. *International Journal of Plant Genomics* 2007: 64358.

Foolad MR, Lin GY. 1997. Genetic potential for salt tolerance during germination in *Lycopersicon* species. *Hortscience* 32:296-300.

Fouqué A. 1973. Espèces fruitières d'Amérique tropicale - Famille des Solanacées. *Fruits* 28:41-49.

García-Mas J, Graziano 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.

Gebhardt C, Blomendahl C, Schachtschabel U, Debener T, Salamini F, Ritter E. 1989. Identification of 2n breeding lines and 4n varieties of potato (*Solanum tuberosum*, ssp. *tuberosum*) with RFLP fingerprints. Theoretical and Applied Genetics 78: 16-22.

Gentile AG, Stoner AK. 1968. Resistance in *Lycopersicon* species to the tobacco flea beetle. Journal Economic Entomology 61:1347-1349.

Gentile AG, Webb R, Stoner AK. 1969. *Lycopersicon* and *Solanum* resistant to the carmine and the two-spotted spider mite. Journal Economic Entomology 62:834-836.

Gragera J. 2006. Mejora de la calidad del tomate para industria. En: Llácer G, Díez MJ, Carrillo JM, Badenes ML (eds.), Mejora Genética de la Calidad en Plantas. Universidad Politécnica de Valencia, 299-332.

Hajjar R, Hodgkin T. 2007. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 156: 1-13.

Heiser CB. 1964. Origin and variability of the pepino (*Solanum muricatum*): A preliminary report. Bailey 12:151-158.

Heiser CB. 1969a. Nightshades: the paradoxical plants. Freesman, San Francisco, CA, USA.

Heiser CB. 1969b. *Solanum caripense* y el origen de *Solanum muricatum*. Revista Politécnica (Quito, Ecuador) 1(3):5-11.

Heiser CB. 1985. Of plants and people. University of Oklahoma Press, Norman, OK, USA.

Hunziker AT. 1979. South American Solanaceae: a synoptic survey. En: Hawkes JG, Lester RN, y Skelding AD (eds.), The biology and taxonomy of the Solanaceae. London: Academic Press, 49-85.

- Hunziker AT. 2001. The genera of Solanaceae. Ruggel, Lichtenstein: A.R.G. Gantner Verlag.
- IPGRI. 1996. Descriptores para el tomate (*Lycopersicon* spp). International Plant Genetic Resources Institute, Rome, Italy.
- IPGRI y COMAV. 2004. Descriptors for Pepino (*Solanum muricatum*). International Plant Genetic Resources Institute, Rome, Italy, and Centro de Conservación y Mejora de la Agrodiversidad Valenciana, Valencia, Spain.
- Kalloo G, Bergh BO. 1993. Genetic Improvement of Vegetable Crops. Pergamon Press, Oxford, UK.
- Kardolus JP, van Eck HJ, van den Berg RG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (*Solanaceae*). Plant Systematics and Evolution 210:87-103.
- Kebede H, Martin B, Nienhuis J, King G. 1994. Leaf anatomy of two *Lycopersicon* species with contrasting gas exchange properties. Crop Science 34:108-113.
- Kennedy GG, Yamamoto RT. 1979. A toxic factor causing resistance in a wild tomato to the tobacco hornworm and some other insects. Entomologia Experimentalis et Applicata 26:121-126.
- Knapp S, Bohs L, Nee M, Spooner DM. 2004. Solanaceae: a model for linking genomics and biodiversity. Comparative and Functional Genomics 5: 285-291.
- Knapp S, Helgason T. 1997. A revision of *Solanum* section *Pteroidea*: *Solanaceae*. Bulletin of the Natural History Museum, Botany Series 27: 31-73.
- Kohler GW, Lincoln RE, Porter JW, Zscheile FP, Caldwell RM, Harper RH, Silver W. 1947. Selection and breeding for high beta-carotene content (provitamin A) in tomato. Botanical Gazette 109:219-225.

