



Universidad Nacional de La Plata
Facultad de Ciencias Exactas
Departamento de Ciencias Biológicas

Tesis Doctoral
**Análisis y Registro de la Biodiversidad de la
Producción Hortícola Platense de la Provincia de
Buenos Aires: Estudio del Tomate 'Platense'**

Lic. Raúl M. Amado Cattáneo

Director: Dr. Andrés N. McCarthy
Co-Directora: Dra. Christina B. McCarthy

2020

A Verito

Agradecimientos

Esta tesis comenzó a gestarse gracias a una beca doctoral en temas estratégicos de CONICET, enmarcada en el plan nacional de Ciencia, Tecnología e Innovación “Argentina Innovadora 2020”, y atravesó diferentes vaivenes no solo de la política nacional sino también, y específicamente, de la política en Ciencia y Tecnología. A pesar de los embates sufridos en esta área, siempre encaré el trabajo desde la firme convicción de la condición social de todo conocimiento, tanto en la academia como en el campo de trabajo. Esta tesis es producto de los múltiples intercambios con profesores, colegas y productores cuyas ideas, observaciones, lecturas y generosidad han contribuido al resultado aquí expuesto.

Agradezco a mi director y a mi co-directora de tesis, Dr. Andrés McCarthy y Dra. Christina McCarthy, y a mi co-directora de beca, Ing. Agr. María Laura Bravo, por su predisposición, por confiar en mí y en el tema de investigación y por haberme guiado durante estos años de trabajo.

Al Centro Regional de Estudios Genómicos (CREG - UNLP), lugar de radicación de mi beca doctoral CONICET, en particular al Dr. Luis Diambra y a los compañeros de laboratorio.

Al laboratorio de Agrobiotecnología del INTA Balcarce, donde realicé mi pasantía, a su director Dr. Sergio Feingold, a la Dra Gabriela Massa, a la Dra. Natalia Norero, al Lic. Martín Carboni y a todos los compañeros de laboratorio. Quiero agradecer en particular a Paula y a Leo por abrir las puertas de su casa y hacer muy amena mi estadía en Balcarce.

A la Chacra Experimental de Gorina del actual Ministerio de Desarrollo Agrario, a su responsable de entonces, el Ing. Agr. Néstor Mezquiriz, que sin dudar puso a disposición la Chacra para realizar parte del proyecto. No puedo dejar de agradecer tampoco el trabajo compartido con los trabajadores del INTA y del Ministerio.

Al Ing. Agr. Juan José (Lolo) Garat y su equipo de UPID y el trabajo en conjunto con el Grupo de Productores del Tomate Platense, por el rescate y valorización cultural del Tomate Platense.

Al Dr. Joaquin Hasperué y a su equipo del Laboratorio de Investigación en Productos Agroindustriales- CIDCA- CONICET, por la ayuda en la determinación de la calidad en poscosecha.

No quiero dejar de agradecer también a la Ing. Agr. María Clara Mediavilla, Ing. Agr. Guillermina Ferrari, Ing. Agr. Camila Gómez, Dr. Guillermo Cap y a Mg. Sergio Durmrauf y tanto otros compañeros que me enseñaron a caminar y estar en las quintas de los productores Familiares. En igual sentido, a todos los compañeros que están y han participado en el proyecto de extensión universitario “Banco de Germoplasma: fortalecimiento de la Agricultura Familiar”, con los cuales buscamos generar espacios para seguir discutiendo con los futuros biotecnólogos su rol social.

A Diego León Viera Nobre que me ha enseñado a observar la naturaleza y compartir su pasión por el mundo de las plantas.

Quisiera agradecer también a los productores de la Agricultura Familiar del Cordón Hortícola Platense, y en especial a la familia Quispe, sin cuyo trabajo y ayuda esta tesis no hubiese sido posible.

A mi familia toda, mis padres, mi hermana, sobrinos y en particular a mis abuelos.

A las autoridades y al personal del Doctorado en Ciencias Exactas de la Universidad Nacional de La Plata por las gestiones y dedicación, más en esta época de pandemia.

Índice

Resumen	12
Objetivos	16
Organización de la tesis	18
Capítulo 1. Aspectos Generales de la Variedad Local: Tomate ‘Platense’	19
1.1. Introducción	20
1.2. Centros de domesticación del tomate cultivado	22
1.2.1. Origen del tomate ‘Platense’	24
1.3. Concepto de variedad local	26
1.4. Mejoramiento participativo de variedades locales	28
1.5. Síntesis	34
Capítulo 2. Evaluación morfoagronómica de la población de Tomate ‘Platense’	35
2.1. Introducción	36
2.2. Materiales y Métodos	45
2.2.1. Cálculo de Grados Días	50
2.3. Análisis de la calidad de la poscosecha de frutos	51
2.3.1. Coordenadas L^* , a^* y b^* y la relación a^*/b^*	52
2.3.2. Porcentaje de pérdida de peso de los frutos	53
2.4. Extracción de las semillas	54
2.5. Resultados	55
2.5.1 Ensayo realizado en la Chacra Experimental de Gorina	55
2.5.2 Ensayo realizado en la quinta de productores familiares	58
2.6. Análisis de vida poscosecha	67
2.7. Conclusiones	75

Capítulo 3. Caracterización de la diversidad genética de las entradas de tomates ‘Platense’ del Banco de Germoplasma - INTA - La Consulta	77
3.1. Introducción	78
3.1.1. Tipos de marcadores moleculares	80
3.1.2. Microsatélites	82
3.1.3. Modelos evolutivos de microsatélites	84
3.1.4. Marcadores moleculares en tomate	88
3.2. Materiales y métodos	89
3.2.1. Extracción y purificación de ADN total	90
3.2.2. Selección de los cebadores	92
3.2.3. Reacciones de amplificación y electroforesis en gel de agarosa	93
3.2.4. Protocolo de electroforesis en geles de poliacrilamida	93
3.2.5. Análisis de la diversidad genética	94
3.3. Resultados	100
3.3.1. Índice de diversidad	106
3.3.2. Capacidad de estimación de identidad	107
3.3.3. Dendrograma	110
3.3.4. Análisis de coordenadas principales	111
3.4. Conclusiones	113
Capítulo 4. Análisis filogenómico de cloroplastos de tomate	116
4.1. Introducción	117
4.2. La nueva era de la filogenómica	117
4.2.1. Alineamiento de secuencias múltiples y sus limitaciones	118
4.2.2. Métodos libres de alineamiento	119
4.3. Origen de las secuencias	121
4.3.1. Regiones nucleares	121
4.3.2. Genomas mitocondriales de planta	122
4.3.3. Origen y evolución de los plástidos	123
4.3.4. Genoma de los plástidos	125
4.4. Código de barras de la vida	126

4.5. Materiales y métodos	128
4.5.1. Conjunto de datos genómicos	128
4.5.2. Topología del árbol	130
4.5.3. Método libre de ensamblado y alineamiento (AAF)	131
4.5.4. Longitud óptima de k-mer	131
4.5.5. Filtrado de k-mers de baja frecuencia	132
4.6. Resultados	134
4.6.1. Filogenómica de lecturas reales y simuladas de 12 datos de secuencia de cloroplastos	135
4.6.2. Estudio filogenómico de 41 datos reales y 3 simulados de secuenciación de cloroplastos	140
4.7. Conclusiones	144
Conclusiones y perspectivas futuras	146
Bibliografía	150
Anexos	177

Resumen

La presente tesis tiene como objetivo general analizar la biodiversidad de las variedades locales del tomate (*Solanum lycopersicum*), denominadas Tomate 'Platense', cultivadas en el Cinturón Hortícola Platense (CHP). Estudiamos su identidad genética, diversidad local, conservación y relaciones filogenómicas, mediante técnicas de Biotecnología vegetal y Bioinformática. Las variedades locales son cultivos tradicionales que han evolucionado con el tiempo a través de su adaptación a su entorno ambiental y cultural, y actualmente son consideradas como un reservorio genético valioso. La variedad local conocida como Tomate 'Platense' se encuentra adaptada al suelo local y al clima del CHP, como resultado de la gradual selección de una serie de cultivares de tomate que fueron introducidos en Argentina hacia fines del siglo XIX después de una inmigración masiva, principalmente europea. En esta tesis evaluamos la identidad genética y la diversidad de esta variedad local y de todas las subpoblaciones registradas en el Banco de Germoplasma - INTA - La Consulta (Mendoza, Argentina). La información genética obtenida a través de 14 marcadores microsatélites polimórficos, permitió la evaluación de la variación genética entre y dentro de las entradas, así como la identificación de todos los genotipos estudiados, con un poder de exclusión de más del 99,99%. Asimismo, se calcularon los valores de diversidad, que resultaron notablemente altos, tanto para la variedad local completa como de varias de sus subpoblaciones. Asimismo, describimos los aspectos generales de la variedad local Tomate 'Platense' mediante una evaluación morfoagronómica de las poblaciones estudiadas y su calidad poscosecha. Por otro lado, analizamos las relaciones filogenómicas de cloroplastos del tomate cultivado y silvestre por medio de estudios de alineamientos libres de alineación y ensamblado, a partir de secuencias disponibles en bases de datos internacionales públicas. Esta información es muy útil en la descripción de la variedad local 'Platense', y constituye un enfoque sólido para la caracterización genética de los materiales disponibles en las accesiones del Banco de Germoplasma y para planes de conservación *in situ* en manos de los productores.

Abstract

The general objective of this thesis is to analyse the biodiversity of the tomato (*Solanum lycopersicum*) landrace called *Tomate 'Platense'*, which is grown in the Horticultural Belt of La Plata (CHP). We studied the genetic identity, local diversity, conservation and phylogenomic relationships of tomatoes using plant biotechnology and bioinformatics techniques. Local varieties or landraces are traditional crops that have evolved over time through their adaptation to their environmental and cultural surroundings, and are currently considered a valuable genetic reservoir. The landrace known as *Tomate 'Platense'* is adapted to the local soil and climate of the CHP as a result of the gradual selection of a series of tomato cultivars, that were introduced in Argentina towards the end of the 19th century after a massive, mainly European, immigration. In this thesis we evaluated the genetic identity and diversity of this local variety and of all the subpopulations registered in the Germplasm Bank - INTA - La Consulta (Mendoza, Argentina). The genetic information obtained using 14 polymorphic microsatellite markers allowed us to evaluate the genetic variation between and within entries and to identify all the studied genotypes, with an exclusion power of more than 99.99%. Likewise, diversity values were calculated, which were remarkably high, both for the all the local variety and for several of its subpopulations. Moreover, we describe the general aspects of the local *Tomate 'Platense'* variety through a morphoagronomic evaluation of the studied populations and of their postharvest quality. Likewise, we analyse the phylogenomic relationships of cultivated and wild tomato chloroplasts by an alignment and assembly free method, based on sequences available in public international databases. This information is very useful for describing the local *Tomate 'Platense'* landrace and constitutes a solid approach for the genetic characterization of material available in the Germplasm Bank accessions and for *in situ* conservation plans for the horticulturists.

Resumo

O objetivo geral desta tese é analisar a biodiversidade de variedades locais de tomate (*Solanum lycopersicum*), denominadas Tomate 'Platense', cultivadas no Cinturón Hortícola Platense (CHP). A sua relevância reside no facto de, através de técnicas de Biotecnologia vegetal e Bioinformática, estudarmos a sua identidade genética, a sua diversidade local, a sua conservação e relações filogenômicas. As variedades locais são cultivos tradicionais que evoluíram ao longo do tempo através de sua adaptação ao seu ambiente ambiental e cultural. Elas são atualmente consideradas um valioso reservatório genético. A variedade local, conhecida como Tomate 'Platense' é adaptada ao solo local e ao clima do CHP, como resultado da seleção gradual de uma série de cultivares de tomate que foram introduzidas na Argentina no final do século XIX após a imigração, principalmente europeia. No presente estudo avaliamos a identidade e diversidade genética desta variedade local e todas as suas subpopulações registradas no Banco de Germoplasma - INTA - La Consulta (Mendoza, Argentina). A informação genética obtida, por meio de 14 marcadores microssatélites polimórficos, permitiu a avaliação da variação genética entre e dentro das entradas, bem como a identificação de todos os genótipos estudados com um poder de exclusão superior a 99,99%. Asimismo, se calcularon los valores de diversidad, que resultaron notablemente altos, tanto para la variedad local completa como de varias de sus subpoblaciones. Na tese, descrevemos os aspectos gerais da variedade local Tomate 'Platense' por meio da avaliação morfoagronômica das populações estudadas e sua qualidade pós-colheita. Da mesma forma, analisamos as relações filogenômicas de cloroplastos de tomate cultivado e selvagem por meio de estudos de alinhamento livre e montagem, com base em sequências disponíveis em bancos de dados públicos internacionais. Esta informação é muito útil na descrição da variedade 'Platense' local e constitui uma sólida abordagem para a caracterização genética dos materiais disponíveis nos acessos do Banco de Germoplasma e para planos de conservação in situ nas mãos dos produtores.

Objetivos

La Biotecnología moderna, como campo específico de aplicación tecnológica, utiliza sistemas biológicos y organismos vivos o sus derivados para la creación o modificación de productos o procesos para usos específicos. Si bien desde la generalidad la Biotecnología se vincula sólo con los organismos modificados genéticamente (o transgénicos), las grandes empresas transnacionales que son dueñas de las semillas¹ o la biopiratería², para nombrar algunos ejemplos, lo cierto es que también se puede relacionar con productos y procesos pensados en conjunto con otro tipo de actores, muchas veces invisibilizados, como la Agricultura Familiar³ y que, igualmente, demandan nuevos desarrollos. Así, la Biotecnología permite pensar otro tipo de innovación tecnológica para procesos locales, y la relevancia de esta tesis se centra en que, mediante técnicas de Biotecnología vegetal y Bioinformática, se puede estudiar la diversidad genética local, su conservación y mejoramiento participativo.

En este marco, la presente tesis tiene como objetivo general analizar la biodiversidad de las variedades locales de tomate (*Solanum lycopersicum*), denominadas Tomate 'Platense' cultivadas en el Cinturón Hortícola Platense (CHP).

Debido a que, hasta la fecha, no se ha iniciado un relevamiento sistemático de las diferentes variedades locales en la región de origen y de producción (La Plata), disponibles en bancos de germoplasma y en territorio, los objetivos específicos incluyen:

¹ A partir de los datos de mercado del 2013 indican que el 55% del comercio mundial de semillas fue controlado por Monsanto, DuPont y Syngenta. En particular, estas son todas las empresas de productos químicos agrícolas que se expandieron a semillas. Hasta el 2017, estas firmas formaban parte del "Big Six" de las compañías de agroquímicos que juntas representaron más de las tres cuartas partes del mercado mundial de plaguicidas. En el 2017 "Big Six" se redujo a "Big Four", cuando DuPont y Dow ganaron la aprobación antimonopolio de EE. UU. y se completó su fusión, renombrándose como Corteva. ChemChina también compra la firma suiza Syngenta en 2017 y en 2018 la firma alemana, Bayer, adquiere la cede en EE. UU. de Monsanto. La firma remanente es la compañía alemana BASF (Howard, 2020).

² La biopiratería es una práctica mediante la cual investigadores o empresas utilizan ilegalmente la biodiversidad de países en desarrollo y los conocimientos colectivos de pueblos indígenas o campesinos para realizar productos y servicios que se explotan comercial o industrialmente sin la autorización de sus creadores o innovadores (Biopiratería.org).

³ La Agricultura Familiar para Balsa (2012) "La familia conforma un equipo de trabajo; en estas unidades no se explota trabajo asalariado, y presentan una racionalidad particular propia de la conjunción de la integración entre unidad productiva y doméstica el papel que juega en la dinámica productiva familiar la conservación del patrimonio familiar, y la existencia de un proyecto de vida vinculado a la actividad agropecuaria y con un cierto modo de vida rural deseable".

- Relevar y registrar el patrón genético de las variedades locales de Tomate ‘Platense’ (*Solanum lycopersicum*) cultivadas en el CHP.
- Identificar las diferentes variedades locales de Tomate ‘Platense’ (*Solanum lycopersicum*) y planificar su eventual mejoramiento por cruza.
- Establecer las relaciones filogenéticas de las diferentes entradas de tomate, en particular la variedad Tomate ‘Platense’, y de especies silvestres, mediante el análisis de las secuencias genómicas de sus cloroplastos, disponibles en bases de datos públicas.
- Conservar el germoplasma de las variedades locales de Tomate ‘Platense’.

La relevancia, entonces, de esta tesis, se centra en que, mediante técnicas de la Biotecnología vegetal y la Bioinformática, se puede estudiar la diversidad genética local, su conservación y mejoramiento participativo.

Organización de la tesis

Esta tesis ha sido organizada en 4 capítulos y “Conclusiones y perspectivas futuras”. El primer capítulo brinda una descripción detallada de los antecedentes disponibles del Tomate ‘Platense’, y en los tres capítulos restantes se describen los resultados obtenidos y la metodología específica. En el segundo capítulo mostramos la evaluación morfoagronómica de las poblaciones estudiadas. En el tercer capítulo realizamos la caracterización de la diversidad genética de las entradas del Tomate ‘Platense’ almacenadas en el Banco de Germoplasma INTA- La consulta (Mendoza, Argentina). Por último, en el cuarto capítulo, hicimos un análisis filogenómico del cloroplasto del tomate a partir de secuencias disponibles en bases de datos públicas. En las “Conclusiones y perspectivas futuras” recuperamos los resultados centrales y más significativos de cada uno de los capítulos. Finalmente, exponemos las perspectivas futuras y las diferentes líneas temáticas que surgieron a partir de estos resultados, las cuales podrán ser retomadas en futuras investigaciones.

Capítulo 1

Aspectos generales de la variedad local: tomate 'Platense'



1.1 Introducción

El tomate cultivado *Solanum lycopersicum* L. pertenece a la familia de las Solanaceae, la cual contiene a más de 3.000 especies que se encuentran en una gran extensión de hábitats (Knapp, 2002). A su vez, esta gran familia se caracteriza por incluir especies de valor comercial: en alimentos: tomate (*S. lycopersicum*), papa (*S. tuberosum*), berenjena (*S. melongena*) y pimientos (*Capsicum*); en medicinales: *Datura*, beleño negro (*Hyoscyamus niger*) y belladona (*Atropa belladonna*); y con fines ornamentales, las petunias (*Petunia*).

El género *Solanum* sect. *Lycopersicon* consiste de 13 taxones muy relacionados entre sí. El tomate, *Solanum lycopersicum*, es la única especie que puede encontrarse domesticada o en estado silvestre (Peralta, Spooner, y Knapp, 2008), y sus 12 especies hermanas sólo existen en forma silvestre: *S. arcanum*, *S. cheesmaniae*, *S. chilense*, *S. chmielewskii*, *S. corneliomulleri*, *S. galapagense*, *S. habrochaites*, *S. huaylasense*, *S. neorickii*, *S. pennellii*, *S. peruvianum* y *S. pimpinellifolium* (Peralta y Spooner, 2005; Spooner, Peralta, y Knapp, 2005; Peralta y cols., 2008). Todas las especies silvestres se encuentran en las laderas occidentales de los Andes (desde el Ecuador, Perú, hasta el norte de Bolivia y Chile), en ambientes de desierto seco o pre-desérticos, exceptuando dos de ellas que son endémicas de las islas Galápagos (*S. galapagense* y *S. cheesmaniae*). La Figura 1 muestra la distribución geográfica de los tomates silvestres elaborada por Knapp y Peralta (2016). Además, *S. galapagense*, planta de frutos amarillo/ naranja, fue apartado de *S. cheesmaniae*, ambas especies endémicas de Galápagos (Darwin, Knapp y Peralta, 2003). Lucatti y cols. (2013) han sugerido que las especies *S. galapagense* y *S. cheesmaniae* deberían ser consideradas conespecíficas, pero Knapp y Peralta (2016) argumentan que, según la evidencia morfológica y molecular, no deberían estar agrupadas en un mismo taxón. Estas útiles diferencias observadas ya son utilizadas por los fitomejoradores (Grandillo y cols., 2011).

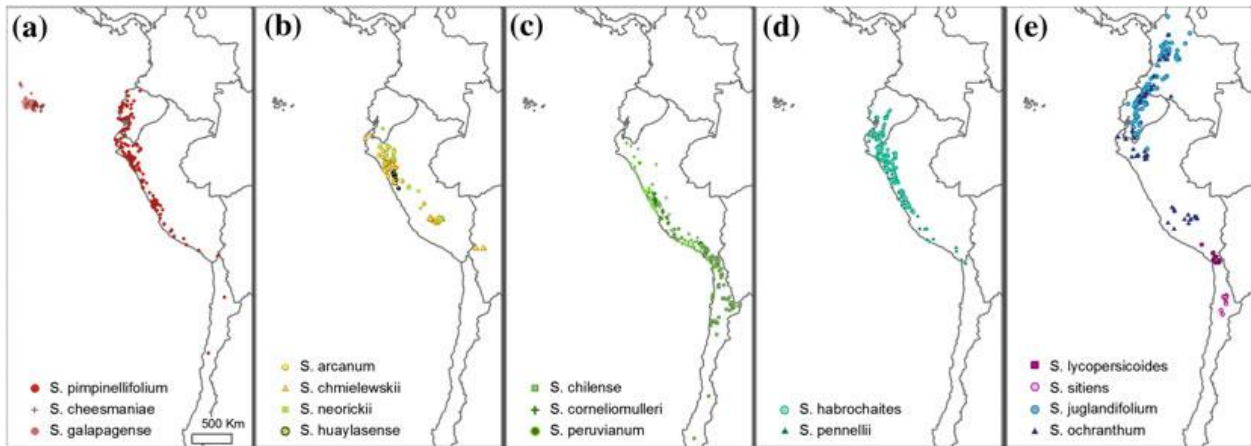


Figura 1.1: Distribución geográfica de los tomates silvestres (Knapp, 2016).

Peralta y cols. (2008) pudieron clasificar las 13 especies en 4 grupos taxonómicos informales, ‘Lycopersicon’, ‘Arcanum’, ‘Eriopersicon’ y ‘Neolycopersicon’, basados en la combinación de datos morfológicos y moleculares (Cuadro 3.1). Esta clasificación informal es de gran utilidad a la hora de sacar conclusiones en análisis filogenéticos.

Grupo Informal	Especie
Lycopersicon	<i>S. arcenum</i> <i>S. lycopersicum</i> <i>S. cheesmaniae</i> <i>S. pimpinellifolium</i> <i>S. lycopersicum</i> var. <i>cerasiforme</i> <i>S. galapagense</i>
Arcanum	<i>S. neorickii</i>
Eriopersicon	<i>S. habrochaites</i> <i>S. chilense</i> <i>S. peruvianum</i>
Neolycopersicon	<i>S. pennellii</i>

Cuadro 1.1: Clasificación en grupos informales dentro de la sección *Lycopersicon*.

1.2. Centros de domesticación del tomate cultivado

En general, es difícil determinar el origen de las plantas domesticadas, en parte por el movimiento de la flora debida a la apropiación europea de la biodiversidad del continente americano que comenzó en los siglos XV - XVI. Aún hoy en día, con las actuales herramientas moleculares, no se puede determinar con seguridad sus posibles orígenes. A esto se suma la falta de materiales silvestres locales para poder comparar, y el movimiento de estas plantas por el mundo con sus posteriores cruces para lograr nuevas variedades mejoradas (Grandillo y cols., 2011).

El tomate tuvo su primer centro de domesticación en América y, aunque el sitio original no se conoce con certeza, los dos posibles centros serían Perú o México. Basado en razones botánicas, lingüísticas e históricas, Candolle (1886) propuso que el tomate fue introducido al continente europeo desde Perú, hipótesis que fue apoyada por otros científicos (Muller, 1940; Luckwill, 1943). Por el contrario, Jenkins (1948) sugirió que fue introducido desde México, basándose principalmente en la lingüística del Náhuatl, en la que el nombre vulgar para *S. lycopersicum* es *jitomatl*, muy parecido a “tomate”. Peralta y Spooner (2007) consideran que el origen del tomate cultivado es todavía incierto y que no hay suficiente evidencia para determinar cuál de los dos países fue el centro de domesticación. Esta incertidumbre es generada debido a la poca variabilidad genética de *S. lycopersicum*, la dificultad de encontrar hoy en día materiales silvestres en esas regiones, y el desconocimiento de cuáles fueron los primeros tomates introducidos al continente europeo (Peralta y cols., 2008).

Los primeros tomates introducidos en Europa fueron clasificados según las referencias de las plantas conocidas en el continente y utilizando la clasificación griega para plantas. Siguiendo esta tradición, Mattioli (1544) describió al tomate en la sección *Della Mandragore*, indicando que el tomate era una especie de mandrágora. Mattioli viajó mucho por el norte de Italia y estudió tanto las plantas cultivadas como las que crecían en forma silvestre en esa región (Raphael, 1989). Su descripción del tomate fue del conocimiento de

primera mano de las plantas vivas. En sus primeros escritos se refirió al tomate como *pomi d' oro* (manzana de oro), en referencia al color de los frutos amarillos (Mattioli, 1544).



Figura 1.2: Ilustración de *Solanum lycopersicum* de Mattioli 1590, en el cual se observa la fasciación de las flores y los frutos largos y multiloculares presente en los primeros tomates europeos (Peralta y col 2016)

El tomate cultivado fue clasificado por Linnaeus (1753) como *Solanum lycopersicum*, pero un año más tarde Miller incorporó a los tomates en el género *Lycopersicum* (Peralta y cols., 2008) y clasificó a la especie cultivada como *Lycopersicum esculentum*. Esto facilitó su aceptación como un alimento ya que su parecido a otras especies venenosas la ubicaban como una posible planta tóxica. En la actualidad, los tomates están incluidos en el género *Solanum*, como originalmente propuso Linnaeus, en base a características morfológicas y análisis moleculares del ADN nuclear y de las organelas (Spooner, Anderson, y Jansen, 1993; Spooner y cols., 2005; Bohs, 1997; Olmstead y Palmer, 1997; Olmstead, Sweere, Spangler, Bohs, y Palmer, 1999; Marshall y cols., 2001; Peralta y Spooner, 2001; Bohs, 2005). Como

ya se mencionó, Peralta y cols. (2008) reconocen el grupo informal ‘Lycopersicon’ dentro de este género a partir de análisis morfológicos, moleculares y geográficos.

En Europa el tomate tuvo mayor éxito en los países mediterráneos como España e Italia (Soressi, 1969; García-Martínez, Andreani, Garcia-Gusano, Geuna, y Ruiz, 2006). Allí *S. lycopersicum* encontró su segundo centro de diversificación (Bailey, Tracy, Kyle, y Watts, 1960), logrando importantes variaciones en las formas de los frutos que hoy en día se pueden seguir encontrando en los mercados locales (Soressi, 1969; García-Martínez y cols., 2006; Ruiz, García-Martínez, Picó, Gao, y Quiros, 2005). El tomate alcanzó un gran avance en su domesticación durante los siglos XVIII y XIX (Nuez, 1996; Sims, 1979) aunque generó un cuello de botella que estrechó aún más la base genética de los materiales cultivados (Rick, 1976). En el siglo XIX y XX comenzaron los programas de mejoramiento y el tomate pasó a ser económicamente importante a nivel mundial.

1.2.1. Origen del tomate ‘Platense’

La ciudad de La Plata (Buenos Aires) fue fundada en 1882 y en el trazado del casco urbano se reservó una superficie considerable para quienes proveyeran de verduras frescas, frutas y leche a los futuros habitantes. De esta forma, la producción hortícola en el “Cinturón Verde” se remonta a los orígenes mismos de la ciudad (Morosi y Terán, 1981). Es así que nació, junto con la ciudad, la producción hortícola local y el “quintero”, denominando así al productor de hortalizas en esta región (Garat, 2002). De origen italiano, español o portugués, y más recientemente de la vecina Bolivia, los quinteros supieron, junto a su familia, crear una cultura alrededor del trabajo en la quinta.

Fueron más de 100 años a través de los cuales la producción hortícola produjo y reprodujo las especies mejor adaptadas a esta zona y a la demanda de los consumidores de la región y del país (Garat, 2002). Es por ello que las variedades actuales son consideradas productos de la cultura local, ya que sintetizan los conocimientos y las experiencias acumuladas durante varias generaciones (Garat, 2002). Entre estas variedades se encuentran las del Tomate ‘Platense’, apio “fajado”, hinojo, sandía “cuarentina”, cardo blanco, alcaucil

“ñato”, grillo nabo, nabiza, ají vinagre, ají morrón, ají “putito”, ají “diablito”, acelga “penca verde” y brócoli “italiano”, “criollo” o “calabrés”, cubriéndose una cantidad importante de especies hortícolas de consumo cotidiano (Garat, 2007). Sin embargo, en la actualidad, la conservación de estas variedades tradicionales se encuentra en riesgo. En la década del ’80 comenzó una transformación en la horticultura local debido a un cambio en el origen de los materiales genéticos utilizados para la producción. Esto causó una reducción paulatina de las variedades hortícolas locales que fueron sustituidas por otras más rentables o con mayores rendimientos (Benencia, Cattaneo y Fernández, 1997). Este proceso se profundizó con los cambios en el modo de producción que ocurrieron una década más tarde (incorporación masiva del invernáculo, riego localizado, etc.). Es así que muchas variedades adaptadas a la zona y apreciadas por sus características organolépticas sobresalientes, se han perdido o se encuentran seriamente amenazadas. El caso del Tomate ‘Platense’, producto típico del CHP, resulta un ejemplo emblemático de esta situación ya que su presencia en el mercado se redujo hasta prácticamente desaparecer.

El tradicional tomate de esta zona está constituido por una gama relativamente amplia de variedades de polinización abierta, más o menos mejoradas, y las características morfológicas del fruto lo aproximan al tipo morfológico “Marmande” (Nuez y cols., 1996). El Tomate ‘Platense’ reconoce un origen incierto en cultivares introducidos desde Francia o Italia a partir de 1930 (Garat, 2003). La hipótesis más fidedigna sostiene que procede del cultivar francés introducido en Argentina con el nombre de “Perdigreon” (Folquer, 1979). Sus diferentes características agronómicas y su rusticidad contribuyeron a su amplia difusión, convirtiéndolo en la variedad de tomate redondo cultivada casi con exclusividad en todas las áreas productoras de la Argentina. Hacia la década del ’80, la irrupción de híbridos comerciales de alto rendimiento dependientes de tecnología de altos insumos, determinó el comienzo de un largo y sostenido proceso de sustitución del Tomate ‘Platense’. Sumado a esto, la adopción masiva del invernadero plástico por parte de los productores y la aparición de las grandes cadenas de mayoristas que exigían materiales de larga vida, terminaron por marginar al Tomate ‘Platense’ casi por completo a la huerta familiar. Sin embargo, y como un ejemplo paradigmático de conservación *in situ* de germoplasma, las variedades locales de Tomate ‘Platense’ se mantuvieron, a pesar de las exigencias fluctuantes del mercado, gracias a los horticultores que producen verdura para consumo propio (Castro, Nico y Garat, 2004).

Por otro lado, ciertos caracteres agronómicos como su tolerancia al virus *tomato spotted wilt virus* (TSWV) (Von der Pahlen, 1970), que produce la comúnmente llamada “peste negra”, y caracteres de calidad comercial como su alto contenido en sustancias responsables de sabor (von der Pahlen, 1970; Boggio, Palatnik, Heldt y Valle, 2000; D’Angelo, 2019)), promovieron su revalorización.

Según el testimonio recogido entre los horticultores, durante la época de mayor difusión comercial del Tomate ‘Platense’ en el CHP habrían existido hasta 15 poblaciones diferentes, reconocidas según el apellido del horticultor responsable de su selección y conservación (‘Carcione’, ‘Bustos’, ‘Del Manso’, ‘Gentile’, ‘Prieto’, ‘Grasso’, ‘Molinaro’, ‘Tomaíno’, ‘Alborghetti’, ‘Luna’, ‘Paolettich’, ‘Sinópoli’, ‘Volpi’, ‘Cataldo’ y ‘Breccia’) (Nico y cols., 2006). Sin embargo, debido al proceso de sustitución por variedades comerciales y al fallecimiento o retiro laboral de los horticultores que las conservaban, no se continuó con la preservación de muchas de estas poblaciones. A la fecha, sólo han podido ser rescatadas las poblaciones ‘Carcione’, ‘Bustos’, ‘Del Manso’, ‘Gentile’, ‘Grasso’ y ‘Luna’ (Nico y cols., 2006). La aplicación de descriptores permitió discriminar esas poblaciones sobre la base de caracteres tales como forma del fruto, número promedio de flores por racimo y frecuencia de defectos comerciales como fasciación, agrietado, etc. Los resultados preliminares de una caracterización agronómica y morfológica de 6 variedades locales del Tomate ‘Platense’ (‘Bustos’, ‘Carcione’, ‘Del Manso’, ‘Luna’ y ‘Gentile’), indicarían que los conceptos particulares de calidad agronómica y comercial que cada productor aplicó a sus selecciones derivaron en materiales de comportamiento diferente (Castro, Nico y Garat, 2004).

1.3. Concepto de Variedad Local

Las variedades locales (*landraces* en la bibliografía inglesa) han tenido un papel fundamental en la historia de la agricultura. Su definición engloba varios conceptos que han variado durante el tiempo ya sea según el interés de la promoción de su uso, por su resguardo y conservación, o como materiales de partida para posibles programas de mejoramiento

vegetal. Hoy en día estos materiales han adquirido importancia no solamente como fuente de biodiversidad, sino también como productos con valores nutricionales y organolépticos. En el trabajo de revisión de Zeven (1998) se describe cómo fue evolucionando el concepto de *landraces* en las publicaciones científicas a lo largo de casi 90 años (1909-1998). El primero que utilizó el término *landrace* fue von Rümker (1908) para variedades de cereales, las cuales llevan el nombre de la región donde se cultivaban desde tiempos inmemoriales sin ningún tipo de selección por parte del/de la productor/a. Zeven (1998) propone una nueva y concisa descripción de lo que se puede considerar *landrace* a partir de la modificación de la definición de Mayr (1937), la cual tenía en cuenta la posibilidad de contaminación de las variedades locales con material foráneo. Para Zeven una variedad autóctona es la que es cultivada durante un período largo de tiempo en un determinado sistema agrícola. A medida que el ambiente se va modificando y que la variedad local puede contaminarse con algunos pocos genotipos externos, como otra(s) variedad(es) o cultivar(es)⁴, esta se irá adaptando continuamente. Por otro lado, Zeven define a una variedad alóctona como aquella que proviene de una región extranjera, la cual es recientemente introducida en la región en estudio. Este último caso puede llegar a ser poco típico, ya que después de su introducción es frecuente que se contamine con algún(os) genotipos de la variedad(es) autóctona(s) o cultivar(es). Dependiendo del número de generaciones posteriores y de la frecuencia de intercambio de las semillas, puede llegar a convertirse en una variedad autóctona (Casañas, Simó, Casals, y Prohens, 2017). Más allá de estas definiciones modificadas, Zeven considera que la mejor definición es la de Mansholt (1909), la cual describe a las variedades locales como aquellas que poseen la capacidad para tolerar altos niveles de estrés biótico y abiótico. Esto resulta en una alta estabilidad en campo y un nivel de rendimiento intermedio en sistemas agrícolas de bajos insumos. Finalmente, Zeven señala que las variedades locales están, a nivel genético, en continua evolución y cruzándose con otras variedades locales o cultivares. La selección natural y artificial, junto a la migración de materiales por medio de las prácticas de intercambio de semillas, puede contribuir a crear diferentes variedades locales (o poblaciones), así como grupos de variedades locales interrelacionadas (que podrían

4 El Código Internacional de Nomenclatura para Plantas Cultivadas define al cultivar como aquel que posea las siguientes características: *distintivos* (que caracterizan al cultivar, que lo diferencien de los demás), *homogéneos* (que se encuentren en todas las plantas del cultivar) y *estables* (que sean heredables), por el método de propagación indicado.

considerarse como metapoblaciones) (Zeven, 1999). Por lo tanto, las variedades locales deben ser vistas como entidades en evolución en contraste con lo que define a un cultivar moderno, de los cuales se espera que se mantengan estables y homogéneos de acuerdo con las normas de la Unión para la Protección de las Obtenciones Vegetales (UPOV).

En un trabajo más actual realizado por Casañas (2017), se advierte que todas las definiciones antes mencionadas se basan más en las propiedades de la población que en los eventos o métodos que condujeron al desarrollo de las mismas (su historicidad) y, por lo tanto, no tienen en cuenta quién ha multiplicado las semillas ni si se aplicó algún criterio de selección o alguna metodología más moderna. Cita a Villa (2005) quien define a la variedad local como una población dinámica con un origen histórico, identidad distintiva, la cual carece de un mejoramiento formal y que a menudo pueden ser genéticamente diversa, adaptada localmente y asociada a sistemas más tradicionales de agricultura. Retoma el enfoque de Negri (2009) sobre la conservación en las unidades productivas, proponiendo que el concepto de selección debe ser eliminado de la definición de variedades locales: la conservación en las unidades productivas debe reformularse como el manejo de la diversidad genética de los cultivos desarrollados localmente (variedades locales) por los agricultores dentro de sus propios sistemas productivos (agrícolas, hortícolas o agro-silviculturales). Esta perspectiva considera que las variedades locales pueden evolucionar dentro de cualquier sistema productivo (deja afuera de la definición el término de sistema tradicional de agricultura), por lo que se podría utilizar la genética para obtener nuevas variedades.

En síntesis, Casañas (2017) propone una definición más amplia de *landrace* que abarca variedades cultivadas que han evolucionado y pueden seguir evolucionando mediante reproducción convencional o técnicas modernas, en ambientes agrícolas tradicionales o nuevos, dentro de un área ecogeográfica definida, y bajo la influencia de la cultura local. Esto incluye la adaptación de variedades locales a los sistemas de manejo de trabajo, y a la selección inconsciente o consciente realizada por agricultores o fitomejoradores con la tecnología de la que disponen. Esta coevolución ha estado modelando variedades locales como resultado de varios factores: la diversidad genética inicial; la generación de nueva variación a través de mutación, migración, recombinación, y cruce con otras poblaciones; el suelo, el clima y otras condiciones ecológicas de la región de cultivo; y la influencia humana,

que incluye técnicas de cultivo (evolucionando con el tiempo) y sus preferencias (sensoriales, nutricionales, religiosas, etc.). Esta definición enfatiza el papel de la humanidad en la evolución de las variedades locales ya que la intervención humana ha sido un factor clave (Knapp, 2019). En este sentido, Casañas cuestiona definiciones más restrictivas que plantean que las variedades locales deben mantener cierta estabilidad en el tiempo, su adaptación y conservación por métodos tradicionales de agricultura. Una definición que no tenga en cuenta los cambios tecnológicos (mejoramiento agronómico y biotecnológico), relega a las variedades locales al estado de “curiosidades” mantenidas en existencia mediante su conservación en zonas productivas como una mera actividad subsidiada, que puede ser inviable bajo las restricciones de fondos, desastres ambientales o para la conservación *in situ*. De esta manera, sugiere que algunos problemas relacionados con los métodos tradicionales, como el arrastre de genes no deseados por retrocruzamiento, pueden ser superados utilizando enfoques biotecnológicos disponibles en la actualidad (Nogué y cols., 2016) y otros que, seguramente, se diseñarán en el futuro.

Casañas (2017) remarca las siguientes debilidades de las variedades mejoradas (cultivares):

(a) La casi o nula ausencia de los lazos culturales con los consumidores, especialmente en frutas y verduras.

(b) Su producción requiere grandes inversiones en energía (el caso extremo es la alta tecnología e invernaderos).

(c) Su logística de distribución para la comercialización (por ejemplo, cosechar fruta inmadura), puede disminuir la calidad sensorial.

Por otra parte, los materiales que cumplen con la definición de variedades locales también sufren de inconvenientes serios, como bajo rendimiento productivo y vulnerabilidad a nuevas plagas introducidas como consecuencia de la globalización. Sumado a esto, la creciente preocupación de nuestra sociedad por el medio ambiente (uso de agroquímicos, reciclado, etc.) y el creciente interés en cuestiones de identidad cultural (valor sensorial de la comida, cocina local y étnica, etc.), son aspectos que las variedades tradicionales poseen y que pueden ser apreciados por los consumidores y, por lo tanto, ser utilizados como herramienta para el desarrollo y progreso de los agricultores.

1.4. Mejoramiento participativo de variedades locales

Es posible ampliar la propuesta de un programa de mejoramiento participativo de variedades locales (Figura 1.3) descrito por Castañas (2017) a partir tanto de las experiencias logradas durante esta tesis doctoral como, y en paralelo, del proyecto de extensión universitaria “Banco de Germoplasma: fortalecimiento de la Agricultura Familiar”.

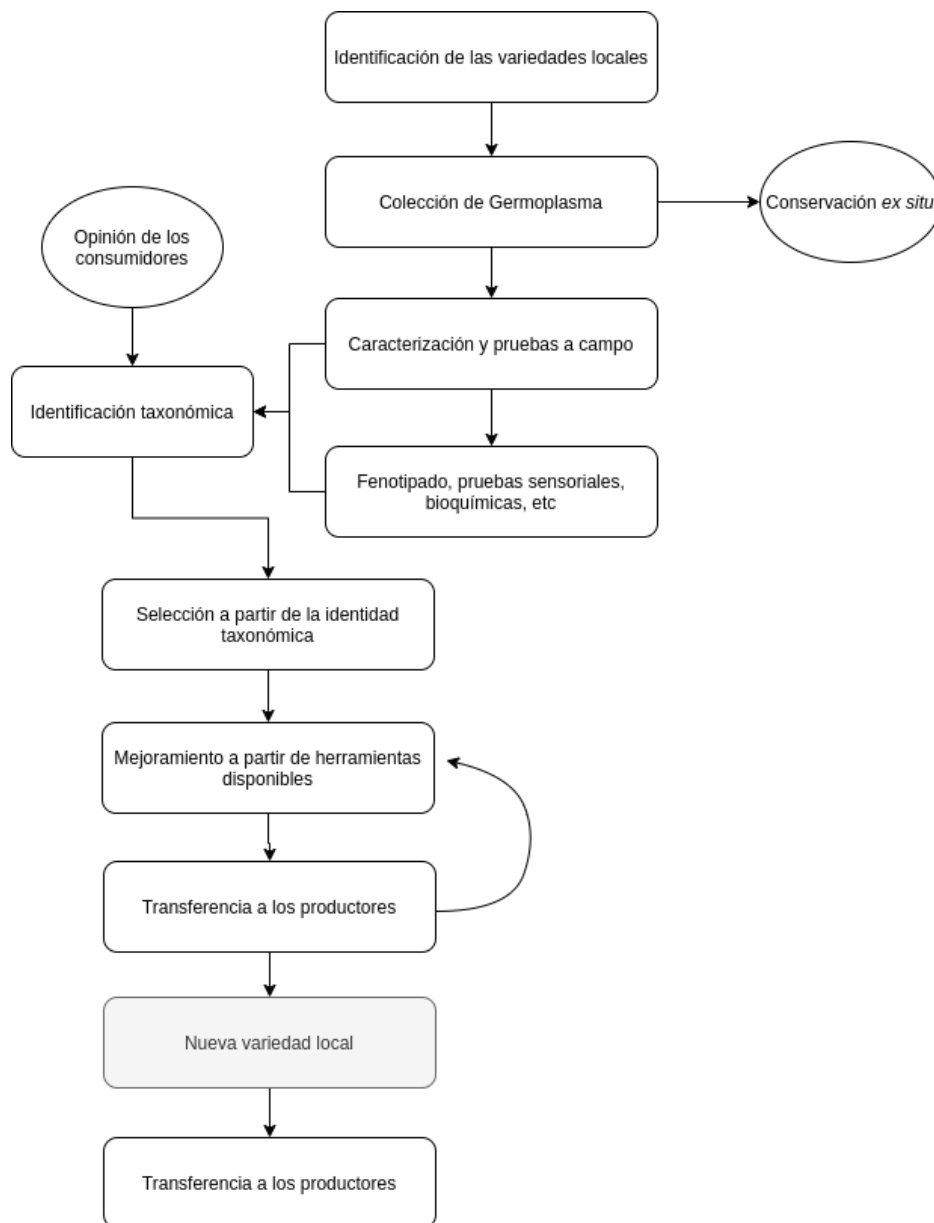


Figura1.3: Diagrama de mejoramiento de variedades locales propuesto por Castañas (2017).

Dentro de un programa de mejoramiento se debe plantear una primera etapa que identifique variedades locales con ciertas características valoradas por la sociedad, tales como el valor nutricional, sensorial y/o cultural. Las variedades locales no necesariamente tienen que ser endémicas, sino que su adaptación al clima y suelo local es en sí una característica que las define como “locales”. Los territorios sin variedades propias deberían poder buscarlas en otras regiones y convertirlas en alóctonas según la definición de Zeven (1998). En cuanto a la conservación y almacenamiento de variedades locales, es importante tener actividades especiales para las organizaciones y grupos de productores, para lograr una política de conservación *in situ*. Muchos de estos grupos ya poseen experiencias de conservación, algunos denominados “guardianes de semillas” y, a su vez, varias organizaciones poseen bancos propios catalogados como “casas de las semillas” (Figura 1.4 (a) y (b)). Sumado a este tipo de conservación en la unidad productiva, se deben tener en cuenta las “ferias de intercambio de semillas” como práctica cultural de los productores familiares para incrementar la diversidad genética de sus cultivos. Estas ferias tienen que estar acompañadas por diferentes instituciones públicas (como el INTA, INASE, universidades, ministerios, etc.), no solo para fortalecerlas como política pública, sino también para prevenir el movimiento de enfermedades o insectos plaga entre regiones.

Por otro lado, es estratégico mantener un resguardo de las diferentes variedades en forma *ex situ*, por ejemplo, a partir de la red de bancos de germoplasma del INTA o las colecciones o bancos de las instituciones públicas como las Universidades Nacionales (Figura 1.4 (c) y (d)). En nuestra experiencia con el proyecto de extensión, y al igual que remarca Bonicatto (2018) en su tesis doctoral, muchos productores no cuentan con el tiempo y las condiciones de infraestructura para asegurar la correcta conservación de las semillas por largos períodos de tiempo. Es por ello que es primordial tener políticas públicas de protección y reproducción de la agrobiodiversidad de nuestras diferentes regiones.

En este punto no hay que dejar de tener una mirada crítica sobre la problemática de los bancos de germoplasma en general, como se describe en el Atlas del Agronegocio (2017). En particular, cómo muchas veces los productores han entregado confiadamente sus semillas a diversas instituciones, pero nunca más volvieron a saber sobre esas investigaciones, o incluso muchos siguen a la espera de semillas prometidas disponibles en esas mismas

instituciones. Es importante que se firmen acuerdos donde las instituciones sean solamente custodios por un tiempo determinado de los materiales de las organizaciones de productores. Conforme a los acuerdos internacionales, los bancos de germoplasma tienen la obligación de entregar materiales solicitados para investigación mediante resoluciones de transferencia de material (ATM), pero cuyos resultados pueden terminar en productos registrados por empresas privadas.

En una segunda etapa, debería evaluarse el potencial de las variedades locales utilizando técnicas de cultivo aceptadas socialmente. Hoy en día las sociedades están tomando una mayor conciencia de los efectos de los agroquímicos, sus efectos sobre el ambiente y la salud de las personas, y se están revalorizando los procesos agroecológicos de producción. Las evaluaciones de estos materiales pueden desarrollarse en centros de investigación, estaciones experimentales o en las mismas unidades productivas de los productores, pero los agricultores siempre deben ser partícipes en la evaluación de estos procesos.

