UNIVERSIDAD DE ALMERÍA

Genetic and molecular control of monoecy stability, fruit set and fruit quality in watermelon

Encarnación Aguado Donaire

Almería, 2020

Control genético y molecular de la estabilidad de la monoecia, el cuajado y la calidad del fruto en sandía

Genetic and molecular control of monoecy stability, fruit set and fruit quality in watermelon

Encarnación Aguado Donaire

ALMERÍA, June 2020

Ph.D. DISSERTATION

Genetic and molecular control of monoecy stability, fruit set and fruit quality in watermelon

Control genético y molecular de la estabilidad de la monoecia, el cuajado y la calidad del fruto en sandía

Área de Biología y Geología

Departamento de Biología y Geología

Escuela Superior de Ingeniería

Universidad de Almería

Fdo. Encarnación Aguado Donaire

ALMERIA, June 2020

Dpto. de Biología y Geología

Área de Genética

Universidad de Almería

Dr. Manuel Jamilena Quesada, Doctor del Área de Genética de la Universidad de Almería y Dra. Susana Manzano Media.

HACEN CONSTAR:

Que el presente trabajo se ha realizado bajo nuestra dirección y recoge la labor realizada por la Ingeniero Superior Encarnación Aguado Donaire para optar al Grado de Doctor Ingeniero Agrónomo.

Fdo. Susana Manzano Medina

Fdo. Manuel Jamilena Quesada

Almería, junio de 2020

Scientific contributions

Scientific contributions directly related to the doctoral thesis

Peer-Reviewed Articles

Manzano S., **Aguado E.**, Martínez C., Megías Z., García A., Jamilena, M. (2016). The Ethylene Biosynthesis Gene *CitACS4* Regulates Monoecy/Andromonoecy in Watermelon (*Citrullus lanatus*). PLoS One, 11, e0154362. doi:10.1371/journal.pone.0154362.

Aguado E., García A**.**, Manzano S., Valenzuela J. L., Cuevas J., Pinillos V., Jamilena M. (2018). The sex-determining gene *CitACS4* is a pleiotropic regulator of flower and fruit development in watermelon (*Citrullus lanatus*). Plant Reproduction, 1–16. doi:10.1007/s00497-018-0346-1.

Aguado E., García A., Iglesias Y., Romero J., Wehner T.C., Gómez-Guillamón M.L., Picó B, Garcés-Claver A., Martínez C., Jamilena M. (2020). Mapping a *partial andromonoecy* locus in *Citrullus lanatus* using BSA-seq and GWAS approaches. Submitted to Frontiers in Plant Science.

Proceedings and conference communications

Aguado E., Manzano S., García A., Jamilena M. (2016). The andromonoecious allele of CitACS4 reduces seed and fruit set in watermelon. Cucurbitaceae, XIth Eucarpia Meeting on Cucurbit Genetics and Breeding, July 24-28, 2016, Warsaw, Poland, July 24-28. 29-33. ISBN: 978-83-7987-896-3.

Aguado E., García A., Manzano S., Martínez C., Romero J., Cebrián G., Jamilena M. (2018). The inheritance of trimonoecy in *Citrullus lanatus.* 1er Congreso de Jóvenes Investigadores en Ciencias Agroalimentarias, Almería, Spain, Dec 20, 2018.

Aguado E., García A., Manzano S., Romero J., Cebrián G., Jamilena M. (2018). Genetic analysis of trimonoecy in watermelon. Cucurbitaceae, Davis, California, USA, Nov 12-15, 2018.

Aguado E., Manzano S., García A., Cebrián G., Romero J., Valenzuela J. L., Jamilena M. (2018). Análisis genético del papel del gen *CitACS4* en la maduración y calidad de la fruta de sandía. XII Simposio Nacional y X Ibérico sobre maduración y postcosecha, Badajoz, España, Jun 4-7, 2018.

Aguado E., García A., Martínez C., Iglesias Y., Romero J., Cebrián G., Jamilena M. (2019). Generation, selection and characterization of ethylene insensitive mutants in watermelon. VI International Symposium on Cucurbits, Gent, Belgium, Jun 30-July 4, 2019.

Other contributions

Other associated Peer-Reviewed Articles

García A., **Aguado E.**, Parra G., Manzano S., Martínez C., Megías Z., Cebrián G., Romero J., Beltrán S., Garrido D., Jamilena M. (2018). Phenomic and Genomic Characterization of a Mutant Platform in *Cucurbita pepo*. Frontiers in Plant Science. 9, 1049. doi:10.3389/fpls.2018.01049.