Labate JA, Grandillo S, Fulton T, Munos S, Caicedo A, Peralta I, Ji Y, Chetelat R, Scott JW, Gonzalo MJ, Francis D, Yang W, Van Der Knaap E, Baldo AM, Smith-White B, Mueller L, Prince J, Blanchard N, Storey D, Stevens M, Robbins M, Wang J, Liedl B, Oconnell M, Stommel JR, Aoki K, Iijima Y, Slade A, Hurst S, Loeffler D, Steine M, Vafeados D, Mcguire C, Freeman C, Amen A, Goodstal J, Facciotti D, Van Eck J, Causse M. 2007. Tomato. En: Kole C. (ed.) Genome mapping and molecular breeding in plants. Volume 5 Vegetables. NY. Springer Publishing Co. 1-125.

Lawson DM, Lunde CF, Mutschler MA. 1997. Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, *Lycopersicon pennellii*, to the cultivated tomato, *Lycopersicon esculentum*. Molecular Breeding 3: 307-317.

Lijavetzky D, Cabezas JA, Ibáñez A, Rodríguez V, Martínez-Zapater JM. 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. BMC Genomics 8:424.

Lincoln RE, Zscheile FP, Porter JW, Kohler GW, Cladwell RM. 1943. Provitamin A and vitamin C in the genus *Lycopersicon*. Botanical Gazette 105:113-115.

Linneo C. 1753. Species Plantarum. The Ray Society, London.

Marshall JA, Knapp S, Davey MR, Power JB, Cocking EC, Bennett MD, Cox AV. 2001. Molecular systematics of *Solanum* section *Lycopersicum* (*Lycopersicon*) using the nuclear ITS rDNA region. Theoretical and Applied Genetics 103: 1216-1222.

Martin B, Thorstenson Y. 1988. Stable carbon isotope composition ($d^{13}C$), water use efficiency and biomass productivity of *Lycopersicon esculentum*, *Lycopersicon pennellii* and the F₁ hybrid. Plant Physiology 88:213-217.

Martín I. 2001. Conservación de recursos fitogenéticos. Ministerio de Agricultura, Pesca y Alimentación.

McClearn PE, Hanson MR. 1986. Mitochondrial DNA sequence divergence among *Lycopersicon* and related *Solanum* species. Genetics 112: 649-667.

Miller JC, Tanksley SD. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics* 80: 437-448.

Miller P. 1754. *The gardeners dictionary abridged*, 4th ed. London.

Mione T, Anderson GJ. 1992. Pollen-ovule ratios and breeding system evolution in *Solanum* section *Basarthurum* (*Solanaceae*). *American Journal of Botany* 79: 279-287.

Momotaz A, Scott JW, Schuster DJ. 2005. Searching for silverleaf whitefly and begomovirus resistance genes from *Lycopersicon hirsutum* accession LA1777. *Acta Horticulturae* 695: 417-422.

Muller CH. 1940a. The taxonomy and distribution of the genus *Lycopersicon*. *The National Horticultural Magazine* 19: 157-160.

Muller CH. 1940b. A revision of the genus *Lycopersicon*. Washington DC, USDA Misc. Publ. 382.

NAS. 1972. *Genetic Vulnerability of Major Crops*. National Academy of Sciences, Washington DC.

Nee M. 1999. Synopsis of *Solanum* in the New World. En: Nee M, Symon DE, Lester RN, y Jessop JP (eds.), *Solanaceae IV: advances in biology and utilization*. Kew: Royal Botanic Gardens, 285-333.

Nesbitt TC, Tanksley SD. 2002. Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162: 365-379.

Nuez F. 1995. *El cultivo del tomate*. Madrid: Mundi-Prensa.

Nuez F, Díez MJ, Prohens J, Blanca JM, Sifres A, Picó B, Cordero L, Zuriaga E. 2008. The study of molecular diversity in natural populations of wild and weedy tomatoes and its implications in tomato breeding. *Acta Horticulturae* 789:249-256.

Nuez F, Prohens J, Blanca JM. 2004. Relationships, origin, and diversity of Galapagos tomatoes: implications for the conservation of natural populations. *American Journal of Botany* 91:86-99.