La tercera etapa consistiría en determinar cuáles son las debilidades de las variedades sin perder de vista sus fortalezas. Como resalta Casañas (2017), es esencial aprovechar al máximo todas las herramientas de mejoramiento disponibles, tanto clásicas como las más actuales en biotecnología. Este enfoque requiere la intervención de instituciones de ciencia y tecnología, dados los costos y capacidades técnicas que se precisan. Sin embargo, es necesario también que los agricultores sean copartícipes de estas nuevas variedades que los centros de innovación puedan generar. En este punto queremos remarcar la importancia de la participación de diferentes actores sociales en la obtención de nuevos recursos genéticos, más allá de los resultados prácticos de aplicar alguna técnica novedosa. Es así que, si no existe una participación pública y social, siempre serán favorecidos los sistemas hegemónicos de producción.



Figura 1.4: Diferentes tipos de bancos *in situ* y *ex situ* Casa de las semillas - parcela agroecológica de CEDEPO - Florencio Varela, Buenos Aires (a y b). Banco de germoplasma proyecto de extensión- Fac. de Cs. Exactas - UNLP (c y d).

La cuarta y última etapa que se debe resolver es la propiedad intelectual de las nuevas variedades locales, su certificación, su relación comercial y su mantenimiento (muchas semillas locales o mejoradas se van perdiendo si no se encuentran en bancos de semillas). La ley de Reparación histórica de la Agricultura Familiar para la construcción de una nueva ruralidad en la Argentina (Ley 27.118) del año 2014, fomenta la preservación y recuperación, multiplicación artesanal y en escala, provisión y acceso de las semillas nativas y criollas dentro de los diferentes programas estatales. Sin embargo, no se puede dejar por fuera de estos programas a los sistemas de desarrollo de innovación en semillas, los cuales generan

nuevas variedades que son también materiales demandados por la heterogeneidad de productores de nuestro país.

1.5. Síntesis

En este capítulo se realizó una descripción de la historia de cómo el tomate cultivado fue evolucionando a través de su recorrido por diferentes territorios y posibles centros de domesticación y diversificación, con particular énfasis en lo que hoy se conoce como variedad local tomate ‘Platense’. También se hizo una revisión de cómo fue cambiando el concepto de variedades locales (*o landraces*) a lo largo de la historia, y su complejidad en contraste con los cultivares. Por último, a partir de experiencias locales, se planteó una ampliación del programa de mejoramiento participativo de variedades locales propuesto por Casañas (2017), situando la presente investigación en contacto estrecho con prácticas de extensión en territorio.

Capítulo 2

Evaluación morfoagronómica de la población de Tomate ‘Platense’



2.1. Introducción⁵

La domesticación de las plantas cultivadas fue generada por sucesivos cambios esenciales como consecuencia de las variaciones genéticas de sus especies ancestros que se encontraban en la naturaleza. El tomate cultivado (*Solanum lycopersicum* L.) deriva de su ancestro silvestre *Solanum pimpinellifolium* L., ambas especies nativas de América del Sur (Knapp y Peralta, 2016). En Europa adquirió diferentes centros de diversificación secundarios, principalmente en España y en Italia. A finales del siglo XIX y principios del XX se volvió a introducir en la Argentina a partir de la inmigración. Los quinteros locales fueron seleccionando variedades que aún hoy siguen presentes en diversos mercados locales, como fue descrito en el Capítulo 1.

El tomate cultivado tiene una limitada variabilidad genética debido principalmente a eventos relacionados con la domesticación y evolución de los cultivares modernos (Rick, 1976). Se estima que *S. lycopersicum* aporta sólo cerca del 5% del total de la variación genética del género *Solanum* (Miller y Tanksley, 1990; Rick y Fobes, 1975), mientras que las especies silvestres del género presentan una abundante fuente de genes deseables y características para el mejoramiento de cultivos. Las especies con la mayor variabilidad son *S. chilense*, *S. hirsutum*, *S. peruvianum* y *S. pennellii*, mientras que las especies con menor variabilidad son *S. cheesmanii* y *S. pimpinellifolium* (Miller y Tanksley, 1990; Rick y Fobes, 1975). Durante los últimos 70 años estas especies silvestres se han utilizado en programas de cruzamiento para mejorar el tomate cultivado (Rick y Fobes, 1975; Rick, 1982). En las últimas décadas, los protocolos tradicionales de genética de plantas y cruzamientos, basados en la selección fenotípica y testeo de progenie, han sido muy eficientes en mejorar la productividad y calidad (Rick, 1973; Warren, 1998).

Las variedades comerciales modernas fueron desarrolladas para obtener altos rendimientos, resistencia a plagas y enfermedades, homogeneidad en el tamaño, forma y color del fruto, firmeza y calidad industrial. Una consecuencia no deseada de los procesos de selección fue que la mayoría de los genes responsables del sabor y el aroma fueron atenuados

⁵ Parte de los resultados fueron presentados en el 40º Congreso de la Asociación Argentina de Horticultura (ASAHO).

(o eventualmente perdidos) en materiales comerciales (Tieman y cols., 2017; Gao y cols., 2019). Actualmente, los consumidores reconocen la falta de calidad organoléptica de los frutos, haciendo que el mejoramiento genético de esas características se vuelva un desafío importante y complejo (Saliba-Colombani y cols. 2001). En este sentido, al tener adaptación local al ambiente, el germoplasma del tomate criollo tiene un alto potencial para los programas de mejoramiento. Por ello su rescate constituye un reservorio de variabilidad genética fundamental como patrimonio biológico y cultural. En particular, en los frutos de ciertas poblaciones locales ('Bustos' y 'Carcione') se destacó la presencia de caracteres morfológicos asociados a una condición "salvaje" atractiva a la vista (frutos achatados, acostillados, con la inserción del cáliz hundida e intenso "hombro verde"). Estas características contrastan marcadamente con la morfología típica de los frutos de híbridos comerciales (tomates isodiamétricos lisos y de color uniforme a la maduración) (Castro, Nico y Garat, 2004).

Por otra parte, se han obtenido líneas mejoradas de este cultivar a través de programas oficiales de mejoramiento. Las accesiones que corresponden a estas líneas mejoradas se encuentran en diversos bancos de germoplasma dentro y fuera de Argentina, entre ellas las correspondientes a cultivares inscriptos por INTA en el registro nacional de cultivares: Platense 'Línea 9' (INTA San Pedro), 'J.J Gomez' (INTA Alto Valle), 'Carneuco', 'UcoPlata' (INTA La Consulta), 'Raúl INTA' (INTA San Juan), y 'Línea 7' (Estación Experimental Gorina, Ministerio de Desarrollo Agrario de la Pcia. de Buenos Aires). Cabe destacar que en un estudio donde se comparó el rendimiento agronómico de las variedades de Tomate 'Platense' con las variedades inscriptas por INTA, se encontró que algunas variedades locales ('Simonetti', 'Gentile' y 'Del Manso') eran significativamente superiores a algunas de las variedades mejoradas inscriptas por INTA (Raúl INTA y Línea 7) (Ullé, 2007).

Análisis morfoagronómico

El tomate es una planta perenne de porte arbustivo que se cultiva como anual, aunque su duración vegetativa en condiciones climáticas favorables puede prolongarse varios años (Argerich y cols, 2010).

Como describe Vallejo Cabrera (1999), cuando la planta crece directamente de la semilla sin sufrir trasplantes, desarrolla una potente raíz principal que le permite adaptarse a ecosistemas semidesérticos, pero cuando la raíz principal se daña, por ejemplo, a consecuencia del trasplante, se desarrolla un sistema de raíces laterales adventicias (Chamarro, 1994). El tomate puede crecer, florecer y fructificar en una gran variedad de condiciones ambientales que incluyen un amplio rango de latitudes, alturas, temperaturas, suelos, métodos de cultivo, etc. La temperatura ideal para la germinación es de 25 °C; por debajo de 15 °C y superior a 35 °C se ve afectada negativamente. Para el buen cuajamiento de frutos se requieren temperaturas diurnas de 24-25 °C y nocturnas de 18 °C. Temperaturas nocturnas excesivamente altas (superiores a los 25 °C) perjudican significativamente el cuajamiento de frutos y en consecuencia la producción (Vallejo Cabrera, 1999).

El tallo principal tiene 2-4 cm de diámetro en la base y está cubierto por pelos glandulares y no glandulares que salen de la epidermis. Forma de seis a doce hojas que crecen lateralmente, con una filotaxia de 2/5, antes de que la yema principal se transforme en una inflorescencia. El crecimiento subsiguiente se produce a partir de la yema axilar de la última hoja, la cual desarrolla un tallo secundario que crece como una prolongación del tallo primario y desplaza lateralmente la inflorescencia. Los sucesivos segmentos del tallo se desarrollan de forma similar, produciendo una inflorescencia cada tres hojas. El aspecto es el de un tallo principal, que crece de forma continua, con inflorescencias internodales laterales cada tres hojas. Cuando este proceso se repite indefinidamente los cultivares se denominan indeterminados (Figura 2.1) y el seudo tallo principal puede crecer más de 10 m por año (Rick, 1978), con un porte rastrero o trepador. Estos cultivares son adecuados para recolección continua ya que florecen y fructifican de forma regular y uniforme. Los brotes laterales, que se desarrollan en las axilas de las hojas, se eliminan y el tallo principal se amarra a una estaca o tutor. Los cultivares determinados tienen un crecimiento limitado que puede

extenderse unos 2 metros. Los segmentos sucesivos del eje principal soportan, de forma progresiva, un número inferior de hojas y terminan en una inflorescencia (Figura 2.1). El sistema de ramificación lateral experimenta un crecimiento limitado dando a la planta un aspecto arbustivo con simetría circular que requiere de menos espacio que los cultivares indeterminados. La floración y la fructificación se producen en un período limitado, lo que provoca la concentración de la producción permitiendo efectuar la cosecha mecánica. A su vez, es posible reconocer algunas plantas con hábitos de crecimiento semi-determinado (Figura 2.1). Variantes alélicas del gen *sp* (*selfprunning*) asociado a los individuos con hábitos determinados, resultan en una mayor concentración de la biomasa hacia los frutos y consecuentemente mayor producción (Vallejo Cabrera, 1999). El resultado son plantas con hábitos de crecimiento arbustivo, con floración y producción de frutos concentradas y con un crecimiento del tallo principal que finaliza con la aparición de una inflorescencia terminal.

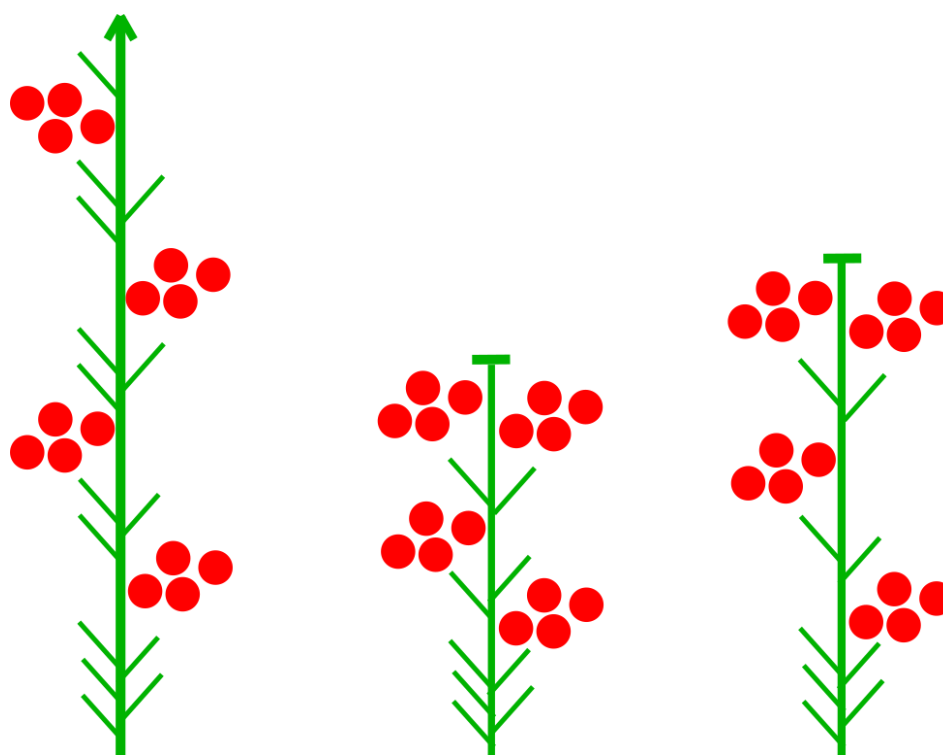


Figura 2.1: Esquemas de una planta de tomate de crecimiento indeterminado (izquierda), de una de crecimiento determinado (centro) y otra de crecimiento semideterminado (derecha).

Vallejo Cabrera (1999) describe a las hojas del tomate como pinnado compuestas. Una hoja típica de las plantas cultivadas tiene unos 50 cm de largo, algo menos de anchura, con un gran foliolo terminal y hasta ocho grandes foliolos laterales, que pueden a su vez ser compuestos (Figura 2.2). Los foliolos son generalmente peciolados y lobulados irregularmente con bordes dentados. Las hojas están recubiertas de pelos del mismo tipo que los del tallo. Las hojas del tomate son de tipo dorsiventral o bifacial.



Figura 2.2: Hoja adulta de tomate (Chamarro, 1994).

La flor del tomate es perfecta, regular e hipógina. El cáliz posee cinco o más sépalos verdes que se prenden al fruto incluso hasta después de la maduración. La corola está formada por cinco o más pétalos amarillos y encurvados cuando la flor está abierta. El número de estambres es de cinco, los cuales se encuentran soldados formando un cono. Las anteras son cortas y anchas. La flor está unida al eje floral por un pedicelo articulado que contiene la zona de abscisión, la cual se distingue por un engrosamiento. Durante la cosecha, la separación del fruto puede producirse por la zona de abscisión o por la inserción del fruto al pedicelo. En los frutos destinados a la industria, la presencia de parte del pedicelo es indeseable y por ello se prefieren las variedades en las que la separación se produce en la

zona de unión al fruto. En determinadas condiciones ambientales desfavorables, la flor se separa de la planta antes de la apertura de los pétalos (aborto floral), en otros casos después de la apertura de los pétalos (caída de la flor), lo que ocasiona grandes pérdidas de producción. Las flores están agrupadas en inflorescencias que pueden ser simples, bifurcadas o ramificadas (Figura 2.3). Las formas simples se presentan con mayor frecuencia en la parte inferior de la planta y las ramificadas en la parte superior. El número de flores por inflorescencia es muy variable, habiéndose descrito algunas con más de 300 flores.

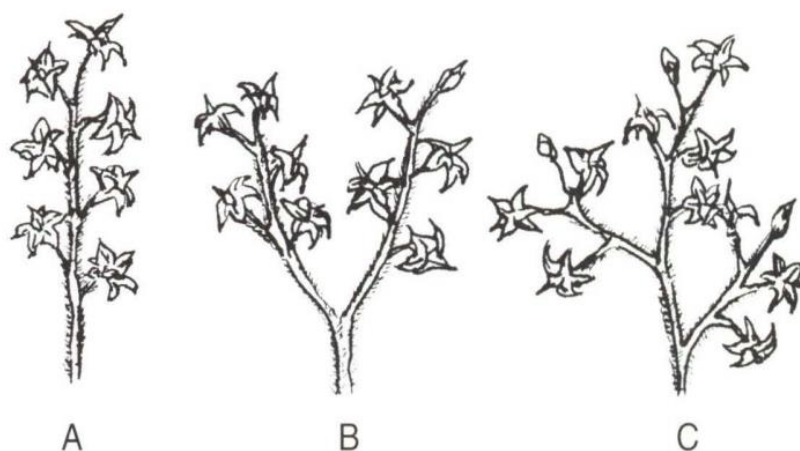


Figura 2.3: Inflorescencia del tomate, simple (A), bifurcada (B) y ramificada (C) (Melo, 1989).

Los frutos se presentan agrupados en racimos, de forma y tamaño muy variables (Figura 2.4). Internamente, los frutos se encuentran divididos en lóculos, sitios donde se ubican las semillas inmersas en el mucílago placentario. De acuerdo con el número de lóculos, los frutos pueden ser clasificados en bi, tri, tetra o pluriloculares.

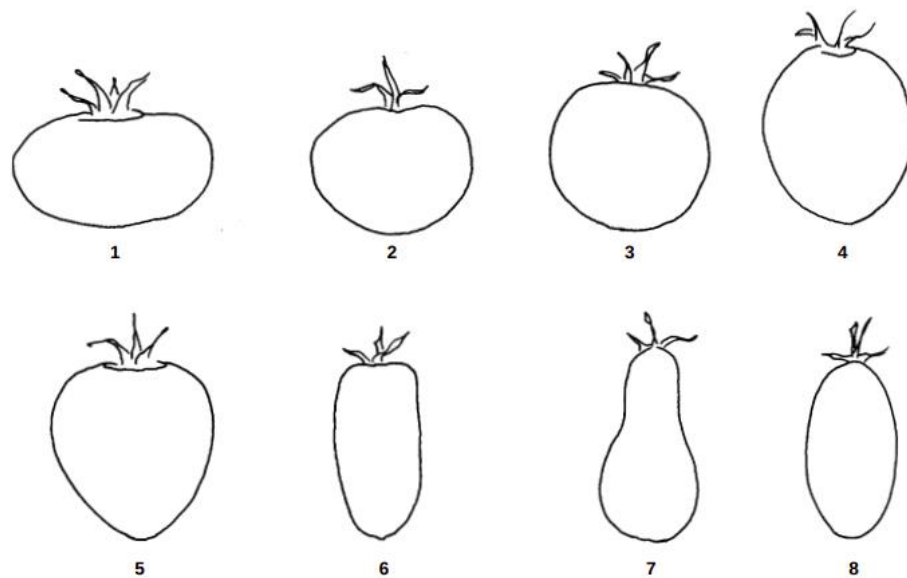


Figura 2.4: Forma predominante del fruto, 1- Achatado, 2- Ligeramente achatado, 3- Redondeado, 4- Redondo-alargado. 5- Cordiforme, 6- Cilíndrico (oblongo-alargado), 7- Piriforme, 8- Elipsoide (forma de ciruela) (Descriptorios para el tomate *Lycopersicon* spp. - International Plant Genetic Resources Instit., (IPGRI), Rome (Italy).

Para la obtención de nuevos cultivares con una base genética amplia, es necesario tener programas de mejoramiento sostenidos en el tiempo que logren la investigación sistemática en la caracterización morfológica y evaluación agronómica de las colecciones de germoplasma de tomate, tanto lo conservado en los bancos de germoplasmas como de las variedades locales que se encuentran, hoy en día, en los diferentes territorios. Es ejemplo de este tipo de tarea el trabajo realizado por la Dra. Peralta (2007) logrando un Catálogo de poblaciones criollas de pimiento, tomate y zapallo, colectadas en valles andinos de Argentina. Una vez que los recursos genéticos disponibles han sido caracterizados, es posible decidir cuáles serían los cruzamientos que más contribuirían a la expansión de la base genética.

Pratta y cols. (2003) proponen que si bien algunos caracteres morfovegetativos (longitud de entrenudos, perímetro del tallo en las partes basal, media y apical, número de flores por racimos, entre otros) son importantes *per se* para la determinación de la aptitud agronómica de una variedad y por lo tanto estar también asociados a los rendimientos finales de cada uno de los genotipos. Por otro lado, el tamaño de los frutos es una de las características fenotípicas que posee una gran influencia en la determinación de la calidad

comercial de los tomates, tanto para consumo en fresco como industrial. Otro factor adicional a tener en cuenta, es el que está ligado a la evolución de la madurez del fruto, el cual determinará su vida en la estantería.

Calidad de poscosecha

Los componentes de la calidad del tomate incluyen apariencia (como color, tamaño, forma, ausencia de defectos y descomposición), la firmeza, el sabor y el valor nutricional. La apariencia está relacionada con su estadio en el momento de la cosecha y los cambios de composición durante el manejo poscosecha (Kader, 1986). Existen grandes variaciones genotípicas en los atributos de la calidad del tomate y es posible desarrollar nuevos cultivares que tengan una buena calidad de consumo y mantengan su firmeza cuando están completamente maduros para que puedan soportar los procedimientos de manipulación poscosecha. Las pérdidas en poscosechas en calidad y cantidad se relacionan con inmadurez en la cosecha, control de calidad inicial inadecuado, incidencia y gravedad de daños físicos, exposición a temperaturas inadecuadas y retrasos entre la cosecha y el consumo. Acortar el tiempo entre la cosecha y el consumo puede minimizar la pérdida del aroma característico del tomate.

El color externo de los tomates es el resultado de los colores de la piel y la pulpa. Se han desarrollado varias escalas de calificación subjetiva y tablas de colores para clasificar las etapas de madurez del tomate. En el estándar estadounidense, para la clasificación de tomates frescos se utilizan distintos conceptos para indicar el estado de maduración, tales como:

- *green* completamente verde;
- *breakers* con ruptura definitiva en el color que va de verde a amarillo oscuro, rosado o rojo en no más del 10% de la superficie;
- *turning* con más del 10% pero no más del 30% de la superficie en conjunto que muestra un definitivo cambio de color de verde a amarillo oscuro, rosado, rojo o una combinación;

- *pink* con más del 30% pero no más del 60% de la superficie, muestra un color rosado o rojo;

- *light red* con más del 60% de la superficie, en conjunto muestra un color rosado-rojizo o rojo y no más del 90% es de color rojo;

- *red* con más del 90% de la superficie de color rojo;

- *mixed color* cualquiera que no cumpla las designaciones anteriores en la gama de colores (USDA 1997).

Los métodos objetivos de evaluación del color del tomate incluyen la medición de la reflectancia y las técnicas de transmitancia de la luz. La determinación del contenido de pigmento (por ejemplo, clorofila, licopeno, β -caroteno) también se puede utilizar para indicar cambios de color en el tomate. La cantidad predominante de licopeno, que causa el color rojo de frutas, se mide mediante el parámetro a^* . El β -caroteno es un pigmento que determina el color naranja de las frutas, se mide por el parámetro b^* en el sistema de color *CIELab* (Sacks y Francis, 2001). El proceso de maduración de los tomates se caracteriza bien por la evolución del color de la superficie del fruto (Hertog y cols, 2007), donde la clorofila se descompone y los carotenoides, principalmente el licopeno, se acumulan durante la maduración (Brandt y cols., 2006).

La calidad del tomate estándar se basa principalmente en la uniformidad de la forma (redondo, aplanado, de pera u ovalado, entre otras) y en la ausencia de defectos de crecimiento y manejo. El tamaño no es un factor que defina el grado de calidad, pero puede influir de manera importante en las expectativas de su calidad comercial. De la misma forma que se analiza el color por medio de índices subjetivos, se pueden desarrollar escalas a partir de un análisis visual de daños, dados por la presencia de hongos sobre su superficie, la firmeza del fruto, entre otros.

Como objetivo de este capítulo nos proponemos realizar una primera aproximación a las características morfoagronómicas de las subpoblaciones de tomate 'Platense' y su comportamiento de poscosecha.

2.2. Materiales y Métodos

Búsqueda y pedido de los diferentes materiales

Durante el desarrollo de esta tesis se logró formar una colección de 42 muestras de tomate, obtenidas de distintas procedencias, entre ellas diferentes subpoblaciones de tomate ‘Platense’ (*Solanum lycopersicum* L.) y otras variedades de tomate mejoradas por el INTA - La Consulta (Mendoza) y por fitomejoradores de la Chacra Experimental de Gorina. A su vez, se utilizaron como materiales de referencia un híbrido comercial y un tomate silvestre (*Solanum pimpinellifolium* L.).

Los materiales tienen los siguientes orígenes:

De la Colección de semillas de la estación Experimental de Gorina dependiente del Ministerio de Desarrollo Agrario de la provincia de Buenos Aires:

-Tomate ‘Platense’: ‘Carcione’, ‘Gentile’ y ‘Del Manso’.

-Mejorados de Tomate ‘Platense’ por la estación experimental de Gorina: denominados como ‘LT’ y ‘J’.

De la “Unidad Promocional de Investigación y Desarrollo Semillas Hortícolas Locales del Cinturón Verde Platense” (UPID) de la Facultad de Ciencias Agrarias y Forestales - UNLP. El director de esta unidad, Ing. Agrónomo Juan José Garat, colaboró con las siguientes subpoblaciones de tomate ‘Platense’: ‘Luna’, ‘Gentile’ y ‘Del Manso’.

Del Banco de Germoplasma Hortícola del INTA - La Consulta: se realizó una búsqueda de todas las entradas del banco y se firmó un Acuerdo de Transferencia de Material (ATM) para retirar todas las entradas de tomate platenses disponibles, entre ellas: ‘Bustos’, ‘Carcione’, ‘Gentile’, ‘Luna’, ‘Prieto’, ‘Paolettich’, ‘Molinaro’, ‘Grasso’ y ‘Del Manso’. Dentro de los materiales solicitados se encontraba una entrada de platense, pero sin dato de cual subpoblación se trataba, por lo tanto, fue denominado como “sin dato”.

A pesar de tener diferentes orígenes, se mantuvo el nombre de las diferentes subpoblaciones, pero se distinguen entre ellas a partir del código interno que se describe en el Anexo I

Cultivares obtenidos a través de INTA - La Consulta: Uco Plata INTA, Uco 14 INTA, Uco 15 INTA, Uco 18 M INTA y Lomitero.

También se tomaron muestras de los tomates que son cultivados por el Grupo de Productores de Tomate Platense presentes durante la Fiesta del Tomate Platense realizada en 2015, denominadas como Grupo Tomate Platense.

Se compraron semillas comerciales de las empresas *Bonanza Seeds* y *Florensa* rotuladas en el envase como ‘Platense’ y el híbrido comercial, Elpida, en comercios de insumos agrícolas de La Plata.

De la Facultad de Ciencias Agrarias de la Universidad Nacional de Cuyo, gracias a la Dra. Iris Peralta, se obtuvieron las semillas del silvestre *S. pimpinellifolium* LA1589, que fue utilizado como *outgroup* de la filogenia realizada por medio de marcadores moleculares.

El total de las muestras obtenidas suman 42 entradas:

Tipo de entrada	Cantidad de entradas
Subpoblaciones de tomate ‘Platense’	27
Materiales seleccionados a partir de tomate ‘Platense’	4
Cultivares de tomate	6
Comerciales	4
Silvestre	1

Cuadro 2.1: Tipo y cantidad de entradas estudiadas.

En el Anexo I, se completa la información de origen del material utilizado.

Evaluación del germoplasma

Durante el segundo semestre del 2016, en los invernáculos de la Facultad de Ciencias Agrarias y Forestales - UNLP se realizaron las germinaciones de 25 semillas de cada una de las 42 entradas, excepto las semillas de *S. pimpinellifolium* LA1589 que se germinaron en primavera del 2017. La siembra se realizó en recipientes plásticos de 150 cc conteniendo un sustrato compuesto por tierra, turba y perlita en proporciones iguales. Los plantines fueron cultivados en invernadero durante los 30 días siguientes a su siembra, manteniendo una temperatura que no superó los 35 °C y realizando un riego diario (Figura 2.5.)



Figura 2.5: Siembra de las diferentes entradas evaluadas en invernadero vidriado - Facultad de Ciencias Agrarias y Forestales - UNLP.

De cada entrada obtenida se planificaron dos ensayos:

-Bajo cubierta en la Chacra Experimental de Gorina dependiente del actual Ministerio de Desarrollo Agrario (provincia de Buenos Aires). El objetivo del mismo era además de tener una primera aproximación de las características fenotípicas de estos materiales en condiciones bajo cubierta, era poder obtener semillas de plantas individuales (denominadas a, b, c y d)

-A campo en una quinta de una productora de la localidad de Romero (partido de La Plata). Este ensayo además de tener datos de las características fenotípicas en condiciones a

campo, era poder evaluar la calidad de poscosecha de los frutos y extraer semillas.

Las plantas fueron trasplantadas bajo cubierta en la Chacra Experimental de Gorina, se midieron las diferentes aptitudes agronómicas presentes en el Cuadro 2.2 para poder tener una primera aproximación de las características fenotípicas de estos materiales. Cada entrada fue considerada como un tratamiento con 4 repeticiones, las cuales fueron rotuladas como a, b, c y d. Antes de su siembra, se les extrajo ADN de sus hojas, como se describe en el Capítulo 3). Al final del ensayo, se cosecharon semillas de cada una de las plantas (a, b, c, d) (Figura 2.6 a y b).

El segundo ensayo fue sembrado a campo bajo un diseño de bloques aleatorios de todas las entradas (tratamientos) y con 6 plantas (repeticiones) (Figura 2.6 c y d). El acuerdo con los productores fue que establecieran el manejo del cultivo durante todo el ensayo de acuerdo a su experiencia y, a cambio, ellos podrían vender parte de la producción en el Paseo de la Economía social del Rectorado de la UNLP. En este cultivo se realizaron medidas agronómicas (descriptas en Cuadro 2.2) y se realizó una cosecha de frutos para medir tanto la productividad por entrada como su clasificación según el peso, es decir en tomates de primera (>150 gr), segunda ($150 - 90$ gr) y descarte (< 90 gr). Se escaneó parte de los frutos cosechados para poder analizarlos con el software *Tomato Analyzer* (Rodríguez y cols., 2010). Con este programa se pueden medir alrededor de 40 parámetros morfológicos de los frutos, entre ellos color (L^* , a^* , b^*) y diferentes relaciones de tamaño y forma. Por último, durante este ensayo se realizó un análisis de poscosechas en el laboratorio del CIDCA con el acompañamiento del Dr. Joaquín Hasperué y su equipo de investigación. De este ensayo también se obtuvieron semillas de todos los tratamientos, es decir de todos los genotipos evaluados.

Estadio fenológico	Variable (magnitud/estado)
Vegetativo	<ul style="list-style-type: none"> -Altura de la planta (cm) -Hábito de crecimiento (determinado, semideterminado, indeterminado) -Diámetro del tallo (mm)
Reproductivo	<ul style="list-style-type: none"> -Número de coronas por planta -Número de flores por planta
Productivo	<ul style="list-style-type: none"> -Número de frutos cuajados -Peso del fruto (gr) -Productividad estimada -Ancho del fruto -Parámetro de color L^* -Parámetro de color a^* -Parámetro de color b^* -Relación a^*/b^* -Medidas con programa <i>Tomato analyzer</i> -Número de lóculos
Poscosecha (en función del tiempo y 10 °C)	<ul style="list-style-type: none"> -Parámetro de color L^* -Parámetro de color a^* -Parámetro de color b^* -Relación a^*/b^* -Índice de color (<i>Hue</i>) -Índice de daño -Pérdida de peso de fruto promedio (gr)

Cuadro 2.2: Características morfoagronómicas y de calidad poscosecha medidas en los diferentes materiales estudiados.



Figura 2.6: Diseño y siembra de los dos ensayos en la Chacra Experimental de Gorina cubierta N° 1 y (a, b) en la quinta de los productores familiares de la Coop. Moto Méndez (c, d).

2.2.1. Cálculo de Grados Días

El crecimiento de las plantas está condicionado por una serie de factores ambientales como por ejemplo la luz, la temperatura, la humedad y la nutrición que reciben (Sobrinio Illescas y Sobrinio Vesperinas 1989). Por ello, se analizaron las temperaturas máxima y mínima durante el período de cultivo a fin de determinar lo que se denomina como grados día (llamado también suma de calor, grados día de desarrollo, unidades de calor o tiempo fisiológico) (Ruiz-Corral y col., 2002). Esta medida se utiliza para objetivar un determinado proceso de desarrollo o fase fenológica (Trudgill y cols., 2005). Entre otros, se puede utilizar para calcular la tasa de aparición de nudos, hojas, inflorescencias y desarrollo de frutos (Rodríguez y Flórez, 2006; NeSmith, 1997; Normand y Léchaudel, 2006; Almanza y cols.,

2010), y para estimar la producción potencial (Salazar y col., 2007). Para el cálculo de la acumulación de grados día ($^{\circ}D$) se utilizó la siguiente fórmula (Rodríguez y Flórez, 2006):

$$GDC(^{\circ}D) = (T^{\circ}máx. + T^{\circ}mín.) / 2 - T^{\circ}base$$

dónde $T^{\circ} máx.$ es la temperatura máxima diaria del aire, $T^{\circ} mín.$ es la temperatura mínima diaria del aire y $T^{\circ} base$ es la temperatura en la que el proceso metabólico del tomate es mínimo. Esta temperatura se estableció en 10 °C, de acuerdo a lo que recomiendan Rangel y Silvas (1987) y Zotarelli y cols. (2009). Debido a que uno de los ensayos fue a campo, en la quinta de un productor familiar, y a que no contamos con equipos de medición para dejar a campo, se utilizaron los datos climatológicos suministrados por el Departamento de Sismología e Información Meteorológica - FCAG - UNLP. La comparación entre $^{\circ}D$ y días se observa en el Cuadro 2.3 para el período de diciembre 2016 a marzo 2017.

Días	$^{\circ}D$
15	191,85
30	414,1
45	629,25
60	841,35
75	1035,3
90	1285,55
105	1461,65
120	1631,35

Cuadro 2.3: Equivalencia entre los días durante los que se realizó el ensayo y la acumulación de grados día calor ($^{\circ}D$).

2.3. Análisis de la calidad de poscosecha

La comercialización del tomate en fresco está condicionada principalmente por la aceptación del tamaño de los frutos y el color (Riquelme-Ballesteros, 1995). La estrecha relación que existe entre la evolución de los pigmentos y su estado de madurez, permite diferenciaciones en base a los cambios que experimenta el color. Se han desarrollado

numerosas escalas y cartas de color para realizar una clasificación subjetiva, de forma gráfica o descriptiva, del estado de maduración en una progresión de tomate verde a tomate rojo. Por otra parte, los colorímetros son instrumentos portátiles que han permitido ampliar las mediciones que ofrecen exactitud y de forma objetiva. Es instrumento puede medir distintos índices de color tanto en tomates frescos y salsas, incluyen el índice a^*/b^* y ofrecen lecturas en coordenadas de espacios de color CIE (L^* , a^* , b^*), entre otros.

2.3.1. Coordenadas L^* , a^* y b^* y la relación a^*/b^*

El color se evaluó en tres puntos, en lados opuestos de la región ecuatorial del fruto, con un colorímetro manual *MiniScan XE Plus* (HunterLab, serie 5348), que proporciona los registros Hunter de L^* , a^* y b^* , con los cuales se calculó el ángulo de tono (*Hue*) mediante la siguiente fórmula (McGuire, 1992):

$$Hue = \arctan(b^*/a^*)$$

Los resultados fueron expresados en grados ($^{\circ}$).

Los parámetros L^* , a^* y b^* delimitan una coordenada en espacio que corresponde al valor de un color, como se observa en la Figura 2.2. La descomposición del color se hace en las coordenadas que contempla los valores de L^* positivos hacia el blanco y valores negativos hacia el negro. Los valores positivos de a^* marcan rojo y los negativos el verde. Los valores positivos de b^* marcan amarillo y los negativos azul. Una relación a^*/b^* con valores negativos indica presencia del color verde (coordenada a^* con valores negativos) y la tendencia hacia valores positivos implica disminución del verde e incremento de los colores amarillo (coordenada b^* con valores positivos) y rojo (coordenada a^* con valores positivos), propios de la maduración progresiva de los frutos. Valores de ángulo del tono (*Hue*) de 0° a 90° indican una tendencia que va, paulatinamente, del color rojo cambiando por combinación con el amarillo. De 90° a 180° la tendencia es del amarillo al verde como en la Figura 2.2.

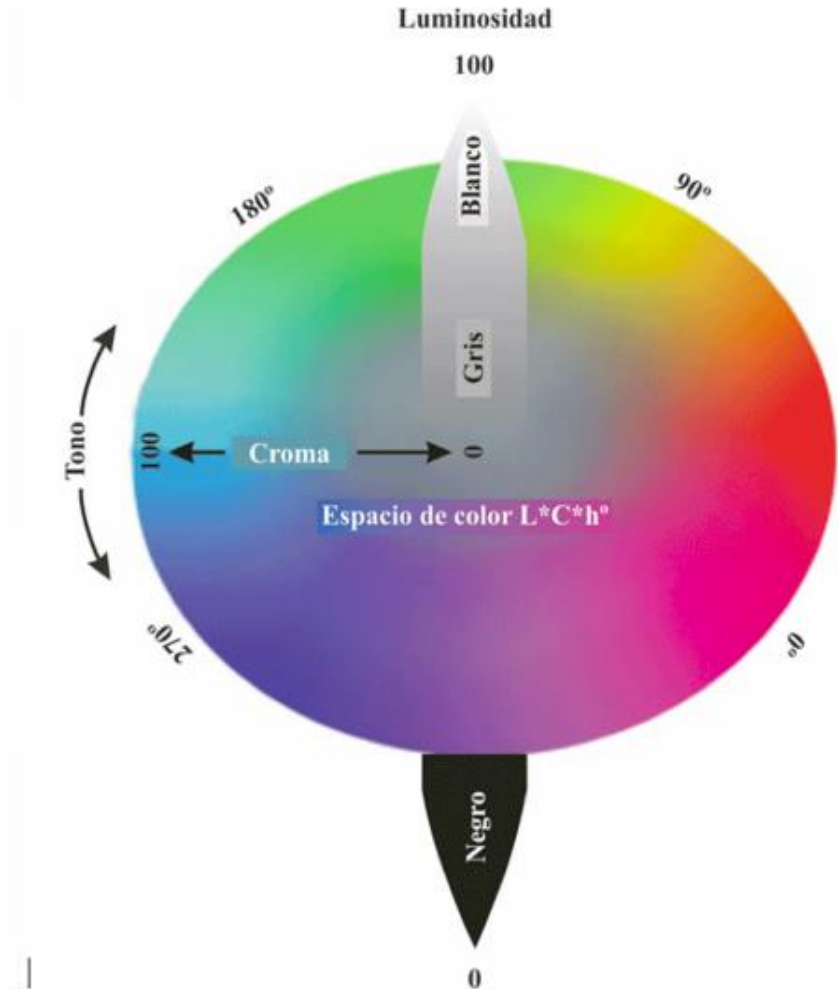


Figura 2.2: Representación del espacio cromático cilíndrico CIE-L*a*b*. Imagen tomada de Padrón Pereira y cols., 2012.

2.3.2 Porcentaje de pérdida de peso de los frutos

Se ha demostrado que la duración del almacenamiento, la temperatura de acopio y el tratamiento tienen efectos significativos sobre la pérdida de peso (Kumar y col, 1999). Posiblemente, una mayor tasa de transpiración en tomates almacenados a temperatura ambiente en comparación con tomates almacenados a temperaturas más bajas podría ser la causa principal de una mayor pérdida de peso. El peso del fruto se evaluó con una balanza granataria digital. Los frutos fueron pesados antes de su almacenamiento en cámara de frío y a los 6 y 16 días, para obtener la pérdida de agua acumulada en relación con su peso inicial.

Los resultados se expresaron en porcentaje de peso perdido en referencia al inicial.

Entre el período de 1289,8 y 1362,25 grados días se cosecharon todos los frutos en estado rojo (con más del 90% de la superficie de color rojo) de los siguientes materiales (tratamientos): ‘Uco Plata INTA’, ‘Prieto’, ‘Paolettich’, ‘Molinario’, ‘Luna’, ‘Grasso’, ‘Gentile’, ‘Del Manso’, ‘Carcione’, ‘Minuchi’, ‘Bustos’ y ‘Ventura’, con un número de muestra no menor a 6 frutos por tratamiento. Los materiales fueron analizados en el laboratorio de poscosecha del CIDCA. Las muestras fueron sumergidas previamente en agua clorada 100 PPM para bajar la carga microbiana y se dejaron secar para luego ser almacenadas en cámara de frío a temperatura de 10 °C.

A los 6 y 16 días de almacenamiento en cámara se evaluaron las siguientes variables de calidad: porcentaje de pérdida de peso, parámetros de color (a^* , b^* , L^* , relación a^*/b^* , *Hue*), índice de daño (*ID*) y porcentaje de frutos con presencia de hongos sobre su superficie.

2.4. Extracción de las semillas

La extracción de semillas de todos los frutos maduros se realizó mediante el método tradicional de fermentación. La pulpa y jugo de los tomates se dejó fermentar durante dos días a temperatura de 25-27 °C. Con esta fermentación, el gel que recubre la semilla es degradado y se reduce el inóculo de microorganismo patógenos. Las semillas se lavaron con abundante agua y se secaron por 10 hs a una temperatura de 32-35 °C, por medio de un deshidratador eléctrico. De esta manera, se obtuvieron semillas con baja humedad para ser almacenadas correctamente. Todos estos procedimientos se encuentran descriptos en el Manual de Producción de Semillas Hortícolas de Argerich y Gaviola⁶. Luego se tomaron al azar 10 entradas y se realizaron pruebas de germinación utilizando las condiciones estándares de la *International Seed Testing Association* - ISTA⁷. Finalmente, las semillas fueron guardadas en sobres herméticos de polietileno a 5 °C.

⁶ Ver en: <http://inta.gob.ar/documentos/manual-de-produccion-de-semillas-hortícolas.-tomate>

⁷ Ver en: <https://www.seedtest.org/en/home.html>

2.5. Resultados

2.5.1 Ensayo realizado en la Chacra Experimental de Gorina

Luego de acumular 489.95 °D en invernadero (35 días post siembra bajo cubierta) se midieron las variables vegetativas y reproductivas de las plantas sembradas por variedad (N=4) (Figura 2.3).



Figura 2.3: Ensayo bajo cubierta en la Chacra Experimental de Gorina en Cubierta N°1.

Los datos obtenidos se muestran en la Tabla 2.1. El monitoreo de plagas y riego estuvo a cargo de los trabajadores de la Chacra Experimental.

Accesión	Altura (cm)	Ø (mm)	Nºcoronas	Nºflores	Nºfrutos
G1-Uco Plata INTA	34.5 ^a ± 13.86	8,3 ^{abcde} fg ± 3.51	1	3.25	0
G2-Lomitero INTA	34,75 ^a ± 6.02	10,4 ^{cdef} gh ± 2,43	0,5	1,5	0
G3-Uco 14 INTA	39 ^{abc} ± 4.55	6,99 ^{abc} ± 1,85	1,15	7	1,5
G4-Río Grande	36.75 ^{ab} ± 16.5	7 ^{abc} ± 3,13	2,25	8	0,5
G5-Platense-Fecoagro	44.5 ^{abcde} fg ± 3.87	7,41 ^{abcde} f ± 1,71	3,5	12,75	3,5
G6-Platense- Bonanza	46,50 ^{abcde} fg ± 1,91	11,14 ^f gh ± 1,18	4	13,5	4,5
G7-Platense- Florensa	41,67 ^{abcde} ± 4,93	5,02 ^a ± 1,51	2	9,67	1,67
G8-Carcione	50,25 ^{bcde} fg ± 3,86	8,43 ^{abcde} fg ± 3,29	4	13,75	0,25
G9-Gentile	64 ⁱ ± 6,27	11,5 ^{gh} ± 3,42	4	17,25	1,5
G10-Del Manso	58,25 ^{ghi} ± 4,57	10,08 ^{bcde} fg ± 1,56	4	12,75	2,25
G11-Luna	50,5 ^{bcde} fg ± 10,02	8,84 ^{bcde} fg ± 0,84	3,25	9	1,5
G13-Del Manso	54,5 ^{efghi} ± 3,7	9,73 ^{bcde} fg ± 1,02	3,75	11,5	4,5
G14-Gentile	53,25 ^{efgh} ± 4,19	10,56 ^{efgh} ± 2,23	3,75	10,25	3
G15-Prieto	56,33 ^{efghi} ± 4,51	12,5 ^h ± 1,5	3	9,5	1,75
G16-Paolettich	46,25 ^{abcde} fg ± 12,71	8,01 ^{abcde} fg ± 1,48	2,25	6	1,75
G17-Molinaro	48 ^{abcde} fg ± 8,41	9,06 ^{bcde} fg ± 1,26	2,25	7,5	0,75
G18-Luna	53,5 ^{defghi} ± 13,43	9,92 ^{bcde} fg ± 3,62	2,5	8	2,5
G19-sin dato	58,75 ^{hi} ± 8,06	8,02 ^{abcde} fg ± 1,13	3	10,5	2
G20-Grasso	44,00 ^{abcde} f ± 2,94	7,06 ^{abc} ± 1,19	2,75	6,75	0,25
G21-Grasso	40,33 ^{abcd} ± 5,03	6,37 ^{ab} ± 1,31	0,67	2,67	0
G22- Gentile	42,75 ^{abcde} f ± 7,46	7,04 ^{abc} ± 2,61	1	3,25	0
G23-Del Manso	49,75 ^{bcde} fg ± 5,03	7,98 ^{abcde} fg ± 2,04	1,5	2,25	0
G24-Carcione	49 ^{bcde} fg ± 0,82	8,02 ^{abcde} fg ± 1,33	2,25	5	0
G25-Carcione	46,75 ^{abcde} fg ± 3,59	7,32 ^{abcde} ± 1,23	1,25	3	0
G26- Bustos	47,25 ^{abcde} fg ± 11,47	7,13 ^{abcd} ± 2,6	2	4	0
G28- GTP	47,50 ^{cde} fg ± 0,58	7,18 ^{abcde} ± 1,79	2	4,75	0
G29- Minuchi	49,67 ^{bcde} fg ± 3,51	8,3 ^{abcde} fg ± 0,95	2,33	5,33	0,33
G30- GTP	52,75 ^{cde} fghi ± 2,75	9,09 ^{bcde} fg ± 1,66	2	5,5	0,25
G31-Bustos	45,50 ^{abcde} fg ± 8,81	9,21 ^{bcde} fg ± 1,66	2	6,5	0
G32-Ventura	53 ^{defghi} ± 6,88	7,84 ^{abcde} fg ± 1,98	3,5	8,75	0,25
G33-Grasso	42,50 ^{abcde} f ± 5,45	8,6 ^{abcde} fg ± 0,73	0,5	3,25	0
G34- Carcione	45,50 ^{abcde} fg ± 5,2	7,34 ^{abcde} ± 1,85	1,25	5	0,25

Tabla 2.1: Valores medidos y desvíos estándar de las variables morfo-agronómicas evaluadas bajo cubierta en la Chacra Experimental de Gorina - Campaña 2017. Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Duncan ($p > 0.05$).

En esta primera toma de datos, pudimos observar que los genotipos platenses de las empresas comerciales de Fecoagro y Bonanza (G5 y G6) fueron los más precoces en cuajado de frutos, junto a ‘Del Manso’ y ‘Gentile’ (G10, G13 y G14). En un segundo grupo, se pueden separar el tomate, de crecimiento determinado y forma de fruto perita, Uco 14 INTA (G3),

el platense de *Florensa* (G7) y las subpoblaciones ‘Gentile’, ‘Prieto’ y ‘Paolettich’ (G9, G15, G16). El genotipo ‘Gentile’ a su vez presentó la mayor cantidad de flores y una relación promedio de 4 flores por corona. Los materiales platense ‘Gentile’ (G9), ‘Del Manso’ (G10, G13), ‘Prieto’ (G15), ‘Luna’ (G18), sin dato (G19), grupo de tomate platense (G30) y ‘Ventura’ (G32) fueron los que se destacaron por altura.