García A., Manzano S., Megías Z., **Aguado E.**, Martínez C., Garrido D., Rebolloso M.M., Valenzuela J.L., Jamilena M. (2018). Use of mutant platforms to discover novel postharvest fruit-quality traits in *Cucurbita pepo*. Acta Horticulturae. 367–373. doi:10.17660/ActaHortic.2018.1194.53.

García A., **Aguado E.**, Martínez C., Loska D., Beltrán S., Valenzuela J. L., Garrido D., Jamilena M. (2019). The ethylene receptors *CpETR1A* and *CpETR2B* cooperate in the control of sex determination in *Cucurbita pepo*. Journal of Experimental Botany. doi:10.1093/jxb/erz417.

García A., Valenzuela J. L., Manzano S., Cebrián G., Romero J., **Aguado E.**, Garrido D., Jamilena M. (2019). Postharvest fruit quality in ethylene insensitive mutants of zucchini squash. Acta Horticulturae. 217–222. doi:10.17660/ActaHortic.2019.1256.30.

García A., **Aguado E.**, Garrido D., Martínez C., Jamilena M. (2020). Two androecious mutations reveal the crucial role of ethylene receptors in the initiation of female flower development in *Cucurbita pepo.* The Plant Journal. doi: 10.1111/tpj.14846

Manzano S., Megías Z., Martínez C., García A., **Aguado E.**, Chileh T., López-Alonso D., García-Maroto F., Kejnovský E., Široký J., Kubát Z., Králová T., Vyskot B., Jamilena M. (2017). Overexpression of a flower-specific aerolysin-like protein from the dioecious plant *Rumex acetosa* alters flower development and induces male sterility in transgenic tobacco. Plant Journal, 89, 58–72. doi:10.1111/tpj.13322.

Megías Z., Manzano S., Martínez C., García A., **Aguado E**., Garrido D., Rebolloso M.M., Valenzuela J.L., Jamilena M. (2017). Postharvest cold tolerance in summer squash and its

association with reduced cold-induced ethylene production. Euphytica, 213. doi:10.1007/s10681-016-1805-0.

Megías Z., González-Rodríguez L.J., **Aguado E**., García A., Manzano S., Rebolloso M.M., Valenzuela J.L., Jamilena M. (2018). Effect of cold storage time on chilling injury in two zucchini cultivars. Acta Horticulturae, 1194, pp. 479-485. doi: 10.17660/ActaHortic.2018.1194.70.

Megías Z., Manzano S., Martínez C., García A., **Aguado E.**, Garrido D., Rebolloso M. M., Valenzuela J. L., Jamilena M. (2018). Breeding for postharvest cold tolerance in zucchini squash. Acta Horticulturae, 357–362. doi:10.17660/ActaHortic.2018.1194.51.

Resumen

Citrullus lanatus y otras especies de la familia Cucurbitaceae han evolucionado hacia el desarrollo de flores unisexuales femeninas y masculinas a partir de especies con flores hermafroditas. La distribución y ocurrencia de flores femeninas, masculinas, bisexuales y hermafroditas a lo largo de los tallos principales y secundarios definen tres morfotipos sexuales en sandía: monoecia (la planta produce flores masculinas y femeninas), andromonoecia (la planta produce flores masculinas y hermafroditas), y andromonoecia parcial (la planta produce flores masculinas, femeninas, bisexuales y hermafroditas). El objetivo principal de esta tesis doctoral ha sido estudiar el control genético y molecular de la monoecia y la andromonoecia total y parcial en sandía.