Nuez F, Ruiz JJ, Prohens J. 1997. Mejora genética para mantener la diversidad en los cultivos agrícolas. Estudio Informativo 6. Comisión de Recursos Genéticos para la Alimentación y la Agricultura. FAO.

Olmstead RG, Palmer JD. 1992. A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. *Annals of the Missouri Botanical Garden* 79:346-360.

Olmstead RG, Palmer JD. 1997. Implications for phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Systematic Botany* 22: 19-29.

Olsen KM, Schaal BA. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences, USA* 96: 5586-5591.

Palmer JD, Zamir D. 1982. Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. *Proceedings of the National Academy of Sciences USA* 79: 5006-5010.

PBI Solanum Project. 2009. Solanaceae Source. <http://www.nhm.ac.uk/solanaceaesource/>. Comprobado acceso: junio 2009.

Peralta IE, Knapp S, Spooner DM. 2005. New species of wild tomatoes (*Solanum* section *Lycopersicon*: Solanaceae) from northern Peru. *Systematic Botany* 30(2): 424-434.

Peralta IE, Spooner DM. 2001. Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). *American Journal of Botany* 88 (10): 1888-1902.

Peralta IE, Spooner DM. 2005. Morphological characterization and relationships of wild tomatoes (*Solanum* L. Section *Lycopersicon*). En: Keating RC, Hollowell VC, Croat TB (eds.) A festschrift for William G. D'Arcy: the legacy of a taxonomist. *Monographs in Systematic Botany from the Missouri Botanical Garden* 104: 227-257.

Peralta IE, Spooner DM, Knapp S. 2008. Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). Systematic Botany Monographs 84.

Pérez-Benlloch L, Prohens J, Soler S, Nuez F. 2001. Yield and fruit quality losses caused by ToMV in pepino and search for sources of resistance. Euphytica 120: 247-256.

Péron JY, Demaure E, Hamnetel C. 1989. Les possibilités d'introduction et de développement de solanacées et de cucurbitacées d'origine tropicale en France. Acta Horticulturae 242:179-186.

Pluda D, Rabinovitch HD, Kafkafi U. 1993. Fruit set and yield of pepino dulce response to nitrate-nitrogen and salinity levels and thinning of side branches and trusses. Journal of Plant Nutrition 16:2121-2133.

Prohens J. 1997. Mejora del cuajado y de la precocidad en pepino dulce (*Solanum muricatum* Aiton). Tesis doctoral. Universidad Politécnica de Valencia.

Prohens J, Anderson GJ, Rodríguez-Burruezo A, Nuez F. 2003. Exploiting wild species for the genetic improvement of the pepino (*Solanum muricatum*). Journal of Applied Botany 77: 21-27.

Prohens J, Nuez F. 2001. The use of wild species in the improvement of pepino (*Solanum muricatum*). En: van den Berg RG, Barendse GWM, van der Weerden GM, y Mariani C (eds.), Solanaceae V: advances in taxonomy and utilization. Nijmegen, The Netherlands: Nijmegen University Press, 297-306.

Prohens J, Rodríguez-Burruezo A, Nuez F. 2005. Utilization of genetic resources for the introduction and adaptation of exotic vegetable crops: The case of pepino (*Solanum muricatum*). Euphytica 146, 133-142.

Prohens J, Ruiz JJ, Nuez F. 1996. The pepino (*Solanum muricatum*, Solanaceae): a “new” crop with a history. Economic Botany 50: 255–268.

Rick CM. 1963. Barriers to interbreeding in *Lycopersicon peruvianum*. Evolution: International Journal of Organic Evolution 17:216-232.

Rick CM. 1971. *Lycopersicon*. En: Wiggins IL, Porter DM (eds) Flora of the Galapagos Islands. Stanford University Press, Stanford, 468-471.

Rick CM. 1973. Potential genetic resources in tomato species: clues from observations in native habitats. En: Srb A (ed.) Genes, Enzymes and Populations. Plenum, New York, 255-269.