Este ensayo se vio interrumpido por una gran tormenta con ráfagas de viento que azotó a La Plata en febrero de 2017⁸ y que dejó al invernadero sin plásticos (Figura 2.4). Sin embargo, más allá de este inconveniente climático, las plantas se pudieron replantar y se continuó con su cultivo para una posterior cosecha de semillas, pero no se pudieron realizar los análisis de poscosecha de sus frutos.



8 Ver en: <https://www.eldia.com/nota/2017-2-6-una-tormenta-de-viento-desato-el-caos-con-rafagas-de-hasta-110-kilometros-por-hora>



Figura 1.4: Cubierta N°1 de la Estación Experimental de Gorina sin plásticos después de la tormenta del 5 de febrero 2017.

2.5.2 Ensayo realizado en la quinta de productores familiares

Como se mencionó anteriormente, el resto de las plantas que no se sembraron bajo cubierta en la Chacra Experimental de Gorina, se llevaron a la quinta de una productora, Virginia Quispe, que participa activamente del proyecto de extensión “Banco de Germoplasma: fortalecimiento de la Agricultura Familiar”. Se realizó un diseño experimental de bloques al azar y el cuidado del cultivo lo manejaron los productores, según su experiencia (Figura 2.5.). En charlas durante las recorridas del cultivo, nos contaron que ellos habían tenido experiencia en la producción de tomate tipo ‘Platense’ a campo. La conducción de las plantas se realizó en forma de V y encañadas, permitiendo obtener plantas robustas que no se vieron afectadas por la tormenta del 5 de febrero, más allá de que en la zona hubo importantes destrozos en los invernaderos.



Figura 2.5: Cultivo a campo en la quinta de los productores familiares.

A los 742,2 °D se realizaron las medidas a campo de los caracteres vegetativos y reproductivos (Tabla 2.2.).

Accesión	Altura (cm)	Ø (mm)	Nº coronas	Nº flores	Nº frutos
G1-Uco Plata INTA	73,47 ^{cdefg} ± 8,98	10,9 ^{bc} ± 1,82	6,87	14,6	3,4
G2-Lomitero INTA	54,2 ^{abcd} ± 9,43	11,07 ^c ± 2,68	4,8	11,7	2,5
G3-Uco 14 INTA	45,36 ^a ± 6,15	7,19 ^a ± 0,52	6,82	20,36	5,82
G4-Rio Grande	55,75 ^{abcde} ± 5,75	7,54 ^{ab} ± 1,76	9	18,75	13,08
G5-Platense-Fecoagro	59,75 ^{abcde} ± 8,62	9,89 ^{abc} ± 2,37	6,17	17,17	9,75
G6-Platense- Bonanza	61,2 ^{abcdefg} ± 9,38	9,46 ^{abc} ± 0,73	6,3	12,3	10,8
G7-Platense- Florensa	48,6 ^{ab} ± 3,65	9,32 ^{abc} ± 1,37	7,4	8,8	8,2
G8-Carcione	80,88 ^{fg} ± 10,86	10,1 ^{abc} ± 1,23	7,76	21,59	7,82
G9-Gentile	74,28 ^{defg} ± 11,22	11 ^c ± 1,92	7,89	23,78	8,56
G10-Del Manso	71,24 ^{cdefg} ± 9,97	11,06 ^c ± 1,94	6,65	21,53	6,88
G11-Luna	68,42 ^{bcddefg} ± 8,54	9,02 ^{abc} ± 1,11	6,33	14,58	6,5
G13-Del Manso	74,17 ^{defg} ± 4,36	10,59 ^{bc} ± 2,03	6,5	13,83	5,67
G14-Gentile	69,33 ^{bcddefg} ± 10,56	11,21 ^c ± 1,27	6,42	15,75	4,92
G15-Prieto	71,78 ^{cdefg} ± 11,3	11,39 ^c ± 2,38	6,44	23,72	6,5
G16-Paolettich	67 ^{bcddefg} ± 11,14	11,9 ^c ± 2,06	7,76	23,71	5
G17-Molinario	67,91 ^{bcddefg} ± 12,96	10,79 ^{bc} ± 2,66	6,64	18,36	4,45
G18-Luna	73,11 ^{cdefg} ± 12,51	10,12 ^{abc} ± 1,24	6,28	22,06	6,11
G19-sin dato	76,41 ^{efg} ± 15,47	10,63 ^{bc} ± 1,51	6,29	19,47	5,18
G20-Grasso	76,06 ^{efg} ± 11,42	10,48 ^{abc} ± 1,28	6,41	20,59	5,65
G21-Grasso	67,31 ^{bcddefg} ± 13,47	10,24 ^{abc} ± 1,54	4,63	17,31	3,69
G22- Gentile	60,06 ^{abcdef} ± 11,26	10,57 ^{bc} ± 1,33	5,63	21,13	3,88
G23-Del Manso	64,25 ^{abcdefg} ± 10,82	12,23 ^c ± 1,34	4,42	11,83	3
G24-Carcione	82,08 ^g ± 15,63	11,39 ^c ± 2,06	5,67	16,83	3,58
G25-Carcione	66,1 ^{bcddefg} ± 9,27	11,39 ^c ± 1,04	4,4	13,9	0,8
G26- Bustos	67,27 ^{bcddefg} ± 14,05	11,29 ^c ± 1,41	7,33	24,53	4,47
G28- GTP	62,25 ^{abcdefg} ± 12,14	10,83 ^{bc} ± 1,36	4,19	15,63	1,44
G29- Minuchi	60,82 ^{abcdef} ± 11,9	11,8 ^c ± 1,52	4,53	15,41	1,53
G30- GTP	58,88 ^{abcde} ± 7,61	10,89 ^c ± 2,16	6,13	23,44	3,88
G31-Bustos	68,36 ^{bcddefg} ± 7,17	11,38 ^c ± 1,33	6,73	25,27	4,27
G32-Ventura	57,31 ^{abcde} ± 7,41	10,13 ^{abc} ± 1,71	6,88	17,19	4,44
G33-Grasso	52,8 ^{abc} ± 9,35	10,98 ^c ± 1,76	3,86	14,5	3
G34- Carcione	58,53 ^{abcde} ± 9,73	10,35 ^{abc} ± 1,61	5	16,18	0,76

Tabla 2.2: Valores medidos y desvíos estándar de las variables morfo-agronómicas evaluadas en cultivo a campo en la quinta de productor familiar de La Plata. Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Duncan ($p > 0.05$).

En este ensayo se observó que ‘Rio Grande (G4), de crecimiento determinado y forma de fruto perita, junto al comercial de Fecoagro (G5) y Bonanza (G6) fueron los materiales con más frutos cuajados. En cuanto a floración ‘Gentile’ (G9), ‘Prieto’ (G15), ‘Paolettich’ y ‘Bustos’ (G26 y G31) fueron los más productivos. En cuanto al crecimiento de altura de plantas fue parejo para todos los materiales, seguramente influenciado también por la forma de conducción del cultivo en forma de V.

Si comparamos los dos ensayos, tanto bajo cubierta como a campo, se puede observar que el segundo posee una mayor cantidad de frutos cuajados y número de flores (más allá de que el ensayo a campo posee una mayor acumulación de grados días, ya que se midió semanas más tarde). Esto va relacionado directamente también con el tamaño de la planta y diámetro. De estos dos ensayos no se pueden sacar conclusiones estadísticas de comparación entre los mismos, sino tener una primera caracterización de estos materiales. Sin dudas, la forma de cultivo en V por parte de los productores y en forma lineal en la Chacra Experimental, no facilita la comparación entre los ensayos, pero muestra las ventajas de rendimiento para los productores.

A los 916,55 °D se cosecharon todos los frutos de dos plantas por tratamiento (de las 6 disponibles) (Figura 2.7.). Los frutos fueron clasificados según su peso y se obtuvieron los siguientes resultados:

Accesión	Peso 1 ^{era} (gr.)	Peso 2 ^{da} (gr.)	Peso descarte (gr.)	Peso total (gr.)
G1-Uco Plata INTA	1020	103	160	1283
G2-Lomitero INTA	1320	220	0	1540
G3-Uco 14 INTA	300	250	40	590
G4-Rio Grande	895	500	30	1425
G5-Platense-Fecoagro	525	1110	330	1965
G6-Platense- Bonanza	690	1155	485	2330
G7-Platense- Florensa	230	390	170	790
G8-Carcione	946	516	120	1582
G9-Gentile	1000	965	95	2060
G10-Del Manso	815	390	45	1250
G11-Luna	1845	240	0	2085
G13-Del Manso	515	0	0	515
G14-Gentile	1360	0	45	1405
G15-Prieto	540	185	403	1128
G16-Paolettich	1323	210	0	1533
G17-Molinaro	1190	65	0	1255
G18-Luna	1036	216	50	1302
G19-sin dato	1013	116	103	1232
G20-Grasso	2330	280	0	2610
G21-Grasso	1736	283	0	2019
G22- Gentile	1090	95	0	1185
G23-Del Manso	645	0	0	645
G24-Carcione	806	125	40	971
G25-Carcione	530	95	0	626
G26- Bustos	583	270	0	853
G28- GTP	480	0	0	480
G29- Minuchi	1545	110	45	1700
G30- GTP	1083	133	0	1216
G31-Bustos	1530	360	36	1926
G32-Ventura	736	470	45	1242
G33-Grasso	886	185	0	1116
G34- Carcione	476	96	0	572

Tabla 2.3: Pesos de los frutos clasificados por peso, cosechados a campo en la quinta de productor familiar de La Plata.

Cuando se clasifican los frutos por peso (figura 2.6), se pudo observar que la mayoría de las subpoblaciones de ‘Platense’ se caracterizan por tener frutos con peso mayor a 150 g y casi sin descarte (menor a 90 g), salvo para el caso ‘Prieto’ y ‘Ventura’ (G15 y G32) donde hay una mayor variación de tamaños. El material G7 es un platense comercial (*Florensa*), en el cual se observa que fue seleccionado a partir de un tamaño de sus frutos de peso intermedio

(entre 90-150 g). Es común que en los procesos de selección se busquen tamaños de frutos intermedios (que no superen los 150 g) por “gusto de los consumidores” en la relación de cuántos tomates entran en un kilo y que sean de buen tamaño.

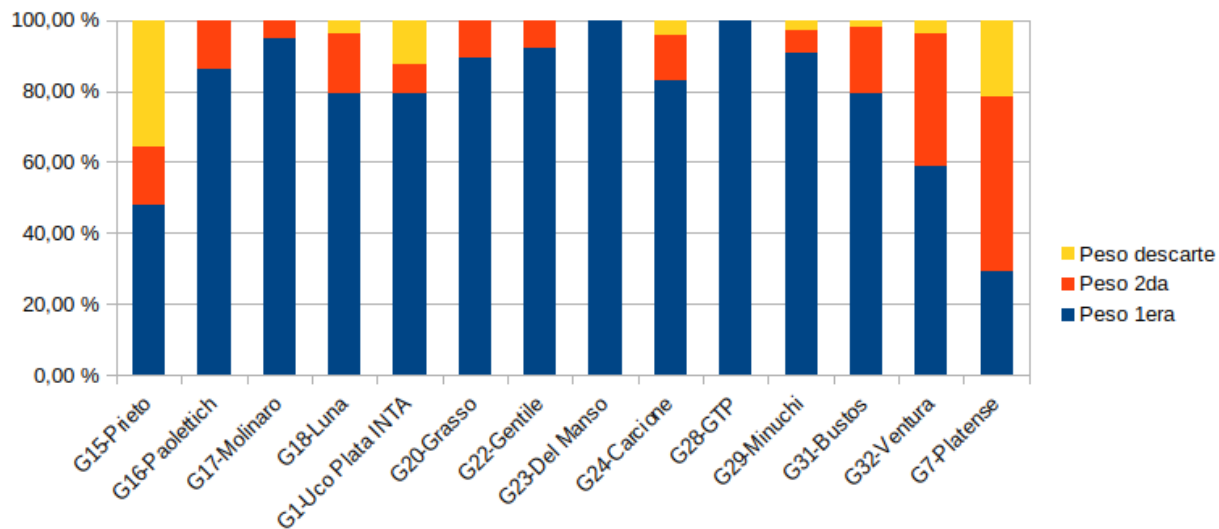

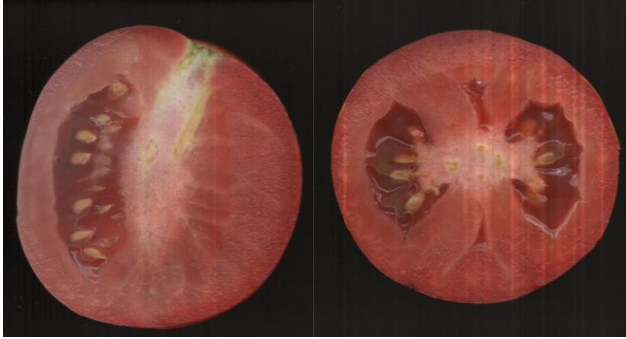


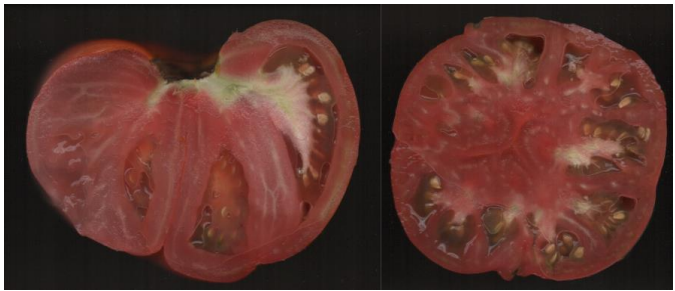



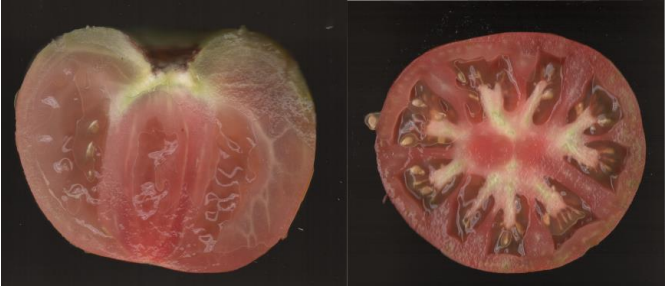


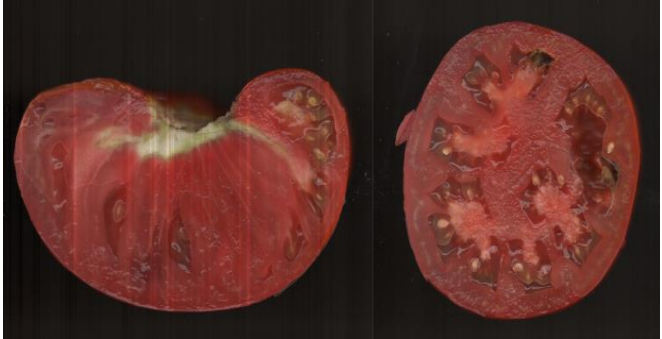
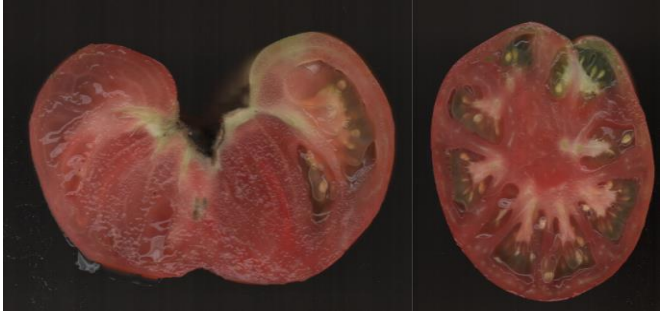


Figura 2.6: Gráfico de barras con porcentaje de peso de frutos según de algunas de las variedades según la clasificación según los pesos.

Antes de extraer las semillas de los frutos cosechados y pesados, éstos se escanearon y se procesó su imagen utilizando el programa *Tomato analyzer*. Los resultados se pueden observar en la Tabla 2.4:

Variedad	Característica	Valor	
G1- UCO PLATA INTA 	Peso promedio fruto (g)	148,22 ± 37,16	
	Perímetro (cm)	28,09 ± 3,56	
	Área (cm ²) N° de lóbulos	50,07 ± 9,46	
	Max. Ancho (cm)	8,2 ± 0,42	
	<i>L*</i>	7,97 ± 0,41	
	<i>a*</i>	39,55 ± 1,49	
	<i>b*</i>	25,08 ± 3,51	
	<i>Hue</i>	23,85 ± 2,06	
			43,74 ± 2,87

<p style="text-align: center;">G3-Uco 14 INTA</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>79,44 ± 16,36 17,98 ± 0,59 22,36 ± 1,32 2 ± 0 5,57 ± 0,24 43,43 ± 2,15 34,4 ± 2,73 28,41 ± 2,73 39,44 ± 1,48</p>
<p style="text-align: center;">G7-Platense</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>180 ± 70,71 16,49 ± 1,11 21,35 ± 3,32 7,67 ± 0,52 6 ± 1,41 46,05 ± 0,45 26,82 ± 0,51 27,44 ± 0,15 45,66 ± 0,71</p>
<p style="text-align: center;">G15- Prieto</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>171,89 ± 56,64 26,27 ± 2,05 47,65 ± 7,09 7,5 ± 0,58 7,71 ± 0,99 42,57 ± 4,22 23,72 ± 9,74 24,79 ± 3,93 47,93 ± 11,74</p>
<p style="text-align: center;">G16 – Paolettich</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>143,34 ± 40,73 30,19 ± 2,01 58,37 ± 5,81 10,33 ± 0,52 8,73 ± 0,49 38,39 ± 1,14 28,30 ± 2,52 23,63 ± 1,76 39,88 ± 0,77</p>

<p style="text-align: center;">G17- Molinaro</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>231,29 ± 67,29 27,29 ± 1,54 49,45 ± 5,52 9,2 ± 1,1 8,04 ± 0,52 40,77 ± 1.,56 25,8 ± 4,84 24,17 ± 4,84 43,41 ± 3,3</p>
<p style="text-align: center;">G18- Luna</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>151,02 ± 37,45 37,21 ± 18,93 51,72 ± 6,47 8,5 ± 0,58 8,78 ± 0,69 39,4 ± 4,77 33,03 ± 4,77 27,29 ± 3,32 39,6 ± 0,37</p>
<p style="text-align: center;">G20- Grasso</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>182,05 ± 51,44 29,92 ± 1,92 58,32 ± 4,98 12,5 ± 2,45 8,61 ± 0,33 41,59 ± 1,91 29,93 ± 4,39 24,79 ± 3,51 39,66 ± 2,37</p>
<p style="text-align: center;">G22- Gentile</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>158,64 ± 10,65 30,31 ± 4,22 54,99 ± 7,77 11,83 ± 1,85 8,38 ± 0,66 40,90 ± 4,92 25,32 ± 8,45 25,08 ± 8,46 46,23 ± 12,12</p>

<p style="text-align: center;">G23- Del Manso</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>193,60 ± 55,28 29,10 ± 0,47 54,71 ± 0,99 10 ± 0 8,54 ± 0,01 42,04 ± 1,19 35,94 ± 0,3 29,68 ± 0,12 39,55 ± 0,35</p>
<p style="text-align: center;">G24- Carcione</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>219,61 ± 75,17 28,33 ± 2,57 52,99 ± 5,76 11 ± 1,15 8,34 ± 0,44 32,95 ± 21,57 19,90 ± 12,61 18,33 ± 11,73 32,09 ± 21,25</p>
<p style="text-align: center;">G28- Grupo Tomate Platense</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>240 ± 11,55 30 ± 3,21 49,87 ± 2,03 9 ± 1,15 8,40 ± 0,38 42,99 ± 0,82 26,85 ± 1,42 24 ± 1,09 41,80 ± 0,21</p>
<p style="text-align: center;">G29- Minuchi</p> 	<p>Peso promedio fruto (g) Perímetro (cm) Área (cm²) Nº de lóbulos Max. Ancho (cm) <i>L*</i> <i>a*</i> <i>b*</i> <i>Hue</i></p>	<p>258,36 ± 92,95 31,14 ± 3,34 60,08 ± 7,62 10,75 ± 1,75 8,33 ± 0,63 40,86 ± 3,55 31,56 ± 3,3 26,19 ± 3,02 39,68 ± 1,7</p>



<p style="text-align: center;">G31- Bustos</p> 	<p>Peso promedio fruto (g)</p> <p>Perímetro (cm)</p> <p>Área (cm²)</p> <p>Nº de lóbulos</p> <p>Max. Ancho (cm)</p> <p><i>L*</i></p> <p><i>a*</i></p> <p><i>b*</i></p> <p><i>Hue</i></p>	<p>166,67 ± 32,27</p> <p>28,26 ± 1,23</p> <p>52,86 ± 1,56</p> <p>9,67 ± 0,52</p> <p>8,24 ± 0,13</p> <p>39,87 ± 1,37</p> <p>28,67 ± 2,76</p> <p>24,03 ± 2,46</p> <p>39,95 ± 1,13</p>
<p style="text-align: center;">G32- Ventura</p> 	<p>Peso promedio fruto (g)</p> <p>Perímetro (cm)</p> <p>Área (cm²)</p> <p>Nº de lóbulos</p> <p>Max. Ancho (cm)</p> <p><i>L*</i></p> <p><i>a*</i></p> <p><i>b*</i></p> <p><i>Hue</i></p>	<p>117,59 ± 42,45</p> <p>27,54 ± 2,55</p> <p>48,53 ± 9,2</p> <p>9,29 ± 1,33</p> <p>7,95 ± 0,66</p> <p>41,95 ± 2,9</p> <p>27,32 ± 3,53</p> <p>25,03 ± 2,9</p> <p>42,54 ± 2,19</p>

Tabla 2.4: Escaneado y análisis de la imagen por medio del programa *Tomato analyzer*.

2.6 Análisis de vida poscosecha

Pasada la acumulación de los 900 °D, se tomaron muestras de frutos, del ensayo de la quinta de los productores, para ser analizados en el Laboratorio de Investigación en Productos Agroindustriales (CIDCA-UNLP-CONICET) (Figura 2.7), donde se caracterizaron los parámetros más relevantes para el análisis de vida poscosecha.



Figura 2.7: Cosecha de todos los frutos maduros de dos plantas por tratamiento en la quinta de la productora familiar.

Peso promedio de frutos

Como se observa en los resultados de la Tabla 2.5, el peso promedio más alto (mayor a 150 g) fue del grupo conformado por ‘Minuchi’, ‘Carcione’, ‘Del Manso’, ‘Grasso’, ‘Luna’ y ‘Molinaro’ (G29, G24, G23, G20, G18 y G17) pero con alto desvío estándar. El menor pesaje correspondió al grupo ‘Ventura’, ‘Bustos’, ‘Gentile’, ‘Paolettich’, ‘Prieto’, Uco Plata-INTA (G32, G31, G22, G16, G15 y G1). Este grupo también se correlaciona con el resultado de tamaño de frutos graficados en la Figura 2.6 donde se encuentran los genotipos con un cierto porcentaje de frutos de segunda (pesos menores a los 150 g).

Accesión	Peso promedio (g)	D.E.
G1-Uco Plata-INTA	148,22 ^{ab}	37,16
G15-Prieto	171,89 ^{abcd}	56,64
G16-Paolettich	143,34 ^{abc}	40,73
G17-Molinaro	231,29 ^e	67,29
G18-Luna	151,02 ^{de}	37,45
G20-Grasso	182,05 ^{bcd}	51,44
G22-Gentile	158,64 ^{abc}	10,65
G23-Del Manso	193,60 ^{bcd}	55,28
G24-Carcione	219,61 ^{cde}	75,17
G29-Minuchi	258,36 ^e	92,95
G31-Bustos	166,67 ^{abc}	32,27
G32-Ventura	117,59 ^a	42,45

Tabla 2.5: Peso promedio de frutos recolectados en la quinta del productor familiar. Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Fisher ($p > 0.05$).

Pérdida de peso durante el almacenamiento

La pérdida de peso durante el almacenamiento de poscosecha, de los frutos cosechados de la quinta del productor, se determinó restando los pesos de la muestra de sus pesos registrados previamente, y se presentó como porcentaje de pérdida de peso en comparación con el peso inicial, a los 6 y 16 días almacenados en frío (10 °C) (Tabla 2.6).

Accesión	N	Día 6	D.E.	Día 16	D.E.
G1-Uco Plata-INTA	9	0,540 ^a	0,008	1,877 ^{cde}	0,26
G15-Prieto	12	0,567 ^a	0,18	1,307 ^{ab}	0,44
G16-Paolettich	14	0,402 ^a	0,091	1,304 ^{abc}	0,279
G17-Molinaro	6	0,87 ^b	0,43	2,26 ^e	0,21
G18-Luna	9	0,496 ^a	0,13	1,348 ^{abcd}	0,38
G20-Grasso	12	0,568 ^a	0,175	1,770 ^{bcd}	0,183
G22-Gentile	12	0,453 ^a	0,11	1,506 ^{abcd}	0,37
G23-Del Manso	11	0,504 ^a	0,1	1,276 ^a	0,28
G24-Carcione	10	0,542 ^a	0,12	1,347 ^{ab}	0,31
G29-Minuchi	11	0,443 ^a	0,11	1,259 ^{abc}	0,36
G31-Bustos	16	0,481 ^a	0,153	1,616 ^{abcd}	0,367
G32-Ventura	9	0,627 ^a	0,13	1,934 ^{de}	0,61

Tabla 2.6: Porcentaje de pérdida de peso de los frutos a los 6 y 16 días almacenados en cámara de frío (10 °C). Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Fisher ($p > 0,05$).

La pérdida de peso de los tomates en almacenamiento durante los primeros 6 días a 10 °C no fue significativa en la mayoría de los casos, excepto para la población ‘Molinario’ (G17) que a los 16 días de almacenamiento continuó con una tasa de pérdida de peso de 0,14% por día. El grupo de materiales más aptos para almacenamiento fueron ‘Prieto’, ‘Paolettich’, ‘Luna’, ‘Gentile’, ‘Carcione’, ‘Minuchi’ y ‘Bustos’ (G15, G16, G18, G22, G24, G29 y G31) con una tasa de pérdida de peso promedio de 0,09% por día. Dentro de este grupo, la variedad denominada ‘Minuchi’ presentó el menor valor con 1,259% a los 16 días de poscosecha.

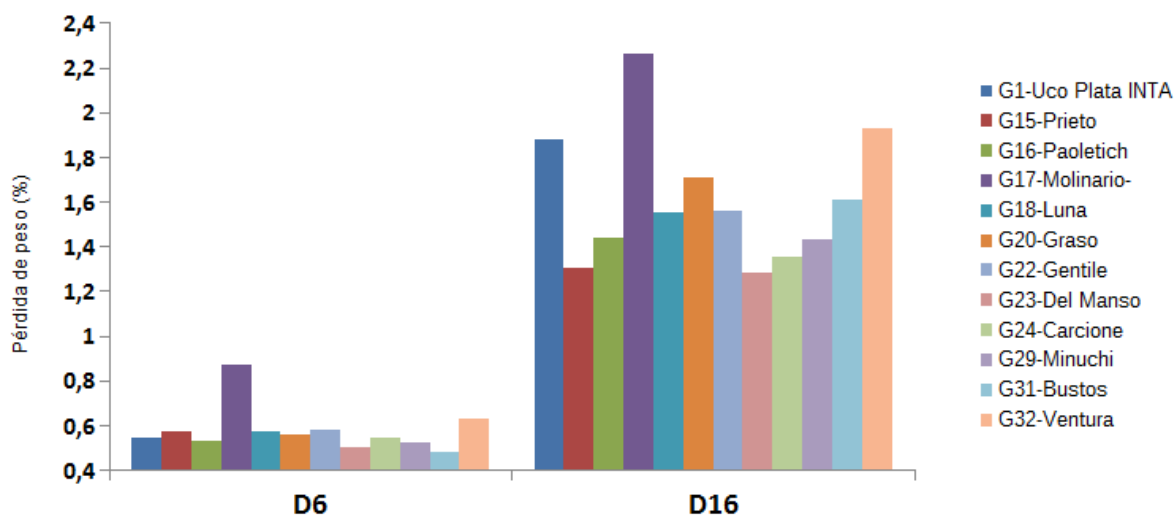


Figura 2.8: Gráfico de barras con el porcentaje de pérdida de peso (%), al día 6 (D6) y día 16 (D16) almacenados en cámara de frío a 10 °C. Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Fisher ($p > 0.05$).

Valores de *Hue* y relación a^*/b^*

Se tomaron medidas de L^* , a^* y b^* , para calcular la relación a^*/b^* y los de *Hue* de los materiales estudiados en cámara de frío, a día 0 (D0), 6 (D6) y 16 días (D16) de almacenamiento. Los resultados se observan en la siguiente tabla (2.7):

Accesión	<i>L</i> *	D.E.	<i>a</i> *	<i>b</i> *	Hue	D.E.	<i>a</i> */ <i>b</i> *	D.E.
D0								
G1- Uco Plata Inta	47,57	2,44	16,59	32,36	62,92 ^b	6,28	0,51 ^{abc}	0,14
G15- Prieto	48,31	3,57	17,11	30,75	61,34 ^{abcd}	10,25	0,55 ^{abcd}	0,22
G16- Paolettich	47,56	2,71	16,06	31,18	63,06 ^{abcd}	3,07	0,51 ^{abcd}	0,21
G17-Molinaro	51,13	2,51	16,52	38,44	66,77 ^d	3,48	0,43 ^a	0,17
G18- Luna	47,49	3,19	17,46	32,29	62,06 ^{abcd}	2,98	0,54 ^{abcd}	0,21
G20- Grasso	47,31	3,19	17,33	31,56	61,29 ^{abc}	3,7	0,54 ^{cd}	0,03
G22- Gentile	49,16	3,23	15,79	32,73	65,21 ^d	2,9	0,48 ^a	0,28
G23- Del Manso	49,49	2,27	17,43	32,40	62,01 ^{abcd}	3,12	0,53 ^{abcd}	0,17
G24- Carcione	47,25	3,6	17,01	30,07	60,79 ^{abcd}	2,98	0,56 ^{abcd}	0,2
G29- Minuchi	48,20	3,46	15,96	32,35	64,12 ^{cd}	3,7	0,49 ^{ab}	0,22
G31- Bustos	48,69	3,18	19,09	32,35	59,74 ^{abc}	2,93	0,59 ^{bcd}	0,18
G32- Ventura	47,82	2,77	18,03	31,36	60,24 ^a	2,5	0,57 ^d	0,13
Elpida	45,51	1,7	20,87	33,27	58,16 ^{ab}	6,24	0,62 ^{cd}	0,15
D6								
G1- Uco Plata Inta	45,14	2,62	17,3	33,59	62,73 ^d	2,94	0,51 ^a	0,12
G15- Prieto	46,33	2,89	19,96	31,88	57,95 ^{abc}	2,47	0,62 ^{cd}	0,12
G16- Paolettich	45,6	2,68	18,81	30,82	58,64 ^{abc}	3,47	0,62 ^{cd}	0,18
G17-Molinaro	46,55	2,84	18,9	33,37	59,81 ^{cd}	2,6	0,59 ^{abc}	0,13
G18- Luna	45,27	2,67	19,94	31,77	57,93 ^{abc}	3,33	0,62 ^{cd}	0,12
G20- Grasso	45,36	3,26	19,97	31,92	58,00 ^{ab}	3,71	0,62 ^d	0,16
G22- Gentile	47,27	3,12	17,51	32,46	62,16 ^d	2,55	0,53 ^{ab}	0,22
G23- Del Manso	46,90	2,88	19,52	35,0	60,85 ^{abc}	3,84	0,55 ^{bcd}	0,12
G24- Carcione	43,41	3,01	21,16	31,77	56,53 ^a	3,55	0,66 ^d	0,14
G29- Minuchi	46,32	3,53	19,24	32,77	59,79 ^{bcd}	3,22	0,58 ^{bcd}	0,21
G31- Bustos	45,47	3,21	22,16	33,92	57,05 ^{abc}	3,0	0,65 ^{cd}	0,16
G32- Ventura	45,16	2,51	19,67	31,73	58,29 ^{abc}	2,21	0,61 ^{cd}	0,10
Elpida	44,87	1,24	21,64	33,38	57,29 ^{abc}	5,67	0,64 ^{cd}	0,14
D16								
G1- Uco Plata Inta	44,63	2,71	16,48	34,02	64,17 ^e	3,06	0,48 ^a	0,1
G15- Prieto	43,69	3,13	22,15	34,34	57,03 ^{bcd}	3,28	0,64 ^{bcd}	0,09
G16- Paolettich	42,95	2,57	22,66	33,74	55,98 ^{bcd}	3,99	0,67 ^{bcd}	0,11
G17-Molinaro	44,71	2,98	21,46	33,36	56,85 ^{bcd}	5,41	0,66 ^{bcd}	0,14
G18- Luna	43,38	2,85	21,54	33,65	57,29 ^{cd}	4,19	0,64 ^{bc}	0,12
G20- Grasso	43,61	3,45	20,74	33,14	57,86 ^{ab}	3,78	0,62 ^{de}	0,11
G22- Gentile	44,00	2,65	21,76	33,95	57,29 ^d	3,19	0,64 ^{bc}	0,16
G23- Del Manso	44,89	2,93	22,68	36,30	58,02 ^d	3,74	0,62 ^{bc}	0,15
G24- Carcione	42,63	3,17	20,69	30,58	56,19 ^{abc}	3,79	0,67 ^{cd}	0,12
G29- Minuchi	44,69	2,44	22,36	35,27	57,60 ^{cd}	3,80	0,63 ^{bc}	0,12
G31- Bustos	44,74	2,59	22,16	33,22	56,34 ^d	2,80	0,66 ^b	0,13
G32- Ventura	44,24	2,27	19,96	32,18	58,21 ^{bcd}	3,20	0,62 ^{bcd}	0,13
Elpida	44,65	2,22	25,93	33,89	52,64 ^a	6,20	0,76 ^e	0,17

Tabla 2.7: Valores medidos y desvíos estándar de los parámetros de color en frutos a los 0, 6 y 16 días almacenados en cámara de frío (10 °C), los cuales fueron cosechados a campo en la quinta de la familia Quispe. Los promedios con letras iguales no presentan diferencias significativas, según la prueba de Fisher ($p > 0,05$).

Al día 0 de cosecha se midieron los parámetros de los frutos, antes de ser ingresados a la cámara de frío. Los valores de *Hue* van de un rango rojo/amarillo de 58,16 (Elpida) a 66,77 (Molinario). El grupo con tonos más rojos pertenece a Elpida, G32 (Ventura), G31 (Bustos), G23 (Del Manso), G24 (Carcione), G18 (Luna), G20 (Grasso), G15 (Prieto) y G16 (Paolettich). Con tonos más amarillos se encuentra el grupo G15 (Prieto), G16 (Paolettich), G17 (Molinario), G18 (Luna), G23 (Del Manso) y G24 (Carcione). En el mismo sentido, la relación a^*/b^* del híbrido comercial Elpida tomó el valor más alto del rango observado (0,62) y Molinaro el valor más bajo (0,43). A los 16 días se observó que la mayoría de los materiales fueron aumentando el valor de la relación a^*/b^* (valores de 0,62-0,67), es decir que fueron tomando tonalidades al rojo, excepto el Uco Plata INTA que casi no cambió de valores. Dos materiales se diferenciaron del resto, Elpida con un valor significativo de 0,76, y G1 (Uco Plata INTA) con el valor más bajo (0,48), como se observa en la Figura 2.10.

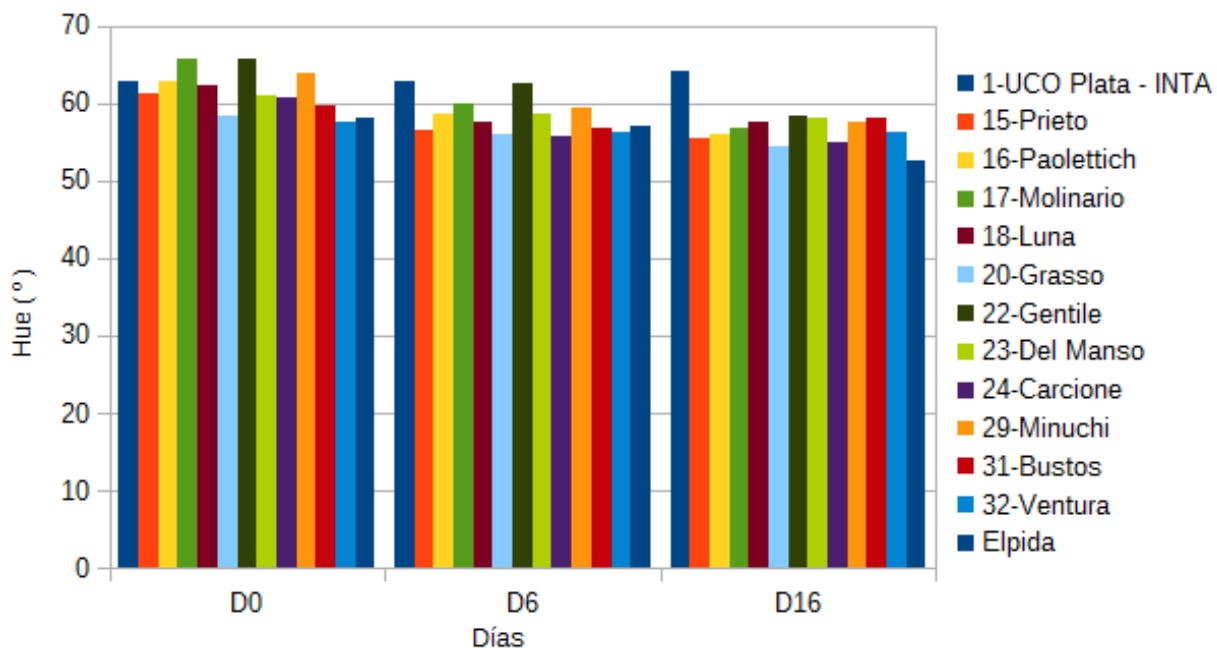


Figura 2.9: Gráfico de barras de los valores de *Hue* en función de los días 0, 6 y 16 en almacenamiento en cámara de frío de 10 °C.

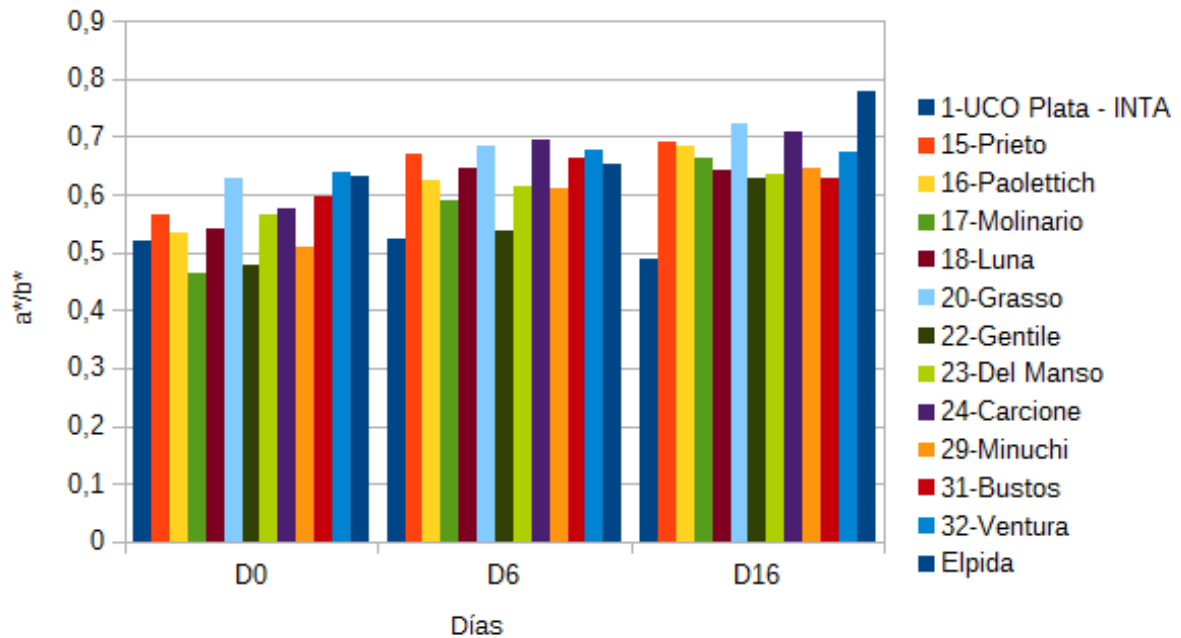


Figura 2.10: Gráfico de barras de la relación a^*/b^* en función de los días 0, 6 y 16 en almacenamiento en cámara de frío de 10 °C.

Índice de daño

Los frutos seleccionados luego de la cosecha, fueron tratados según lo descripto en Materiales y Métodos (2.3.2 Porcentaje de pérdida de peso de los frutos). Se determinó una escala relativa de valores de daño como: 1: incipiente, 2: moderado y 3: grave, y se calculó un Índice de Daño (*ID*) como el promedio de estos valores relativos. También se calculó el porcentaje de materiales con presencia de hongos. Los resultados de este ensayo se muestran a continuación:

Día 6 en cámara de 10 °C

Variedad	N	1	2	3	% dañados	% hongos	ID
G16-Paolettich	14	2	1	0	21,4	0	0,3
G22- Gentile	12	7	0	0	58,3	0	0,6
G18- Luna	9	1	1	0	22,2	11,1	0,1
G32- Ventura	10	1	0	0	10,0	0	0,1
G15- Prieto	12	1	2	0	25,0	8,3	0,4
G20- Graso	12	0	1	0	8,3	0	0,2
G31- Bustos	16	1	0	1	12,5	0	0,3
G1- Uco Plata	9	1	0	1	11,1	0	0,4
INTA	11	0	2	0	18,2	0	0,2
G29- Minuchi	10	0	0	0	0	0	0
G24- Carcione	11	0	0	0	0	0	0
G23- Del Manso	7	0	4	0	57,1	28,6	1,1
G17- Molinaro							

Día 16 en cámara de 10 °C.

Variedad	N	1	2	3	% dañados	% hongos	ID
G16-Paolettich	14	3	1	5	64,23	42,9	1,4
G22- Gentile	12	1	2	9	100,0	33,3	2,7
G18- Luna	9	0	3	5	88,9	66,7	2,3
G32- Ventura	10	3	1	5	90,0	60,0	2,0
G15- Prieto	12	2	4	5	91,7	50,0	2,1
G20- Graso	12	0	2	6	91,7	66,7	1,8
G31- Bustos	16	2	3	6	68,8	50,0	1,6
G1- Uco Plata INTA	9	3	0	5	88,9	44,4	2,0
G29- Minuchi	11	5	2	1	72,7	45,5	1,1
G24- Carcione	10	1	4	1	60,0	40,0	1,2
G23- Del Manso	11	4	2	0	54,5	27,3	0,7
G17- Molinaro	7	0	0	4	57,1	57,1	1,7

Tabla 2.8: Medidas de Índice de Daño (*ID*) y porcentajes de frutos dañados y con presencia de hongos de los frutos cosechados en la quinta de productores familiares, a los 6 y 16 días de almacenamiento en frío (10 °C).

A los 6 días de almacenamiento en cámara de frío ya se observa claramente que el genotipo G17 (Molinaro) casi alcanza un valor superior a 1 (incipiente) (Figura 2.11). El resto de los materiales están entre 0,1-0,6, excepto G24 y G23 (Carcione y Del Manso) que

no tienen daños. A los 16 días de almacenamiento, el G22 (Gentile) presenta un índice moderado llegando a grave (3). Con índice moderado (2) se encuentran los frutos de G18, G31, G32, G15, G20, G1 y G17 (Luna, Bustos, Ventura Prieto, Grasso, Uco Plata INTA y Molinaro). Por debajo del índice incipiente se encuentran G23 (Del Manso), G24, G29 y G16 (Carcione, Minuchi y Paolettich).

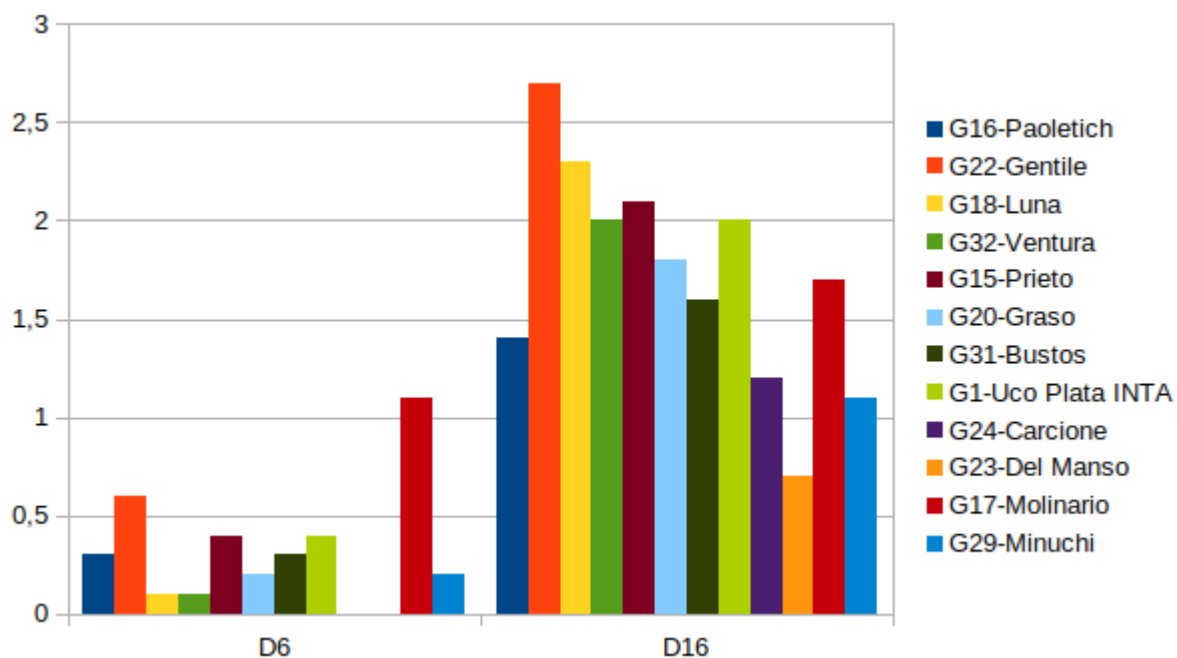


Figura 2.11: Gráfico de barras de los Índices de Daño (*ID*), a los días 6 y 16 (D6 y D16) durante su almacenamiento en cámara de frío (10 °C).

2.5. Conclusiones

En este capítulo, logramos tener un primer análisis de las características morfoagronómicas de las subpoblaciones de tomate ‘Platense’ y el comportamiento de poscosecha de sus frutos. Como fue descrito durante el desarrollo de este trabajo, para lograr obtener nuevos cultivares, es necesario contar con una base genética amplia y desarrollar de manera sistemática evaluaciones agronómicas de las colecciones de germoplasma, tanto de

los conservados en los bancos como de las variedades locales que se encuentran, hoy en día, en las unidades productivas.

Por medio de los dos ensayos se pudieron caracterizar las condiciones morfoagronómicas, tanto bajo cubierta como a campo. No se pudieron sacar conclusiones estadísticas robustas de comparación, pero sí se pudo realizar una primera caracterización de las diferentes subpoblaciones de Tomate 'Platense' y las ventajas en el manejo que poseen los productores a campo y sus ventajas en cuanto a rendimientos.