Para estudiar el control genético de la andromonoecia completa se realizó un análisis de segregación de tres poblaciones F2 independientes procedentes del cruce entre tres líneas monoicas y una línea andromonoica. Se ha demostrado que la andromonoecia está controlada por un solo gen con dos alelos, *M* que confieren monoecia, y *m* que confiere andromonoecia. El alelo *M* es semi-dominante con respecto al alelo *m*, pues las plantas heterocigotas (*Mm*) tienen un fenotipo parcialmente andromonoico y no solo producen flores femeninas sino también flores bisexuales y hermafroditas. En otras especies de cucurbitáceas, la andromonoecia está causada por mutaciones de falta de función en los genes de biosíntesis de etileno *CmACS7*, *CsACS2* y *CpACS27A* de melón, pepino y calabacín, respectivamente. En sandía hemos clonado y caracterizado el gen homólogo *CitACS4*, que codifica para una enzima ACS tipo III que se expresa predominantemente en flores pistiladas. En la línea andromonoica detectamos una mutación sin sentido en un residuo altamente conservado de CitACS4 (C364W) que cosegrega con el fenotipo andromonoico en las tres poblaciones F2 estudiadas. Estos datos indican que *CitACS4* debe estar involucrado en la biosíntesis de etileno requerida para la inhibición del crecimiento del estambre durante el desarrollo de las flores femeninas. La mutación C364W reduce la producción de etileno en los botones florales pistilados, promoviendo la conversión de las flores femeninas en flores hermafroditas, por lo tanto, de monoecia (genotipo *MM*) en andromonoecia parcial (genotipo *Mm*) o andromonoecia (genotipo *mm*).

De forma paralela hemos estudiado si el gen *CitACS4* podría también estar involucrado en otros caracteres regulados por etileno durante el desarrollo de flores y frutos, incluyendo la

transición de la floración pistilada, el número de flores femeninas por planta, el desarrollo de los órganos florales, y el cuajado y desarrollo de frutos y semillas. Para ello, estos caracteres se han fenotipado en las tres poblaciones segregantes para los dos alelos del gen *CitACS4* (*M*, monoico; *m*, andromonoico). El alelo *m* del gen *CitACS4*, además de conferir el tipo sexual andromonoico, cosegregaba con una transición floral más temprana, un aumento del porcentaje de flores pistiladas por planta, una reducción del crecimiento y maduración de los carpelos y pétalos, así como con un retraso de la maduración y apertura de las flores pistiladas. Se encontró que el alelo de la andromonoecia estaba también asociado con una reducción del cuajado de frutos y semillas, y que esa falta de cuajado no estaba causada por una deficiencia en la polinización. El gen *CitACS4* también afectó las tasas de crecimiento longitudinal y transversal del ovario y fruto, lo que significa que los frutos de plantas andromonoicas (*mm*) eran más redondos que las de las plantas monoicas (*MM*). En su conjunto, todos estos datos demostraban que el locus definido por la biosíntesis de etileno y el gen determinante del sexo *CitACS4* actúa como un regulador pleiotrópico del desarrollo completo de la flor pistilada y el desarrollo temprano del fruto.

En la última parte de este trabajo de tesis, se realizó un estudio de la andromonoecia parcial en sandía. Se ha fenotipado y genotipado (para el locus *M/m*) un panel de 207 accesiones de *Citrullus lanatus* de todas partes del mundo, incluyendo entre ellas cinco líneas puras e híbridos comerciales. Encontramos varias accesiones que a pesar de ser *MM* tenían un fenotipo parcialmente andromonoico en varios lugares y años diferentes. Un análisis de cosegregación entre una variante alélica de *CitACS4* y el fenotipo parcialmente andromonoico, demostró que la aparición de flores bisexuales y hermafroditas en una línea parcialmente andromonoica de sandía no dependía de *CitACS4*, sino que estaba conferida por un gen recesivo no ligado al que hemos llamado *pa*. Se llevaron a cabo dos estrategias diferentes para mapear el gen *pa* en el genoma de *C. lanatus*: la secuenciación masiva de bulks segregantes (BSA-seq) y el estudio de asociación del genoma completo (GWAS). El estudio BSA-seq se realizó utilizando el ADN de dos grupos contrastantes de una población F2, el grupo monoico *M* y el grupo parcialmente andromonoico PA, cada uno generado a partir de ADN de 20 plantas F2 monoicas o parcialmente andromonoicas, respectivamente. Los análisis de GWAS se realizaron con datos de genotipado por secuenciación GBS en 122 accesiones procedentes del USDA. La combinación de las dos estrategias indicaba que el locus *pa* mapea en una región genómica que se expande entre 32.24-36.44 Mb en el cromosoma 1 de sandía. El mapeo fino redujo el locus *pa* a una región genómica de 867 Kb

que contiene 101 genes. Finalmente, se seleccionaron varios genes candidatos, por su función en la biosíntesis y la señalización de etileno, así como en el desarrollo de las flores y la determinación del sexo, pero también por el impacto de las mutaciones detectadas entre los dos grupos de ADN secuenciados.