Rick CM. 1979. Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. En: Hawkes JG, Lester RN, Skelding AD (eds.), The biology and taxonomy of the Solanaceae. Academic Press, New York, USA, 667-678.

Rick CM. 1986. Reproductive isolation in the *Lycopersicon peruvianum* complex. En: D'Arcy WG (ed.) Solanaceae: biology and systematics. Columbia University Press, New York, 477-495.

Rick CM, Fobes JF. 1975. Allozyme variation in the cultivated tomato and closely related species. Bulletin of the Torrey Botanical Club 102:376-384.

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, Holle M. 1990. Andean *Lycopersicon esculentum* var. *cerasiforme*. Genetic variation and its evolutionary significance. Economic Botany 44:69-78.

Rick CM, Holle M, Thorp RW. 1978. Rates of cross-pollination in *Lycopersicon pimpinellifolium*: impact of genetic variation in floral characters. Plant Systematics and Evolution 129:31-44.

Rick CM, Kesicki E, Fobes JF, Holle M. 1976. Genetic and biosystematic studies on two new sibling species of *Lycopersicon* from Interandean Peru. Theoretical and Applied Genetics 47: 55-68.

Rick CM, Lamm R. 1955. Biosystematic studies on the status of *Lycopersicon chilense*. American Journal of Botany 42: 663-675.

Rick CM, Zobel RW, Fobes JF. 1974. Four peroxidase loci in red-fruited tomato species: genetics and geographic distribution. *Proceedings of the National Academy of Sciences, USA* 71:835-839.

Rodríguez-Burruezo A, Prohens J, Nuez F. 2002. Genetic analysis of quantitative traits in pepino (*Solanum muricatum*) in two growing seasons. *Journal of the American Society for Horticultural Science* 127:271-278.

Rodríguez-Burruezo A, Prohens J, Nuez F. 2003. Performance of hybrid segregating populations of pepino (*Solanum muricatum*) and its relation to genetic distance among parents. *Journal of Horticultural Science and Biotechnology* 78:911-918.

Roselló S, Nuez F. 2002. Mejora de la calidad nutritiva del tomate: estudio de la estructura de la variación del contenido en licopeno y β -caroteno en entradas de *Lycopersicon*. *Actas de Horticultura* 34:287-292.

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*. Universidad Politécnica de Valencia, 333-360.

Russell RJ, Fuller JD, Macaulay M, Hats BG, Jahoor A, Powell W, Waugh R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics* 95:714-722.

Sakamoto K, Taguchi T. 1991. Regeneration of intergeneric somatic hybrid plants between *Lycopersicon esculentum* and *Solanum muricatum*. *Theoretical and Applied Genetics* 81:509-513.

Sakata Y, Nishio T, Mon'ma S. 1989. Resistance of *Solanum* species to *Verticillium* wilt and bacterial wilt. En: VIIth EUCARPIA Meeting on Genetics and Breeding of Capsicum and Eggplant, Kragujevac, Yugoslavia, 171-181.

Sánchez Peña P, Fender SE, O'Connell MA. 1995. Leaf water relations of *Lycopersicon chilense* during a drought cycle. *Tomato Genetics Cooperative Report* 45:40-41.

Schultes RE, Romero-Castaneda R. 1962. Edible fruits of *Solanum* in Colombia. Harvard University Botanical Museum Leaflets 19: 235-286.

Seithe A. 1962. Die Haararten der Gattung *Solanum* L. und ihre taxonomische Verwertung. Botanische Jahrbücher für Systematik. 81:261-336.

Simpson MJA. 1979. A chemotaxonomic study of *Solanum* section *Basarthurum*. M.Sc. thesis. Birmingham: University of Birmingham.

Song J, Bradeen JM, Naess SK, Raasch JA, Wielgus SM, Haberlach GT, Liu J, Kuang H, Austin-Phillips S, Buell CR, Helgeson JP, Jiang J. 2003. Gene RB cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. Proceedings of the National Academy of Sciences USA 100: 9128-9133.

Spooner DM, Anderson JG, Jansen RK. 1993. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes and pepinos (Solanaceae). American Journal of Botany 80(6): 676-688.