Cuando se clasificaron los frutos por peso se observó que la mayoría de las subpoblaciones de 'Platense' se caracterizaron por tener frutos con peso mayor a 150 g y casi sin descarte (menor a 90 g), salvo para el caso 'Prieto' y 'Ventura', donde hay una mayor variación de tamaños. A su vez en los materiales comerciales, se observó que su selección fue a partir del tamaño de fruto intermedio (entre 90-150 g), característica muchas veces elegida por los consumidores.

En relación a la calidad de poscosecha de los frutos almacenados en frío, se pudieron observar cuáles materiales poseen mejores aptitudes para mantener las características visuales, como lo demostraron las subpoblaciones de 'Del Manso', 'Carcione', 'Minuchi' y 'Paolettich'. En cuanto a la pérdida de contenido de agua, los materiales más aptos para almacenamiento fueron 'Prieto', 'Paolettich', 'Luna', 'Gentile', 'Carcione', 'Minuchi' y 'Bustos'. El híbrido comercial se destacó por seguir madurando los colores rojos durante su almacenamiento.

Este primer análisis de estos genotipos, sumado a los datos obtenidos por marcadores moleculares del siguiente capítulo, son el primer paso para establecer un programa de mejoramiento y revalorizar la fuente de diversidad fenotípica, genómica y cultural de estas variedades locales, muchas veces subestimadas por el actual modelo productivo.

Capítulo 3

Caracterización de la diversidad genética de las entradas de tomates 'Platense' del Banco de Germoplasma - INTA - La Consulta



3.1 Introducción⁹

Estudio de variedades hortícolas locales

Como ya se mencionó en el Capítulo 1, las variedades locales son poblaciones de plantas cultivadas, genéticamente heterogéneas, con una fuerte identidad cultural y agronómica que se encuentran en un estado de evolución constante, pero por fuera de programas de mejoramiento formal. Esta complejidad favorece la biodiversidad agrícola y es de primordial importancia para la producción local y la soberanía alimentaria (Villa y cols., 2005).

El Cordón Hortícola Platense (CHP), como el resto de los cinturones verdes productivos de la Argentina, es fuente de una gran agrobiodiversidad. El trabajo de localización, preservación y valorización de las hortalizas típicas desarrollado por el Curso de Extensión Rural de la Facultad de Ciencias Agrarias y Forestales (UNLP) junto a otros actores locales (Grupo de Productores de Tomate Platense y la Escuela Agropecuaria N°1 Alejandro Korn), permitió identificar 35 poblaciones de hortalizas típicas locales. Fue a partir de este trabajo de rescate y valorización que tomamos como caso de estudio la variedad de tomate ‘Platense’, que se estima que desde los años ’30 del siglo XX se encuentra adaptada al clima y al suelo de las quintas (unidades productivas) del CHP.

Varios bancos de germoplasma dentro y fuera de la Argentina contienen accesiones de variedades locales argentinas (tanto de tipo ‘Platense’ como otras) y cultivares, provenientes de programas de mejoramiento de instituciones públicas como el INTA y Universidades Nacionales. Dentro de país existe una Red de Bancos de Germoplasma (RBG) del INTA, que cuenta con más de 30.000 entradas, nueve bancos activos de recursos fitogenéticos, 12 colecciones distribuidas en diversas áreas ecológicas y un Banco Base que guarda un duplicado de todo. En el banco activo de INTA La Consulta es donde se encuentran las variedades hortícolas disponibles del país. La *Northeast Regional PI Station* (Geneva,

⁹ Los resultados de este capítulo fueron publicados en: Amado Cattáneo, R., McCarthy, A.N. y Feingold, S.E. Evidence of genetic diversity within *Solanum Lycopersicum* L. ‘Platense’ landrace and identification of various subpopulations. *Genet Resour Crop Evol* (2020). Ver en: <https://doi.org/10.1007/s10722-020-00959-8>

New York, Estados Unidos) representa un ejemplo fuera del país. Allí se encuentran más de sesenta entradas de tomate recolectadas entre los años 1932 a 1974 a partir de mercados locales argentinos (desde la provincia de Jujuy hasta la localidad de Los Talas en la provincia de Buenos Aires). Estas entradas fueron utilizadas en estudios de fenotipado y diversidad genética (Hu, Wang, Chen y Yang, 2012a; Facultad de Agronomía y Biotecnología, Universidad de Agricultura, China). El banco *Tomato Genetics Research Center* (TGRC) representa otro ejemplo fuera del país. Allí se encuentra una entrada de tomate ‘Platense’ con código de acceso LA3243 (TGRC, s/f), que fue secuenciada por un consorcio internacional (Lin y cols., 2014). Sin embargo, hasta la fecha no se ha realizado un estudio sistemático referido a las entradas de los bancos que contienen las diferentes poblaciones locales que les dieron origen y que son de interés para nuestra región.

Como ya se mencionó en el Capítulo 1, según el registro de los testimonios de los horticultores, durante el período de mayor expansión comercial del ‘Platense’ dentro del CHP hubo hasta 15 subpoblaciones diferentes del tipo ‘Platense’, de las cuales sólo 9 (‘Carcione’, ‘Bustos’, ‘Del Manso’, ‘Gentile’, ‘Grasso’, ‘Prieto’, ‘Paolettich’, ‘Molinario’ y ‘Luna’) se conservan actualmente en el Banco de Germoplasma de INTA - La Consulta (Mendoza, Argentina). La caracterización del germoplasma de tomate, tanto para su identificación o registro como cultivar, siempre es evaluada por medio de caracteres morfológicos, fenotípicos, y productivos, entre otros (Arens, Bredemeijer, Smulders y Vosman, 1995), y así se clasificaron las subpoblaciones de ‘Platense’. Sin embargo, la influencia del ambiente y el tiempo necesario para medir esos datos son las principales limitantes de esta caracterización (Becerra y Paredes, 2000).

Los marcadores moleculares tienen varias ventajas sobre los marcadores morfológicos, tanto para la identificación de variedades, como para el estudio de la diversidad genética. Entre otros, presentan independencia de factores ambientales, alto nivel de polimorfismo, mayor sensibilidad para detectar diferencias genotípicas entre individuos, y mayor reproducibilidad (Spooner, 2005). El uso de estas técnicas en las políticas de mantenimiento y optimización de los recursos dentro de los bancos de germoplasma es de gran impacto para el diseño de las colecciones base, ya que puede identificar las diferentes entradas almacenadas, confirmar su identidad taxonómica, evaluar su diversidad genética, etc. (Spooner, 2005).

A continuación, se describen los diferentes tipos de marcadores moleculares existentes y, en particular, los microsatélites que fueron los utilizados en esta tesis.

3.1.1. Tipos de marcadores moleculares

Los marcadores moleculares pueden ser clasificados en tres grupos (Kalia y cols., 2011):

Marcadores basados en la hibridación del ADN

Los más utilizados en planta son los polimorfismos en la longitud de fragmentos de restricción de ADN o RFLPs (*Restriction Fragment Length Polymorphisms*). El segmento genómico en estudio, amplificado o no por medio de PCR, es tratado con enzimas de restricción específicas que producen fragmentos de diferentes longitudes. Estos son separados mediante electroforesis y visualizados por medio de hibridación con sondas específicas.

Marcadores basados en reacción de amplificación por PCR

Los primeros en desarrollarse fueron los fragmentos polimórficos de ADN amplificados al azar o RAPDs (*Random Amplification of Polymorphic DNAs*). Esta técnica se basa en utilizar un único cebador de 10 pb que hibrida aleatoriamente en el ADN en estudio. Es necesario que la hibridación se dé lo suficientemente cerca y en orientación opuesta para que haya amplificación. Otra técnica es la que genera los polimorfismos en la longitud de fragmentos amplificados o AFLPs (*Amplified Fragment Length Polymorphisms*), la cual se basa en una primera digestión del ADN genómico mediante enzimas de restricción y la subsecuente amplificación de algunos de esos fragmentos mediante PCR. Por último, los microsatélites, como los SSR o STR (*Simple Sequence Repeat* y *Short Tandem Repeats*), están conformados por un motivo repetitivo en el cual se encuentra contenida la secuencia

repetida y dos regiones flanqueantes que incluyen la secuencia de los cebadores, los cuales se ubican a ambos lados del motivo repetitivo. Una variante son las inter-secuencias simples repetitivas o ISSRs (*Inter Simple Sequence Repeats*), las cuales utilizan como cebador las secuencias degeneradas de los motivos repetidos antes descritos.

Los SSR han ganado considerable importancia en los estudios genéticos en plantas (por ejemplo, en estudios de estructura poblacional, flujo génico, procesos evolutivos, etc.) y en programas de mejoramiento (mapas de ligamiento, QTLs, selección por marcadores, etc.), ya que poseen ciertas ventajas en comparación con los otros marcadores nombrados. Los microsatélites son altamente informativos, poseen herencia codominante, su hipervariabilidad es reproducible experimentalmente y transferible a otras especies relacionadas. Poseen abundancia relativa, cobertura sobre todo el genoma (incluidos los genomas de las organelas), ubicación específica, susceptibilidad a la automatización y genotipificación de alto rendimiento (Parida y cols., 2009).

En el Cuadro 3.1 se resumen las principales características y ventajas de los diferentes tipos de marcadores moleculares:

Característica	Marcadores Moleculares		
	SSR	RFLP	RAPD/AFLP
Secuencia blanco	Requerida	No requerida	No requerida
Polimorfismo	Alto	Bajo	Moderado
Dominancia	Codominante	Codominante	Dominante
Utilidad en MAS*	Alto	Moderado	Moderado
Costos y tiempo	Alto	Alto	Moderado

Cuadro 3.1: Comparación de las principales características entre diferentes tipos de marcadores moleculares. Modificado de Kalia y cols., 2011. (*) MAS, *Marker Assisted Selection*.

Marcadores basados en secuencias

Los SNPs, *Single Nucleotide Polymorphisms*, son el tipo de marcador que permite distinguir mutaciones a nivel de un solo nucleótido e *indels* (inserción o delección). Estos son la base de la mayoría de las diferencias entre alelos. Esta tecnología sigue en evolución para simplificar su uso y reducir los costos.

Por todas las características de los SSR antes nombradas para el estudio genético en plantas, es que decidimos realizar nuestra investigación con este tipo de marcador molecular.

3.1.2. Microsatélites

El término de ADN satélite fue originado en la década de 1960 a partir de la observación de que una fracción de ADN poseía una densidad distintiva cuando la muestra era sometida a una centrifugación de gradiente de densidad. Luego, esta fracción de ADN fue reconocida como la región repetitiva en tándem de los centrómeros. Los microsatélites (Litt y Luty, 1989), también conocidos como *Short Tandem Repeats* (STRs) (Edwards, Civitello, Hammond y Caskey, 1991), *Simple Sequence Repeats* (SSRs) (Jacob y cols., 1991), son motivos repetidos en tándem de 1-6 pb que se encuentran tanto en procariotas como en eucariotas (Zane, Bargelloni y Patarnello, 2002).

En plantas, estas repeticiones son generalmente más ricas en secuencias en adenina y timina (AT) mientras que, en animales, las más comunes son en repeticiones en adenina y citosina (AC). Parecen ser una característica general distintiva entre los genomas vegetales y animales (Powell, Machray y Provan, 1996).

Los microsatélites están distribuidos por todo el genoma, no necesariamente al azar, especialmente en eucromatina en eucariotes, en zonas codificantes y no codificantes, en ADN nuclear y en organelas (Pérez-Jiménez y cols., 2013; Phumichai y cols., 2015). Históricamente, las secuencias repetidas en tándem eran consideradas ADN no funcional debido a su alta variabilidad y, más allá de algunas enfermedades neurodegenerativas en humanos, se creía eran neutrales y sin consecuencias fenotípicas (Gemayel y cols., 2012). Esto cambió cuando se detectaron microsatélites en regiones transcriptómicas y reguladoras

del genoma. Diferentes estudios han demostrado que los satélites desempeñan un rol en procesos como la regulación de la transcripción y traducción, la estructura del ARN, organización de la cromatina, tamaño del genoma y el ciclo celular (Nevo, 2001; Li y cols., 2002-2004; Haas y cols., 2013; Gao y cols., 2013).

En plantas, a pesar del hecho de que existe una alta densidad de SSR, estos han sido detectados en regiones 5'-UTR (Fujimori y cols., 2003; Tranbarger y cols., 2012; Zhao y cols., 2014) y hay pocos estudios que verifiquen su efecto sobre la regulación de la expresión génica. Además, la codificación de tri- y hexanucleótidos parece estar controlada por una mayor presión de mutación en regiones codificantes que en otras regiones génicas. Por lo tanto, en las plantas hay menos variabilidad de alelos de SSR en zonas exónicas que en las intrónicas. La distribución sesgada de microsatélites y motivos de microsatélites también sugieren que los microsatélites de diferentes tipos desempeñan distintos roles en diferentes regiones génicas, como el caso de promotores, intrones y exones en plantas (Li y cols., 2004; Gemayel y cols., 2012; Gao y cols., 2013).

Los microsatélites se pueden clasificar con diferentes criterios: tamaño, tipo de unidad de repetición o su localización genómica (Cuadro 3.2) (Kalia y cols., 2011). Según el número de nucleótidos por unidad de repetición se pueden definir como mono-, di-, tri-, penta- o hexa nucleótidos y dependiendo del arreglo en el motivo de las repeticiones, Weber (1990) utiliza el término de perfecto, imperfecto y compuesto, mientras que Wang y cols. (2009a) los clasifica como perfecto simple, imperfecto simple e imperfecto compuesto. Las repeticiones perfectas son arreglos en tándem de un motivo simple repetido, mientras que las repeticiones imperfectas son interrumpidas por motivos no repetidos en alguna zona. Los compuestos son dos motivos repetitivos presentes juntos en varias configuraciones en un mismo motivo. Jarne y Lagoda (1996) acuñaron los términos puros e interrumpidos para perfecto e imperfecto. La mayoría de los marcadores SSRs son nucleares (nuSSRs), aunque los microsatélites también pueden estar distribuidos en mitocondrias (mtSSRs) y cloroplastos (cpSSR). Weising y Gardner (1999) descubrieron cpSSRs en *Nicotiana*, mientras que mtSSRs fueron descritos por Soranzo y cols. (1999) en *Pinus*. En años recientes los cpSSRs han sido ampliamente utilizados para estudiar variaciones genéticas en plantas y en flujo génico en poblaciones naturales (Provan y cols. 2001) debido a su herencia uniparental. Por

otro lado, comparar patrones de variabilidad con marcadores biparentales (nuSSR) y uniparentales (de organelas, es decir, cpSSR y mtSSR), puede proporcionar información complementaria para la población y la biología evolutiva (Powell y cols. 1996).

Basado en el número de nucleótidos
Mononucleótido (A) _n
Dinucleótido (AT) _n
Trinucleótido (CGT) _n
Tetranucleótido (CAGA) _n
Pentanucleótido (AAATT) _n
Hexanucleótido (CTTTAA) _n
Basado en el arreglo en el motivo de repetición
Puro o perfecto o perfecto simple (AT) _n
Imperfecto simple (AAC) _n ACT (AAC) _{n+1}
Compuesto o compuesto simple (CA) _n (GA) _n
Interrumpido o imperfecto o compuesto imperfecto (CCA) _n TT (CGA) _{n+1}
Basado en su ubicación genómica
Nuclear (nuSSRs)
Cloroplasto (cpSSRs)
Mitocondrial (mtSSRs)

Cuadro 3.2: Clasificación de los tipos de microsatélites. *n* = número de repeticiones.

3.1.3. Modelos evolutivos de microsatélites

La génesis de microsatélites como su evolución, en el sentido de aumento o disminución de secuencias repetidas, es un proceso dinámico y ha demostrado ser extremadamente complejo (Ellegren, 2004; Pearson, Edamura y Cleary, 2005). Posibles explicaciones propuestas para su génesis incluyen el deslizamiento de ADN monocatenario, errores en la recombinación de ADN, falta de coincidencia en la reparación por la rotura de doble cadena y retrotransposición. Durante la replicación del ADN, el deslizamiento de la

ADN polimerasa III sobre el ADN cadena molde en la región de repetición puede generar, en la nueva cadena, una expansión o reducción. Estos errores son corregidos por el sistema de reparación de desapareamiento (MMR, *mismatch repair*), por lo tanto, la variabilidad de los SSR depende de un equilibrio entre el deslizamiento de la polimerasa y la efectividad de reparación del mecanismo MMR (Figura 3.1) (Wang, Barkley y Jenkins, 2009).

Por otro lado, los daños causados en el ADN por factores de estrés externo como la radiación UV, γ -irradiación y estrés oxidativo, entre otros, pueden inducir los errores por deslizamiento y aumentar las tasas de mutación en los SSR (Jackson, Chen y Loeb, 1998; Chang y cols., 2002; Li, Korol, Fahima, Beiles y Nevo, 2002; Trifonov, 2003). El deslizamiento en la replicación favorece el crecimiento, mientras que la mutación puntual puede romper una repetición larga en dos o más repeticiones cortas. Por ello los cambios en las frecuencias de deslizamiento y en las mutaciones puntuales podrían tener un efecto directo en la distribución de SSR en un genoma. La interacción del deslizamiento y la recombinación también podrían afectar la estabilidad de los SSR (Figura 3.1) (Li y cols., 2002).

Por último, otra fuente de origen son los retrotransposones, fragmentos de ADN repetitivos con capacidad de insertarse en los cromosomas después de ser transcritos de forma inversa a partir de una molécula de ARN. Durante el estudio del genoma del arroz y del genoma humano se demostró que la generación de microsatélites era acompañada por eventos de retrotransposición (Nadir, Margalit, Gallily, y Ben-Sasson, 1996; Temnykh y cols., 2001). La inserción y acumulación de los retrotransposones entre los genes posiblemente haya tenido un rol importante en la expansión del genoma en las plantas (Bennetzen, 2000).

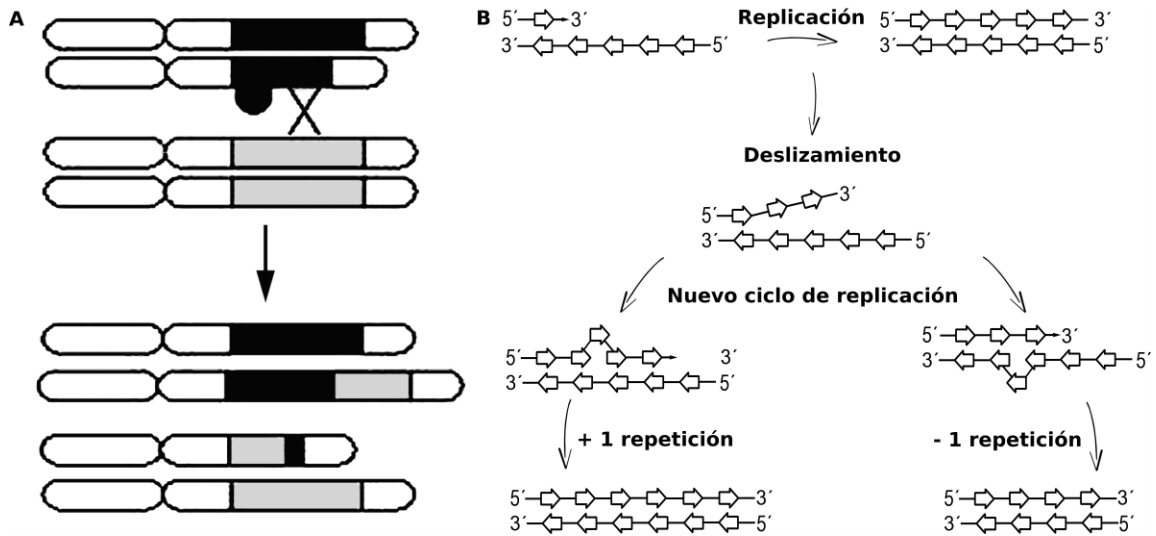


Figura 3.1: A- Cruce desigual entre cromosomas homólogos. Las regiones negras y grises corresponden a secuencias de repetición de microsatélites. B- Deslizamiento durante la replicación del ADN. Supongamos que la molécula de ADN original tenía 5 repeticiones del motivo, simbolizadas por una flecha. El deslizamiento conduce a la formación de nuevos alelos con 6 y 4 repeticiones, dependiendo en cuál de las hebras contenga el error de la polimerasa. Figura modificada de Olivera y cols. (2006).

Es importante tener en cuenta qué modelo teórico se aplicará con los datos obtenidos a partir de microsatélites, para determinar correctamente los parámetros de estructura poblacional o la distancia genética entre individuos. Estos modelos mutacionales son utilizados para estimar el número esperado de alelos en una población a partir de la heterocigosis observada y su análisis estadístico. En general se utilizan cuatro modelos (Olivera y cols., 2006):

Modelo de alelos infinitos

En este modelo, cada mutación crea al azar un nuevo alelo. Aplicando este modelo a loci del microsatélite, las mutaciones alteran el número de repeticiones. Con este modelo un alelo con 10 repeticiones no se diferencia genéticamente de un alelo con 15 repeticiones como con uno de 16, es decir, la proximidad en términos del número de repeticiones no indica

una mayor relación filogenética. Este es el modelo clásico de Wright (1931) en el que usa F -statistics donde la probabilidad de que un individuo sea homocigoto:

$$F = \frac{1}{4Neu + 1}$$

donde u es la tasa de mutación y Ne es el tamaño efectivo de población.

Modelo de mutación por pasos

Cuando un locus de microsatélite muta, ganando o perdiendo una repetición, implica que dos alelos que se diferencian por un solo motivo están más relacionados (es decir, comparten un ancestro común) que los alelos que se diferencian en varias repeticiones. Slatkin (1995) propuso una medida de diferenciación genética (R_{ST}) basada en este modelo, similar a la F_{ST} de Wright (1951), y se define:

$$R_{st} = \frac{S - S_w}{S}$$

donde S_w es la suma promedio de cuadrados de las diferencias en el tamaño de los alelos dentro de cada población y S las probabilidades de elegir dos copias diferentes del locus del mismo población y dos copias de diferentes poblaciones.

Este modelo suele ser el preferido cuando se estiman las relaciones entre los individuos y la estructura de la población, excepto en presencia de homoplasia (es decir, cuando dos alelos son idénticos en estado (mismos tamaños de pares de bases) pero no comparten descendencia común). La homoplasia puede influir seriamente en los estudios de población que involucran altas tasas de mutación y grandes tamaños de población junto con fuertes limitaciones de tamaño de alelos (Estoup y cols., 2002).

Modelo de dos fases

Di Rienzo y col. (1994) introdujeron este modelo como ampliación del modelo de mutación por pasos para estudios sobre microsatélites. Propone que la mayoría de los eventos

mutacionales resultan en un aumento o disminución de una unidad de repetición. Aunque también se pueden producir alteraciones en más de una unidad de repeticiones, estas son consideradas poco frecuentes. Estos autores proponen que a mayor número de repeticiones es mayor la frecuencia de mutación. Por el contrario, Murray (1996) propone que, para este modelo, el número de mutaciones disminuye al aumentar la complejidad de las repeticiones.

Modelo de k -alelos

Este modelo propuesto por Crow y Kimura (1970) asume que, si hay exactamente k posibles alelos en un locus dado, la probabilidad de que un alelo dado pueda mutar en cualquier otro es $\mu / k - 1$, donde μ es la tasa de mutación.

3.1.4. Marcadores moleculares en tomate

Los microsatélites han mostrado ser de gran utilidad para el estudio de cultivares de tomate, por ser altamente informativos (polimórficos) (Smulders, Bredemeijer, Rus-Kortekaas, Arens, y Vosman, 1997; He, Poysa y Yu, 2003;) en comparación con otras técnicas moleculares como los RFLPs (Miller y Tanksley, 1990), sondas y RAPD (Rus-Kortekaas y col., 1994), o su combinación (Van der Beek, Zabel y Lindhout, 1992). Los SNPs e *indels* son marcadores cuyo uso ha ido aumentando, debido a su alto rendimiento en la selección asistida por marcadores, identificación de marcadores de secuencia expresada o EST (*Expressed Sequence Tags*), y para generar cartografía e integración de mapas genéticos y físicos. Mediante SNPs, Ranc y cols. (2008) y Blanca y cols. (2013) pudieron hacer análisis taxonómicos y evolutivos sobre el origen del tomate cultivado y sus especies hermanas. Los trabajos de Corrado y cols. (2013 y 2014) lograron caracterizar la estructura de la población e identificar loci putativos bajo selección positiva dentro de las colecciones de tomates cultivados y *landraces*, lo que proporciona un marco útil no sólo para la caracterización y conservación eficiente de las variedades locales, sino también para su posible uso en genética y mejoramiento. Por último, también se han utilizado para construir mapas de ligamiento y genotipificación (Shirasawa y cols., 2010; Sim 2012).

Más allá de la falta de variabilidad del tomate cultivado como consecuencia de su naturaleza autopolinizante en combinación con la estrecha base genética utilizada para lograr los cultivares modernos (Rick 1976), los marcadores microsatélites han sido utilizados exitosamente para estudiar la diversidad genética de diferentes cultivares (Smulders y cols., 1997; Bredemeijer y cols., 2002; He y cols., 2003; Frary y cols., 2005; Garcia-Martinez y cols., 2006; Song y cols., 2006; Ispizúa y cols., 2007; y Monte y cols., 2018), para el mejoramiento asistido por marcadores (MAS) y QTLs (Frary y cols., 2000). Esta técnica también tiene el potencial de identificar a las variedades locales y distinguirlas de las variedades comerciales, pudiendo facilitar las gestiones de la Denominación de Origen para su revalorización cultural y comercial (Caramante y cols., 2009; Corrado y cols., 2014; Mazzucato y cols., 2010; Castellana y cols., 2020). En este mismo sentido, posee la capacidad de rastrear cultivares de tomate en cadenas alimenticias para control del producto (Caramante y cols., 2011).

En este capítulo, nos proponemos evaluar la identidad genética y la diversidad de la variedad local tomate 'Platense' y de todas las subpoblaciones de las entradas disponibles en el banco activo de germoplasma hortícola del INTA - La Consulta (Mendoza, Argentina).

3.2. Materiales y métodos

De toda la colección de germoplasma de tomate *S. lycopersicum* L. evaluada durante el desarrollo de esta tesis (Anexo I), estudiamos las 9 entradas provenientes del Banco de Germoplasma del INTA La Consulta (Mendoza, Argentina), un híbrido comercial, tres cultivares comerciales del INTA La consulta y, como *outgroup*, el tomate silvestre, *S. pimpinellifolium* L. La información general de estas entradas se detalla en el Cuadro 3.3.

Entrada	Nombre Común	Código Banco/Origen
Uco Plata INTA		Comercial INTA La Consulta
Uco 14 INTA		Comercial INTA La Consulta
Uco 18 M INTA		Comercial INTA La Consulta
Platense	‘Del Manso’	3820-INTA La Consulta
Platense	‘Paolettich’	3834-INTA La Consulta
Platense	‘Grasso’	3825-INTA La Consulta
Platense	‘Luna’	3832-INTA La Consulta
Platense	‘Carcione’	3812-INTA La Consulta
Platense	‘Bustos’	3811-INTA La Consulta
Platense	‘Prieto’	3837-INTA La Consulta
Platense	‘Gentile’	3824-INTA La Consulta
Platense	‘Molinaro’	3833-INTA La Consulta
Elpida		Comercial
<i>S. pimpinellifolium</i>		LA1589*

Cuadro 3.3: Información general de las 14 entradas estudiadas. (*) Entrada originalmente proveniente del C. M. Rick *Tomato Genetics Resource Center, University of California, Davis, USA.*

3.2.1. Extracción y purificación de ADN total

Durante la primavera (octubre del 2016) las semillas fueron sembradas en bandejas de germinación con un sustrato mezcla para germinación comercial. Las plantas crecieron en un invernadero de la Facultad de Ciencias Agrarias y Forestales (UNLP). Por cada entrada se sembraron 25 semillas, a los 30 días post siembra se tomaron muestras de hojas (100 mg) de 4 plantas individuales (a, b, c y d, las cuales fueron rotuladas), y las 21 restantes se trataron como una única muestra (200 mg) denominada *pool* (Figura 3.2). De esta manera, por cada entrada se extrajo ADN de 5 muestras.

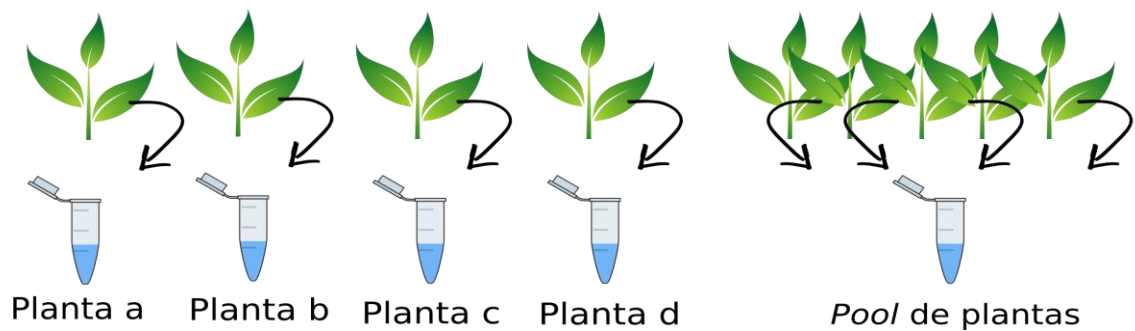


Figura 3.2: Esquema del diseño de muestras de ADN por cada entrada

La extracción de ADN fue realizada mediante el método modificado de Doyle y Doyle (1987), el cual consta de los siguientes pasos: se cortaron entre 100-300 mg de hojas de cada muestra en tubos de 1.5 ml; se almacenaron durante 24 hs en *freezer* de $-80\text{ }^{\circ}\text{C}$, y luego fueron liofilizadas (Heto FD4, Denmark). Luego de moler manualmente con una micromaja en microtubo, hasta obtener un polvo fino, se agregaron $700\text{ }\mu\text{l}$ de una solución 2% de CTAB en 1.4 M de NaCl, 20 mM de EDTA y 100 mM de Tris-HCl, agregando luego $2\text{ }\mu\text{l}$ de β -mercaptoetanol a cada tubo. Se mezcló por inversión e incubó por 45 minutos en baño térmico a $65\text{ }^{\circ}\text{C}$, mezclando cada 15 minutos por inversión. Los tubos se dejaron enfriar a temperatura ambiente; se agregaron $700\text{ }\mu\text{l}$ de cloroformo:alcohol isoamílico (C:IAA) en proporción de 24:1, y se homogeneizó por inversión repetida. Los tubos se centrifugaron por 10 minutos a 13200 rpm. La fase superior se extrajo a un nuevo tubo y se agregaron $80\text{ }\mu\text{l}$ de una solución de limpieza de ADN compuesta por CTAB 10% en NaCl 0.7 M. A continuación, se agregaron nuevamente $700\text{ }\mu\text{l}$ de C:IAA 24:1, se homogeneizó por inversión repetida y se centrifugó durante 10 minutos a 13200 rpm. El sobrenadante se transfirió a un tubo nuevo. La concentración de los ácidos nucleicos extraídos se realizó por precipitación agregando a cada muestra $500\text{ }\mu\text{l}$ de isopropanol a $-20\text{ }^{\circ}\text{C}$ y dejando reposar durante 15 minutos a $-20\text{ }^{\circ}\text{C}$. Pasado este tiempo, las muestras se centrifugaron a 13200 rpm durante 20 minutos, se descartó el sobrenadante y se dejó secar el precipitado durante 2 minutos a temperatura ambiente. El primer lavado del precipitado de ADN se realizó agregando 1 ml de etanol 70%, mezclando por inversión durante 3 minutos, centrifugando las muestras

durante 20 minutos a 13200 rpm, y luego descartando el sobrenadante. El segundo lavado del precipitado se realizó de la misma manera, pero utilizando etanol 96%. Los tubos con el precipitado limpio se dejaron abiertos a temperatura ambiente hasta la evaporación completa del etanol. La resuspensión del precipitado se realizó con 75 µl de agua ultra pura. La determinación de la calidad del ADN total extraído de todas las muestras fue realizada mediante electroforesis en gel de agarosa al 1% en buffer TBE (89 mM Tris-Borato, 20 mM ácido etilendiaminotetraacético), utilizando 1 µl de muestra directa en 4 µl de solución de carga como indicador de frente de corrida. Las condiciones de electroforesis fueron de 60 V constante durante 40 minutos. El revelado se hizo por medio de luz U.V. y colorante GelRed (GenBiotech, Argentina). Para cuantificar ADN y pureza, se midió espectrofotométricamente (SmartSpec™3000 Spectrophotometer, BIORAD, USA) absorbancia a 260 nm y se calculó la relación de absorbancia 260 nm/280 nm. Para todas las muestras se observó una pureza óptima con valores de 1.6-2.0. A partir de estas medidas se realizaron diluciones para alcanzar una concentración de 10 ng/µl.

3.2.2. Selección de los cebadores

Los cebadores (*primers*) fueron seleccionados a partir de la literatura de acuerdo a su contenido de información polimórfica y de manera de obtener una cobertura uniforme y completa del genoma del tomate. Así, los locus seleccionados fueron los siguientes: LECHI3, LELEUZIP , LE21085, SSR627 (LEACS4A), Tom152-153 (LEMDD-Na), LELE25, Tom144-145 (LE20592), LESSRPSPGb, LEEF1Aa (Smulders y cols., 1997); SSR96, SSR327, SSR383 (Frary y cols., 2005); Tom47-48, Tom59-60, Tom236-237 (Suliman-Pollatschek, Kashkush, Shats, Hillel, y Lavi, 2002); TMS63, TMS58, TMS60, TMS42, TMS52, TMS59 (Mazzucato y cols., 2008; Areshchenkova y Ganal, 2002); TMS42 (Areshchenkova, 2000); SSR5, SSR111, SSR43, SSR638 (TMS22 - LEac002) (He y cols., 2003). SR192, SSR45, SSR69, SSR295, SSR162, SSR63, SSR318 (Francis, s/f) y SSR20 (Hu, Wang, Chen, y Yang, 2012b). Detalles de los cebadores seleccionados se describen en el Anexo II.

3.2.3. Reacciones de amplificación y electroforesis en gel de agarosa

Las reacciones de amplificación fueron realizadas en un volumen total de 20 μ l, conteniendo 25 ng de ADN genómico como molde, 50 pmol de cada uno de los cebadores, 200 μ l dNTPs, 2 mM MgCl₂, 1 X de *buffer* Taq polimerasa y 1 U Taq ADN polimerasa (Promega Madison, WI, USA). La amplificación se llevó a cabo en placas de 96 pocillos por medio del termociclador MJ (PTC-100 MJ Research, Watertown, Mass., USA) con un ciclo inicial de desnaturalización de 2 minutos a 94 °C, 35 ciclos con un segundo paso de 15 segundos a 92 °C y un tercero de 15 segundos a la temperatura de unión de los cebadores Anexo II y 30 segundos a 72 °C para extender el fragmento. Se finalizó con 5 minutos de extensión a 72 °C y conservando el producto de amplificación a 16 °C. Todos los productos de las reacciones de amplificación y sus pruebas negativas fueron revisados por medio de una electroforesis en gel de agarosa al 1% en *buffer* TBE. Las condiciones de electroforesis fueron de 80 V constante durante 25 minutos. El revelado se realizó por medio de luz U.V. y colorante *GelRed* (GenBiotech, Argentina).

3.2.4. Protocolo de electroforesis en geles de poliacrilamida

Los productos de la PCR (reacción de la cadena de la polimerasa *-polymerase chain reaction-*) fueron resueltos en electroforesis desnaturalizantes en geles de poliacrilamida al 6% y tinción con plata, como detalla Benbouza (2006). El procedimiento descrito es el siguiente: limpiar los vidrios con abundante agua y detergente utilizando un cepillo; enjuagarlos con agua ultra pura. Secar con alcohol 96% y papel. Una vez limpios y secos, distribuir *RAIN X* por toda la superficie de una cara, del vidrio más chico (cara hidrofóbica). Preparar una solución de *bind silane* (1 ml etanol 95%: ácido acético 5% + 2 μ l de *bind silane*) para la superficie hidrofílica del otro vidrio. Prepara el gel con poliacrilamida fría al 6% (p/v) y agregar al mismo tiempo 360 μ l de APS 10% (persulfato de amonio 1 g en 10 ml de agua ultrapura) + 60 μ l de TEMED (tetrametiletildiamina). Volcar sobre el vidrio grande y deslizar el vidrio chico cuidando de no dejar burbujas entre los vidrios. Se fijan los

vidrios con trabas y se coloca un peine que marque las calles de corrida, durante una hora y media hasta que polimerice el gel. Las muestras se preparan agregando *buffer* de carga (0.25% azul de bromofenol + 30% glicerol +3% formamida + agua). Se usan 12 μ l del *buffer* de corrida, por cada 20 μ l de reacción de PCR. Una vez montados los vidrios con el gel sobre la cuba electroforética, agregar TBE 1 X. Antes de sembrar la muestra, el ADN se desnaturaliza durante 5 minutos a 96 °C y mantiene en frío hasta el momento de la siembra. Sembrar 3,5 - 4 μ l en el gel. Realizar la corrida electroforética a 4000 V, 200 mA y 80 W, hasta que el primer frente de corrida del *buffer* alcance el borde del vidrio. La tinción del gel para su revelado se realiza fijándolo con 20 ml de etanol + 1 ml de ácido acético + 179 ml de agua ultrapura. La solución debe enfriarse a 10 °C. Verter sobre el gel durante 5 minutos, con agitación suave. Luego impregnar el gel con 0,3 g de nitrato de plata + 0,375 ml de formaldehído + 200 ml de agua ultrapura (proteger de la luz). La reacción tarda de 6 a 7 minutos. Realizar un lavado rápido con agua ultrapura y proceder con el revelado por medio de 3 g de hidróxido de sodio en 200 ml agua ultrapura + 0,4 ml de formaldehído durante 3 a 5 minutos. La reacción se frena con la misma solución utilizada como fijador durante 2 minutos. Lavar con abundante agua ultrapura y secar a temperatura ambiente.

El tamaño de banda fue determinado por comparación con un patrón de peso molecular de 25 pb (Promega, USA) y utilizando el programa *Gel Pro Analyser V3.1* (Media Cybernetics, USA).

3.2.5. Análisis de la diversidad genética

Todas las accesiones estudiadas fueron calificadas con respecto a la presencia y ausencia de las bandas correspondientes entre los genotipos. Cada alelo de microsatélite se consideró un marcador dominante, que se codificó como '1' o '0' para indicar la presencia o ausencia de bandas para cada alelo del marcador SSR, respectivamente.

Medidas de distancia genética

Un marcador puede ser considerado altamente informativo para estudios de vínculos genéticos si se toma al azar un individuo y éste resulta heterocigoto para ese marcador. Los marcadores con muchos alelos o polimórficos, tienden a ser altamente informativos. El cálculo estadístico llamado contenido de información de polimorfismo o *PIC* (*Polymorphism Information Content*) permite cuantificar cuán informativos son.

El *PIC* se calcula de acuerdo a la siguiente fórmula:

$$PIC = 1 - \sum_{i=1}^n p_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i^2 p_j^2 = 1 - \sum_{i=1}^n p_i^2 - \left(\sum_{i=1}^n p_i^2 \right)^2 + \sum_{i=1}^n p_i^4$$

donde p_i es la frecuencia del marcador alélico (a_i) y n es el número de alelos diferentes (Hildebrand y cols., 1994). Supone la existencia de equilibrio Hardy-Weinberg y que el par de alelos ocurre con una frecuencia igual al producto de las dos frecuencias, con lo que podemos determinar la frecuencia de cada pareja a partir de la frecuencia p_i de cada alelo (a_i).

Teóricamente los valores *PIC* pueden variar de 0 a 1. Con un *PIC* de 0, el marcador sólo tiene un alelo, mientras que un *PIC* de 1 indica que el marcador tendría un número infinito de alelos. Un valor *PIC* mayor que 0,7 se considera altamente informativo y un valor de 0,44 se considera moderadamente informativo. Un gen o marcador con sólo dos alelos tiene un *PIC* máximo de 0,375. Los marcadores con mayor número de alelos tienden a tener mayores valores de *PIC* y por lo tanto son más informativos.

Una medida alternativa del grado de polimorfismo de un marcador es su heterocigosidad. Esta medida se puede calcular como la probabilidad de que un individuo elegido al azar sea heterocigoto para cualesquiera de sus dos alelos en un locus del marcador que tenga frecuencias alélicas p_i . Analíticamente, *heterocigosidad* = $1 - \sum_{i=1}^n p_i^2$, donde $\sum_{i=1}^n p_i^2$ es la *homocigosidad*. Por lo tanto, el valor de *PIC* siempre será inferior a la heterocigosidad (Hildebrand y cols., 1994).

Heterocigosidad esperada y observada

Los valores de heterocigosidad media calculados a través de varios loci (es decir, H_o : frecuencia observada de heterocigosis y H_e : frecuencia esperada bajo el apareamiento aleatorio de Hardy-Weinberg), son parámetros valiosos utilizados para estimar el grado de variación genética dentro de una población. La estructuración de la población ocurre cuando las frecuencias de genotipo se desvían de las proporciones esperadas de Hardy-Weinberg (HW). Si se produce endogamia o selección, entonces las poblaciones pueden considerarse estructuradas de alguna manera (Sbordoni, Allegrucci y Cesaroni, 2012).

Índices de fijación

Wright (1978) desarrolló tres índices de fijación (F) para evaluar la estructuración de las poblaciones: F_{is} (interindividual), F_{st} (subpoblaciones) y F_{it} (población total).

F_{is} representa la endogamia de un individuo o la fijación de un alelo respecto de la población total, y se conoce como coeficiente de endogamia. Aunque el apareamiento no aleatorio no tiene ningún efecto sobre la frecuencia de los alelos, la endogamia hace que cambien las frecuencias de los genotipos, dado que la frecuencia de los homocigotos aumenta y la frecuencia de los heterocigotos disminuye. Por lo tanto, dado que el apareamiento no aleatorio solo reorganiza las frecuencias de genotipo con respecto a sus expectativas de HW, podemos usar la desviación de las frecuencias de genotipo de sus valores esperados como una medida de endogamia. Si H_o es la frecuencia observada de heterocigotos y H_e es la frecuencia esperada en el apareamiento aleatorio del total de la población, entonces el coeficiente de endogamia es:

$$F_{local} = 1 - H_o/H_e$$

Si no hay heterocigotos en la población, entonces el coeficiente de endogamia es 1. Cuando la frecuencia de heterocigotos es igual a la expectativa de HW, entonces el

coeficiente de endogamia es 0. En los casos en que hay un exceso de heterocigotos observados, entonces el coeficiente de endogamia puede ser negativo.

El índice F_{st} mide el efecto de la subpoblación comparado con la población total, es decir la reducción de la heterocigosidad en una subpoblación debido a la deriva genética. F_{st} es la medida más inclusiva de la subestructura de la población y es más útil para examinar la divergencia genética general entre las subpoblaciones. También se llama coeficiente de coancestría (Q) (Weir y Cockerham, 1984). Se calcula utilizando la heterocigosidad de subpoblación (promedio) y la heterocigosidad esperada de la población total. Los valores de F_{st} siempre son positivos; oscilan entre 0 (sin subdivisión, apareamiento aleatorio, sin divergencia genética dentro de la población) y 1 (aislamiento completo, subdivisión extrema). Valores de F_{st} de hasta 0,05 indican una diferenciación genética insignificante, mientras que un valor de 0,25 indica una gran diferenciación genética dentro de la población analizada.

El índice F_{it} raramente se usa y representa la endogamia de un individuo en relación con la población total.

Para poder calcular los diferentes índices F , definimos tres H , calculadas simplemente como promedios dependiendo del nivel de la estructura de la población analizada. Por lo tanto, H_i toma el valor promedio de la heterocigosis observada de los individuos dentro de la subpoblación. H_s toma el valor promedio de la heterocigosis esperada de los individuos dentro de la subpoblación, suponiendo equilibrio HW. Por último, H_t es el promedio de la heterocigosis esperada en la población completa, también asumiendo equilibrio HW. De esta manera, los índices se calculan:

$$F_{it} = 1 - \frac{H_i}{H_t}; F_{is} = 1 - \frac{H_i}{H_s}; F_{st} = 1 - \frac{H_s}{H_t}$$

Índice de diversidad

El índice de diversidad genética es calculado de acuerdo a la propuesta de Weir (1996):

$$D = 1 - \left(\frac{1}{m}\right) \sum_{l=1}^n \sum_{u=1}^n p_{lu}^2$$

donde P_{lu}^2 es la frecuencia del alelo u^{th} en el l^{th} locus y m es el número de loci. Las frecuencias se calcularon en base a los seis loci que resultaron informativos para la totalidad de los individuos de la variedad local tomate ‘Platense’ estudiados. Se utilizaron los mismos loci para calcular el índice de diversidad de cada una de las subpoblaciones de ‘Platense’.

Estimación de identidad

La probabilidad de identidad aleatoria o *PRI* (*Probability of Random Identity*) para cada una de las poblaciones y con todos los marcadores SSR utilizados, se calculó de la siguiente manera:

$$PRI = \prod_{i=1}^n p_i$$

donde p_i es la frecuencia del alelo i y n es el número de marcadores evaluados. En base a los valores de *PRI*, el porcentaje de potencia de exclusión (*EP*, *Excluding Power*) se calculó de la siguiente forma:

$$EP = (1 - PRI) \times 100$$

que indica el poder de identificar inequívocamente a un individuo o subpoblación dada, entre todos los demás en el conjunto total, utilizando los 16 marcadores SSR descritos en este estudio (Ribeiro y cols., 2017). Por lo tanto y para ejemplificar, si para una entrada se obtiene un perfil molecular generado por un marcador específico o la suma de varios marcadores, que tiene una probabilidad de identidad aleatoria del 0,0001%, la probabilidad de exclusión será del 99,9999%. Cuando se obtiene este perfil molecular en otra muestra, esto indica que la probabilidad de esta identidad no es aleatoria y que las muestras corresponden a la misma entrada.

Dendrograma

La similitud genética se estimó a partir de matrices binarias utilizando el coeficiente de *Jaccard* (Sneath y SoIkal, 1973). Se realizó un análisis de conglomerados (construcción del dendrograma) sobre la base de los datos binarios utilizando el software *NTSys 2.02*. (Rohlf, 2000) con el método de grupo de pares no ponderados que utiliza el método de promedios aritméticos (UPGMA).

Análisis de coordenadas principales

Este análisis permite ver la interdependencia entre variables categóricas y encontrar una representación gráfica de los n individuos y subpoblaciones estudiadas, tal que se refleje la distancia entre ellos. Estas distancias fueron calculadas a partir de la matriz de similitud según el coeficiente de *Jaccard* y por medio del programa *NTSys pc 2.02* (Rohlf, 2000). A diferencia del análisis de componentes principales que requiere variables cuantitativas, el análisis de coordenadas principales (PCoA, *Principal Coordinates Analysis*) se puede hacer con cualquier tipo de variables, incluso con una mezcla de variables.