Palabras clave: etileno, determinación del sexo, monoecia/ parcial andromonoecia/ andromonoecia/ cuajado, polinización, secuenciación masiva de bulks segregantes (BSA-seq), estudio de asociación del genoma completo (GWAS), genotipado por secuenciación (GBS), mapeo fino.

Summary

As in other species of the Cucurbitaceae family, the different sex morphotypes of *Citrullus lanatus* can be differentiated according to the distribution and occurrence of male, female, bisexual and hermaphrodite flowers on the main and secondary shoots. Depending on the distribution of the four types of flowers, the plants can be monoecious (producing male and female flowers), andromonoecious (producing male and hermaphrodite flowers), or partially andromonoecious (producing male, female, bisexual and hermaphrodite flowers) within the same plant. The main objective of this thesis is to analyze the genetic and molecular control of monoecy, andromonoecy and partial andromonoecy in watermelon.

To study the genetic control of andromonoecy, a segregation analysis of three independent F2 populations derived from the crosses between three monoecious lines and one andromonoecious line was performed. It has been demonstrated that the andromonoecious trait is controlled by a single gene with two alleles, the monoecious M allele, and the andromonoecious *m* allele. The *M* allele is semidominant respect to *m*, in such a way that the heterozygous *Mm* genotype is partially andromonoecious. In other cucurbit species andromonoecy is conferred by mutations in the ethylene biosynthesis genes *CmACS7*, *CsACS2* and *CpACS27A* in melon, cucumber, and squash, respectively. The gene with the highest homology in watermelon is *CitACS4*, which encodes for an ACS type III enzyme that is predominantly expressed in pistillate flowers. In the andromonoecious line, we detected a missense mutation in a very conserved residue of CitACS4 (C364W) that cosegregated with the andromonoecious phenotype in three independent F2 populations. These data indicated that *CitACS4* is likely to be involved in the ethylene biosynthesis required for stamen arrest during the development of female flowers. The C364W mutation would reduce the production of ethylene in pistillate floral buds, promoting the conversion of female into hermaphrodite flowers, and therefore of monoecy (genotype *MM*) into partial andromonoecy (genotype *Mm*) and andromonoecy (genotype *mm*).

In parallel we have studied whether the *CitACS4* gene could be also involved in other ethylene-regulated traits during flower and fruit development, including pistillate flowering transition and the number of female flowers per plant, the development of floral organs other than stamens, as well as fruit and seed set, and fruit development. A linkage analysis approach was performed in three independent F2 populations segregating for the two alleles

of the gene (*M*, monoecious; *m*, andromonoecious), and the different traits under study. The CitACS4 *m* allele not only co-segregated with andromonoecy, but also with earlier pistillate transition, an increased number of pistillate flowers per plant, and a slower growth and maturation of petals and carpels, which delayed anthesis time in hermaphrodite flowers. The *m* allele was also found to be linked to a reduced fruit set, which was not caused by a deficiency in pollination or fertilization. The gene also affected the longitudinal and transverse growth rates of the ovary and fruit, which means that fruits from andromonoecious plants (*mm*) were rounder than those from monoecious (*MM*) ones. Taken together, these data indicate that the locus defined by the ethylene biosynthesis and sex determining gene *CitACS4* acts as a pleiotropic regulator of the complete development of the pistillate flower and the earlier development of the fruit.

In the last part of this thesis we have performed genetic and molecular analysis of the partially andromonoecious phenotypes of watermelon. We phenotyped and genotyped, for the *M/m* locus, a panel of 207 *Citrullus lanatus* accessions from all parts of the world, including 5 inbreds and hybrids, and found several accessions that were repeatedly phenotyped as partially andromonoecious in several locations and different years, despite being *MM*. A cosegregation analysis between a SNV in *CitACS4* and the partially andromonoecious phenotype, demonstrated that the occurrence of bisexual and hermaphrodite flowers in a partially andromonoecious line is not dependent on *CitACS4*, but conferred by an unlinked recessive gene which we called *pa*. Two different strategies were performed to map the *pa* gene in the genome of *Citrullus lanatus*: bulk segregant analysis sequencing (BSA-seq) and genome wide association analysis studies (GWAS). The BSA-seq study was performed using two contrasting bulks, the monoecious M-bulk and the partially andromonoecious PA-bulk, each one generated by pooling DNA from 20 F2 plants possessing the most contrasting phenotypes for monoecy and PA, respectively. For GWAS, 122 accessions from USDA gene bank, already re-sequenced through genotyping by sequencing (GBS), were used. The combination of the two strategies indicates that *pa* maps onto a genomic region expanding across 32.24-36.44 Mb in chromosome 1 of watermelon. Fine mapping narrowed down the *pa* locus to a 867 Kb genomic region containing 101 genes. Finally, several candidate genes were selected, for their function in ethylene biosynthesis and signaling, as well as in flower development and sex determination, but also by the impact of the SNPs and indels differentially detected in the two sequenced bulks.