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ. 2005a. A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. Proceedings of the National Academy of Sciences USA 102: 14694 -14699.

Spooner DM, Nunez J, Trujillo G, Herrera MR, Guzman F, Ghislain M. 2007. Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. Proceedings of the National Academy of Sciences USA 104:19398-19403.

Spooner DM, Peralta IE, Knapp S. 2005b. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes [*Solanum* L. section *Lycopersicon* (Mill.) Wettst.]. Taxon 54: 43-61.

Stevens MA, Rick CM. 1986. Genetics and breeding. En: Atherton JG, Rudich J (eds.), The tomato crop: a scientific basis for improvement. Chapman and Hall, London, 35-105.

Stockman AK, Bond JE. 2008. A taxonomic review of the trapdoor spider genus *Promyrmekiaphila* Schenkel (Araneae, Mygalomorphae, Cyrtaucheniidae, Euctenizinae). *Zootaxa* 1823: 25-41.

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, 1-34.

Trognitz F, Trognitz B. 2004. Mapping genes of *Solanum caripense* involved in resistance to *Phytophthora infestans*, the causal agent of potato late blight. En: Vollman H, Grausgruber H, Ruckenbauer P (eds.), *Genetic Variation for Plant Breeding*. EUCARPIA & BOKU, University of Natural Resources and Applied Life Sciences, Vienna, 249-253.

USDA. 2002. Nutritive value of foods. Home and Garden Bulletin 72.

Van der Berg RG, Bryan GJ, del Rio A, Spooner DM. 2002. Reduction of species in the wild potato *Solanum* section *Petota* series *Longipedicellata*: AFLP, RAPD and chloroplast SSR data. *Theoretical and Applied Genetics* 105: 1109-1114.

Van der Knaap E, Tanksley SD. 2003. The making of a bell pepper-shaped tomato fruit: Identification of loci controlling fruit morphology in Yellow Stuffer tomato. *Theoretical and Applied Genetics* 107:139-147.

Vázquez JF, Sánchez-Yélamo MD, Carrillo JM. 2000. Marcadores morfológicos y bioquímicos. En: Nuez F, Carrillo JM (eds). *Los marcadores genéticos en la mejora vegetal*. Universidad Politécnica de Valencia, 23-89.

Villand J, Skroch PW, Lai T, Hanson P, Kuo CG, Nienhuis J. 1998. Genetic variation among tomato accessions from primary and secondary centers of diversity. *Crop Science* 38:1339-1347.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23:4407-4414.

Wang Y, Diehl A, Wu F, Vrebalov J, Giovannoni J, Siepel A, Tanksley SD. 2008. Sequencing and Comparative Analysis of a Conserved Syntenic Segment in the Solanaceae. *Genetics* 180:391-408.

Warnock SJ. 1988. A review of taxonomy and phylogeny of the genus *Lycopersicon*. *HortScience* 23(4):669-673.

Warnock SJ. 1991. Natural habitats of *Lycopersicon* species. *HortScience* 26: 466-471.

Weese TL, Bohs L. 2007. A three-gene phylogeny of the genus *Solanum* (Solanaceae). *Systematic Botany* 32(2):445-463.

Welles GWH. 1992. Experiences with growing and consumer appreciation of pepino fruits (*Solanum muricatum* Ait.) in the Netherlands. *Acta Horticulturae* 318:211-212.

Welsh J, McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18:7213-7218.

Williams CE, St. Clair DA. 1993. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome* 36:619-630.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18:6531-6535.

Williams WG, Kennedy GG, Yamamoto ET, Thacker JD, Borner J. 1980. 2-Tridecanone – a naturally occurring insecticide from the wild tomato *Lycopersicon hirsutum* f. *glabratum*. *Science* 207:888-889.

Wyman AR, White R. 1980. A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences USA* 77:6754-6758.

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, and Cregan PB. 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123-1134.

Zimmer C. 2008. ¿Qué es una especie?. *Investigación y Ciencia* 383:66-73.