Por último, se realizó un análisis de escalado multidimensional clásico basado en la matriz de disimilitud, y se generó una gráfica para construir la configuración de puntos en el plano para las dos primeras coordenadas. También se calcularon elipses de confianza para cada subpoblación de las variedades locales ‘Platense’ en el gráfico PCoA, suponiendo que los datos se encuentran normalmente distribuidos a lo largo de ambos ejes de coordenadas principales. Estas elipses de error definen la región que contiene el 95% de todas las muestras que pertenecen a la distribución gaussiana esperada.

3.3. Resultados

En una primera etapa se utilizaron los 34 marcadores seleccionados (ver 3.2.2. Selección de los cebadores) pero solo en las muestras denominadas como *pool* de las subpoblaciones de ‘Platense’, para poder así determinar cuáles de estos marcadores son polimórficos para las mismas. En este primer ensayo, se encontró que un 15% (5) de los marcadores evaluados fueron polimórficos para las subpoblaciones de tomate ‘Platense’. Otros marcadores (6) dieron bandas inespecíficas (18%), las cuales se pueden deber a la falta de optimización de las temperaturas de *annealing* de la reacción de PCR. Más allá de que se probaron diferentes temperaturas, quedó pendiente probar temperaturas de *annealing* de forma escalonadas (Don, Cox, Wainwright, Baker y Mattick, 1991).

Nº	Locus	Cromosoma	Subpoblación ‘Platense’
1	<u>TMS63</u>	1	polimórfico
2	SSR192	1	monomórfico
3	<u>SSR96</u>	2	polimórfico
4	SSR5	2	monomórfico
5	SSR295	2	monomórfico
6	TMS58	3	monomórfico
7	Tom59-60	3	<i>n.e.</i>
8	<u>SSR111</u>	3	polimórfico
9	Tom47-48	3	monomórfico
10	SSR43	4	monomórfico
11	LE21085	4	<i>n.e.</i>
12	SSR627	4	monomórfico
13	SSR638	4	<i>n.e.</i>
14	SSR162	5	monomórfico
15	Tom152-153	5	monomórfico
16	EST258529	5	<i>n.e.</i>
17	SSR47	6	monomórfico
18	TMS60	7	monomórfico
19	SSR45	7	monomórfico
20	SSR327	8	monomórfico
21	TMS59	8	monomórfico

22	SSR63	8	monomórfico
23	LELEUZIP	8	monomórfico
24	SSR69	9	monomórfico
25	<u>Tom236-237</u>	9	polimórfico
26	LELE25	10	monomórfico
27	<u>SSR318</u>	10	polimórfico
28	TMS42	11	monomórfico
29	Tom144-145	11	monomórfico
30	SRR20	12	monomórfico
31	TMS52	12	monomórfico
32	LECHI3	<i>n.d.</i>	monomórfico
33	LESSRPSPGb	<i>n.d.</i>	<i>n.e.</i>
34	LEEF1Aa	<i>n.d.</i>	<i>n.e.</i>

Cuadro 3.4: Evaluación de 34 marcadores seleccionados en las muestras denominadas como *pool* de las subpoblaciones de ‘Platense’. *n.e.* = banda no específica. *n.d.* = no determinado. Locus polimórficos subrayados.

A partir de estos resultados (Cuadro 3.4), se evaluó el resto de las accesiones. En total formaron parte del estudio 13 accesiones de *S. lycopersicum*, a saber, los 3 cultivares comerciales del INTA, un híbrido comercial y las 9 subpoblaciones de tomate ‘Platense’ (‘Bustos’, ‘Prieto’, ‘Grasso’, ‘Molinario’, ‘Paolettich’, ‘Luna’, ‘Gentile’, ‘Carcione’ y ‘Del Manso’), y *S. pimpinellifolium* L. (LA1589) como *outgroup*. De los marcadores evaluados se seleccionaron un total de 16 (TMS63, SSR5, SSR96, TMS58, SSR111, Tom47-48, SSR43, Tom152-153, TMS60, LELEUZIP, SSR327, Tom236-237, SSR383, LECHI3, TMS42, TMS52).

Del total de SSR evaluados el 87,5% (14 SSR) fueron polimórficos para todas las accesiones y solamente 2 fueron monomórficos (algunos de estos resultados en Figura 3.2). El promedio por locus fue de 2,71, con un total de 41 alelos para todas las accesiones estudiadas. Los marcadores SSR111 y Tom236-237 presentaron el máximo número de alelos, con 5 por locus (Cuadro 3.5).

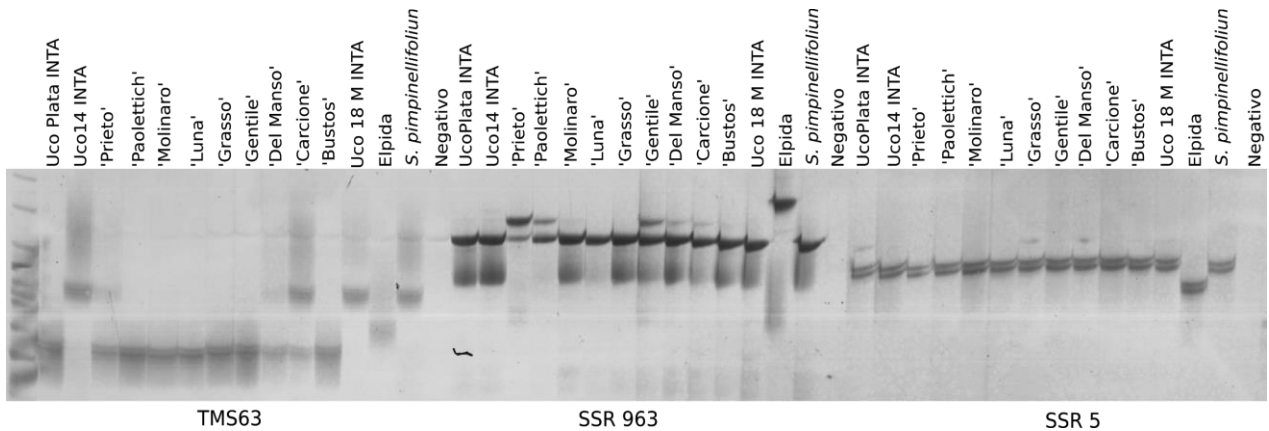


Figura 3.2: Corrida electroforética en gel de policridamida de los productos de PCR de los marcadores TMS63, SS963 y SSR 5 de las muestras *pool* de Uco Plata INTA, Uco 14 INTA, ‘Prieto’, ‘Paolettich’, ‘Molinaro’, ‘Luna’, ‘Grasso’, ‘Gentile’, ‘Del Manso’, ‘Carcione’, ‘Bustos’, Uco 18 M INTA y Elpida.

Los valores de *PIC* variaron en un rango de 0,12 (locus TMS58, Tom47-48 y SSR5) a 0,56 (locus Tom236-237), con un promedio de 0,28. Los marcadores SSR altamente informativos (*PIC* alto) como SSR111 y Tom236-237 demostraron ser muy útiles tanto para la identificación de variedades como para la evaluación genética de estas accesiones de tomate.

SSR	Nº alelos	H_o	H_e	F_{local}	PIC
TMS63	3	0,08	0,31	0,74	0,27
SSR5	2	0	0,13	1	0,12
SSR96	3	0,14	0,34	0,59	0,30
TMS58	2	0	0,13	1	0,12
SSR111	5	0,76	0,59	- 0,29	0,52
Tom47-48	2	0	0,13	1	0,12
SSR43	1	0	0	-	0
Tom152-153	3	0	0,26	1	0,24
TMS60	3	0,93	0,56	- 0,66	0,47
LELEUZIP	2	0,76	0,47	- 0,62	0,36
SSR327	2	0	0,13	1	0,13
Tom236-237	5	0,20	0,56	0,64	0,56
SSR383	2	0,08	0,42	0,81	0,42

LECHI3	2	1	0,50	-1	0,38
TMS42	2	0	0,24	1	0,41
TMS52	2	0,86	0,50	-0,72	0,38
promedio	2.56	0,32	0,35		0.28

Cuadro 3.5: Heterocigosidad esperada (H_e), coeficiente de endogamia (F_{local}), heterocigosidad observada (H_o) para 16 loci SSR estudiados en genotipos de tomate (cultivares comerciales, subpoblaciones de variedades locales e híbrido).

Para los marcadores que presentaban heterocigosis en su muestra *pool*, se realizó la reacción de PCR con los mismos marcadores sobre los 4 individuos de esas mismas entradas (a, b, c y d), como se muestra en la Figura 3.3. De esta forma, pudimos ver si la entrada estaba compuesta por individuos con diferentes tamaños de alelos o heterocigotas o ambas. Para el marcador TMS63 se analizaron los individuos de las poblaciones ‘Prieto’, ‘Del Manso’ y ‘Carcione’. Para el marcador SRR96 se analizaron los individuos de las poblaciones ‘Prieto’, ‘Paolettich’, ‘Gentile’, ‘Del Manso’ y ‘Carcione’. Para el marcador SRR111 se analizaron los individuos de las poblaciones ‘Prieto’, ‘Paolettich’, ‘Luna’, ‘Grasso’, ‘Gentile’, ‘Del Manso’, ‘Carcione’ y ‘Bustos’. Para el marcador LELEUZIP se analizaron los individuos de las poblaciones ‘Molinaro’, ‘Luna’, ‘Grasso’, ‘Gentile’, ‘Del Manso’, ‘Carcione’ y ‘Bustos’. Para el marcador TOM236-237 se analizaron los individuos de las poblaciones ‘Paolettich’, ‘Luna’, ‘Grasso’ y ‘Del Manso’. Para el marcador SSR383 se analizaron los individuos de las poblaciones ‘Paolettich’, ‘Grasso’ y ‘Del Manso’.

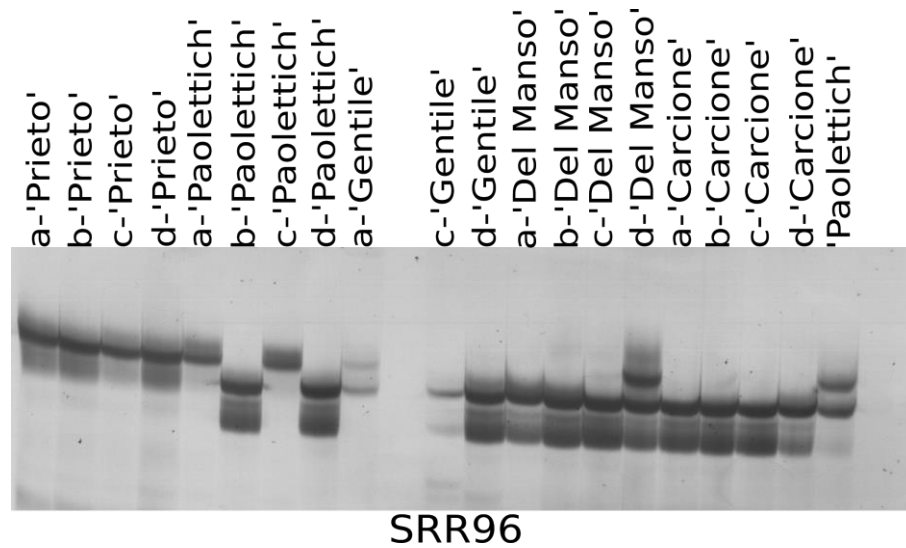


Figura 3.3.: Corrida electroforética en gel de poliacrilamida del marcador SRR96 y de los individuos de las poblaciones ‘Prieto’, ‘Paolettich’, ‘Gentile’, ‘Del Manso’ y ‘Carcione’ y el *pool* de ‘Paolettich’.

Cuando se analizaron estos índices por subpoblaciones, se vio que contenían un alto índice de heterocigosis observada promedio (H_o), como la subpoblación ‘Bustos’ con 0,31 y un coeficiente de endogamia (F_{local}) promedio de -1. El resto de los valores se pueden observar en el Cuadro 3.6.

En las variedades cultivadas de tomate se esperaría encontrar que los alelos se encuentren fijados en un estado homocigota, ya que son autógamas, pero esto no siempre puede llegar a tener un beneficio genético de selección para permanecer en dicho estado. Esta alta heterocigosis observada se puede explicar por el cruzamiento entre plantas por medio de insectos polinizadores, aunque esto es muy poco frecuente por la anatomía de la flor. Otra posible fuente de cruce es a partir de la práctica de los productores familiares de la región, los cuales seleccionaban los frutos denominados “florones” para sacar semillas. El florón es el fruto que se forma a partir de flores fasciadas, lo que da la mayor probabilidad de cruzamiento (Ahumada, Otero y Garat, 2011). Por último, estos cultivos anuales suelen ser difundidos por el intercambio de semillas entre productores, lo que dificulta la identificación de límites claros para una “estructura” genética (Zeven, 1999).

Subpoblaciones	H_o	H_e	F_{local}
'Del Manso'	0,30	0,22	- 0,37
'Paolettich'	0,20	0,19	- 0,05
'Grasso'	0,31	0,21	- 0,5
'Luna'	0,28	0,16	- 0,8
'Carcione'	0,27	0,17	- 0,6
'Bustos'	0,31	0,16	- 1
'Prieto'	0,19	0,12	- 0,60
'Gentile'	0,19	0,11	- 0,64
'Molinaro'	0,25	0,13	- 1

Cuadro 3.6: Heterocigosis observada (H_o) Heterocigosis esperada (H_e) y coeficiente de endogamia (F_{local}) de todas las subpoblaciones del tomate 'Platense'.

También se calcularon los índices, F_{is} , F_{it} , F_{st} a partir de las heterocigosis esperada y observada promedio (H_i y H_s) y la heterocigosis global esperada (H_t) (Cuadro 3.7). Anteriormente, utilizamos F_{local} para medir la deficiencia de heterocigosis debido a la endogamia dentro de las subpoblaciones. El apareamiento no aleatorio también puede surgir debido a la estructura geográfica de la población. Por ejemplo, las subpoblaciones pueden estar mezcladas (intercambio) con semillas de productores vecinos cercanos, o estar suficientemente alejadas para generar subpoblaciones estructuradas.

El valor de F_{st} calculado para nuestra población 'Platense' fue de 0,33, indicando una alta diferenciación genética entre las subpoblaciones y en concordancia con el hecho de que cada familia poseía su colección propia de semillas.

H_i	H_s	H_t
0,26	0,16	0,24
F_{is}	F_{it}	F_{st}
- 0,63	- 0,08	0,33

Cuadro 3.7: Índices de heterocigosis media observada y esperadas (H_s , H_i y H_t) y coeficientes F (F_{st} , F_{it} y F_{is}) de la población tomate ‘Platense’.

3.3.1. Índice de diversidad

Los valores del índice de diversidad (D) se calcularon para las accesiones de subpoblación y el índice osciló entre 0,08 y 0,41, con un valor medio de 0,23. El valor mínimo de D , se encontró para la subpoblación ‘Molinario’ (0,08), lo que podría indicar que esta accesión en particular ha sido sometida a un mayor grado de selección que sus otras contrapartes locales, por lo que se podría considerar más como un cultivar que como una subpoblación diversa desde un punto de vista genético. Por otro lado, la mayor diversidad intra subpoblación (0,41) correspondió a la subpoblación ‘Del Manso’, casi tan alta como el valor de diversidad alcanzada por la población ‘Platense’ (0,42). El índice de diversidad para la variedad local ‘Platense’ (calculado sobre la base de todas las accesiones de ‘Platense’ estudiadas) fue sólo ligeramente más bajo que el índice de diversidad calculado para todas las accesiones en el presente estudio (0,46). Estos resultados indicarían que la variedad local de tomate ‘Platense’ está compuesta por una población general en la que se mantiene una diversidad sustancial. Sin embargo, cuando se analizó a nivel de subpoblación se observaron grados de diversidad muy diferentes.

Grupo/Subpoblación	<i>D</i>
Todas las accesiones	0,46
Todas las subpoblaciones de 'Platense'	0,42
'Del Manso'	0,41
'Paolettich'	0,35
'Grasso'	0,31
'Luna'	0,25
'Carcione'	0,23
'Bustos'	0,17
'Prieto'	0,15
'Gentile'	0,14
'Molinaro'	0,08

Cuadro 3.8: Índice de diversidad (*D*) de todas las subpoblaciones del tomate 'Platense'.

3.3.2. Capacidad de estimación de identidad

El conjunto de 14 marcadores informativos de SSR (para todas las accesiones estudiadas) generó una combinación alélica única para cada accesión evaluada, con una probabilidad de valores de identidad aleatorios que fueron desde $10e-11$ a $10e-15$, para el híbrido comercial Elpida y *S. pimpinellifolium*, hasta $10e-4$ para UCO Plata INTA y 'Molinaro'. Esto corresponde a valores de poder de exclusión de más del 99,99% para cada accesión estudiada (Cuadro 3.9).

Accesiones	<i>PRI</i>	<i>EP (%)</i>
Uco Plata INTA	1,3784e-04	99,986216
Uco 14 INTA	4,5577e-08	99,999995
Uco 18 M INTA	1,4137e-05	99,998586
'Prieto'	2,3172e-05	99,997683
'Paolettich'	3,7225e-07	99,999963
'Molinaro'	1,0452e-04	99,989548
'Luna'	1,1840e-05	99,998816
'Grasso'	9,0614e-07	99,999909
'Gentile'	5,2492e-05	99,994751
'Del Manso'	7,6519e-07	99,999923
'Carcione'	1,1717e-06	99,999883
'Bustos'	1,4768e-05	99,998523
Elpida	5,6952e-11	>99,999999
<i>S. pimpinellifolium</i>	5,3686e-15	>99,999999

Cuadro 3.9: Probabilidad de identidad aleatoria (*PRI*) y el porcentaje de poder de exclusión (*EP*) de los 14 marcadores informativos SSR para cada una de las accesiones evaluadas (Subpoblaciones de tomate 'Platense', cultivares del INTA, un híbrido y un silvestre).

Para evaluar todos los marcadores SSR utilizados, se calculó el valor de *PRI* por marcador, obteniendo en consecuencia el *PRI* promedio por marcador para todas las accesiones. Con estos valores se realizó un rango del *PRI* promedio por marcador (Cuadro 3.10.). Este rango de marcador se utilizó posteriormente para calcular el *EP* frente al número de marcadores utilizados. Los resultados se muestran en la Figura 3.3. De esta representación se puede concluir fácilmente que, además del hecho de que el uso de los 14 marcadores informativos nos permite identificar todas las accesiones con un *EP* de al menos 99,99%, el uso de los 9 marcadores mejor clasificados representa la identidad con un *EP* de más del 99,5%, mientras que el uso de los primeros 6 marcadores clasificados permite la identificación con un *EP* de más del 95%. Por lo tanto, este conjunto de 14 marcadores SSR confiere una alta capacidad de distinción y podría recomendarse como una base de datos molecular para las subpoblaciones de variedades locales de tomate 'Platense' evaluadas en este estudio, ya sea para la gestión del banco de semillas o para fines de programas de mejoramiento participativo. Además, un subconjunto de estos marcadores podría usarse en eventuales ensayos de identificación de variedades locales 'Platense' no exhaustivos.

SSR	Promedio <i>PRI</i> / <i>EP</i> (%)	rango
TMS63	0,52 / 48,25	3
SSR96	0,52 / 47,82	5
SSR5	0,87 / 13,27	10
SSR111	0,18 / 82,19	2
TMS58	0,87 / 13,27	10
Tom47-48	0,87 / 13,27	10
Tom152-153	0,74 / 25,51	6
TMS60	0,21 / 79,47	9
LELEUZIP	0,27 / 72,72	11
Tom236-237	0,29 / 71,50	1
SSR383	0,48 / 52,27	4
LECHI3	0,30 / 69,64	12
TMS42	0,76 / 24,49	7
SSR327	0,87 / 13,27	10
TMS52	0,31 / 68,75	8
SSR43	1,00 / 0,00	12

Cuadro 3.10: Cálculo de valores de *PRI* por marcador y los *PRI* promedio por marcador para todas las accesiones, estableciendo así un rango de marcador de mayor a menor *PRI* promedio.

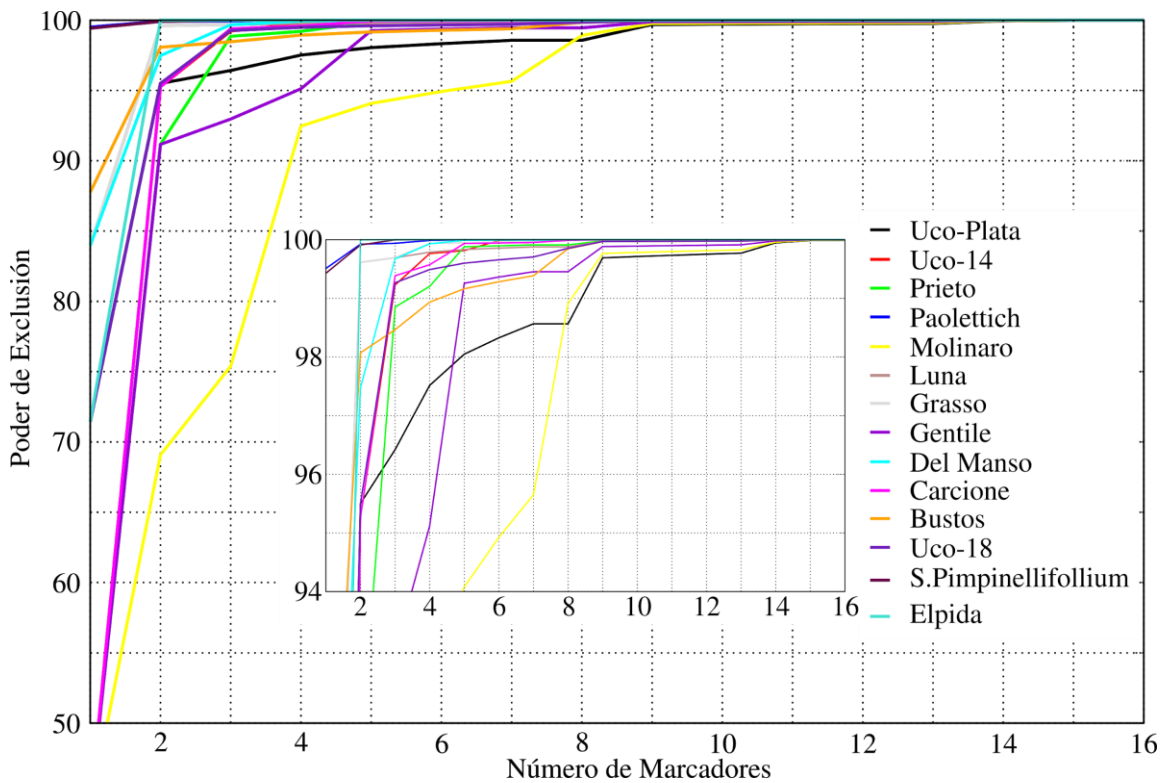


Figura 3.3: Poder de exclusión frente al número de marcadores SSR mejor clasificados utilizados.

3.3.3. Dendrograma

El análisis de agrupamiento se realizó en todos los individuos y muestras de subpoblación, calculando un dendrograma de similitud UPGMA (Método de grupo de pares no ponderados usando promedios aritméticos) basado en polimorfismos SSR. Los resultados se muestran en la Figura 3.4, y, en general, están de acuerdo con los valores de diversidad obtenidos para todas las accesiones estudiadas, así como para toda la variedad local 'Platense' y sus subpoblaciones. Es decir, tres de las subpoblaciones con el índice de diversidad más bajo ('Molinaro', 'Bustos' y 'Prieto') aparecen estrechamente agrupadas en el dendrograma, mientras que la subpoblación con el índice de diversidad más alto ('Del Manso') aparece completamente desagrupada. El segundo índice de diversidad más alto ('Paolettich') muestra, igualmente, desagrupación casi completa. Las otras tres subpoblaciones con valores medios de índice de diversidad ('Grasso', 'Luna' y 'Carcione') muestran un comportamiento de agrupamiento intermedio.

Con respecto a las accesiones no 'Platense', el cultivar de tomate redondo Uco 18 M INTA aparece en el medio del grupo 'Platense'. Sin embargo, se conecta a una gran distancia (coeficiente *Jaccard* 0,79) con la accesión más cercana de esta variedad local, casi en la raíz del grupo 'Platense' (coeficiente *Jaccard* 0,74). Además, esta distancia es 1,615 veces mayor que la mayor distancia que separa dos accesiones Platense dentro del grupo (es decir, el coeficiente de *Jaccard* de 0,87 para las subpoblaciones 'Paolettich' y 'Del Manso'). Por otro lado, el cultivar comercial Uco Plata INTA muestra exactamente el mismo patrón de alelo SSR que el individuo b-Luna (coeficiente *Jaccard* de 1,00) y hermano de la subpoblación 'Luna', lo que indica un grado máximo de similitud genética con esta accesión. Esto no es sorprendente dado que el INTA describe este cultivar como una variedad de tomate de tipo platense alto (Lepez, n/d). Además, este resultado confirma el pedigrí general 'Platense' de UCO Plata INTA y sitúa a la subpoblación 'Luna' como su pariente más cercano entre las accesiones 'Platense' disponibles en el banco de semillas INTA- La Consulta. Finalmente, Uco 14 INTA (un tomate con frutos de forma elipsoide) aparece como hermano de todo el grupo platense, mientras que el híbrido comercial Elpida es hermano de todo este último, y *S. pimpinellifolium* aparece como un grupo externo (*outgroup*).

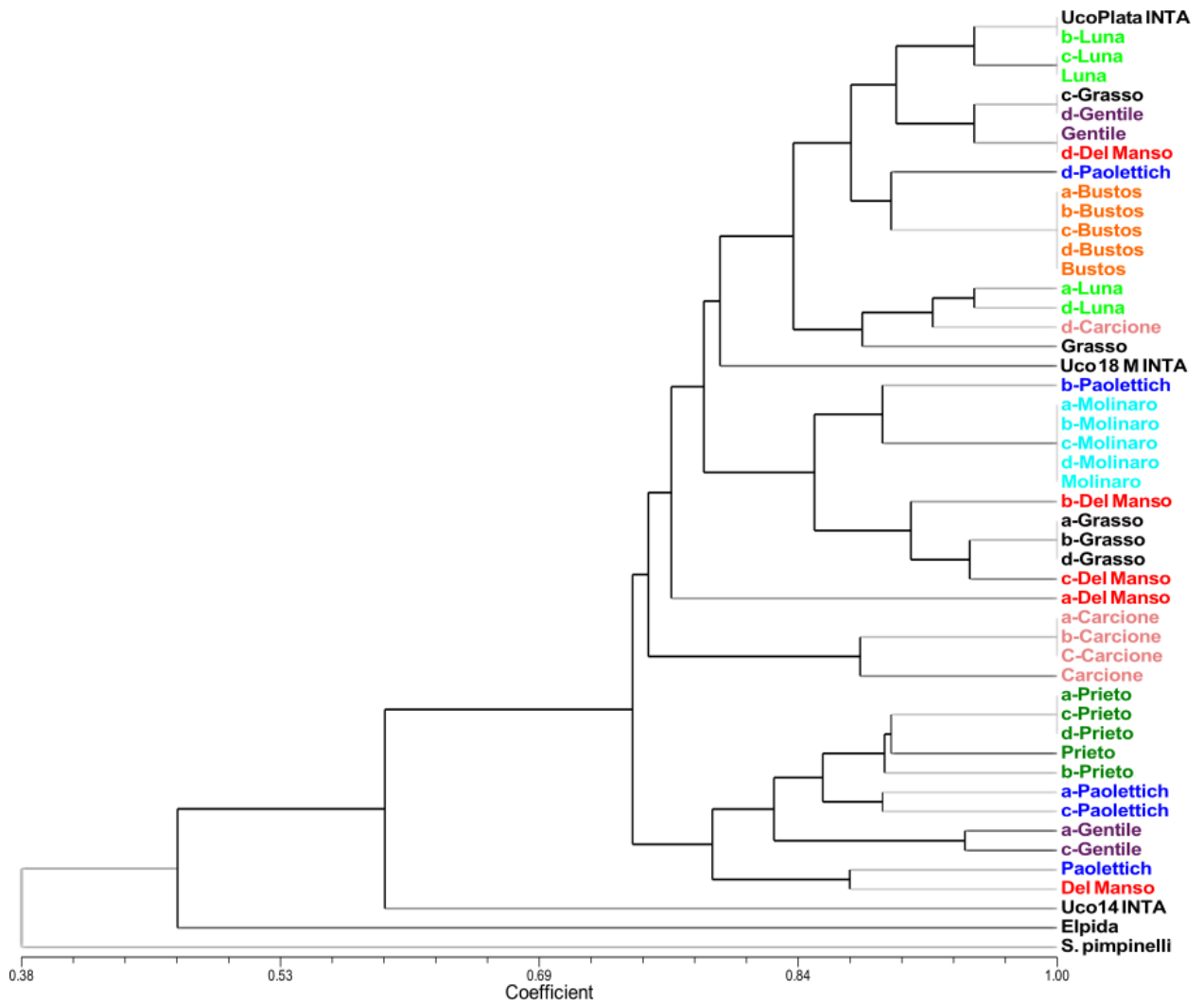


Figura 3.4: Dendrograma para las subpoblaciones de tomate ‘Platense’ y sus muestras individuales, cultivares del INTA, híbridos y de *S. pimpinellifolium*, construidas utilizando el método UPGMA. La escala se basa en el coeficiente de similitud de *Jaccard*.

3.3.4. Análisis de coordenadas principales

Se realizó un análisis de coordenadas principales (PCoA) para todos los individuos y subpoblaciones de las 14 accesiones estudiadas. Los resultados se muestran en la Figura 3.5. Las dos primeras coordenadas del análisis de PCoA representan el 36,8% de la variación total revelada por los 14 marcadores polimórficos SSR entre todas las accesiones estudiadas, con la primera y segunda coordenadas capturando el 23,1% y el 13,7% de la variación total, respectivamente (Figura 3.5.). Además, se muestra una elipse de confianza para cada

subpoblación de variedades locales 'Platense' en el gráfico PCoA, suponiendo datos 2D normalmente distribuidos. Estas elipses de error representan un iso-contorno de la distribución gaussiana para cada subpoblación y permiten la visualización de un intervalo de confianza 2D del 95%, que define la región que contiene 0,95 de todas las muestras que podrían extraerse de una distribución gaussiana subyacente. El área rodeada por cada óvalo da una idea de la variabilidad capturada por las dos primeras coordenadas principales para cada subpoblación. El 'Manso' muestra la elipse de confianza más grande, seguida de 'Paolettich', 'Grasso' y 'Luna', mientras que 'Molinaro' y 'Bustos' muestran variabilidad cero en este plano, y 'Carcione', 'Prieto' y 'Gentile' muestra valores intermedios. Esto confirma las observaciones generales previamente establecidas por el análisis de diversidad de la Tabla 3, así como el análisis de agrupamiento de la Figura 3.5. Es decir, que las subpoblaciones 'Platense' muestran grados de variabilidad muy diferentes, que abarcan desde casos muy diversos genéticamente (como 'Del Manso' o 'Paolettich') hasta otros muy homogéneos (como 'Molinaro' y 'Bustos'). La unión de todas las elipses de subpoblación explica la variabilidad de la variedad local 'Platense' total capturada por las dos primeras coordenadas principales, y abarca casi la mitad del plano PCoA. Además, la intersección entre elipses da una idea del grado de similitud genética entre las diferentes subpoblaciones. En particular, las nueve subpoblaciones 'Platense' constituyen un grupo común. Sin embargo, se observan grados variables de conectividad entre sí, que abarcan subpoblaciones sueltas ('Carcione' conectado solo a 'Del Manso' y 'Luna', o 'Prieto' conectado a 'Paolettich', 'Gentile' y 'Del Manso') a uno totalmente conectado ('Del Manso') y el resto muestra conectividad intermedia. Finalmente, las subpoblaciones menos diversas ('Molinaro' y 'Bustos') aparecen dentro del grupo 'Platense', aunque sin una elipse de confianza, dado que todos sus individuos aparecen exactamente en el mismo punto del plano PCoA. Con respecto a las accesiones no 'Platense' actualmente estudiadas, aunque Uco 18 M INTA aparece casi en el medio del grupo principal 'Platense' sin embargo, no es adyacente a ninguna accesión 'Platense' particular. Por otro lado, Uco Plata INTA aparece en la misma posición que el individuo b-Luna. De manera paralela a estas, las últimas tres accesiones no 'Platense' aparecen claramente fuera de las regiones 'Platense' del plano PCoA, como se espera por los resultados mostrados en las secciones de análisis de diversidad y agrupación.

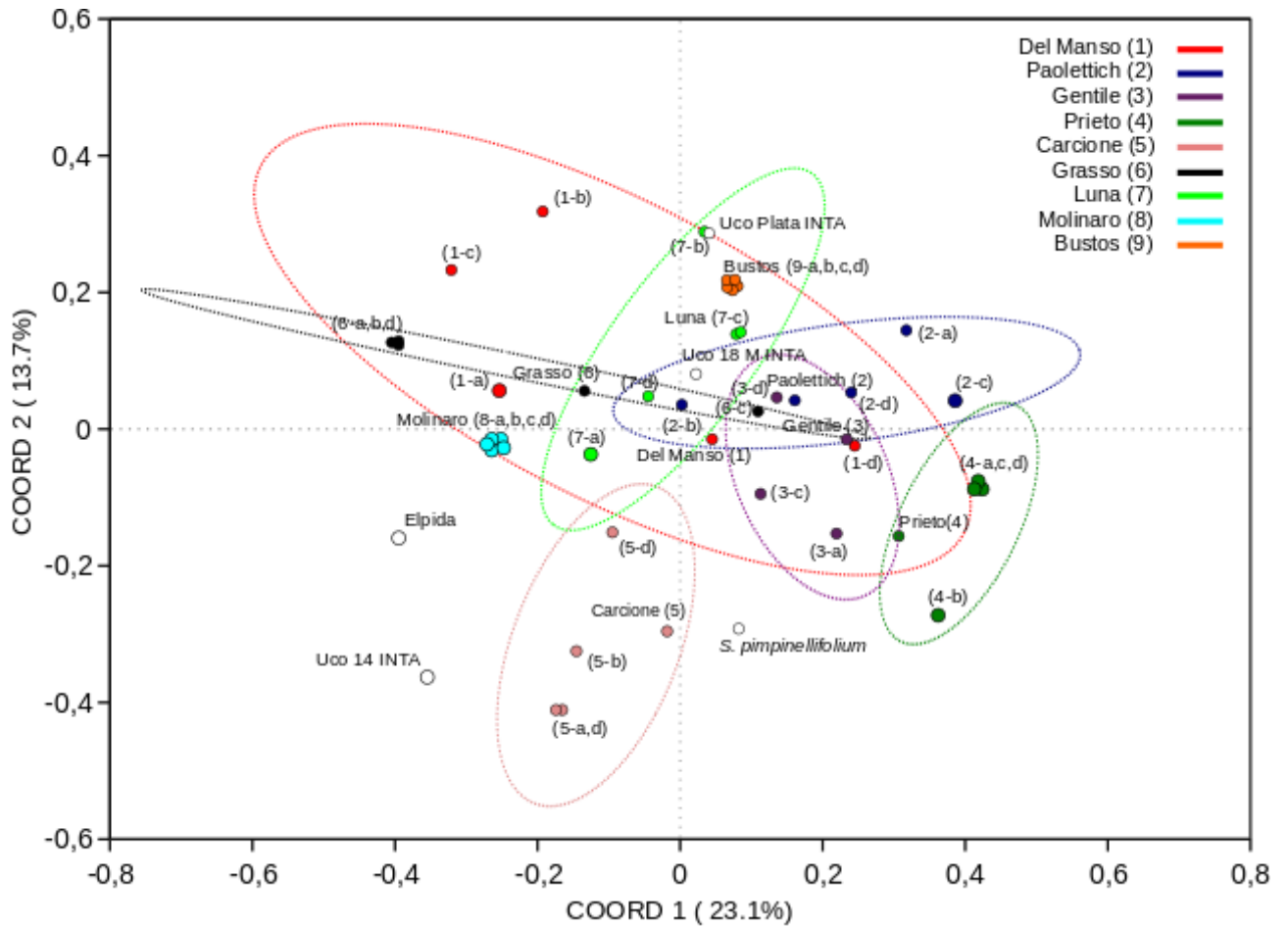


Figura 3.5: Análisis de coordenadas de componentes principales basados en la información obtenida de los marcadores SSR utilizados.

3.4. Conclusiones

A pesar de que un gran número de variedades locales históricamente se han cultivado en la región de La Plata (Nico y cols., 2006), existen pocos estudios centrados en su caracterización genética. En el presente estudio hemos identificado genéticamente la variedad local de tomate 'Platense' y de sus subpoblaciones actualmente disponibles en el Banco de Germoplasma de INTA - La Consulta.

Se considera que el tomate presenta una estrecha base genética. Sin embargo, se considera que las variedades locales de tomate contienen valores de diversidad genética y

fenotípica superiores a los cultivares comerciales e híbridos (Park y cols., 2004; Mazzucato y cols., 2008). Por lo tanto, constituyen una fuente de recursos valiosos para los programas de mejoramiento genético vegetal. En este sentido, el presente estudio ha determinado la variedad local de tomate 'Platense' como un recurso altamente diverso y algunas de sus subpoblaciones con gran parte de la diversidad total observada. Por otra parte, pocas subpoblaciones presentaron homogeneidad genética (para los marcadores utilizados) y, por lo tanto, pueden llegar a considerarse como cultivares.

En general, el reemplazo de las variedades de semillas es común entre los agricultores. Además, el intercambio de semillas es la forma más común por la cual las plantas anuales altamente autopolinizantes se diseminan (Zeven, 1999). Estas prácticas generalmente dificultan la identificación clara de los “límites genéticos” y el tomate 'Platense', como variedad local, no es una excepción a esta regla.

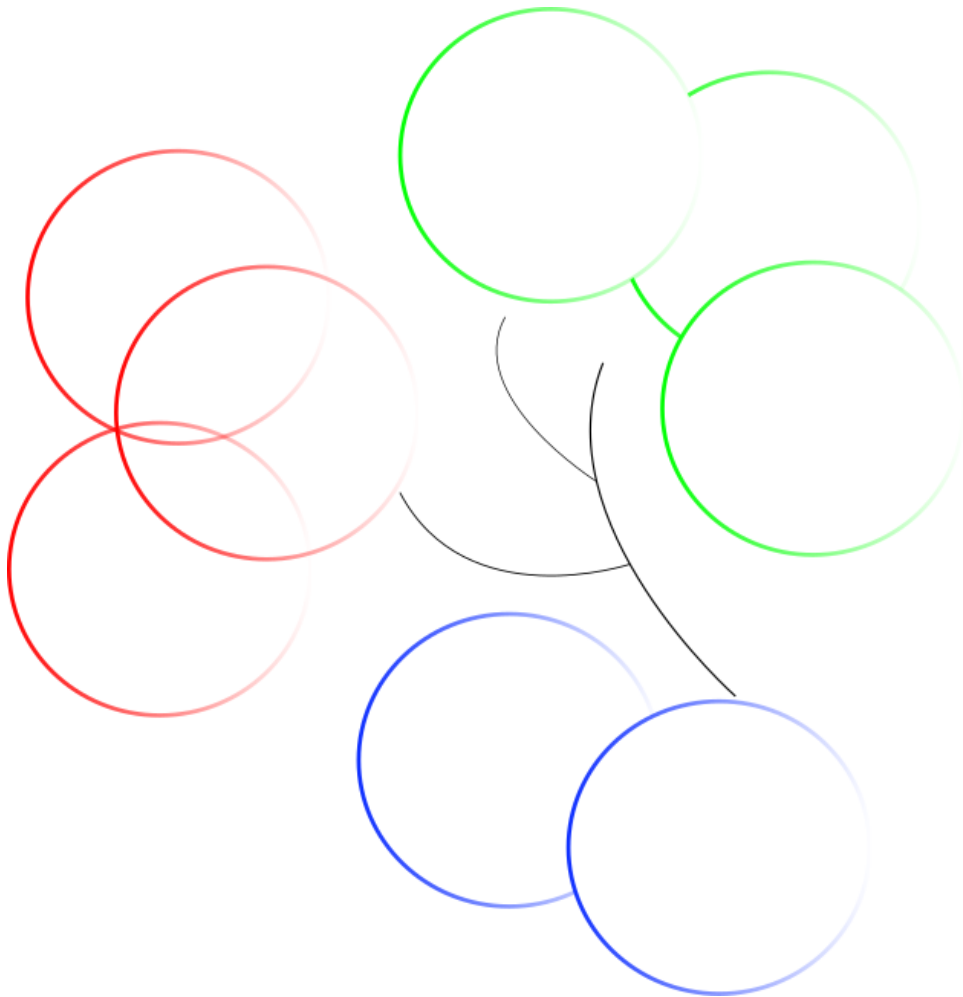
Los resultados obtenidos en este capítulo, pueden ser utilizados como referencia tanto para estudios futuros sobre muestras de germoplasma 'Platense' de origen diverso, como para el estudio de nuevas subpoblaciones autóctonas que se encuentran actualmente en uso por los horticultores locales, semillas comerciales o accesiones 'Platense' actualmente disponibles en otros bancos de semillas de todo el mundo. Esta alternativa puede resultar de relevancia dado que, durante los últimos años, la producción y comercialización regional de esta variedad local ha sido revalorizada socioculturalmente tanto por su calidad organoléptica como por su adaptabilidad a crecer en condiciones agroecológicas (Garat y Otero, 2012).

Desde principios del siglo XX, los agricultores locales del cinturón hortícola de La Plata han realizado progresivamente procesos de selección que dieron origen a las actuales y diversas variedades de tomate, muy probablemente de origen francés o italiano (Folquer, 1976; Garat, 2002), produciendo así subpoblaciones del tomate 'Platense'. Este procedimiento complejo e irregular puede, sin embargo, ser considerado como una selección y mejora informal. Tan rico patrimonio hortícola, combinado con los resultados aquí presentados, abre la posibilidad de un novedoso esquema de mejora participativa, integrando horticultores, consumidores e investigadores locales bajo un objetivo común: la consolidación del 'Platense' como *landrace*. Al mismo tiempo, estos resultados reafirman el papel de la conservación *in situ* de los horticultores locales en la preservación de los recursos

genéticos del tomate, permitiendo que concurren sinérgicamente la selección natural y la humana.

Capítulo 4

Análisis filogenómico de cloroplastos de tomate



4.1. Introducción¹⁰

Todo organismo vivo comparte un ancestro común y el objetivo de la taxonomía es reconstruir el árbol o la red de relaciones entre los diferentes taxones y su historia evolutiva. Poco después de estar disponible la primera proteína secuenciada, se desarrollaron enfoques filogenéticos moleculares para comprender las relaciones evolutivas entre las proteínas (o genes). Quedó claro que los árboles genéticos también ayudarían a desentrañar la filogenia de las especies. El desarrollo del método de secuenciamiento de Sanger y de la PCR y su disponibilidad de estas tecnologías en toda la comunidad científica, contribuyó al surgimiento de la filogenética. A finales de la década de 1990, los resultados de estudios filogenéticos desafiaron las relaciones evolutivas de muchos grupos (por ejemplo, en animales y plantas).

Hoy en día, las filogenias moleculares son ampliamente aceptadas para representar las relaciones evolutivas entre los organismos. A mediados del 2000, el camino de la secuenciación se ha visto revolucionado, dado el desarrollo de técnicas de secuenciación de nueva generación (*next-generation sequencing*, NGS), la caída en los costos, la cantidad de datos de secuencia accesibles en bases de datos públicos y la secuenciación de organismos no modelos. Este desarrollo transformó el campo de la filogenética molecular dando lugar a la filogenómica, en la que se pueden utilizar datos de secuencia a escala genómica.

4.2. La nueva era de la filogenómica

La principal distinción entre filogenética y filogenómica es la escala de estudio. La filogenómica vincula la biología evolutiva y los estudios a nivel genómico (Chan y Ragan, 2013). Como este campo de estudio requiere nuevas formas para analizar datos a escala genómica, los métodos desarrollados para realizar análisis filogenéticos han ido mejorando y evolucionando para enfrentar estos desafíos. Las relaciones entre taxones se infieren en función de la homología a través de la comparación a nivel de un gen (heredado a partir de un ancestro común, y generalmente observada como patrones de similitud de secuencias)

10 Los resultados de este capítulo fueron publicados en: Amado Cattáneo, R.M., Diambra, L. y McCarthy, A.N. (2018) Phylogenomics of tomato chloroplasts using assembly and alignment-free method. *Mitochondrial DNA Part A*, 29:7, 1128-1138, DOI: 10.1080/24701394.2017.1419214

(Chan, Beiko, Darling, y Ragan, 2009; Puigbò, Wolf y Koonin, 2010), utilizando múltiples genes concatenados (Burki, Okamoto, Pombert y Keeling, 2012; Yutin, Puigbò, Koonin y Wolf, 2012), o a escala del genoma completo (Rannala y Yang, 2008). Los algoritmos utilizados para los análisis filogenéticos escalan pobremente cuando aumenta el número de secuencias y, por ello, los análisis filogenómicos de grandes cantidades de datos (miles de secuencias o genomas enteros) puede resultar computacionalmente inviable. Sumado a esto, las secuencias obtenidas a partir de tecnologías de SSG pueden ser incompletas y propensas a errores. El análisis también puede resultar complejo debido a la presencia de reordenamientos genómicos (fusión o delección) o transferencia horizontal de genes. Por lo tanto, las secuencias SSG utilizadas en estudios de filogenómica han requerido el desarrollo de nuevas metodologías como los enfoques libres de alineamiento (Chan y Ragan, 2013). A continuación, describiremos brevemente las diferentes metodologías de análisis evolutivo y sus limitaciones.

4.2.1. Alineamiento de secuencias múltiples y sus limitaciones

El alineamiento de secuencias múltiples (*Multiple sequence alignment*, MSA) ha sido durante mucho tiempo la metodología *sine qua non* en filogenética (Notredame, 2007). El objetivo del MSA es organizar las regiones de secuencia entre sí, de manera que representen la hipótesis de homología en todas y cada una de las posiciones. Incluso cuando estas posiciones hayan mantenido su contigüidad y orden relativo a través de la historia evolutiva, la reconstrucción de esta historia requiere suposiciones sobre el modelo de sustitución y la uniformidad del proceso entre sitios. Esta metodología utiliza algoritmos basados en la heurística que requieren gran uso de memoria (Wu, Chatterji y Eisen, 2012; Wong, Suchard y Huelsenbeck, 2008).

Los procesos de recombinación, duplicación, ganancia o pérdida, pueden producirse dentro de los genes y generan regiones que sólo se alinean de manera ambigua o de ninguna forma. La variación estructural local puede conducir a alineaciones con *gaps* que bajan la calidad de la resolución, generando un sesgo en la inferencia filogenética (Figura 4.1. A).

Dada la naturaleza heurística de los pasos clave en los flujos de trabajo filogenómicos estándar, la relevancia de las puntuaciones de alineamiento para la homología

puede ser difícil de evaluar estadísticamente (Mitrophanov y Borodovsky, 2006). Todos estos problemas se intensifican a escala de genoma completo, y pocos son resolubles con un mayor poder de cómputo o mejores modelos de sustitución.