Keywords: ethylene, sex determination, monoecy/ partially andromonoecy/ andromonoecy, fruit set, pollination, bulk segregant analysis sequencing (BSA-seq), genome wide association analysis studies (GWAS), genotyping by sequencing (GBS), fine mapping.

Abbreviation list

A

A: Andromonoecious

ABA: Abscisic acid

ACC: 1-aminocyclopropane-1-carboxylic acid

ACO: ACC Oxidase

ACS: ACC Synthase

A.D.: Anno Domini

AF: Allele Frequency

AI: Andromonoecy Index

AM: Apical meristem

Andro: Andromonoecious

AUX: Auxins

AVG: Aminoethoxyvinylglycine

B

Bi: bisexual flowers

BGH-CITA: Centro de Investigación y Tecnología Agroalimentaria de Aragón

Bp: Base pair

BR: brassinosteroids

BSA-seq: bulk segregant analysis sequencing

C

C: cysteine

°C: Degrees Celsius

cDNA: complementary DNA

CDPK: Calcium Dependent Protein Kinase

CM: CentiMorgan

CO_2: Carbon Dioxide

COMAV: Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana

CTL1: Chitinase Like 1

CTR1: Constitutive Triple Response 1

CuGenDB: Cucurbit Genomics Database

D

DNA: Deoxyribonucleic acid

DNAse: Deoxyribonuclease

DP: Read Depth

DPA: Days Post Anthesis

DDF1-1: Dwarf and Deformed Flowers 1-1

E

EBF: EIN3 Binding F-BOX

EDF: Ethylene-Response DNA-binding Factors

EIN2: Ethylene Insensitive 2

EIN3: Ethylene Insensitive 3

EIN4: Ethylene Insensitive 4

ER: Endoplasmic reticulum

ERF: Ethylene Response Factor

ERS: Ethylene Response

EST: Expressed Sequence Tag

ET: Ethylene

F

FD: Fruit Diameter

Fe: Female Flowers

FL: Fruit Length

FM: Floral Meristem

FS: Fruitlet Shape

Fo: Fertilized Ovule

G

GA: Gibberellin

GATK: Genome Analysis Toolkit

GBS: Genotyping by Sequencing

GLM: Generalized Linear Model

GPG: Germinated Pollen Grain

GQ: Genotype Quality

GRXs: Glutaredoxin Family Protein

GWAS: Genome Wide Association Analysis Studies

H

H: hours

I

IDA: Inflorescence Deficient in Abscission

K

K: Kinship Matrix

Kb: Kilobase

L

L: Litres

Le: Leaves

LOD: logarithm of the odds

M

M: monoecious

Ma: Male Flowers

MAPK: Mitogen Activated Protein Kinase

MAPKKK: MAPK Kinase Kinase

Mb: Megabase

1-MCP: 1-methylcyclopropene

ml: milliliters

MLM: linear mixed model

mm: millimeters

Mono: monoecious

mRNA: messenger RNA

Myr: Million years

N

n: number

NCBI: National Center for Biotechnology Information

ng: nanograms

NGS: Next Generation Sequencing

NPGS: USDA National Plant Germplasm System

O

O_2: oxygen

P

PA: Partially Andromonoecious

PCA: Principal Component Analysis

PCR: Polymerase Chain Reaction

Pg: Pollen Grain

PIs: Plant Introductions

PLP: Pyridoxal 5`Phosphate

Pt: Pollen Tubes

Q

q-PCR: quantitative PCR

Q: Corrected for Population Structure

Q-Q: Quantile-quantile

QTL: Quantitative Trait Locus

R

RNA: Ribonucleic Acid

RT-PCR: Reverse transcription polymerase chain reaction

S

S-AdoMet: S-adenosylmethionine

s: Second

SAM: S-adenosyl methionine

SNP: Single Nucleotide Polymorphism

SNV: Single Nucleotide Variant

St: Stem

STS: Silver Thiosulphate

Subsp.: Subspecies

T

Tª: Temperature

TIR1: Transport Inhibitor Response 1

U

UAL: University of Almería

UCO: Unequal Crossing Over

U.S. National Plant Germplasm System

USDA: United States Department of Agriculture

UTR: Untranslated region

V

W: Tryptophan

VCF: Variant Call Format

WGS: Whole Genome Sequencing

WOX: Wuschel-related homeobox

WT: Wild Type

Table of contents

1. Introduction

1.1. Watermelon

1.1.1. Taxonomy