4.2.2. Métodos libres de alineamiento

Los enfoques basados en subsecuencias definidas, típicamente cortas, conocidas como k -mers, o enfoques libres de alineamiento, ofrecen una alternativa interesante al MSA. En estos métodos los k -mers son extraídos y sus recuentos o distribuciones de frecuencia (histogramas) son calculados. Las regiones reguladoras de genes, elementos repetitivos o regiones transferidas lateralmente pueden exhibir patrones específicos de k -mer, mientras que los errores de secuenciación pueden distinguirse como k -mers inesperados (únicos) (Vinga y Almeida, 2003). Algunos estudios que han utilizado datos simulados sugieren que, al menos en determinadas circunstancias, los árboles basados en distancias calculadas a partir de las estadísticas k -mer pueden ser más precisos que los basados en MSA (Yang y Zhang, 2008). Al desacoplar el carácter de homología de la contigüidad de secuencia (alineamiento de secuencias múltiples), más allá de la longitud del k -mer, evitan simplemente la complejidad computacional de MSA y al mismo tiempo capturan la señal que de otro modo se perdería por el *gap*, la recombinación o la combinación aleatoria (Figura 4.1. B).

Al ampliar el campo de aplicación de los enfoques de k -mer a los conjuntos de datos que contienen *contigs* mal o no ensamblados, transcripciones superpuestas, fragmentos de genes o secuencias de baja calidad de secuenciación, existe una gran oportunidad de aprovechar la experiencia en campos menos familiares para los biólogos evolutivos, como son las metodologías de extracción y recuperación de información de texto (*text-minning*), (Krallinger y Valencia, 2005).

En un enfoque sin alineamiento la relación de dos secuencias se basa en el número y valor de las subsecuencias (k -mers) que comparten. Esta medida se puede transformar para estimar una distancia evolutiva. Una matriz de distancias por pares, construida sobre las estadísticas de un conjunto de secuencias, se puede utilizar para generar una estructura filogenética. La medida de la relación se puede basar en la frecuencia (número de

ocurrencias) de k -mers, o en las posiciones relativas dentro de las secuencias. Las correlaciones entre los conjuntos comunes de k -mer, a veces ajustados por el valor de k (longitud) en relación con la secuencia completa, se pueden incorporar como un factor de corrección. Esta medida puede alternativamente ser normalizada por la probabilidad a la que k -mers correspondientes ocurren en las secuencias, o extendidas para incluir partidas imperfectas de k -mer (Göke, Lasserre y Vingron, 2012).

Recientemente, Fan y cols. (2015) desarrollaron un método de libre ensamblado y alineamiento o AAF (*Assembly and Alignment-Free*) para reconstruir la filogenia. Este método en primer lugar calcula las distancias genéticas por pares, es decir entre dos muestras, de lecturas de secuencia corta. Esta distancia entre muestras o especies se basa en la estimación del parámetro de tasa de un proceso de *Poisson* para una mutación que ocurre en un solo nucleótido bajo el supuesto (modelo evolutivo) de que la tasa de mutación es la misma para todos los nucleótidos a través de los genomas. Esto no sólo incluye mutaciones causadas por sustituciones de nucleótidos, sino también inserciones y deleciones (*indels*). Las relaciones filogenéticas entre las muestras se reconstruyen a partir de la matriz de distancia por pares.

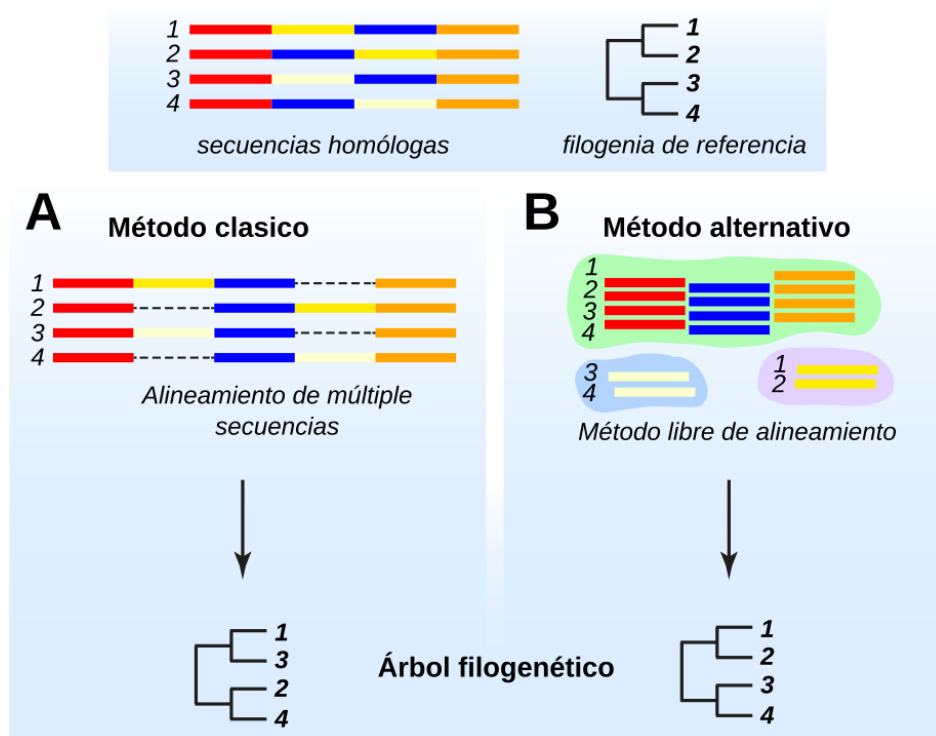


Figura 4.1: Flujo de trabajo simplificado de enfoques filogenómicos. (A) el enfoque clásico basado en el alineamiento de secuencias múltiples, y (B) un enfoque alternativo basado en métodos

sin alineamiento, para un ejemplo de análisis simple de secuencias homólogas 1, 2, 3 y 4, con una filogenia conocida como referencia (se muestra en la parte superior). Los fragmentos de secuencia que comparten la misma ascendencia en las cuatro secuencias (es decir, son muy similares entre sí) se indican con el mismo color (regiones roja, azul, amarilla y naranja en cada secuencia). En este ejemplo, las regiones amarilla y azul de las secuencias 2 y 4 se han reordenado en relación con 1 y 3. Las regiones amarillo oscuro (en 1 y 2) y amarillo claro (en 3 y 4) son similares entre sí.

Mientras que el enfoque clásico basado en el alineamiento de secuencias múltiples (espacios introducidos como líneas discontinuas) produce una filogenia inexacta, el enfoque alternativo sin alineamiento (agrupación de subsecuencias) no se ve afectado por el reordenamiento de secuencias en 2 y 4, y produce la correcta filogenia. La diferencia entre los dos árboles filogenéticos resultantes se resalta en rojo. Modificado de Chan y Ragan, 2013.

4.3. Origen de las secuencias

El análisis directo de datos genómicos no ensamblados, tanto de los tres orgánulos (cloroplasto, mitocondrias y núcleo) en conjunto o por separado, permite a la genómica poder comparar organismos sin una referencia completa disponible. Sin embargo, aún no está claro qué grado de profundidad y calidad de datos de secuenciamiento es necesaria para una reconstrucción filogenética confiable.

4.3.1. Regiones nucleares

Se han utilizado varias regiones nucleares en estudios filogenéticos de plantas, como los genes *waxy*, *leafy*, alcohol deshidrogenasa y fitocromo (Small, Cronn y Wendel, 2004b). Gracias a las plataformas bioinformáticas de transcriptómica y de secuenciamiento de genomas completos, se han ido identificando regiones genéticas que tienden a ser copias únicas en linajes divergentes y, por lo tanto, representan una fuente prometedora de marcadores (Duarte y cols., 2010; Wu, Mueller, Crouzillat, Petiard y Tanksley, 2006). A medida que se obtengan más datos a partir de consorcios de secuenciamiento masivo, como los proyectos 1KP¹¹ y 1000 Genomas de plantas y animales¹², mejorarán las perspectivas de recuperar datos de secuencias nucleares eficientemente. Sin embargo, hay importantes desafíos vinculados a la obtención de un conjunto común de marcadores nucleares que puedan amplificarse y secuenciarse fácilmente a partir de muestras filogenéticamente

11 1000 transcriptomas de plantas terrestres filogenéticamente divergentes. Ver en: www.onekp.com/index.html.

12 Ver en www.ldl.genomics.cn/page/proposingplant.jsp

divergentes. Entre ellos, las mutaciones en el sitio del cebador, duplicación de genes, recombinación, inserción de elementos transponibles, poliploidía y heterocigosidad. Se deben identificar regiones codificantes que tengan sitios variables adecuados para diferenciar a nivel de especie, o que contengan tamaños y posiciones de intrón suficientemente conservados para el diseño de marcadores (o cruce de intrón con exón).

Ha habido avances en este campo gracias a los consorcios antes mencionados, los cuales han logrado secuenciar masivamente múltiples loci nucleares, proporcionando resolución a nivel de especie para muchos grupos de plantas donde las clasificaciones son complejas y/o donde la especiación es reciente.

4.3.2. Genomas mitocondriales de plantas

Mientras que los genomas mitocondriales de animales están organizados de manera uniforme, los genomas mitocondriales de las plantas exhiben una gran diversidad en tamaño, estructura y contenido genético (Mower, Sloan, y Alverson, 2012; Liu y cols., 2012). Hasta la fecha se han secuenciado los genomas mitocondriales de los principales clados de plantas (Rodophytes, Chlorophyta, Charophyta, Anthocerotophyta, Hepaticophyta, Bryophyta, Filicopsida, Lycophyta, Gymnospermae, Angiospermae), mostrando un rango muy amplio en el tamaño del genoma, desde 13 Kb en Chlorophyta hasta 11.3 Mb en Angiospermae (Mower y cols., 2012; Sloan y cols., 2012). Los genomas mitocondriales de las plantas terrestres (embriófitas) poseen características que los hacen singulares, ya que son propensos a la recombinación, edición de ARN, empalme alternativo, inserción del ADN tanto del cloroplasto como del genoma nuclear y de taxones distantes, y la transferencia de genes hacia el núcleo (Knoop, 2012). A su vez, el material genético mitocondrial de las plantas terrestres es muy variable en tamaño, contenido y estructura. El tamaño del genoma de la mayoría supera los 200 Kb y, en general, menos del 20% codifica para proteínas o ARN. El contenido restante del genoma está dominado por la presencia de intrones de grupos I y II y por grandes regiones intergénicas. El número de intrones varía de 19 a 37 y su posición está relativamente conservada dentro de los clados (Mower y cols., 2012). Las regiones intergénicas incluyen elementos repetitivos, como integraciones del núcleo y cloroplasto de sus propios genomas y foráneos (Chaw y cols., 2008; Rice y cols., 2013). Sin embargo, el origen de la mayoría de las regiones intergénicas sigue siendo una incógnita (Sloan y cols.,

2012). Las plantas tienen un mayor número de genes codificantes en sus genomas mitocondriales que los animales. El número de genes identificados en la mayoría de las plantas varía entre 42 a 69, aunque algunas algas verdes (Chlorophyta) sólo tienen alrededor de 10 genes (Mower y cols., 2012; J. Fan y Lee, 2002). Entre ellos se encuentran los genes para los ARN ribosómicos, los ARNt, las proteínas ribosómicas, los genes *tatC* (una subunidad de la translocación de gemelos de arginina), *nad* (deshidrogenasa de NADH), *sdh* (succinato deshidrogenasa), *cob* (citocromo b), *cox* (citocromo c oxidasa), *ccm* (maduración del citocromo c) y las subunidades *atp* (ATP sintasa) (Knoop, 2012).

Curiosamente, aunque los genomas mitocondriales de las plantas muestran una alta variabilidad estructural, las tasas de sustitución en estos genes codificantes son bastante bajas en la mayoría de las especies (Christensen, 2013). Esto los hace menos apropiados como marcadores moleculares a nivel de población o para estudios de códigos de barras.

4.3.3. Origen y evolución de los plástidos

Los plástidos se originaron más tarde que las mitocondrias en la evolución de los eucariotas (algunos eucariotas tienen plástidos, mientras que otros no). La fotosíntesis anoxigénica, la cual genera la conversión de H₂O y CO₂ en azúcares ricos en energía y O₂, evolucionó en el linaje de las cianobacterias hace más de 3.500 millones de años, las cuales transformaron radicalmente la vida en la tierra debido al enriquecimiento de oxígeno en la atmósfera (Gould, Waller, y McFadden, 1999; Hohmann-Marriott y Blankenship, 2011). Los primeros plástidos provienen de la endosimbiosis de una cianobacteria con Archaeplastida, un clado que une Glaucophyta, Rhodophyta y Viridiplantae (Gray, 1999). Los eucariotas eran capaces de cooptar la fotosíntesis integrando un endosimbionte cianobacteriano. Curiosamente, aunque parece que los plástidos están estrechamente relacionados, los organismos que los contienen son de diversos clados eucariotas (Keeling, 2010). Esto puede explicarse por múltiples eventos endosimbióticos, conocidos como endosimbiosis secundaria. En este caso, un eucariota que llevaba un cloroplasto fue fagocitado por otro eucariota con una posterior integración del nuevo endosimbionte. Tales eventos se pueden distinguir de la endosimbiosis primaria por la morfología de los plástidos, ya que todavía portan membranas celulares adicionales derivadas de la fagocitosis. Los endosimbiontes primarios tienen dos membranas plástidas, mientras que los endosimbiontes

secundarios tienen cuatro membranas que pueden reducirse a tres, como en el caso de Euglenophyta y Dinoflagellata (Keeling, 2013). Además, se han documentado algunos casos de eventos de endosimbiosis terciaria para algunos linajes de Dinoflagellata. En el caso de endosimbiosis secundaria, el núcleo del alga generalmente se reduce por completo quedando sólo el cloroplasto. Sin embargo, en algunos casos queda un núcleo vestigial.

El papel principal del plástido es llevar a cabo la fotosíntesis, en cuyo caso se llaman cloroplastos. Sin embargo, como en Apicomplexa, algunos plástidos parecen haber perdido su capacidad fotosintética (Köhler y cols., 1997). Como es típico para endosimbiontes obligados, muchos genes innecesarios se perdieron mientras que otros fueron transferidos al núcleo del huésped. No obstante, los plástidos retienen una pequeña parte de su genoma ancestral, pudiendo deberse a que las proteínas hidrofóbicas son difíciles de transportar al orgánulo, o a que los orgánulos deben tener el control de la expresión de genes que forman parte de la cadena de transporte de electrones, como la regulación *redox* (Allen, 2015; Timmis, Ayliffe, Huang y Martin, 2004). A su vez, se han documentado algunos casos de transferencia de genes nucleares a los genomas de los plástidos (Keeling, 2009). Mientras que los plástidos primarios se encuentran en el citoplasma, los plástidos secundarios se encuentran dentro del sistema de endomembranas. Todos los genes necesarios para el funcionamiento del plástido que se encuentran codificados en el núcleo, tienen un sistema de selección para poder cruzar sus envolturas internas y externas (Strittmatter, Soll y Bölder, 2010). Alrededor del 40% del proteoma del plástido consiste en proteínas que parecen derivarse del genoma nuclear del huésped o varios linajes bacterianos fuera de las cianobacterias (Suzuki y Miyagishima, 2010).

Al igual que las mitocondrias, los genomas de los plástidos generalmente se organizan como moléculas circulares, y la molécula muestra una estructura cuatripartita debido a la presencia de dos grandes repeticiones invertidas que dividen a la molécula en una región de copia única grande y una pequeña (Figura 4.2). Por lo general, los cloroplastos codifican alrededor de 60 a 250 genes normalmente organizados como operones. Las repeticiones invertidas incluyen los genes de ARN ribosómico (16S, 23S y 5S ARNr) y algunos otros genes. El número de ARNt varía entre 27 y 31, y generalmente hay un número variable de genes de proteínas ribosómicas. Los genes codificantes son parte de los fotosistemas I y II, el complejo del citocromo b6f y la ATP sintasa (Green, 2011).

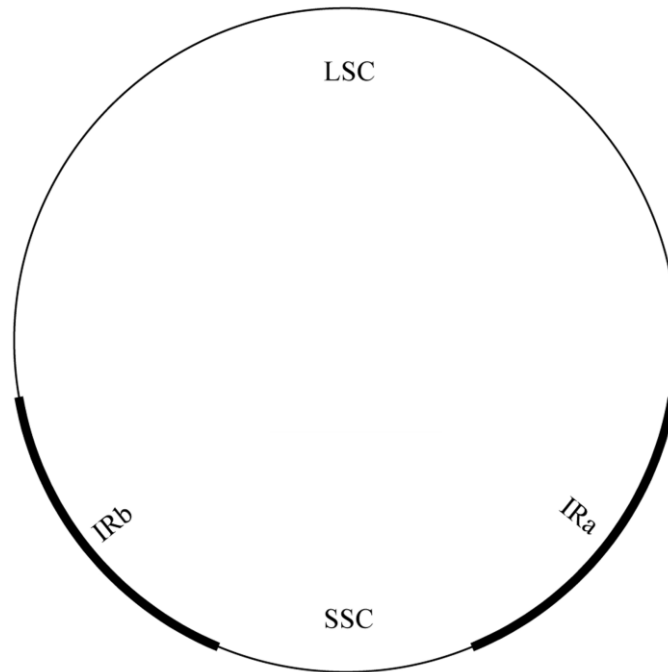


Figura 4.2: Estructura conservada en la mayoría de los genomas de cloroplastos. Poseen cuatro regiones: Dos regiones con secuencias repetitivas idénticas, pero con orientación invertida que se conocen como *IRa* e *IRb*, una región corta denominada SSC (*short single copy*) y una región más larga llamada LSC (*long single copy*).

4.3.4. Genoma de los plástidos

La composición y el tamaño de los plástidos pueden variar drásticamente entre los diferentes taxones. Los genomas más grandes de cloroplastos se encuentran en las algas verdes, con un genoma de aproximadamente 500 Kb en *Floydiella terrestris* (Brouard, Otis, Lemieux y Turmel, 2010). Este gran tamaño se debe principalmente a las regiones repetitivas (alrededor de 100) y no a la cantidad de genes, como pasa en la mayoría de las plantas terrestres. En contraste, los genomas de cloroplasto de algas rojas (Rhodophyta) contienen el mayor número de genes (220-250) de los cloroplastos secuenciados hasta el momento (Janouškovec y cols., 2013). Al parecer, las algas rojas muestran una tasa mucho más lenta de transferencia de genes del plástido al núcleo y, además, los genomas de estos plástidos carecen de repeticiones invertidas.

Por otra parte, los cloroplastos de la mayoría de los dinoflagelados fotosintéticos contienen el pigmento de captación de luz peridina y están rodeados por tres membranas, lo que sugiere una simbiosis secundaria, supuestamente de un huésped de algas rojas. La organización de estos cloroplastos es muy inusual y solo retienen un pequeño número de

genes (17) que se encuentran en diferentes minicírculos. A su vez, este taxón presenta otros tipos de cloroplastos derivados de diferentes eventos de simbiosis que, por lo tanto, difieren en tamaño y organización (Dorrell y Howe, 2015). La mitad de los linajes de dinoflagelados descritos perdieron sus plástidos (Green, 2011). Apicomplexa, el taxón hermano de los dinoflagelados, incluye eucariotas parásitos (incluyendo importantes patógenos de humanos o ganado como *Plasmodium*, *Toxoplasma* y *Eimeria*) que poseen un orgánulo que da origen a su nombre, apicoplasto. Éste derivaría de una endosimbiosis secundaria con un cloroplasto de un alga roja. Sin embargo, a diferencia de los cloroplastos, estos orgánulos no poseen actividad fotosintética, pero conservan su propio genoma y maquinaria de expresión, y están involucrados en la síntesis de ácidos grasos, isoprenoides, grupos hierro-azufre y grupo hemo (Lim y McFadden, 2010). En cloroplastos la transcripción puede estar mediada por dos tipos diferentes de ARN polimerasa o por una combinación de ambos. Según dónde está codificada la polimerasa correspondiente, estos tipos se abrevian como PEP (ARN polimerasa de plástido) o NEP (ARN polimerasa de plástido codificada en el núcleo) (Yagi y Shiina, 2014).

4.4. Código de barras de la vida

La identificación de especies basada en ADN mediante un “código de barras universal” es decir, regiones de secuencias cortas de ADN estandarizadas, se determinó por primera vez para animales (Hebert, Cywinska, Ball y deWaard, 2003) y luego para plantas, hongos y otros eucariotas (Hollingsworth y cols., 2009; Schoch y cols., 2012; Saunders y McDevit, 2012). Existen varias bases de datos (por ejemplo, *Barcode of Life Database*, BOLD) y consorcios internacionales (por ejemplo, *Consortium for the Barcode of Life*, CBOL), que promueven la estandarización de secuencias y la posibilidad de comparar la información almacenada.

El uso de marcadores moleculares para la identificación de especies se utiliza hace décadas para procariotas (Tindall, Rosselló-Móra, Busse, Ludwig y Kämpfer, 2010), de los cuales el más utilizado es el gen 16S rRNA.

En el caso de eucariotas, el código de barras de ADN utilizado con más éxito en animales es el gen citocromo oxidasa 1 (*cox1*) mitocondrial. Las ventajas de este marcador

mitocondrial incluyen: (I) los pares de cebadores son casi universales para el fragmento *coxI* disponible (Folmer, Hoeh, Lutz y Vrijenhoek, 1994); (II) el ADN mitocondrial está disponible en un alto número de copias por célula, facilitando así la extracción y amplificación de ADN; (III) se presupone la existencia de una brecha de código de barras, donde la variación interespecífica excede la variación intraespecífica (Hebert, Ratnasingham, y de Waard, 2003).

El ADN mitocondrial de las plantas evoluciona mucho más lentamente que su contraparte animal y, en consecuencia, se han utilizado diferentes regiones para la reconstrucción filogenética a diferentes niveles taxonómicos: alcohol deshidrogenasa y fitocromo (Small, Cronn, y Wendel, 2004a), secuencias de cloroplastos correspondientes a regiones codificantes (ej. *MatK*, *rbcL*, *rpoB* y *rpoCl*), espaciadores no codificantes (ej. *AtpF-atpH*, *trnH-psbA*, y *psbK-PSBI*) (Hollingsworth y cols., 2009), y espaciadores transcritos internos (ITS, *Internal transcribed spacer*) del ADN ribosomal nuclear (Li y cols., 2012).

En hongos, el código de barras depende de las regiones nucleares ITS, denominadas así ya que son transcritos internos que separan los genes del ARN ribosómico que se encuentran repetidos en tándem (Schoch y cols., 2012).

La dependencia de uno o pocos marcadores genera varias críticas al código de barras como “universal”, ya que centrarse únicamente en los marcadores de orgánulos puede generar errores en las relaciones filogenómicas por ejemplo, la introgresión de genes, por herencia solo materna, seudogenización o heteroplasma, entre otras (Galtier, Nabholz, Glemin, y Hurst, 2009). La presencia de endosimbiontes que manipulan la reproducción del huésped y, por lo tanto, altera la herencia del patrón de genes transmitidos por la madre, puede generar complicaciones adicionales (Gerth, Geißler, y Bleidorn, 2011).

La concatenación de muchos genes individuales se debe utilizar para mejorar la resolución del análisis filogenético. En la era genómica actual, las tecnologías de secuenciamiento NGS proporcionan una gran cantidad de datos genómicos que están fácilmente disponibles en las bases de datos públicas. Tales datos permiten el uso de enfoques filogenómicos para establecer relaciones evolutivas.

Este capítulo tiene dos objetivos. En primer lugar, validar el método AAF (*Assembly and Alignment-Free*) utilizando un estudio de caso conocido, en particular el trabajo de Wu (2015), con el fin de establecer los límites y las condiciones en que este método produce resultados confiables. En segundo lugar, aplicar este método para establecer las relaciones filogenómicas entre la mayor cantidad posible de cloroplastos de tomate, cuyas secuencias están actualmente disponibles en bases de datos genómicas públicas. Respecto del segundo objetivo, aplicamos el método AAF a lecturas de secuencia corta de un conjunto de más de 40 especies de tomates silvestres y cultivadas, aprovechando los 360 genomas secuenciados a partir del trabajo de Lin y cols. (2014), entre los cuales se encuentra una entrada de Tomate ‘Platense’. En lugar de utilizar datos de los tres orgánulos (cloroplasto, mitocondrias y núcleo), solamente utilizamos secuencias de cloroplasto. Las secuencias de ADN de cloroplasto (cp) son una herramienta útil para la identificación de plantas y la determinación de la relación filogenética entre especies (Kress y Erickson, 2008; Lahaye y cols., 2008). Se han utilizado diferentes loci de cloroplasto para calcular relaciones filogenéticas cercanas y distantes entre plantas, pero no se ha identificado un código de barras universal para todo el reino de las plantas (Hollingsworth, Graham y Little, 2011; Nock y cols., 2011).

4.5. Materiales y métodos

4.5.1. Conjunto de datos genómicos

El método AAF se utilizó con datos de secuencia de cloroplasto de dos tipos. En una primera instancia lo aplicamos a datos simulados de secuencias de a par (*pair end*, PE) tipo Illumina, para poder comparar este método con un análisis filogenético previo realizado con genomas completos de cloroplastos (Wu, 2015). En una segunda instancia, utilizamos el método AAF con datos reales PE (Illumina) de 45 entradas de tomates silvestres y cultivados (Cuadro 4.1).

Grupo informal	Variedad botánica	Nº de taxones
Lycopersicon	<i>S. lycopersicum</i>	24
	<i>S. cheesmaniae</i>	3
	<i>S. pimpinellifolium</i>	5
	<i>S. lycopersicum</i> var. <i>Cerasiforme</i>	6
	<i>S. galapagense</i>	1
Arcanum	<i>S. neorickii</i>	1
Eriopersicon	<i>S. habrochaites</i>	1
	<i>S. chilense</i>	1
	<i>S. peruvianum</i>	3
Neolycopersicon	<i>S. pennellii</i> *	1
Outgroup	<i>S. tuberosum</i> *	1
	<i>S. bulbocastanum</i> *	1

Cuadro 4.1: Grupos informales taxonómicos dentro de la sección *Lycopersicon*.

Generación de secuencias simuladas

Para generar las secuencias simuladas se descargaron 10 genomas completos de cloroplastos de tomate (números de acceso GenBank: KP117020-KP117027, NC_007898 y NC_024584) y dos genomas de cloroplastos de papa como grupo externo o *outgroup* (números de acceso GenBank: NC_007943 y NC_008096). Utilizamos el paquete *GemSIM* (McElroy, Luciani, y Thomas, 2012) para generar lecturas PE de 100 pb, alcanzando coberturas de 5X y 1000X para los genomas de cloroplastos descargados. Estas lecturas poseen un tamaño de inserto de 500 pb, con una desviación estándar de 60 pb, generado a partir de un modelo de error de secuencia estándar. Las secuencias simuladas se indican en el Cuadro 4.1 con un (*). Como ya se ha mencionado, este conjunto de datos simulados nos permitió comparar el método AFF con un análisis filogenético MSA de referencia (Wu, 2015).

Selección de secuencias del genoma de cloroplasto

Para el segundo objetivo, utilizamos lecturas PE (*Illumina Inc.*) de Lin y col. (2014). Estos datos están disponibles públicamente en la sección SRA (*Short Read Archive*) de la

base de datos del NCBI, y son el resultado del secuenciamiento de una sola ejecución (Illumina HiSeq 2000) de 360 entradas de tomates silvestres y cultivados (Lin y cols., 2014). De este conjunto, seleccionamos aquellas lecturas que presentaban la mayor puntuación general en cuanto a relación de cobertura y profundidad para cada variedad. Es así que seleccionamos un subgrupo de 45 tomates compuesto por 24 *S. lycopersicum*, 6 *S. lycopersicum* var. *cerasiforme*, 5 *S. pimpinellifolium*, completado con 3 *S. cheismaniae*, 3 *S. peruvianum*, 1 *S. chilense*, 1 *S. neorikii*, 1 *S. galapagense* y 1 *S. habrochiaties*. Los números de acceso para los tomates seleccionados están detallados en el Anexo III, donde también se incluyeron las especies de tomate utilizadas por Wu (2015). Al igual que en el Cuadro 4.1, el (*) indica que las lecturas PE reales no están disponibles y que se utilizaron datos PE simulados, mientras que (**) indica genomas de cloroplastos ensamblados por Wu (2015), y (***) indica números de acceso con código PI CGN de la Universidad Wageningen de Holanda.

Debido a que sólo aplicamos el método AAF a las secuencias de cloroplasto y no a la totalidad de las lecturas de cada planta, debimos realizar previamente un proceso de limpieza para quedarnos únicamente con las lecturas de interés. Para seleccionar estas lecturas mapeamos cada conjunto de datos contra el genoma completo de cloroplasto de *S. Lycopersicum* LA3023 (número de acceso: NC_007898) utilizando el software *Bowtie2* (Langmead y Salzberg, 2012). A partir de este mapeo, elegimos todas las lecturas PE que se alineaban al menos una vez con el genoma de cloroplasto de referencia, con un fragmento de PE máximo. Se utilizaron longitudes de alineamiento de 500 pb. La cobertura promedio de las secuencias de cloroplasto alineadas se muestra en el Anexo III. Por último, redujimos estas secuencias procesadas a la cobertura mínima de cloroplasto presente en las muestras seleccionadas (800X), para obtener datos comparables.

4.5.2. Topología del árbol

Empleamos *SPLITSTREE4* (Hudson y Bryant, 2006) para crear una super red filtrada (10% de todos los árboles generados) a partir de 10.000 árboles de entrada producidos por análisis de máxima parsimonia del programa AAF. La super red final que se muestra en la Figura 8 que utilizamos el filtro del 10% de 5.000 árboles no paramétricos generados por AAF.

4.5.3. Método libre de ensamblado y alineamiento (AAF)

El método AAF utilizado aquí cuenta todos los posibles k -mers para cada conjunto de datos genómicos. Un k -mer es una subsecuencia de nucleótidos (A, C, T y G) de longitud k . La cantidad de k -mers contados depende de la cobertura de secuenciamiento y de la distribución de las lecturas en el genoma. Esta tabla de frecuencias luego se convierte en una tabla de presencia/ausencia de k -mers entre los taxones estudiados. La distancia filogenética D entre dos especies se estima utilizando la métrica (Fan y cols., 2015):

$$D = -1/k \cdot (n_s/n_t)$$

donde n_s es el número de k -mers compartidos entre taxones, y n_t es el número total de k -mers (Fan y cols., 2015). El número de ocurrencias para cada k -mer dentro de las lecturas es n_r y el umbral θ indica el número de repeticiones y se puede configurar para eliminar la mayoría de los errores aleatorios en las lecturas. Cuando este filtro está activado, un k -mer sólo se registrará como presente si ocurre más de θ veces en la misma especie.

Antes de calcular distancias, se debe elegir un umbral de frecuencia adecuado (θ) y una longitud óptima de k -mers, que luego serán utilizados en el análisis estadístico. En esta metodología, si un k -mer contiene, por ejemplo, varias sustituciones dentro de su secuencia, contará igualmente como uno que lleva una sola sustitución. En consecuencia, los k -mers más cortos tienen más probabilidades de tener mayor sensibilidad a eventos evolutivos individuales. Por otra parte, los k -mers idénticos podrían derivar de diferentes regiones físicas, funcionales o evolutivas del genoma, y por lo tanto no serían homólogos (homoplasia). Los k -mers más largos son menos propensos a las homoplasias. Por lo tanto, la selección de la longitud de k -mer es un balance entre el problema de la sensibilidad (que requiere un menor k) y de homoplasia (que requiere una k mayor) (Fan y cols., 2015).

4.5.4. Longitud óptima de k -mer

Calculamos la distribución de frecuencia para ocurrencias de k -mer usando las secuencias de lectura simuladas de Illumina de 12 genomas de cloroplastos (ver sección Generación de secuencias simuladas) de baja y alta cobertura (5X y 1000X). En la Figura 4.3 mostramos la distribución de frecuencias para *S. lycopersicum* LA3023 como ejemplo

ilustrativo. Para baja cobertura, un k -mer corto (7 nucleótidos) es incapaz de diferenciar el primer pico correspondiente a los *singletons* (secuencias únicas), debido principalmente a errores de secuenciación, y el segundo pico de datos que corresponde a los datos útiles (Figura 4.3. A). Esta limitación se supera gradualmente a medida que la longitud del k -mer aumenta. A saber, aparecen dos picos distintos a partir de un k -mer de longitud 9. Aunque para una baja cobertura estos dos picos siempre se superponen, la altura y la posición del segundo pico se vuelven óptimos para $k = 25$. Para el caso de una cobertura alta (Figura 4.3. B), el primer pico (propenso a errores) está completamente separado del segundo para k -mers con longitud mayor que 9. El área bajo el segundo pico crece con la longitud del k -mer, alcanzando un valor óptimo para k -mer 25. Por lo tanto, la longitud de 25 se seleccionó como la óptima para los análisis posteriores.

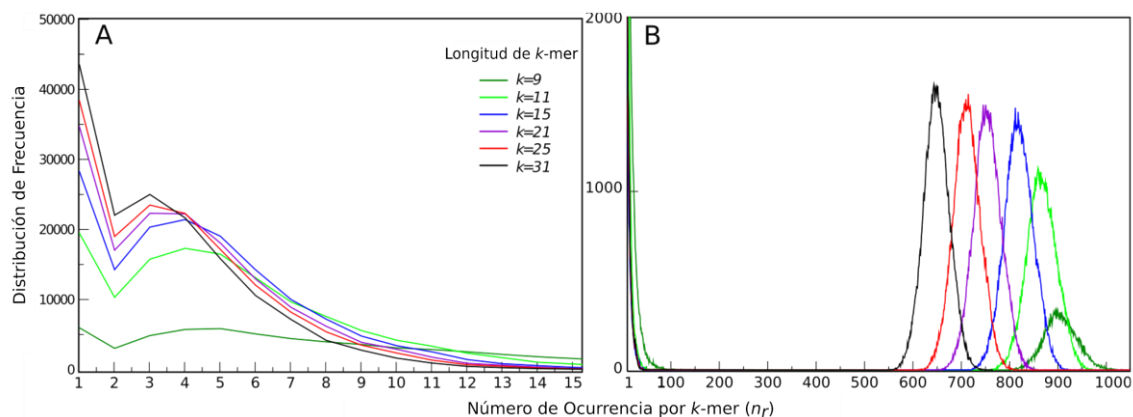


Figura 4.3: Distribución de frecuencia de *S. lycopersicum* LA3023 para diferentes k -mers y coberturas, (A) 5 X y (B) 1000 X.

4.5.5. Filtrado de k -mers de baja frecuencia

Los posibles errores introducidos por la falta de alineamiento generalmente se observan en la inferencia final de las relaciones evolutivas reales entre las especies. Además, la falta de montaje genera principalmente errores de muestreo causados por una baja cobertura del genoma completo y por errores de secuenciamiento (Song, 2013; Yi, 2013). Algunos estudios han propuesto filtrar todos los k -mers cuyas frecuencias están por debajo de un umbral dado llamado θ . Por ejemplo, al eliminar k -mers que presentan menos de tres copias ($\theta = 3$), estos pueden reducir el impacto de los errores de secuenciamiento (Fan,

2015). Sin embargo, cuando la cobertura de secuenciación disminuye, una fracción mayor de k -mers reales serán únicos (*singletons*) en el conjunto de datos y, por lo tanto, el filtrado eliminará k -mers reales. Como consecuencia, aunque el filtrado será beneficioso cuando hay una cobertura alta, a baja cobertura el filtrado será perjudicial.

Filtrar *singletons* puede corregir el error de secuenciamiento en muestras con baja cobertura (entre 5 X y 8 X) según el tamaño del genoma (Fan, 2015).

En la Figura 4.4 (A) se pueden observar dos picos con las secuencias simuladas a partir de los 12 genomas de cloroplasto generando una cobertura de 5X y utilizando una longitud de k -mer de 25. El primero, correspondiente a los *singletons* (donde se espera encontrar naturalmente los errores de secuenciación), y el de $n_r = 3$ que corresponde al segundo pico donde se espera principalmente contar con información sin errores de secuenciación del genoma. Por ello, en este caso, se estableció un valor umbral (θ) de 2. En la Figura 4.4. (B), donde la cobertura es de 1000X (también a partir de secuencias simuladas), se observan tres picos. El primero se extingue bien por debajo de $n_r = 100$ y representa los datos de secuencia propensos a errores. El segundo pico, correspondiente a los datos que existen como una sola copia dentro del ADN del cloroplasto (*singletons*), muestra un máximo alrededor de $n_r = 700$. El tercer pico corresponde a la zona de ADN de cloroplasto repetido e invertido (*IRA* e *IRB*), que muestra un máximo en valores superiores a $n_r = 1400$. Las condiciones de alta cobertura permiten diferenciar claramente las secuencias propensas a error del pico de datos con información. Por lo tanto, se espera que estas condiciones de cobertura ofrezcan resultados más sensibles.

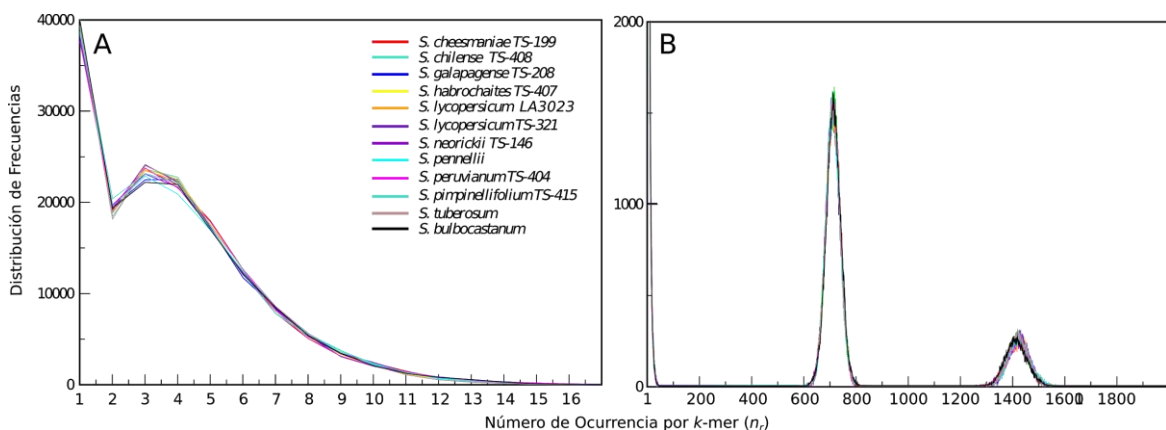


Figura 4.4: Distribución de frecuencia para diferentes secuencias simuladas de cloroplasto para k -mer 25 y coberturas de (A) 5X y (B) 1000X.

4.6. Resultados

Como se indicó anteriormente, antes del análisis filogenómico de las 45 accesiones de tomate, realizamos un estudio sobre un subconjunto de 10 tomates, cuyas secuencias del genoma del cloroplasto ya habían sido ensambladas y establecida su relación filogenética. Esto nos permitió optimizar los parámetros del método AAF y contrastar el resultado obtenido con el publicado previamente por Wu (2015), donde se utilizó un método de alineamiento del cloroplasto completo. Para el análisis y la interpretación de los resultados se utilizó una taxonomía informal en la que se aceptan cuatro grupos informales dentro de la sección *Lycopersicon* (Peralta y cols, 2008):

(i) grupo *Lycopersicon*, el clado de especies con frutos rojos y naranjas que incluye a *Solanum lycopersicum*, *S. cheesmaniae*, *S. galapense* y *S. pimpinellifolium*. Un taxón por debajo del nivel de especie, el tomate de frutos pequeños conocido como *Solanum lycopersicum* var. *cerasiforme*, se ha usado para referirse a formas supuestamente salvajes de *S. lycopersicum* que se consideran progenitoras del tomate cultivado. Es imposible distinguir las formas silvestres de las formas cultivadas o revertantes del cultivo o posiblemente los híbridos de taxones silvestres y malezas (Peralta y cols., 2008).

(ii) Grupo *Arcanum*, el clado de frutas verdes que incluye a *S. arcanum*, *S. chmielewskii* y *S. neorickii*.

(iii) Grupo *Eriopersicon*, que incluye a *S. huaylasense*, *S. chilense*, *S. corneliomulleri*, *S. peruvianum* y *S. habrochaites*.

(iv) Grupo *Neolycopersicon*, que sólo incluye a *S. pennellii* y se considera hermana del resto de la sección debido a la falta de la antera estéril, una sinapomorfia morfológica en *S. habrochaites* y el resto de los tomates centrales (Peralta y cols., 2008).

Estudios más recientes que utilizaron marcadores de secuencia ortóloga conservada (COSII) (Rodríguez, Wu, Ane, Tanksley y Spooner, 2009), polimorfismos de nucleótido único (SNP) de todo el genoma (Aitos et al., 2014; Lin et al., 2014) y elementos de repetición genómica (Dodsworth, Chase, Arkinen, Knapp y Leitch, 2016), en gran medida han apoyado las hipótesis anteriores con respecto a los clados principales dentro de los tomates, aunque las relaciones de especies individuales son menos claras para algunos taxones. Por lo tanto,

dada la aceptación general de esta clasificación informal, aquí la usaremos como referencia para interpretar nuestros resultados.

4.6.1. Filogenómica de lecturas reales y simuladas de 12 conjuntos de datos de secuencia de cloroplastos

Utilizamos el método AAF de dos maneras diferentes para calcular el cladograma y la super red de 10 tomates y 2 papas analizados previamente por Wu (2012). En el primer caso, simulamos datos de secuenciación *Illumina* a partir de las secuencias ensambladas de cloroplasto (cp) de las 12 muestras. En el segundo caso, utilizamos los 8 datos de secuenciamiento reales actualmente disponibles (Lin, 2014) y completamos para facilitar la comparación con 4 datos simulados.

Primer caso de estudio

Datos de secuenciación simulados de cloroplastos:

La Figura 4.5 muestra el árbol más parsimonioso de nuestro análisis con el método AAF a partir de los datos simulados de 12 cp, usando una longitud k -mer de 25 (ver Sección 4.4.4.). La figura 4.5. (A) es el resultado de datos con baja cobertura (5X) y el filtrado de *singletons* ($\theta = 2$) (Sección 4.4.5). Los resultados para el cálculo de alta cobertura (1000X) con $\theta = 550$ se muestran en la Figura 4.5 (C), y con $\theta = 100$ en la Figura 4.5 (E) (Sección 2.2.2).

En este árbol se recuperan dos miembros del grupo Eriopersicon dentro del mismo clado (*S. peruvianum* y *S. chilense*), con alto soporte (> 95%) en los tres cladogramas. Al igual que lo que observa Wu (2015), el tercer miembro, *S. habrochaites*, está separado de este grupo. El grupo Arcanum (*S. neorickii*) se recupera como hermano de los principales miembros del grupo Eriopersicon, con un soporte de más del 95%. El grupo Neolicopersicon, conformado solo por *S. pennellii*, se observa como hermano de los grupos Arcanum y Eriopersicon con un 95% de soporte

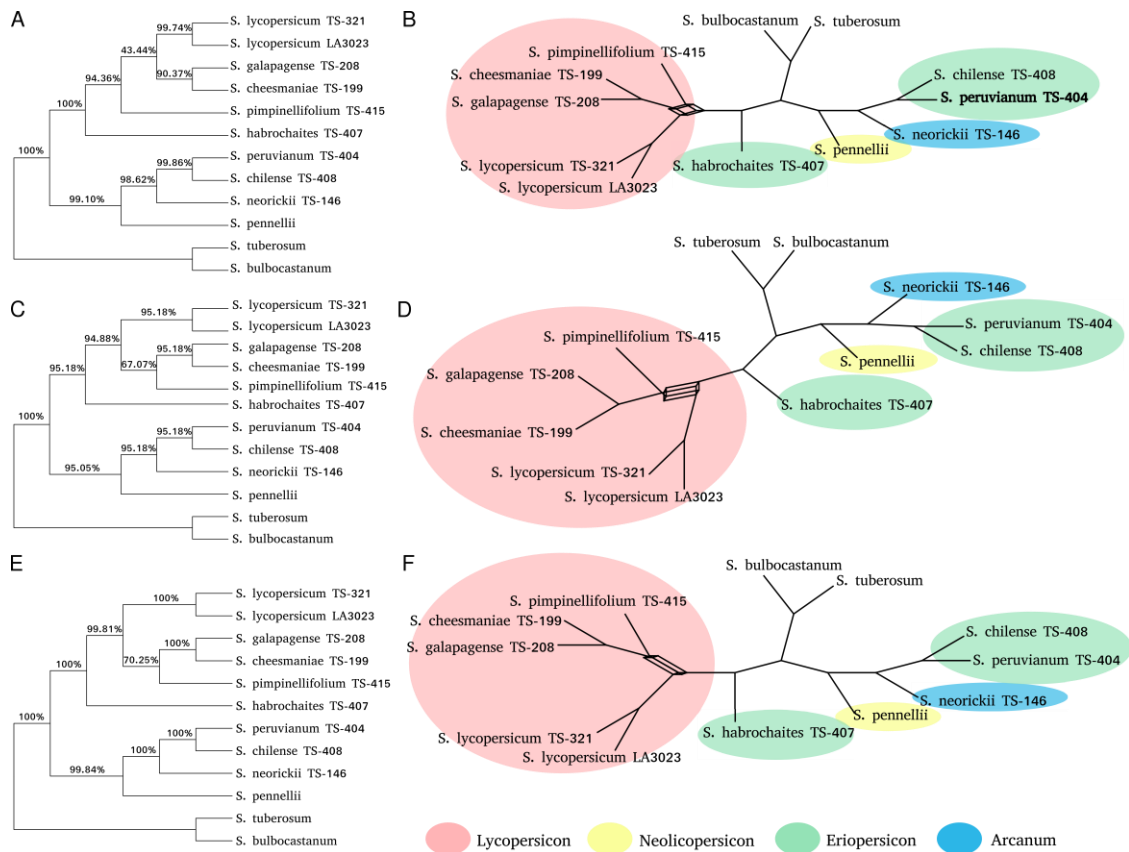


Figura 4.5: Primer caso de estudio de árboles filogenómicos y super-redes filtradas. (A) Árbol filogenómico de secuencias completas del genoma del cloroplasto simuladas con una cobertura de 5X, longitud k -mer de 25 ($k = 25$) y filtro para *singletons* ($\theta = 2$). (B) La red se muestra como una super-red filtrada de secuencias completas del genoma del cloroplasto simuladas con una cobertura de 5X, $k = 25$ y $\theta = 2$. Se muestra el 10% de todos los árboles calculados. (C) Árbol filogenómico de secuencias completas del genoma del cloroplasto simulado con una cobertura de 1000X, $k = 25$ y un filtro de $\theta = 550$. (D) La red se muestra como una super-red filtrada de secuencias completas del genoma del cloroplasto simuladas con una cobertura de 1000X, $k = 25$ y $\theta = 550$. Se muestra el 10% de los árboles calculados. (E) Árbol filogenómico de secuencias completas del genoma del cloroplasto simuladas con una cobertura de 1000X, $k = 25$ y $\theta = 100$. (F) La red se muestra como una super-red filtrada de secuencias completas del genoma del cloroplasto simulado con una cobertura de 1000 X, $k = 25$ y $\theta = 100$. Se muestra el 10% de todos los árboles calculados. Los porcentajes sobre las ramas de la los cladogramas indican los valores de soporte.

Mediante el método AAF se recuperó el clado de frutos rojos/anaranjados, grupo Lycopersicon, conformado por *S. lycopersicum*, *S. pimpinellifolium*, *S. galapagense* y *S. cheesmaniae*, con un soporte fuerte ($> 94\%$) para todos los árboles. En la Figura 4.4 (A), observamos a *S. pimpinellifolium* como hermana de este grupo, mientras que en la Figura 4.4 (C, E) *S. pimpinellifolium* aparece como hermana del grupo formado por *S. galapagense* y *S. cheesmaniae*. Este último resultado es la única diferencia entre los resultados de este análisis y los obtenidos por Wu (2015). Sin embargo, son concordantes con la filogenómica del cloroplasto publicada originalmente por Palmer y Zamir (1982) que, a diferencia de los

resultados de Wu, establece las mismas relaciones filogenómicas de *S. pimpinellifolium* que las observadas en la Figura 4.4 (C, E). Sin embargo, esta discrepancia entre Palmer y Zamir (1982) y Wu (2015) es menor y pueden tenerse en consideración cuando se analicen los resultados mostrados más adelante, por las super-redes correspondientes a las tres condiciones estudiadas.

Las tres super-redes de la Figura 4.4 (B), (D) y (F) muestran la misma topología general, con evidencia de un nodo común de reticulación en el clado Lycopersicon. Cuando se compara las diferentes condiciones observadas en la Figura 4.4 (D) y (F), resulta evidente que ambos estudios recuperan la misma topología de super-redes y, por lo tanto, para el caso de datos simulados, resulta equivalente establecer un corte inmediatamente después el primer pico ($\theta = 100$) o inmediatamente antes del segundo pico ($\theta = 550$). Esto puede ser fácilmente explicado por el hecho de que, para el caso de datos simulados, la distancia que separa el primer y el segundo pico no posee k -mers informativos o propensos a errores. Como conclusión de este primer caso de estudio, los resultados correspondientes a coberturas altas y bajas están en gran correspondencia con la filogenómica de cloroplasto de tomate obtenida por Wu (2015), y no observamos diferencias con respecto a las super-redes obtenidas utilizando diferentes valores de θ .

Segundo caso de estudio

Ocho datos de secuenciación de cloroplastos reales y cuatro simulados:

Cuando se utilizan secuencias reales comienzan a aparecer ciertas restricciones que se deben tener en cuenta para elegir adecuadamente el parámetro de filtrado (θ). La Figura 4.6 ilustra la distribución de frecuencia de 4 secuenciamientos simulados y 8 reales, para coberturas de 5X (Figura 4.6 A) y 1000X (Figura 4.6 B) y una longitud k -mer de 25. Los datos simulados en la Figura 4.6 (B) son los mismos que para el caso anterior, pero en cuanto a los datos reales se puede observar que el segundo y tercer picos (correspondientes a los datos con información) son más amplios y su máximo se desplaza hacia valores más bajos.

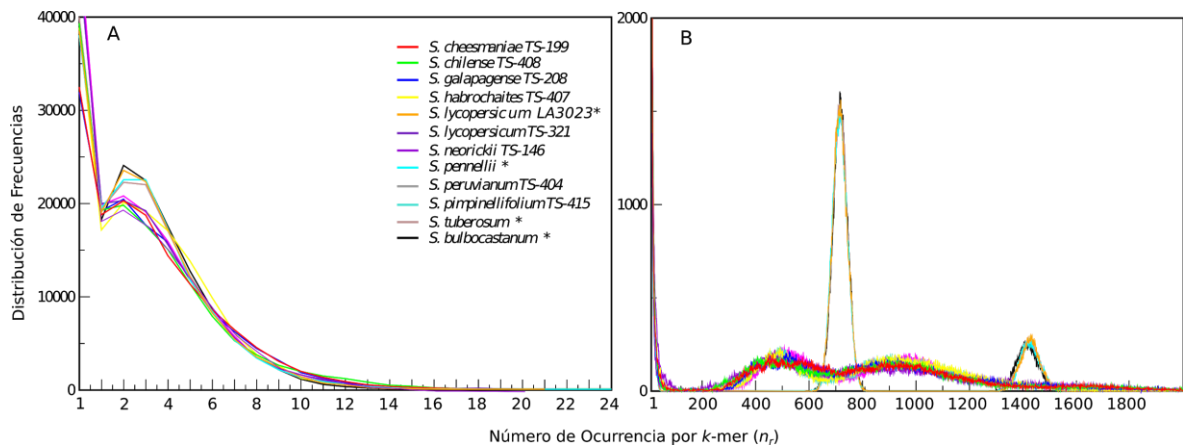


Figura 4.6: Distribución de frecuencia de datos de genomas simulados (*) y datos de secuencias reales, k -mer 25 y coberturas (A) 5X y (B) 1000X.

A modo comparativo, usamos las mismas tres condiciones de filtrado que para el primer caso de estudio, es decir, $\theta = 2$ para 5X de cobertura y $\theta = 550$ y $\theta = 100$ para el caso de cobertura 1000X. En la Figura 4.7 se muestran los árboles y super-redes calculados con las secuencias reales y simuladas. Se observa que se recuperó el Grupo Lycopersicon (clado frutal rojo-anaranjado), incluyendo a *S. lycopersicum*, *S. pimpinellifolium*, *S. galapagense* y *S. cheesmaniae*, con un 62% de soporte (Figura 4.7 (A)) y un soporte fuerte del 99,74% (E). En este grupo *S. pimpinellifolium* TS-415 era hermana de *S. lycopersicum* LA3023 y *S. lycopersicum* TS-321 en (A), mientras que *S. pimpinellifolium* TS-415 era hermana del grupo *S. galapagense* TS-208 y *S. cheesmaniae* TS-199 en (C). En el árbol que se muestra en la Figura 4.7 (C), el grupo Lycopersicon se anidó junto con *S. habrochaites* TS-407 del Grupo Neolicopersicon. El grupo Eriopersicom, con *S. peruvianum* TS-404 y *S. chilense* TS-408, se recuperó con un alto soporte de rama ($> 99,68\%$) en los tres cladogramas. *Solanum neorickii* TS-146 (grupo Arcanum) se recuperó como hermana del grupo Eriopersicon en todos los árboles, aunque con diferentes niveles de soporte. En (A) y (E) supera el 98,16%, mientras que en (C) presenta un valor modesto de 26,13%. Con respecto a las super-redes, (F) es la red más resuelta, a pesar de contener un nodo no resuelto en el grupo Lycopersicon, que también aparece en los tres casos anteriores. Por el contrario, (B) presenta dos nodos no resueltos, mientras que (D) presenta un caso claro de errores debido a la pérdida de datos útiles, debido a la incorrecta selección del valor de θ .

En conclusión, cuando utilizamos el método AAF con datos de secuencia reales de cloroplastos, observamos que se requiere una alta cobertura y datos viables de k -mer en el

rango de $k = 25$, así como el uso de un valor de corte adecuado para k -mers de baja frecuencia (propensos a errores). En nuestro análisis esto no tuvo una limitación real porque los datos genómicos utilizados cumplían con los requisitos mencionados.

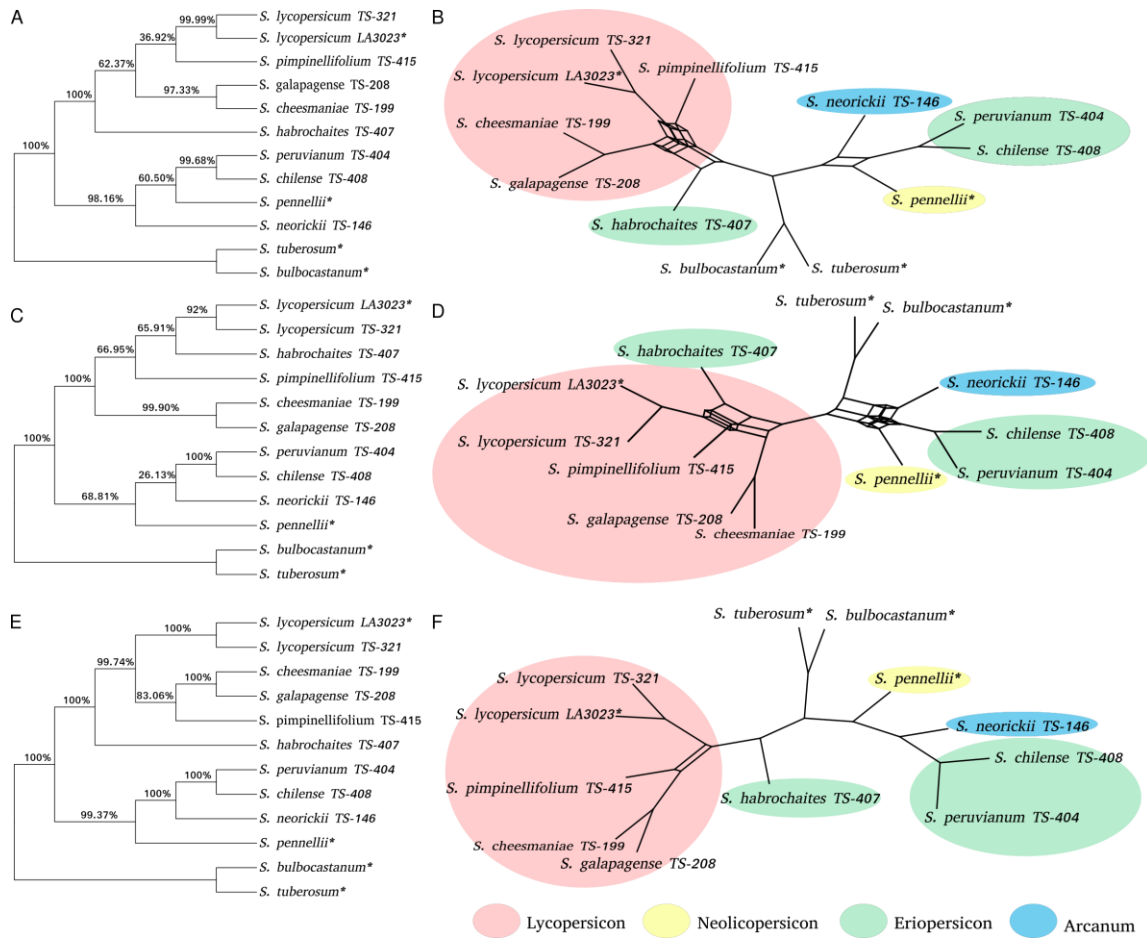


Figura 4.7: Segundo caso de estudio de árboles filogenómicos y super-redes filtradas. (A) Árbol filogenómico de los genomas completos de secuencias de cloroplasto Illumina y de secuencias simuladas, ambos con una cobertura de 5X, longitud k -mer de 25 y filtro para *singletons* ($\theta = 2$). (B) La red se muestra como una super-red filtrada de secuencias de genomas de cp Illumina reales y simuladas, con una cobertura de 5X, $k = 25$ y $\theta = 2$. Se muestra el 10% de los árboles calculados. (C) Árbol filogenómico de los genomas completos de secuencias de cloroplasto Illumina y de secuencias simuladas, con una cobertura de 1000X, $k = 25$ y filtro $\theta = 550$. (D) Red calculada como una super-red filtrada de los genomas completos de secuencias de cloroplasto Illumina y de secuencias simuladas, con una cobertura de 1000X, $k = 25$ y filtro de $\theta = 550$. Se representan el 10% de los árboles calculados. (E) Árbol filogenómico de los genomas completos de secuencias de cloroplasto Illumina y de secuencias simuladas, con una cobertura de 1000X, $k = 25$ y $\theta = 100$. (F) Red calculada como una super-red filtrada de los genomas completos de cloroplasto Illumina y secuencias simuladas con una cobertura de 1000X, $k = 25$ y $\theta = 100$. Se muestra el 10% de los árboles calculados. El número sobre las ramas de los cladogramas indica los valores de soporte.

4.6.2. Estudio filogenómico de 41 datos reales y 3 simulados de secuenciación de cloroplastos

Utilizamos el método AAF para calcular el cladograma y una super-red de 42 tomates y 2 papas como *outgroup*. Las secuencias reales fueron obtenidas por el consorcio internacional que secuenció más de 360 accesiones de tomates recientemente publicados, como se describen en la Sección 4.4.

Lo primero que analizamos fue la calidad de 45 secuencias reales mediante la distribución de frecuencias de los diferentes *k*-mer. En este sentido, descartamos los conjuntos de datos cuya distribución asociada mostraba que el pico de datos propensos a errores se superponía con los datos informativos. La figura 4.8 representa la distribución de frecuencia de cuatro conjuntos de datos (*S. cheesmaniae* TS-217, *S. lycopersicum* TS-237, *S. lycopersicum* TS-267 y *S. lycopersicum* var. *cerasiforme* TS-91) con datos propensos a errores, en comparación con los datos de secuencia informativos de *S. cheesmaniae* TS-207, por lo cual, siguiendo los criterios mencionados, fueron descartados del análisis posterior. En los accesos antes mencionados se encontraba el único que correspondía al tipo ‘Platense’ (TS-237), por lo que no pudimos incluirlo en el análisis filogenómico para ver sus posibles relaciones filogenéticas con los accesos del resto del mundo.

Finalmente, gracias a este criterio de selección nos quedamos con sólo 41 accesiones de las 45 previamente elegidas del consorcio de tomates (360 datos publicados recientemente por Lin y col., 2014) y sumamos un sólo secuenciamiento simulado de tomate silvestre (*S. pennellii*), que no fue secuenciado por el consorcio. Asimismo, se utilizaron dos secuenciamientos simulados de papas como grupo externo (*outgroup*).

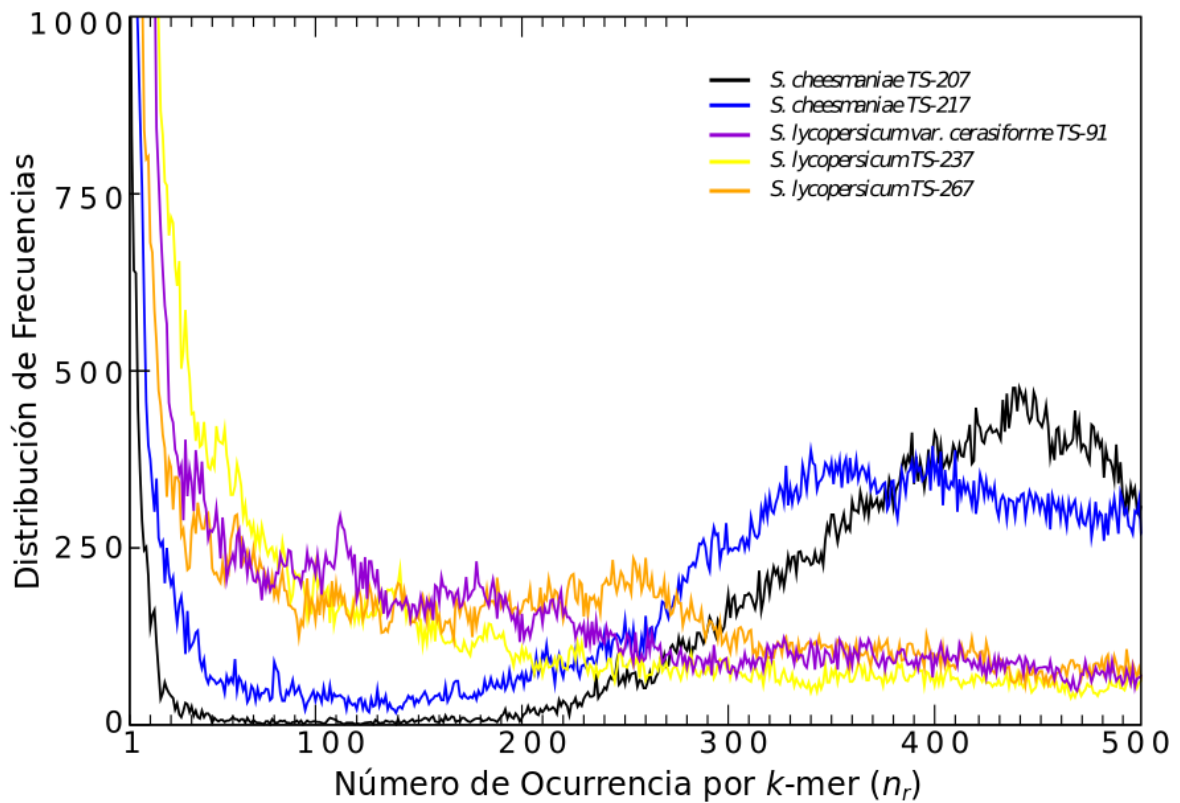


Figura 4.8: Distribución de frecuencias de cinco secuencias reales, que muestran la comparación de las curvas sin picos propenso a errores (*S. cheesmaniae* TS-207) y las distribuciones con picos no resueltos en la zona propensa a ruido (resto). Todos los datos corresponden a una cobertura de 800X y k -mer 25.

El árbol de cloroplasto más parsimonioso de este análisis (Figura 4.9) recuperó el grupo completo de *Lycopersicon* incluyendo *S. lycopersicum*, *S. pimpinellifolium*, *S. galapagense* y *S. cheesmaniae*, con un soporte alto (99,77%). En este grupo podemos distinguir tres subgrupos. El primero, con un soporte de arranque del 99,57%, y hermano de *S. pimpinellifolium* TS-432, compuesto por dos tercios de los 22 *S. lycopersicum* y todos menos uno de los *S. lycopersicum* var. *cerasiforme* analizados. El segundo subgrupo conformado por *S. pimpinellifolium* TS-433, el tercio restante de los *S. lycopersicum* analizados, y el último *S. lycopersicum* var. *cerasiforme*. El tercer subgrupo, con un 70,78% de soporte, recogió ambos *S. cheesmaniae* estudiados, *S. galapagense* y los dos *S. pimpinellifolium* restantes, TS-415 y TS-420.

Tanto *S. cheesmaniae* TS-207 como *S. cheesmaniae* TS-199 aparecieron juntos con un valor de soporte de 99,47%, y como hermanas de *S. galapagense* con un soporte del 100%. Los tres subgrupos están relacionados con al menos una de las cuatro variedades de *S. pimpinellifolium* estudiadas. Este último resultado está en correspondencia con la

variabilidad de las relaciones filogenéticas para *S. pimpinellifolium* mencionada en la literatura actual (Peralta y col., 2008), y en concordancia con los resultados filogenómicos de cloroplastos presentados tanto por Wu (2015) como por Palmer y Zamir (1982).

Con un 98,88% de soporte, los grupos Neolicopersicon (*S. pennellii*), Arcanum (*S. neorickii* TS-146, *S. peruvianum* TS-402) y Eriopersicon (*S. chilense* TS-408, *S. peruvianum* TS-404 y *S. peruvianum* TS-403) se agruparon juntos. Por otro lado, *S. habrochaites* TS-407 apareció como hermana del grupo Lycopersicon con un 100% de soporte. Esto está en completa correspondencia con Wu (2015), así como con nuestros casos de validación antes descritos. El grupo Eriopersicon, con *S. peruvianum* TS-404, *S. peruvianum* TS-403 y *S. chilense* TS-408, se recuperó con altos valores de soporte (100%). *Solanum neorickii* TS-146 y *S. peruvianum* TS-404 (grupo Arcanum) se recuperaron como hermanos del grupo Eriopersicon con 100% de apoyo estadístico. *Solanum peruvianum* puede dividirse entre variedades del norte y variedades del sur, de acuerdo con su capacidad de entrecruzamiento. Peralta y Spooner (2005) incluyeron a *S. peruvianum* norte dentro del grupo Arcanum y a *S. peruvianum* sur dentro del grupo Eriopersicon. Esto no es sorprendente dado que TS-404 y TS-403 corresponden a *S. peruvianum* variedad sur, mientras que TS-402 corresponde a la *S. peruvianum* variedad norte (Jablonska y cols., 2007).

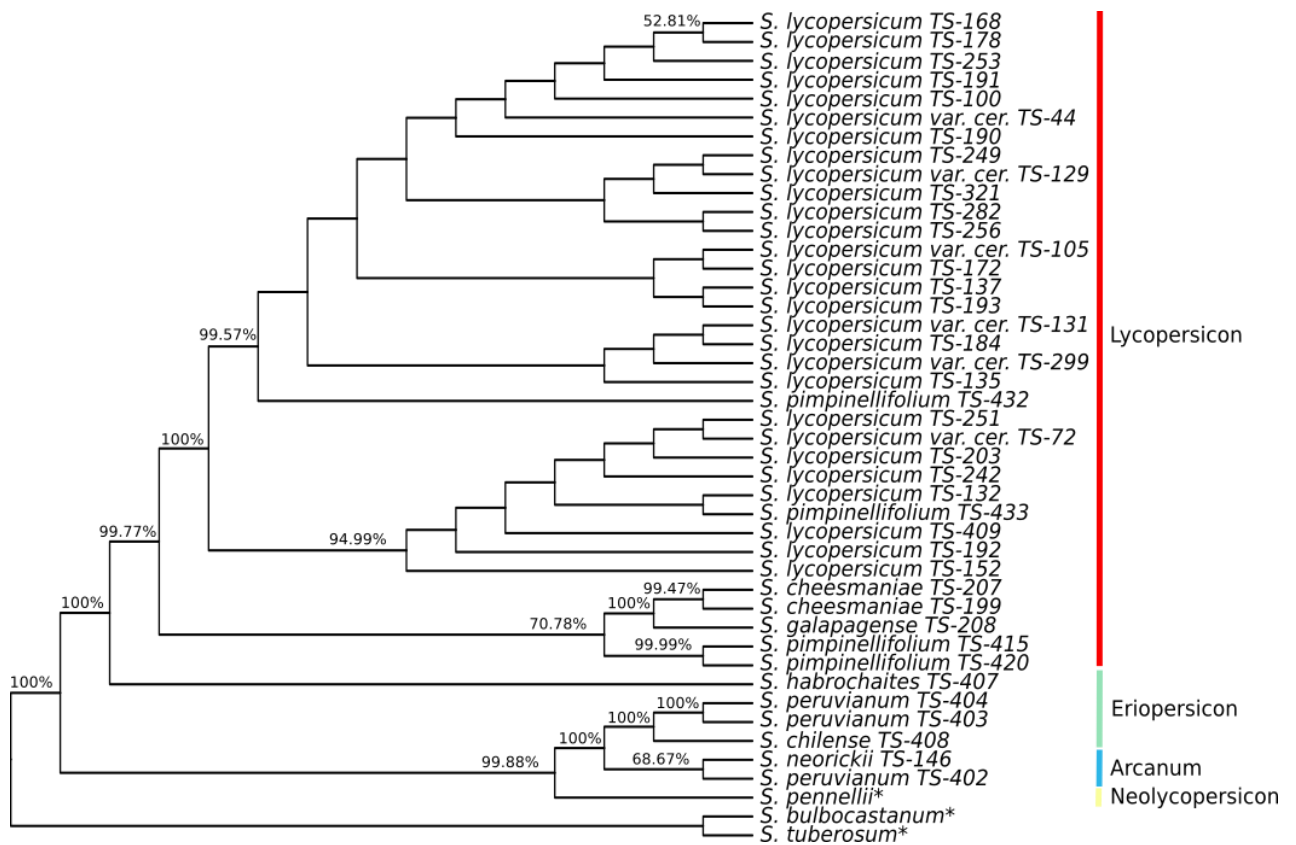


Figura 4.9: Relaciones filogenómicas entre 42 *Solanum*, sección *Lycopersicon*, calculadas con el método AAF. Se muestran valores de soporte superiores al 50% (no paramétricos con 10.000 muestras de cada tabla *k*-mer total).

En la super-red filtrada (Figura 4.10) los tres principales subgrupos de *Lycopersicon* están claramente separados. Los *S. pimpinellifolium* se encuentran distribuidos entre los tres subgrupos, y cuatro de ellos (*S. pimpinellifolium*) están estrechamente conectados con los nodos que separan estos tres subgrupos, indicando una relación cercana entre ellos. Además, cada punto de conexión también es un clúster de indeterminación (nodo no resuelto) en la super-red, que podría explicar razonablemente las relaciones taxonómicas de *S. pimpinellifolium* en la literatura actual (Peralta y col., 2008). Asimismo, la super-red muestra una clara separación entre todos los grupos considerados, es decir, los tres subgrupos de *Lycopersicon* y los grupos de *Neolycopersicon*, *Eriopersicon* y *Arcanum*. Finalmente, *S. habrochaites* aparece conectado a la super-red entre el tercer subgrupo de *Lycopersicon* y la papa (grupo externo), al igual que lo descrito por Wu (2015).

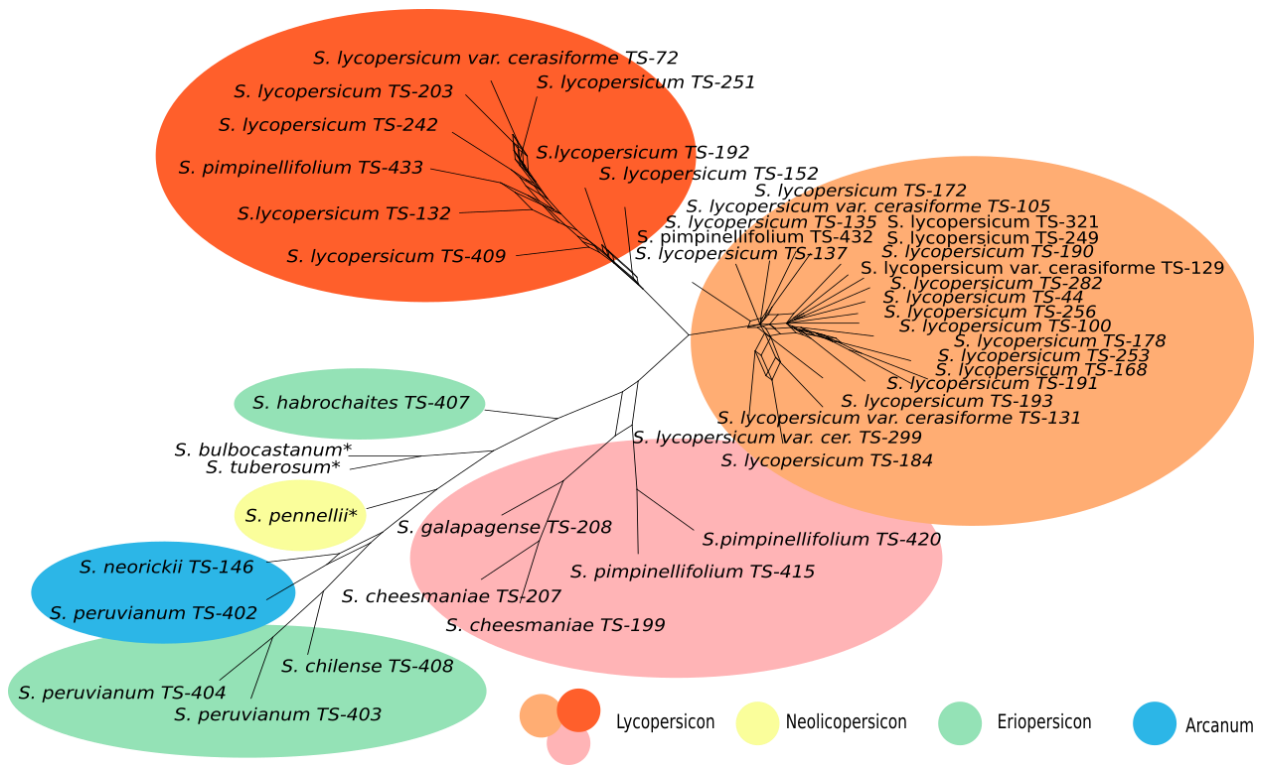


Figura 4.10: Relación entre los genomas de cloroplastos mostrados como una super-red filtrada. Se muestran las divisiones presentes en el 10% de 5.000 árboles con *bootstrap* no paramétricos. Cuatro grupos de conflictos aparecen en las super-redes. Los primeros tres corresponden a cada uno de los clados del subgrupo Lycopersicon, mientras que el cuarto nudo aparece entre los grupos Arcanum y Eriopersicon.

4.7. Conclusiones

En el presente capítulo analizamos las capacidades del método AAF para establecer relaciones filogenómicas confiables para cloroplastos de tomate. El método produce resultados precisos cuando se aplica a datos de secuencia ideales (es decir, datos simulados), tanto para condiciones de baja como de alta cobertura de secuencias en relación al tamaño del genoma. Sin embargo, al analizar datos de secuenciamientos reales, surgen ciertos problemas que deben tenerse en cuenta. Primero, las condiciones de alta cobertura producen mejores resultados que las de baja cobertura. Segundo, un cierto grado de curación de datos es necesario antes de aplicar el método AAF. A saber, los histogramas de distribución de frecuencia de los *k*-mer deben verificarse previamente para examinar la resolución completa entre el primero y el segundo pico (es decir, datos propensos a errores y datos con información). Finalmente, debe establecerse correctamente un valor de corte óptimo de θ ,

común a todos los conjuntos de datos, para descartar datos propensos a errores, sin que conlleve a la pérdida de datos con información. A partir de este estudio pudimos determinar que el método AAF es capaz de establecer correctamente las relaciones filogenómicas del cloroplasto de tomate, abriendo así la posibilidad de realizar más estudios filogenómicos con datos completos de secuenciación genómica.

En las condiciones previamente establecidas, estudiamos las relaciones filogenéticas para 42 cloroplastos de tomate, usando 2 cloroplastos de papa como grupo externo. En este capítulo se obtuvo una estructura de árbol filogenético general compatible con los datos establecidos por estudios previos. A saber, que se agrupan los miembros de los cuatro grupos informales, manteniendo las relaciones esperadas entre sí. Sin embargo, también surgieron ciertas observaciones adicionales, como el hecho de que el grupo *Lycopersicon* aparece en tres subgrupos distintos, dos de los cuales representan todos los *S. lycopersicum* y *S. lycopersicum* var. *cerasiforme*, mientras que el tercer subgrupo está compuesto por *S. galapagense* y ambos *S. cheesmaniae* estudiados, como se esperaría de acuerdo con los datos previamente publicados. Además, los cuatro *S. pimpinellifolium* estudiados aparecen dispersos en estos tres subgrupos de *Lycopersicon*, lo que puede explicar parcialmente la falta de consenso en estudios anteriores sobre su precisa relación filogenética dentro del grupo informal. Sin embargo, es importante tener en cuenta que, para preservar la consistencia metodológica, el presente estudio se restringió a los taxones analizados por el consorcio de secuenciación de 360 genomas, y únicamente se utilizaron datos de la sección SRA del NCBI. Podría esperarse que en futuros estudios utilizando este método ciertas relaciones filogenéticas cambien a medida que se agreguen más taxones.

Resultó interesante observar que *S. peruvianum* TS- 403 se encontraba dentro del grupo informal *Eriopersicon*, mientras que *S. peruvianum* TS-402 estaba dentro del grupo informal *Arcanum*. Esto puede explicarse fácilmente considerando que el primero corresponde a la variedad *S. peruvianum* del norte, mientras que el segundo corresponde a la variedad *S. peruvianum* del sur (Jablonska y cols, 2007). En síntesis, creemos que el presente estudio ha demostrado la capacidad de realizar estudios filogenómicos confiables sin la necesidad de ensamblar o alinear secuencias, lo cual no sólo es una gran ventaja sino una verdadera necesidad considerando la gran cantidad de secuencias producidas en todo el mundo.

Conclusiones y perspectivas futuras

Durante el desarrollo de esta tesis nos propusimos como objetivo general analizar la diversidad genética de las variedades locales de tomate (*Solanum lycopersicum*) conocidas como Tomate ‘Platense’ cultivadas en el Cinturón Hortícola Platense. Pudimos generar un primer relevamiento de forma sistemática de esta variedad local, tanto de los materiales disponibles en el Banco de Germoplasma del INTA como en las unidades productivas que hoy en día los siguen conservando *in situ*. En el desarrollo de la tesis nos planteamos relevar y registrar el patrón genético del Tomate ‘Platense’ y lograr la identificación de las diferentes subpoblaciones locales para poder no solo dar identidad sino mejorar las estrategias para la conservación de su germoplasma. Por último, nos propusimos establecer por medio de la bioinformática las relaciones filogenéticas de las diferentes entradas de tomate, en particular la variedad Tomate ‘Platense’, y de especies silvestres, mediante el análisis de las secuencias genómicas de sus cloroplastos, disponibles en bases de datos públicas. El desarrollo de estos objetivos específicos se plasmó en cuatro capítulos que se sintetizan a continuación.

En el primer capítulo, se realizó una descripción detallada de los antecedentes disponibles del Tomate ‘Platense’, desde sus posibles centros de diversificación, hasta el rescate cultural que hoy en día se ha logrado, en gran medida gracias a la reconocida “Fiesta del Tomate Platense”, la cual ha cumplido ya su vigésima versión anual. Paralelamente, se hizo también una revisión de cómo fue cambiando el concepto general de variedad local (*o landrace*) a lo largo de la historia y su complejidad en contraste con lo que se denominan cultivares. Estos conceptos de variedad local y cultivar son muchas veces utilizados de forma indistinta o de forma poca clara, como sucede también con los términos de semillas nativas o criollas. Es importante debatirlos y definirlos con precisión, ya que los mismos son utilizados para la generación de políticas públicas para la recuperación, producción y valorización de estos materiales y si su terminología no es clara, ello puede comprometer dar respuestas reales a diferentes problemáticas. Por último, a partir de experiencias locales de extensión, se planteó una ampliación del programa de mejoramiento participativo de variedades locales propuesto por Casañas (2017), situando la presente tesis en contacto estrecho con prácticas en territorio. Si bien en este primer capítulo se hace un recorrido histórico-conceptual que nos permite comprender la complejidad de las variedades locales

los resultados obtenidos en el desarrollo en esta tesis, dan evidencia para sostener que el tomate 'Platense' es efectivamente una *landrace*.

En el segundo capítulo, mostramos la evaluación morfoagronómica de las subpoblaciones estudiadas y su vida poscosecha. Como fue descrito durante el desarrollo de esta tesis, para lograr obtener nuevos cultivares es necesario contar con una base genética amplia y desarrollar de manera sistemática evaluaciones agronómicas de las colecciones de germoplasma, tanto de los conservados en los bancos como de las variedades locales que se encuentran, hoy en día, en las unidades productivas de la Agricultura Familiar. A pesar de que los materiales estudiados poseen desventajas comparativas a los híbridos comerciales que los fueron sustituyendo en el mercado, actualmente los consumidores reconocen la falta de calidad organoléptica de los frutos producidos por estos últimos, posibilitando que el mejoramiento genético comience a tener en cuenta la revalorización de las fuentes donde se encuentra la diversidad fenotípica, genómica y cultural de estas variedades locales, muchas veces subestimadas por el actual modelo productivo. La diversidad morfoagronómica que pudimos observar tanto en las plantas como en la calidad de poscosecha de sus frutos en los materiales estudiados de Tomate 'Platense', lo presentan sin dudas como un gran candidato para eventuales programas de mejoramiento en el sentido de lo argumentado.

En el tercer capítulo, realizamos la caracterización de la diversidad genética de las entradas del Tomate 'Platense' almacenadas en el Banco de Germoplasma INTA- La consulta (Mendoza, Argentina). Se considera que el tomate cultivado presenta una estrecha base genética. Sin embargo, se comienza a tener cada vez más evidencia de que las variedades locales de tomate contienen valores de diversidad genética y fenotípica superiores a los cultivares comerciales e híbridos (Park y cols., 2004; Mazzucato y cols., 2008). La presente tesis ha determinado la variedad local de tomate 'Platense' como un recurso genético altamente diverso. Así, los resultados obtenidos, pueden ser utilizados como referencia tanto para estudios futuros sobre muestras de nuevas subpoblaciones autóctonas que se encuentran actualmente en uso por los horticultores locales, semillas comerciales o accesiones 'Platense' actualmente disponibles en otros bancos de semillas. Esta alternativa puede resultar de relevancia dado que, durante los últimos años, la producción y comercialización regional de esta variedad local ha sido revalorizada socioculturalmente tanto por su calidad organoléptica como por su adaptabilidad a crecer en condiciones agroecológicas (Garat y Otero, 2012). Este patrimonio hortícola, combinado con los resultados aquí presentados, abre la posibilidad de un novedoso esquema de mejora

participativa, integrando horticultores, consumidores e investigadores locales bajo un objetivo común: la consolidación del 'Platense' como *landrace*. Al mismo tiempo, estos resultados reafirman el papel de la conservación *in situ* de los horticultores locales familiares en la preservación de los recursos genéticos del tomate. Estos resultados también tienen la capacidad de identificar a las variedades locales y distinguirlas de los cultivares comerciales, pudiendo facilitar las gestiones de la Denominación de Origen para su revalorización cultural y comercial (Caramante y cols., 2009; Corrado y cols., 2014; Mazzucato y cols., 2010; Castellana y cols., 2020). En este mismo sentido, se pueden rastrear cultivares de tomate en cadenas alimenticias para control del producto (Caramante y cols., 2011).

Por último, en el cuarto capítulo, hicimos un análisis filogenómico del cloroplasto del tomate a partir de secuencias disponibles en bases de datos públicas. A partir de realizar un análisis metodológico de las capacidades del método AAF se pudo establecer relaciones filogenómicas confiables para cloroplastos de tomate. El método produce resultados precisos cuando se aplica a datos de secuencia ideales (es decir, datos simulados), tanto para condiciones de baja como de alta cobertura de secuencias en relación al tamaño del genoma. Sin embargo, al analizar datos de secuenciamientos reales, surgen ciertos problemas que se deben tener en cuenta. A partir de este estudio pudimos determinar que el método AAF es capaz de establecer correctamente las relaciones filogenómicas del cloroplasto de tomate, abriendo así la posibilidad de realizar más estudios filogenómicos con datos completos de secuenciación genómica. Se obtuvo una estructura de árbol filogenético general compatible con los datos establecidos por estudios previos en la bibliografía. Esta tesis ha demostrado la capacidad de realizar estudios filogenómicos confiables sin la necesidad de ensamblar o alinear secuencias, lo cual no solo es una gran ventaja sino una verdadera necesidad considerando la gran cantidad de secuencias producidas en todo el mundo.

En síntesis, esta tesis logra mediante técnicas de la Biotecnología vegetal y Bioinformática, estudiar la diversidad genética local, su conservación y propone un mejoramiento participativo.

Finalmente, a continuación, y a partir de los resultados obtenidos, describimos las perspectivas futuras y las diferentes líneas temáticas que surgieron, las cuales podrán ser retomadas en nuevas investigaciones.

En primer lugar, como dijimos previamente, es importante tener documentos académicos que propongan y justifiquen conceptos claros sobre qué denominamos

variedades locales, semillas nativas, criollas, entre otros, para poder implementar políticas públicas con resultados concretos que revaloricen cultivos y permitan su conservación tanto *in situ* como *ex situ*.

En segundo lugar, es preciso generar el trámite de la Denominación de Origen del Tomate 'Platense' para dar valor agregado a la producción local y a su cadena agroindustrial.

Queda también como desafíos futuros poder generar mayor cantidad de datos de fenotipado y de calidad tanto de vida de poscosecha como sensoriales. Este tipo de trabajo sistemático solo se puede alcanzar logrando programas provinciales o nacionales que vuelvan a tener como metas la recuperación y conservación de los recursos genéticos propios.

Estos resultados y metodología de marcadores moleculares permiten seguir sumando más muestras de germoplasma 'Platense' o tomates en general, de origen diverso (nuevas subpoblaciones autóctonas, comerciales o nuevos materiales mejorados).

En el sentido a lo anterior, seguimos aportando evidencia de las ventajas que poseen los marcadores moleculares (en este caso los SSR) para el estudio de variedades. Esta clase de trabajo se puede utilizar en otras plantas hortícolas, medicinales, ornamentales, etc., para seguir aportando información no solo de identidad sino también para repensar programas de mejoramiento regionales. Por otro lado, queda claro que los costos en secuenciamiento genómico masivo son cada vez más accesibles a proyectos de investigación. Como fue descrito, la masividad de datos disponibles y que se pueden generar, necesitan metodologías novedosas de análisis y sin dudas nuestro país posee las capacidades para generar plataformas tecnológicas que las generen y den resguardo a nuestros recursos genéticos (tanto vegetal, animal o de microorganismos).

En suma, la Agrobiotecnología, pensada en contextos de programas públicos y participativos, es una herramienta potente para la generación de posibles respuestas a las demandas locales y para el desarrollo nacional.

Bibliografía

- Adachi, J., Waddell, P., Martin, W. y Hasegawa, M. (2000). **Plastid genome phylogeny and a model of amino acid substitution for proteins encoded by chloroplast DNA.** *Journal of Molecular Evolution*, 50: 348 -358.
- Aflitos, S., Schijlen, E., De Jong, H., De Ridder, D., Smit, S., Finkers, R. y Peters, S. (2014). **Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole - genome sequencing.** *Plant Journal*, 80 (1), 136-148.
- Ahumada, J., Otero, J. y Garat, J. J. (2011). **Las hortalizas típicas del Cinturón Verde de la Plata. Facultad de Ciencias Agrarias y Forestales - UNLP.**
- Allen, J. (2015). **Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocation for redox regulation of gene expression.** *Proc Natl Acad Sci USA*, 112:10231–10238.
- Alvarez, A. E., Van de Wiel, C. C. M., Smulders, M. J. M. y Vosman, B. (2001). **Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*.** *Theoretical and Applied Genetics*, 103 (8), 1283-1292.
- Ardila, G., Fischer, G. y López, H. E. B. (2011). **Caracterización del crecimiento del fruto y producción de tres híbridos de tomate (*Solanum lycopersicum* L.) en tiempo fisiológico bajo invernadero.** *Revista Colombiana de Ciencias Hortícolas*, 5 (1), 44-56.
- Arens, P., Bredemeijer, G., Smulders, M. J. M. y Vosman, B. (1995). **Identification of tomato cultivars using microsatellites.** *In I International Symposium on Solanacea for Fresh Market*, 412 (pp. 49-57).
- Areshchenkova, T. (2000). **Isolation, characterization and mapping of microsatellites from the tomato genome and their application in molecular analysis of centromeric regions.** (Doctoral dissertation, University of Sachsen-Anhalt, Germany).

Areshchenkova, T. y Ganal, M. W. (2002). **Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources.** *Theoretical and Applied Genetics*, 104 (2-3), 229-235.

Argerich, C., Troilo, L., Rodríguez Fazzone, M., Izquierdo, J., Strassera, M. E., Balcaza, L., Dal Santo, S., Miranda, O., Rivero, M. L., González Castro, G. y Iribarren, M. J. (2010). **Manual de buenas prácticas agrícolas en la cadena de tomate.** *Organización de las Naciones Unidas para la Agricultura y la Alimentación - FAO*. Buenos Aires, Argentina. 264p.

Bailey, L. H., Tracy, W. W., Kyle, E. J. y Watts, R. L. (1960). **Tomato. The standard cyclopedia of horticulture.** The Macmillan Company, New York.

Balsa, J. (2012). **Agricultura familiar: caracterización, defensa y viabilidad.** En Revista *Interdisciplinaria de Estudios Agrarios n° 3*.

Becerra, V. V. y Paredes, C. (2000). **Uso de marcadores bioquímicos y moleculares en estudios de la diversidad genética.** *Agricultura Técnica*, 60 (3), 270-281.

Benbouza, H., Jacquemin, J. M., Baudoin, J. P. y Mergeai, G. (2006). **Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels.** *Biotechnologie, agronomie, société et environnement*, 10 (2), 77-81.

Benencia, R., C. Cattaneo y R. Fernández (1997). **Proceso histórico de la conformación del área hortícola.**, in Área Hortícola Bonaerense. Cambios en la producción y su incidencia en los Sectores Sociales. *La Colmena: Buenos Aires*. p. 33-42.

Bennetzen, J. (2000). **Transposable element contributions to plant genome evolution.** *Plant Mol Biol*, 42:251–269.

Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Diez, M.J. y Nuez, F. (2013) **Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato.** *PLoS ONE* 7(10): e48198

Boggio, S. B., Palatnik, J. F., Heldt, H. W. y Valle, E. M. (2000). **Changes in amino acid composition and nitrogen metabolizing enzymes in ripening fruits of *Lycopersicon esculentum* Mill.** *Plant Science*, 159(1), 125-133.

Bohs, L. (2005). **Major clades in *Solanum* based on *ndhF* sequence data.** *Monographs in Systematic Botany*, 104, 27.

Bohs, L. y Olmstead, R. G. (1997). **Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndhF* sequences.** *Systematic Botany*, 5-17.

Bonicatto, M. M. (2018). **Sustentabilidad y agrobiodiversidad.** (Tesis Doctoral, Universidad Nacional de La Plata).

Brandt, S., Z. Pék, Barna, É., Lugasi, A. y Helyes, L. (2006). **Lycopene content and colour of ripening tomatoes as affected by environmental conditions.** *J. Sci. Food Agr.* 86:568–572.

Brewer, M. T., Lang, L., Fujimura, K., Dujmovic, G. S. N. y van der Knaap, E. (2006). **Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species.** *Plant physiology*, 141 (1), 15-25.

Brouard, J.-S., Otis, C., Lemieux, C. y Turmel, M. (2010). **The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the Chlorophyceae.** *Genome Biol Evol*, 2:240–256.

Burki, F., Okamoto, N., Pombert, J. y Keeling, P. (2012). **The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins.** *Proc. R. Soc. B.*, 279:2246–2254.

Candolle, A. (1886). ***Origin of cultivated plants.*** D. Appleton.

Caramante, M., Rao, R., Monti, L. M., & Corrado, G. (2009). **Discrimination of ‘San Marzano’ accessions: a comparison of minisatellite, CAPS and SSR markers in relation to morphological traits.** *Scientia horticultrae*, 120(4), 560-564.

Caramante, M., Corrado, G., Monti, L. M. y Rao, R. (2011). **Simple sequence repeats are able to trace tomato cultivars in tomato food chains.** *Food Control*, 22 (3-4), 549-554.

Casañas, F., Simó, J., Casals, J. y Prohens, J. (2017). **Toward an Evolved Concept of Landrace.** *Frontiers in Plant Science*, 8:145.

Castellana, S., Ranzino, L., Beritognolo, I., Cherubini, M., Luneia, R., Villani, F. y Mattioni, C. (2020). **Genetic characterization and molecular fingerprint of traditional Umbrian tomato (*Solanum lycopersicum* L.) landraces through SSR markers and application for varietal identification.** *Genetic Resources and Crop Evolution*, 1-14.

Castro, A., A. Nico y J.J. Garat (2004). **Proyecto de recuperación y preservación del Tomate Platense en el Cinturón Verde de La Plata, Argentina.** *Cultivar Local*, 4: p. 14-16.

Chamarro, J. **Anatomía y fisiología de la planta.** *Nuez, F. (Ed.). El cultivo del tomate.* Barcelona, MundiPrensa, 1994. pp. 43-91.

Chan, C., Beiko, R., Darling, A. y Ragan, M. (2009). **Lateral transfer of genes and gene fragments in prokaryotes.** *Genome Biol Evol.*, 1:429–438.

Chan, C. y Ragan, M. (2013). **Next-generation phylogenomics.** *Biology Direct.*, 8:3.

Chang, C., Marra, G., Chauhan, D., Ha, H., Chang, D., Ricciardiello, L., ... y Boland, C. (2002). **Oxidative stress inactivates the human DNA mismatch repair system.** *American Journal of Physiology: Cell Physiology*, 283: 148–154.

Chaw, S., Chun-Chieh Shih, A., Wang, D., Wu, Y., Liu, S. y Chou, T. (2008). **The mitochondrial genome of the gymnosperm *Cycas taitungensis* contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites.** *Mol Biol Evol*, 25:603–615.

Christensen, A. (2013). **Plant mitochondrial genome evolution can be explained by DNA repair mechanisms.** *Molecular Biology and Evolution*, 5:1079–1086.

Chu, K., Qi, J., Yu, Z. y Anh, V. (2004). **Origin and phylogeny of chloroplasts revealed by a simple correlation analysis of complete genomes.** *Molecular Biology and Evolution*, 21:200–206.

Corrado, G., Piffanelli, P., Caramante, M., Coppola, M. y Rao, R. (2013). **SNP genotyping reveals genetic diversity between cultivated landraces and contemporary varieties of tomato.** *Bmc Genomics*, 14(1), 835.

Corrado, G., Caramante, M., Piffanelli, P., y Rao, R. (2014). **Genetic diversity in Italian tomato landraces: Implications for the development of a core collection.** *Scientia Horticulturae*, 168, 138-144.

Crow, J. y Kimura, M. (1970). *An Introduction to Population Genetics Theory*.

D'Angelo, M., Zanon, M.I., Burgos, E., Asprelli, P. D., Boggio S. B., Carrari, F., Peralta I. E. y Valle, E. M. **Fruit metabolic and transcriptional programs differentiate among Andean tomato (*Solanum lycopersicum* L.) accessions.** *Planta* 250, 1927–1940 (2019).
<https://doi.org/10.1007/s00425-019-03274-4>

Darrigues, A., Hall, J., van der Knaap, E., Francis, D. M., Dujmovic, N. y Gray, S. (2008). **Tomato analyzer-color test: a new tool for efficient digital phenotyping.** *Journal of the American Society for Horticultural Science*, 133 (4), 579-586.

Darwin, S. C., Knapp, S. y Peralta, I. E. (2003). *Taxonomy of tomatoes in the Galapagos Islands: native and introduced species of *Solanum* section *Lycopersicon* (*Solanaceae*).* *Systematics and Biodiversity*, 12:29–53.

De Las Rivas, J., Lozano, J. y Ortiz, A. (2002). **Comparative analysis of chloroplast genomes: functional annotation, genome-based phylogeny, and deduced evolutionary patterns.** *Genome Research*, 12:567–583.

Di Rienzo, A., Peterson, A., Garza, J., Valdes, A., Slatkin, M. y Freimer, N. (1994). **Mutational processes of simple sequence repeat loci in human populations.** *Proceeding of National Academy of Sciences*, 91:3166-3170.

Dodsworth, S., Chase, M., Arkinen, T., Knapp, S. y Leitch, A. (2016). **Using genomic repeats for phylogenomics: a case study in wild tomatoes (*Solanum* section *Lycopersicon*: *Solanaceae*).** *Biological Journal of the Linnean Society*, 117:96–105.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K. y Mattick, J. S. (1991). **'Touchdown' PCR to circumvent spurious priming during gene amplification.** *Nucleic acids research*, 19 (14), 4008.

Dorrell, R. y Howe, C. (2015). **Integration of plastids with their hosts: lessons learned from dinoflagellates.** *Proc Natl Acad Sci USA*, 112(33):10247–10254.

Duarte, J., Wall, P., Edger, P., Landherr, L., Ma, H., Pires, P. ... y dePamphilis, C. (2010). **Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels.** *BMC Evolutionary Biology*, 10: 61.

Edwards, A., Civitello, A., Hammond, H. A. y Caskey, C. (1991). **DNA typing and genetic mapping with trimeric and tetrameric tan-dem repeats.** *American Journal of Human Genetics*, 49: 746–756.

Ellegren, H. (2004). **Microsatellites: simple sequences with complex evolution.** *Nat. Rev. Genet.*, 5:435–445.

Ellstrand, N. C. (2014). **Is gene flow the most important evolutionary force in plants?.** *Am. J. Bot.*, 101: 737–753.

Ellstrand, N. C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A., de Jong, T. J. y otros. (2013). **Introgression of crop alleles into wild or weedy populations.** *Annu. Rev. Ecol. Evol. Syst.*, 44: 325–345.

Ellstrand, P. H. C., N. C. y Hancock, J. F. (1999). **Gene flow and introgression from domesticated plants into their wild relatives.** *Annual Review of Ecology, Evolution, and Systematics*, 30: 539–563.

Fan, H., Ives, A., Surget-Groba, Y. y Cannon, C. (2015). **An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data.** *BMC Genomics*, 16:522.

Fan, J. y Lee, R. (2002). **Mitochondrial genome of the colorless green alga *Polytomella parva*: two linear DNA Molecules with homologous inverted repeat termini.** *Mol Biol Evol*, 19:999–1007.

Fitz-Gibbon, S. y House, C. (1999). **Whole genome-based phylogenetic analysis of freelifing microorganisms.** *Nucleic Acids Research*, 27:4218–4222.

Folmer, M., O. Black, Hoeh, W., Lutz, R. y Vrijenhoek, R. (1994). **DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates.** *Mol Mar Biol Biotechnol.*, 3:294–299.

Folquer, F. (1979). **El tomate. Estudio de la planta y su producción comercial.** Ed. *Hemisferio Sur*. Buenos Aires, Argentina.

Francis, D. (s/f). Tomato genetics and breeding program - ohio agri-cultural research and development center. Descargado de <https://tomato.cfaes.ohio-state.edu/>

Frary, A., Nesbitt, T. C., Frary, A., Grandillo, S., Van Der Knaap, E., Cong, B., ... & Tanksley, S. D. (2000). **fw2. 2: a quantitative trait locus key to the evolution of tomato fruit size.** *Science*, 289(5476), 85-88.

Frary, A., Xu, Y., Liu, J., Mitchell, S., Tedeschi, E. y Tanksley, S. (2005). **Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for genetics and breeding experiments.** *Theoretical and Applied Genetics*, 111 (2), 291-312.

Galtier, N., Nabholz, B., Glemin, S. y Hurst, G. (2009). **Mitochondrial DNA as a marker of molecular diversity: a reappraisal.** *Mol Ecol.*, 18:4541–4550.

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M. y Thannhauser, T. W. (2019). **The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor.** *Nat Genet.*, 51: 1044–1051.

Garat, J. (2002). **El tomate platense en La Plata, Argentina.** *Biodiversidad*, 34, 19-21.

Garat, J. (2003). **Sobre Tomate Platense.** *Cultivar Local*, 1: p. 23-24.

Garat, J., Ahumada, A., Otero, J., Terminiello, L., Bello, G. y Ciampagna, M. (2009). **Typical local vegetables in the Green Belt of La Plata: its location, preservation and valuation.** *Annu. Rev. Ecol. Evol. Syst.*, 28 (66): 32-39.

Garat, J., et al. (2007). **El rescate de la biodiversidad local y la acción colectiva: Una estrategia de desarrollo a través de la revalorización de hortalizas locales en el Cinturón Verde de La Plata, Buenos Aires, Argentina.** en Congreso Brasileiro de Agroecologia.

García-Martínez, S., Andreani, L., Garcia-Gusano, M., Geuna, F. y Ruiz, J. J. (2006). **Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars.** *Genome.*, 49(6), 648-656.

Gerth, M., Geißler, A. y Bleidorn, C. (2011). **Wolbachia infections in bees (Anthophila) and possible implications for DNA barcoding.** *Syst Biodivers*, 9:319–327.

Gonzalo, M. J., Brewer, M. T., Anderson, C., Sullivan, G. S., D. y van der Knaap, E. (2009). **Tomato fruit shape analysis using morphometric and morphology attributes implemented in Tomato Analyzer software program.** *Journal of the American Society for Horticultural Science*, 134 (1), 77-87.

Gould, S., Waller, R. y McFadden, G. (1999). **Plastid evolution.** *Annu Rev Plant Biol*, 59:491–51.

Grandillo, S., Chetelat, R., Knapp, S., Spooner, D., Peralta, I., Cammareri, M., ... y Pignone, D. (2011). **9. *Solanum* sect. *Lycopersicon*.** In: Kole C. (ed) **Wild crop relatives: genomics and breeding resources.** *Springer, Heidelberg*, Volume 5-vegetables, pp 129-216.

Gray, M. (1999). **Evolution of organellar genomes.** *Curr Opin Genet Dev*, 9:678–687.

Green, B. (2011). **Chloroplast genomes of photosynthetic eukaryotes.** *Plant J*, 66:34–44.

Göke, M., Schulz, J., Lasserre, J., y Vingron, M. (2012). **Estimation of pairwise sequence similarity of mammalian enhancers with word neighbourhood counts.** *Bioinformatics*, 28:656–663.

Haubold, B. (2014). **Alignment-free phylogenetics and population genetics.** *Brief Bioinformatics*, 15:407–418.

He, C., Poysa, V. y Yu, K. (2003). **Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars.** *Theoretical and Applied Genetics*, 106 (2), 363-373.

Hebert, P., Cywinska, A., Ball, S. y deWaard, J. (2003). **Biological identifications through DNA barcodes.** *Proc. R. Soc. Lond. B. Biol. Sci.*, 270:313–321.

Hebert, P., Ratnasingham, S., y de Waard, J. (2003). **Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species.** *Proc. R. Soc. Lond. B Biol. Sci.*, 270: S96–S99.

Hertog, M.L.A.T.M., Lammertyn, J., Scheerlinck, N. y Nicolai, B.M. 2007. **The impact of biological variation on postharvest behaviour: The case of dynamic temperature conditions.** *Postharvest Biol. Technol.* 43:183–192.

Hildebrand, C., David, C., Torney, C., y Wagner, P. (1994). **Informativeness of polymorphic DNA markers.** *Los Alamos Sci*, 20, 100-102.

Hohmann-Marriott, M., y Blankenship, R. (2011). **Informativeness of polymorphic DNA markers.** *Annu Rev Plant Biol*, 62:515–548.

Hollingsworth, P., Forrest, L., Spouge, J., Hajibabaei, M., Ratnasingham, S., van der Bank, M., . . . Fazekas, A. (2009). **A DNA barcode for land plants.** *Proceedings of the National Academy of Sciences*, 106:12794–12797.

Hollingsworth, P., Graham, S., y Little, D. (2011). **Choosing and using a plant DNA barcode.** *PLoS One*, 6: e19254.

Howard PH (2020) **How Corporations Control Our Seeds.** Bite back: *people taking on corporate food and winning.*

Hu, X., Wang, H., Chen, J., y Yang, W. (2012a). **Genetic diversity of Argentina tomato varieties revealed by morphological traits, simple sequence repeat, and single nucleotide polymorphism markers.** *Pak J Bot.*, 44 (2), 485-492.

Hu, X., Wang, H., Chen, J., y Yang, W. (2012b). **Genetic diversity of Argentina tomato varieties revealed by morphological traits, simple sequence repeat, and single nucleotide polymorphism markers.** *Pak J Bot*, 44 (2), 485-492.

Huson, D., y Bryant, D. (2006). **Application of phylogenetic networks in evolutionary studies.** *Molecular Biology and Evolution*, 23:254–267.

Jackson, A., Chen, R., y Loeb, L. (1998). **Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat.** *Proc Natl Acad Sci USA*, 95: 12468–12473.

Jacob, H., Lindpaintner, K., Kusumir, E. L., Bunkery, K., Mao, I., Gantenv, D., . . . Lander, E. (1991). **Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat.** *Cell*, 67: 213–224.

Janouškovec, J., Liu, S.-L., Martone, P., Carré, W., Leblanc, C., Collén, J., y Keeling, P. (2013). **Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers.** *PLoS One*, 8: e59001.

Jarvis, T., D. I. and Hodgkin. (1999). **Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems.** *Mol. Ecol.*, 8: 159–173.

Jenkins, J. A. (1948). **The origin of the cultivated tomato.** *Economic Botany*, 2 (4), 379-392.

Kader, A.A. (1986). **Effects of postharvest handling procedures on tomato quality.** *Acta Hort.* 190, 209-222

Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., y Dhawan, A. K. (2011). **Microsatellite markers: an overview of the recent progress in plants.** *Euphytica*, 177 (3) 309-334.

Keeling, P. (2009). **Role of horizontal gene transfer in the evolution of photosynthetic eukaryotes and their plastids.** In: Gogarten MB, Gogarten JP, Olendzenski LC (eds) *Horizontal gene transfer, Methods in molecular biology*, vol 532. Humana Press, New York, pp 501–515.

Keeling, P. J. (2010). **The endosymbiotic origin, diversification and fate of plastids.** *Philos Trans R Soc Lond Ser B Biol Sci*, 365:729–748.

Keeling, P. J. (2013). **The number, speed, and impact of plastid endosymbioses in eukaryotic evolution.** *Annu Rev Plant Biol*, 64:583–607.

Knapp, S. (2002). **Solanum section Geminata.** *Fl Neotrop*, 84: 1-405.

Knapp, S., y Peralta, I. E. (2016). **The tomato (*Solanum lycopersicum* L., Solanaceae) and its botanical relatives.** In *The Tomato Genome* (pp. 7-21). Springer, Berlin, Heidelberg.

Knapp, S. (2019). **People and plants: The unbreakable bond.** *Plants, People, Planet*, 1(1), 20-26.

Knoop, V. (2012). **Seed plant mitochondrial genomes: complexity evolving.** In: Bock R, Knoop V (eds) *Genomics of chloroplasts and mitochondria, Advances in photosynthesis and respiration*, vol 35. Springer Science + Business Media B.V, Dordrecht, pp 175–200.

Krallinger, M. y Valencia, A. (2005). **Text-mining and information retrieval services for molecular biology.** *Genome Biol.*, 6:224.

Kress, W. y Erickson, D. (2002). **DNA barcodes: genes, genomics, and bioinformatics.** *PNAS*, 105:2761–2762.

Köhler, S., Delwiche, C., Denny, P., Tilney, L., Webster, P., Wilson, R., ... y Roos, D. (1997). **A plastid of probable green algal origin in apicomplexan parasites.** *Science*, 275:1485–1489.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., ... Savolainen, V. (2008). **DNA barcoding the oras of biodiversity hotspots.** *PNAS*, 105:2923–2928.

Langmead, B., y Salzberg, S. L. (2012). **Fast gapped-read alignment with Bowtie 2.** *Nat Methods*, 9:357–359.

Lemieux, C., Otis, C., y Turmel, M. (2000). **Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution.** *Nature*, 403:649–652.

Li, C., Han, G., y Chu, K. (2012). **Analyzing multi-locus plant barcoding data-sets with a composition vector method based on adjustable weighted distance.** *PLoS One*, 7: e42154.

Li, D., Gao, L., Li, H., Wang, H., Ge, X., Liu, J., ... Chen, S. (2012). **From the cover: comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants.** *Proc Natl Acad Sci*, 108:19641–19646.

Li, Y., Korol, A., Fahima, T., Beiles, A., y Nevo, E. (2002). **Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review.** *Mol Ecol*, 11: 2453–2465.

Lim, L. y McFadden, G. (2010). **The evolution, metabolism and functions of the apicoplast.** *Philos Trans R Soc Lond Ser B Biol Sci* , 365:749–763.

Lin, J. y Gerstein, M. (2000). **Whole-genome trees based on the occurrence of folds and orthologs: implications for comparing genomes on different levels.** *Genome Res*, 10:808–818.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., ... Wang, X. (2014). **Genomic analyses provide insights into the history of tomato breeding.** *Nat Genet.*, 46:1220–1226.

Linnaeus, C. (1753). *Species plantarum*. st edn. L. Salvius, Stockholm.

Litt, M. y Luty, J. A. (1989). **A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene.** *American Journal of Human Genetics*, 44: 397–401.

Liu, Y., Wang, B., Cui, P., Li, L., Xue, J., Yu, J., y Qiu, Y. (2012). **The mitochondrial genome of the lycophyte *Huperzia squarrosa*: the most archaic form in vascular plants.** *PLoS One*, 7: e35168.

Lucatti, A. F., van Heusden, A. W., de Vos, R. C. H., Visser, R. G. F., y Vosman, B. (2013). **Differences in insect resistance between tomato species endemic to the Galapagos islands.** *BMC Evol Biol*, 13: 175.

Luckwill, L. (1943). **The Genus Lycopersicon: A Historical, Biological, and Taxonomic Survey of the Wild and Cultivated Tomatoes**: Aberdeen University Press, Aberdeen, Scotland. Macbride, J. F. (1962). Solanaceae. In: Flora of Peru. Field Mus Nat Hist Bot Ser , 13: 3–267.

Mansholt, U. J. (1909). **Van Pesch Plantenteelt, beknopte handleiding tot de kennis van den Nederlandschen landbouw**. 228 pp.: Plantenteelt. Zwolle. 3rd revised edition, pt 2.

Marshall, J. A., Knapp, S., Davey, M. R., Power, J. B., Cocking, E. C., Bennett, M. D. y Cox, A. V. (2001). **Molecular systematics of *Solanum* section *Lycopersicum* (*Lycopersicon*) using the nuclear ITS rDNA region**. *Theoretical and Applied Genetics*, 103(8), 1216-1222.

Martin, B., W. andStoebe, Goremykin, V., Hansmann, S., Hasegawa, M. y Kowallik, K. (1998). **Gene transfer to the nucleus and the evolution of chloroplasts**. *Nature*, 393:162–165.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., . . . Penny, D. (2002). **Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus**. *Proc Natl Acad Sci.*, 99:12246–12251.

Mattioli, P. A. (1544). **Di Pedacio Dioscoride Anazarbeo libri cinquedella historia, et ateria medicinale trodotti in lingua uolgare Italiana**. N. de Bascarini, Venice.

Mayr, E. (1937). **Alpine Landsorten in ihrer Bedeutung für die praktische Züchtung**. *Forschungsdienst*, 4: 162-166.

Mazzucato, A., Papa, R., Bitocchi, E., Mosconi, P., Nanni, L., Negri, V., y Veronesi, F. (2008). **Genetic diversity, structure and marker trait associations in a collection of Italian tomato (*Solanum lycopersicum* L.) landraces**. *Theoretical and Applied Genetics*, 116 (5), 657-669.

McElroy, K., Luciani, F., y Thomas, T. (2012). **GemSIM: general, error-model based simulator of next-generation sequencing data.** *BMC Genomics.*, 13:74.

McGuire, R. G. (1992). **Reporting of objective color measurements.** *HortScience*, 27 (12), 1254-1255.

Melo, P.C.T. **Melhoramento genético do tomateiro.** Campinas. Asgrow do Brasil Sementes. 1989. 55 p.

Messeguer, J. (2003). **Gene flow assessment in transgenic plants.** *Plant Cell Tissue Organ. Cult.*, 73: 201–212.

Miller, J. C., y Tanksley, S. D. (1990). **RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*.** *Theoretical and applied genetics*, 80 (4), 437-448.

Mitrophanov, A., y Borodovsky, M. (2006). **Statistical significance in biological sequence analysis.** *Brief Bioinform.*, 7:2–24.

Morosi, J. y F. Terán (1981). **La Plata, espacio urbano singular.** Ciudad y Territorio, Revista de Ciencia Urbana, 3: p. 81.

Mower, J., Sloan, D., y Alverson, A. (2012). **Plant mitochondrial genome diversity: the genomics revolution.** In: Wendel FJ, Greilhuber J, Dolezel J, Leitch JI (Eds) *Plant genome diversity, Plant genomes, their residents, and their evolutionary dynamics*, vol 1. Springer, Vienna, pp 123–144.

Muller, C. H. (1940). **A revision of the genus *Lycopersicon* (n° 382).** US Department of Agriculture.

Murray, B. W. (1996). **The estimation of genetic distance and population substructure from microsatellite allele frequency data.** McMaster University, Ontario, Canada.

Nadir, E., Margalit, H., Gallily, T., y Ben-Sasson, S. (1996). **Microsa-tellite spreading in the human genome: evolutionary mechanisms and structural implications.** *Proc Natl Acad Sci USA*, 93:6470–6475.

Negri, V., Maxted, N. y Veteläinen, M. (2009). **European landrace conservation: an introduction. 15: 1-22: European landraces: on-farm conservation, management and use.** *Bioversity technical bulletin n° 15.*

Nei, M. (1973). **Analysis of genetic diversity in subdivided populations.** *Proceedings of the National Academy of Sciences*, 70: 3321-3323.

Nico, A., Garat, J. J., Castro, A., Gramuglia, S., Ahumada, A., Gamboa, S., y Sarandon, S. (2006). **Recursos genéticos de plantas hortícolas alóctonas en el cinturón verde metropolitano (La Plata, Buenos Aires, Argentina).** *Noticiero de Recursos Fitogenéticos.*

Nogué, F., Mara, K., Collonnier, C., y Casacuberta, J. M. (2016). **Genome engineering and plant breeding: impact on trait discovery and development.** *Plant Cell Reports*, 35, 1475–1486.

Notredame, C. (2007). **Recent evolutions of multiple sequence alignment algorithms.** *PLoS Comput Biol*, 3: e123.

Nuez, F., et al. (1996). **Banco de Germoplasma de la Universidad Politécnica de Valencia.** *Catálogo de Semillas de Tomate.*, in *Monografías INIA.*, MAPyA-INIA: Madrid, España.

Oliveira, E. J., Pádua, J. G., Zucchi, M. I., Vencovsky, R., & Vieira, M. L. C. (2006). **Origin, evolution and genome distribution of microsatellites.** *Genetics and Molecular Biology*, 29(2), 294-307.

Olmstead, R. G. y Palmer, J. D. (1997). **Implications for the phylogeny, classification, and biogeography of Solanum from cpDNA restriction site variation.** *Systematic Botany*, 19-29.

Olmstead, R. G., Sweere, J. A., Spangler, R. E., Bohs, L. y Palmer, J. D. (1999). **Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA**. *Solanaceae* IV, 1(1), 1-137.

Padrón Pereira, C. A., León, P., Marié, G., Montes Hernández, A. I., y Oropeza González, R. A. (2012). **Determinación del color en epicarpio de tomates (*Lycopersicon esculentum* Mill.) con sistema de visión computarizada durante la maduración**. *Agronomía Costarricense*, Vol. 36 Núm. 1.

Palmer, J. y Zamir, D. (1982). **Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon***. *PNAS*, 79:5006–5010.

Parida, S., Kalia, S., Sunita, K., Dalal, V., Hemaprabha, G., Selvi, A., . . . Mohapatra, T. (2009). **Informative genomic microsatellite markers for efficient genotyping applications in sugarcane**. *Theoretical and Applied Genetics*, 118: 327–338.

Pearson, C., Edamura, N., y Cleary, J. (2005). **Repeat instability: mechanisms of dynamic mutations**. *Nature Reviews Genetics.*, 6: 729–742.

Peralta, I., y Spooner, D. (2005). **Morphological characterization and relationships of wild tomatoes (*Solanum L. sect. Lycopersicon*)**. *Monographs Syst Bot.*, 104:227–257.

Peralta, I. E. y Spooner, D. M. (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, I. E. y Spooner, D. M. (2007). **History, origin and early cultivation of tomato (*Solanaceae*)**. In: Razdan MK, Mattoo AK (eds) Genetic improvement of Solanaceous crops, vol 2, tomato. Science, Enfield, pp 1–27.

Peralta, I. E., Spooner, D. M., y Knapp, S. (2008). **Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; *Solanaceae*)**. *Systematic botany monographs*, 84.

Peralta, I.E.; Makuch, M.; García Lampasona, S.; Occhiuto, P.N.; Asprelli, P.D.; Lorello, I.M. & Togno, L. (2007). **Catálogo de poblaciones criollas de pimiento, tomate y zapallo, colectadas en valles andinos de Argentina.** *Instituto Nacional de Tecnología Agropecuaria*, 128 p.

Piñero-Nelson, A., Van Heerwaarden, J., Perales, H. R., Serratos-Hernández, J. A., Rangel, A., Hufford, M. B., y Álvarez-Buylla, E. R. (2009). **Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations.** *Molecular ecology*, 18 (4): 750-761.

Powell, W., Machray, G. C., y Provan, J. (1996). **Polymorphism revealed by simple sequence repeats.** *Trends Plant Science*, 1: 215–222.

Pratta, G., Cánepa, L., Zorzoli, R. y Picardi, L. (2003). **Efecto del germoplasma silvestre sobre caracteres de interés agronómicos en híbridos intra e interespecíficos del género *Lycopersicon*.** *Revista de Investigaciones de la Facultad de Ciencias Agrarias*. (3):13-21. junio de 2003

Puigbò, P., Wolf, Y. y Koonin, E. (2010). **The tree and net components of prokaryote evolution.** *Genome Biol Evol*, 2:745–756.

Qi, W., Yu, Z.-G., y Yang, J. (2017). **DLTree: a web server for phylogeny reconstruction using dynamical language method.** *Bioinformatics*, 33:2214–2215.

Ranc, N., Muños, S., Santoni, S., Causse, M. (2008) **A clarified position for *Solanum lycopersicum* var. *cerasiforme* in the evolutionary history of tomatoes (Solanaceae).** *BMC Plant Biol* 8:130.

Rannala, B., y Yang, Z. (2008). **Phylogenetic inference using whole genomes.** *Annu Rev Genom Hum Genet.*, 9:217–231.

Raphael, S. (1989). **Matthioli's herbal. In The Matthioli woodblocks. (catalog without pagination).** London: B. Quaritch Ltd.

Ribeiro, C. A. G., Pinto, M. D. O., Maciel, T. E. F., Pastina, M. M., Barros, E. G. D., y Guimarães, C. T. (2017). **Universal tail sequenceSSR applied to molecular characterization of tropical maize hybrids.** *Scientia Agricola*, 74(2), 163-168.

Rice, D., Alverson, A., Richardson, A., Young, G., Sanchez-Puerta, M., Munzinger, J., . . . Palmer, J. (2013). **Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella.** *Science*, 342:1468–1473.

Rick, C.M. (1973). **Potential genetic resources in tomato species: clues from observations in native habitats.** *Genes, Enzymes, and Populations*, A.M. Srb, Editor., Plenum Press: New York, USA. p. 255-269.

Rick, C.M. y J.F. Fobes (1975) **Allozyme variation in the cultivated tomato and closely related species.** *Bulletin of the Torrey Botanical Club*, 102(6): p. 376-384.

Rick, C. (1976). **Tomato (Family Solanaceae).** (In N.W. Simmonds (Ed.) *Evolution of crop plants*. Longman, London.

Rick, C. M. (1978). **Potential improvement of tomatoes by controlled introgression of genes from wild species.** *Proceedings of the Conference on Broadening Genetic Base Crops*, Wagenigen, (pp. 167-173).

Rick, C.M. (1982). **The potential of exotic germplasm for tomato improvement.** *Plant Improvement and Somatic Cell Genetics*, I.K. Vasil, W.R. Scowcroft, and K.J. Frey, Editors., Academic Press: New York, USA. p. 1-28.

Rodríguez, F., Wu, F., Ané, C., Tanksley, S., y Spooner, D. (2009). **Do potatoes and tomatoes have a single evolutionary history, and what proportion of the genome supports this history?.** *BMC Evol Biol.*, 9:191.

Rodríguez, G. R., Moyseenko, J. B., Robbins, M. D., Morejón, N. H., Francis, D. M., y van der Knaap, E. (2010). **Tomato analyzer: a useful software application to collect accurate and detailed morphological and colorimetric data from two-dimensional objects.** *JoVE (Journal of Visualized Experiments)*, (37), e1856.

Rodríguez, W. E. y Flórez, V. J. (2006). **Comportamiento fenológico de tres variedades de rosas rojas en función de la acumulación de la temperatura.** *Agronomía Colombiana*, 24(2), 247-257.

Rohlf, F. (2000). **NTSYSpc. Numerical Taxonomy and Multivariate Analysis System-Version 2.02.** *Applied Biostatistics Inc.*, New York.

Ruiz, J. J., García-Martínez, S., Picó, B., Gao, M. y Quiros, C. F. (2005). **Genetic variability and relationship of closely related Spanish traditional cultivars of tomato as detected by SRAP and SSR markers.** *Journal of the American Society for Horticultural Science*, 130(1), 88-94.

Rus-Kortekaas, W., Smulders, M., Arens, P., y Vosman, B. (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.

Rümker von, K. (1908). **Die systematische Einteilung und Benennung der Getreidesorten für praktische Zwecke.** *Jahrbuch der Deutschen landwirtschafts-Gesellschaft*, 23, 137-167.

s/a. (2017). **AgriFood Atlas. Facts and figures about the corporations that control what we eat.** *Heinrich Böll Foundation and Rosa Luxemburg Foundation and Friends of the Earth Europe.*

Sacks, E.J. y D.M. Francis. 2001. **Genetic and environmental variation for tomato flesh color in a population of modern breeding lines.** *J. Amer. Soc. Hort. Sci.* 126:221-226.

Sankoff, D., Leduc, G., Antoine, N., Paquin, B., Lang, B. y Cedergren, R. (1992). **Gene order comparisons for phylogenetic inference: evolution of the mitochondrial genome.** *Proc Natl Acad Sci.*,89:6575-6579.

Saunders, G., y McDevit, D. (2012). **Methods for DNA barcoding photosynthetic protists emphasizing the macroalgae and diatoms.** Kress JW, Erickson LD (eds) DNA barcodes: methods and protocols. Humana Press, Totowa.

Sbordoni, V., Allegrucci, G. y Cesaroni, D. (2012). **Population Structure**, in Encyclopedia of Caves. (Second Edition).

Shirasawa, K., Isobe, S., Hirakawa, H., Asamizu, E., Fukuoka, H., Just, D., ... y Kohara, M. (2010). **SNP discovery and linkage map construction in cultivated tomato.** *DNA research*, 17(6), 381-391.

Schoch, C., Seifert, K., Huhndorf, S., Robert, V., Spouge, J., Levesque, C. A., Consortium, F. (2012). **Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi.** *Proc Natl Acad Sci USA*, 109:6241–6246.

Sim, S. C., Van Deynze, A., Stoffel, K., Douches, D. S., Zarka, D., Ganal, M. W., ... y Panthee, D. R. (2012). **High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding.** *PLoS one*, 7(9), e45520.

Sims, W. L. (1979). **History of tomato production for industry around the world.** (pp. 25-26): In Symposium on Production of Tomatoes for Processing 100.

Slatkin, M. (1995). **A measure of population subdivision based on microsatellite allele frequencies.** *Genetics*, 139: 457-462.

Sloan, D., Alverson, A., Chuckalovcak, J., Wu, M., McCauley, D., Palmer, J., y Taylor, D. (2012). **Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates.** *PLoS Biol*, 10:e1001241.

Small, R., Cronn, R., y Wendel, J. (2004a). **A measure of population subdivision based on microsatellite allele frequencies.** *Aust Syst Bot.*, 17:145–170.

Small, R., Cronn, R., y Wendel, J. (2004b). **L. A. S. JOHNSON REVIEW No. 2. Use of nuclear genes for phylogeny reconstruction in plants.** *Aust Syst Bot.*, 17:145–170.

Smulders, M. J. M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P. y Vosman, B. (1997). **Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species.** *Theoretical and Applied Genetics*, 94 (2), 264-272.

Sneath, P. y SoIkal, R. (1973). **Numerical taxonomy: The Principles and Practice of Numerical Classification.** *W. H. Freeman*. San Francisco.

Song, K., Ren, J., Zhai, Z., Liu, X., Deng, M. y Sun, F. (2013). **Alignment-free sequence comparison based on next-generation sequencing reads.** *J Comput Biol.*, 20:64–79.

Soressi, G. P. (1969). **Il pomodoro.** Edagricole.

Spooner, D. (2005). **Molecular markers for genebank management.** Bioersity International., (No. 10).

Spooner, D. M., Anderson, G. J., y Jansen, R. K. (1993). **Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (*Solanaceae*).** *American Journal of Botany*, 80(6), 676-688.

Spooner, D. M., Peralta, I. E., y Knapp, S. (2005). **Comparison of AFLPs to other markers for phylogenetic inference in wild tomatoes [*Solanum L.* section *Lycopersicon* (Mill.) *Wettst.* Subsection *Lycopersicon*].** *Taxon*, 54:43–61.

Stiller, J. (2011). **Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer.** *BMC Evol Biol.*, 11:259.

Strittmatter, P., Soll, J., y Bölter, B. (2010). **The chloroplast protein import machinery: a review.** pp 307–321: In: Economou A (ed) Protein secretion, Methods in molecular biology, vol 619. Humana Press, New York.

Stuart, G., Moffett, K. y Baker, S. (2002). **Integrated gene and species phylogenies from unaligned whole genome protein sequences.** *Bioinformatics*, 18:100–108.

Suliman-Pollatschek, S., Kashkush, K., Shats, H., Hillel, J. y Lavi, U. (2002). **Generation and mapping of AFLP, SSRs and SNPs in *Lycopersicon esculentum***. *Cellular and Molecular Biology Letters*, 7 (2A), 583-598.

Suzuki, K. y Miyagishima, S.-Y. (2010). **Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses**. *Mol Biol Evol*, 27:581–590.

Tautz, D. (1989). **Hypervariability of simple sequences as a general source for polymorphic DNA markers**. *Nucleic Acid Research*, 17: 6463–6471.

Tekaia, F., Lazcano, A., y Dujon, B. (1999). **The genomic tree as revealed from whole proteome comparisons**. *Genome Res*, 9:550–557. Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartin hour, S., y

McCouch, S. (2001). **Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential**. *Genome Res*, 11: 1441–1452.

Jablonska, B., Ammiraju, J. S., Bhattarai, K. K., Mantelin, S., de Ilarduya, O. M., Roberts, P. A. y Kaloshian, I. (2007). **The Mi-9 gene from *Solanum arcanum* conferring heat-stable resistance to root-knot nematodes is a homolog of Mi-1**. *Plant Physiology*, 143(2), 1044-1054.

TGRC. (s/f). Tomato genetics resource center at university of california, davis. Descargado de <https://tgrc.ucdavis.edu/>

Tieman, D., Zhu, G., Resende, M., Lin, T., Nguyen, C., Bies, ., D., y Ikeda, H. (2017). **A chemical genetic roadmap to improved tomato flavor**. *Science*, 355: 6323–6391.

Timmis, J., Ayliffe, M., Huang, C. y Martin, W. (2004). **The genomic tree as revealed from whole proteome comparisons**. *Nat Rev Genet*, 5:123–135.

Tindall, B., Rosselló-Móra, R., Busse, H., Ludwig, W. y Kämpfer, P. (2010). **Notes on the characterization of prokaryote strains for taxonomic purposes.** *Int J Syst Evol Microbiol*, 60:249–266.

Trifonov, E. (2003). **Tuning function of tandemly repeating sequences: a molecular device for fast adaptation.** In: Wasser SP (ed) *Evolutionary theory and processes: modern horizons, papers in honor of Eviatar Nevo.* pp 1–24: Kluwer, Amsterdam.

Turmel, M., Otis, C. y Lemieux, C. (1999). **The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes.** *Proc. Natl. Acad. Sci.*, 96:10248–10253.

Turmel, M., Otis, C. y Lemieux, C. (2002). **The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphae-ridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants.** *Proc Natl Acad Sci.*, 99:11275–11280.

Ullé, J. (2007). **Evaluación de variedades locales de tomate conservadas “*in situ*” en el Gran La Plata en comparación con variedades inscriptas por INTA en el registro nacional de cultivares.**, INTA: San Pedro, Buenos Aires.

Van der Beek, H., R, V., Zabel, P., y Lindhout, P. (1992). **Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: CF9 (resistance to *Cladosporium fulvum*) on chromosome 1.** *Theoretical and Applied Genetics*, 84: 106-112.

Vallejo Cabrera, F. A., Pava, J. H., Vargas, J. A. y Araujo A. (1994). **Caracterización morfo-agronómica de especies y variedades botánicas del género *Lycopersicon*.** *Acta Agronómica* 44 (1/4): 37-50.

Vallejo, F.A. (1999). **Mejoramiento genético y producción del tomate en Colombia.** *Universidad Nacional de Colombia. Ed. Feriva.* Palmira.

Villa, T. C. C., Maxted, N., Scholten, M. y Ford-Lloyd, B. (2005). **Defining and identifying crop landraces.** *Plant Genetic Resources*, 3(3), 373-384.

Vinga, S. y Almeida, J. (2003). **Alignment-free sequence comparison - a review.** *Bioinformatics*, 19:513–523.

Von der Pahlen, A. (1970). **Inheritance of resistance of Tomato (cv. Platense) to Tomato spotted wilt virus.** *Revista de Investigaciones Agropecuaria*, Ser. 5, 7(3), 129-134.

Wang, M., Barkley, N. y Jenkins, T. (2009). **Microsatellite markers in plants and insects. Part I. Applications of biotechnology.** *Genes Genomes Genomics*, 3:54–67.

Warren, G.F. (1998). **Spectacular increases in crop yields in the United States in the twentieth century.** *Weed Technology*, 12(4): p. 752-760.

Wegier, A., Piñeyro-Nelson, A., Alarcón, J., Gálvez-Mariscal, A., Alvarez-Buylla, E. R. y Piñero, D. (2016). **Recent long-distance transgene flow into wild populations conforms to historical patterns of gene flow in cotton (*Gossypium hirsutum*) at its centre of origin.** *Molecular Ecology*, 20(19): 4182-4194.

Weir, B. (1996). **Genetic data analysis II. Methods for discrete population genetic data.** 2nd edición. Sinauer Associates, Sunderland.

Weir, B. S. y Cockerham, C. C. (1984). **Estimating F-statistics for the analysis of population structure.** *Evolution*, 38(6), 1358-1370.

Wong, K., Suchard, M. y Huelsenbeck, J. (2008). **Alignment uncertainty and genomic analysis.** *Science*, 319:473–476.

Wright, S. (1931). **Evolution in Mendelian populations.** *Genetics*, 16:97-159.

Wright S (1951) **The genetical structure of populations.** *Annual Eugenics* 15:323-354.

Wright S (1978) *Evolution and the genetics of populations* vol 4. Variability within and among natural populations. Univ. of Chicago Press, Chicago

Wu, F., Mueller, L., Crouzillat, D., Petiard, V., y Tanksley, S. (2006). **Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the eu asteroid plant clade.** *Genetics*, 174: 1407–1420.

Wu, M., Chatterji, S. y Eisen, J. (2012). **Accounting for alignment uncertainty in phylogenomics.** *PLoS One*, 7: e30288.

Wu, Z. (2015). **The completed eight chloroplast genomes of tomato from *Solanum* genus.** *Mitochondrial DNA.*, 27:1–3.

Yagi, Y. y Shiina, T. (2014). **Recent advances in the study of chloroplast gene expression and its evolution.** *Front Plant Sci.*, 5:61.

Yang, K. y Zhang, L. (2008). **Performance comparison between k-tuple distance and four model-based distances in phylogenetic tree reconstruction.** *Nucleic Acids Res.*, 36: e33.

Yi, H. y Jin, L. (2015). **Co-phylog: an assembly-free phylogenomic approach for closely related organisms.** *Nucleic Acids Res.*, 41: e75.

Yutin, N., Puigbò, P., Koonin, E. y Wolf, Y. (2012). **Phylogenomics of prokaryotic ribosomal proteins.** *PLoS One*, 7:5.

Zane, L., Bargelloni, L. y Patarnello, T. (2002). **Strategies for microsatellite isolation: a review.** *Molecular Ecology*, (11), 1-16.

Zeven, A. C. (1998). **Landraces: a review of definitions and classifications.** *Euphytica*, 104: 127–139.

Zeven, A. C. (1999). **The traditional inexplicable replacement of seed and seed ware of landraces and cultivars: A review.** *Euphytica*, 110: 181–191.

Anexo I

Muestras evaluadas y su procedencia. (*) Código de entrada del Banco de Germoplasma - INTA - La Consulta.

Código Interno	Nombre	Origen
G1	Uco Plata INTA	INTA La Consulta
G2	Lomitero	INTA La Consulta
G3	Uco -14-INTA	INTA La Consulta
G4	Rio Grande	ProHuerta-FecoAgro
G5	'Platense'	ProHuerta-FecoAgro
G6	'Platense'	<i>Bonanza Seeds</i>
G7	'Platense'	<i>Florensa</i>
G8	'Carcione'	Chacra Experimental Gorina
G9	'Gentile'	Chacra Experimental Gorina
G10	'Del Manso'	Chacra Experimental Gorina
G11	'Luna'	UPID - UNLP
G13	'Del Manso'	UPID - UNLP
G14	'Gentile'	UPID - UNLP
G15	3837* - 'Prieto' (de Santini)	INTA La Consulta
G16	3834* - 'Paolettich'	INTA La Consulta
G17	3833* - 'Molinaro'	INTA La Consulta
G18	3832* - 'Luna'	INTA La Consulta
G19	3829* - Sin datos	INTA La Consulta
G20	3826* - 'Grasso'	INTA La Consulta
G21	3825* - 'Grasso'	INTA La Consulta
G22	3824* - 'Gentile'	INTA La Consulta
G23	3820* - 'Del Manso'	INTA La Consulta
G24	3813* - 'Carcione'	INTA La Consulta
G25	3812* - 'Carcione'	INTA La Consulta
G26	3811* - 'Bustos'	INTA La Consulta
G27	'Platense'	Grupo Tomate Platense
G28	'Platense'	Grupo Tomate Platense
G29	'Minuchi'	Asociación 1610
G30	'Platense'	Grupo Tomate Platense
G31	'Bustos'	UPID - UNLP
G32	'Ventura'	UPID - UNLP
G33	'Grasso'	UPID - UNLP
G34	'Carcione'	UPID - UNLP
G38	'Platense'	ProHuerta-FecoAgro
HIB	Elpida	Híbrido Comercial
PIM	<i>S. pimpinellifolium</i> LA 1589	Fac. Cs. Agrarias - UNCuyo
Uco18	Uco 18 INTA	INTA La Consulta
LT	Mejorado Platense	Chacra Experimental Gorina
J	Mejorado Platense	Chacra Experimental Gorina

Anexo II

Características de los marcadores moleculares SSR utilizados

SSR	<i>Forward primer</i> (5'-3')	<i>Reverse primer</i> (5'-3')	<i>Temp. anelling</i> (°C)
TMS63	GCAGGTACGCACGCATATAT	GCTCCGTCAGGAATTCTCTC	55
SSR5	TGGCCGGCTTCTAGAAATAA	TGAAATCACCCGTGACCTTT	50
SSR96	GGGTTATCAATGATGCAATGG	CCTTTATGTCAGCCGGTGTT	50
TMS58	CATTTGTTGTATGGCATCGC	CAGTGACCTCTCGCACAAAA	55
SSR111	TTCTTCCCTTCCATCAGTTCT	TTTGCTGCTATACTGCTGACA	50
Tom47-48	CAAGTTGATTGCATTACCTATTG	TACAACAACATTTCTTCTTCCTT	50
SSR43	CTCCAAATTGGGCAATAACA	TTAGGAAGTTGCATTAGGCCA	50
Tom152-153	ATTCAAGGAACTTTTAGCTCC	TGCATTAAGGTTCATAAATGA	50
TMS60	ATGCAGTTCCAAGCATCATT	TTGCCACATTAATGTTGAAG	50
LELEUZIP	GGTGATAATTTGGGAGGTTAC	CGTAACAGGATGTGCTATAGG	50
SSR327	TCAGGATCAGGAGCAGGAGT	TGGACTTGTTCCATGAACCC	50
Tom236-237	GTTTTTCAACATCAAAGAGCT	GGATAGGTTTCGTTAGTGAAC	50
SSR383	ATTGTACAAAGACCCGTGGC	GTTGCACACTGGATCAATGC	50
LECHI3	TAACAATCAAAGAAGAACTTCGC	ATCCCCTTATTGATTACATCC	50
TMS42	AGAATTTTTTCATGAAATTGTCC	TATTGCGTTCCACTCCCTCT	50
TMS52	TTCTATCTCATTGGCTTCTTC	TTACCTTGAGAATGGCCTTG	50

Anexo III

Resumen de la colección de tomate utilizados.

Código	Variedad botánica	Nº accesión SRA	Cob. cloroplasto	Cob. genoma
TS-420	<i>S. pimpinellifolium</i>	SRR1572276	963,72	5,8
TS-267	<i>S. pimpinellifolium</i>	SRR1572259-60-61	6084,02	18,1
TS-433	<i>S. pimpinellifolium</i>	SRR1572285-86	2442,7	5,33
TS-432	<i>S. pimpinellifolium</i>	SRR1572283-84	3272,3	5,4
TS-415**	<i>S. pimpinellifolium</i>	SRR1572271	2615,59	7,7
TS-299	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572435	836	5,5
TS-72	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572344	2655,29	5,8
TS-91	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572349-50	1822,36	6,9
TS-105	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572361	6178,2	4,8
TS-131	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572373	1352,44	5,5
TS-129	<i>S. lycopersicum</i> var. <i>cerasiforme</i>	SRR1572372	1577,4	5,9
TS-44	<i>S. lycopersicum</i>	SRR1572467	4024,4	6,95
TS-100	<i>S. lycopersicum</i>	SRR1572499	4864,82	9,16
TS-132	<i>S. lycopersicum</i>	SRR1572527	1000,69	6,19
TS-135	<i>S. lycopersicum</i>	SRR1572530	4266,05	6,65
TS-137	<i>S. lycopersicum</i>	SRR1572532	4050,88	5,44
TS-152	<i>S. lycopersicum</i>	SRR1572545	1047,16	5,8
TS-168	<i>S. lycopersicum</i>	SRR1572559	4959,9	4,8
TS-172	<i>S. lycopersicum</i>	SRR1572564	3562,4	5,89
TS-178	<i>S. lycopersicum</i>	SRR1572570	3899,5	3,56
TS-184	<i>S. lycopersicum</i>	SRR1572575	900,1	4,8
TS-190	<i>S. lycopersicum</i>	SRR1572582	3907,68	5,68
TS-237	<i>S. lycopersicum</i>	SRR1572619	4451,5	4,03
TS-249	<i>S. lycopersicum</i>	SRR1572626	1112,88	6,06
TS-251	<i>S. lycopersicum</i>	SRR1572627	3038,3	5,2
TS-253	<i>S. lycopersicum</i>	SRR1572628	1512,2	4,5
TS-256	<i>S. lycopersicum</i>	SRR1572630	3782,76	7,44
TS-282	<i>S. lycopersicum</i>	SRR1572654	1903,15	6,1
TS-321**	<i>S. lycopersicum</i>	SRR1572684	4523,09	8,5
TS-409	<i>S. lycopersicum</i>	SRR1572666	5262,07	7,89
TS-242	<i>S. lycopersicum</i>	SRR1572623	885,9	5,35
TS-191	<i>S. lycopersicum</i>	SRR1572583	2865,8	6,1
TS-192	<i>S. lycopersicum</i>	SRR1572584	3321,36	5,8
TS-193	<i>S. lycopersicum</i>	SRR1572585	2613,36	5,6
TS-203	<i>S. lycopersicum</i>	SRR1572594	2599,2	5,2
TS-408**	<i>S. chilense</i>	SRR1572696	5308,99	3,23
TS-407**	<i>S. habrochaites</i>	SRR1572697	1133,56	2,57

TS-404**	<i>S. peruvianum</i>	SRR1572695	3556,01	3,17
TS-403	<i>S. peruvianum</i>	SRR1572694	1342,74	2,83
TS-402	<i>S. peruvianum</i>	SRR1572692-93	1113,13	5,88
TS-146**	<i>S. neorickii</i>	SRR1572685	2924,97	3,46
TS-208**	<i>S. galapagense</i>	SRR1572686	1791,71	2,26
TS-199**	<i>S. cheesmaniae</i>	SRR1572688	3981	3,29
TS-207	<i>cheesmaniae</i>	SRR1572689	4171.3	3.1
TS-217	<i>cheesmaniae</i>	SRR1572690-91	2045.5	2.44