The Cucurbitaceae family includes numerous crops, such as cucumber (*Cucumis sativus*), bitter gourd (*Momordica charantia*), watermelon (*Citrullus lanatus*), citron (*Citrullus amarus*), melon (*Cucumis melo*), zucchini (*Cucurbita pepo*), and bottle gourd (*Lagenaria siceraria)* (Schaefer and Renner, 2011; Renner and Schaefer, 2016; Chomicki et al., 2019).

The Cucurbitaceae family contains 1000 species divided into 96 genera (Renner and Schaefer, 2016). To understand the phylogenetic relationships in this family, Schaefer and Renner (2011) made an extremely comprehensive molecular phylogenetic analysis of the Cucurbitaceae. They used ribosomal DNA, two regions of the mitochondrial genome, and nine regions of the plastidial genome of 664 cucurbit species representing all the genera, to update phylogenetic relationships among species of this family. The Cucurbitaceae family can be divided into two subfamilies: Zanonioideae and Cucurbitoideae, the latter comprising some of the most significant tribes and genera. Among them are the Cucurbiteae and Benincaseae tribes that include the genera considered to be the most important from an economic viewpoint. The *Cucurbita* genus belongs to the Cucurbiteae tribe, while the *Cucumis* and *Citrullus* genera are classified within the Benincaseae tribe (Renner and Schaefer, 2016).

Chomicki and Renner (2015) proposed separating the *Citrullus* genus into seven separate species, as genomic sequencing revealed that the citron, egusi and dessert watermelons differed significantly in genome organization (Guo et al., 2013; Reddy et al., 2013; Chomicki and Renner, 2015). Table 1.1 shows the seven species which currently comprise the *Citrullus* genus: *Citrullus ecirrhosus* Cogn., *Citrullus rehmii* De Winter and *Citrullus naudinianus* (Sond.) Hooker f. (Chomicki and Renner, 2015), which are non-cultivated (Paris, 2015); the rarely cultivated *Citrullus colocynthis* (L.) Schard.; the citron watermelon (*Citrullus amarus* Schrad), the egusi watermelon (*Citrullus mucosospermus* (Fursa) Fursa), which are cultivated regionally; and the dessert watermelon (*Citrullus lanatus*), which is a worldwide cultivated species (Renner et al., 2014; Paris, 2015).

Before Chomicki and Renner's (2015) proposed separation, only four species were considered to be part of the *Citrullus* genus, given that cultivated citron, egusi and dessert watermelons were previously classified as subspecies within the species *C. lanatus* (Thunb.)

Matsum. & Nakai (Table 1.1). The nomenclature was confusing; this classification complexity possibly being due to the intercross ability of the species (Guo et al., 2013; Levi et al., 2013; Reddy et al., 2014, 2015). As an example, *C. lanatus* can cross with *C. mucosospermus* and *C. amarus*; however, *C. mucosospermus* can also cross with *C. amarus*. Although the progeny of *C. lanatus* and *C. mucosospermus* are fertile (Levi et al., 2011) (Table 1.1), populations derived from crosses between *C. amarus* and either *C. lanatus* or *C. mucosospermus* produce skewed (non-Mendelian) segregation ratios for most genomic regions (Levi et al., 2002), probably due to genomic differences observed in *C. amarus* when compared to *C. lanatus* or *C. mucosospermus* (Guo et al., 2013).

Table 1.1. Botanical classification of the genus *Citrullus*.

Common Name	***Citrullus* genus (classification after Renner et al. 2014, and Chomicki and Renner, 2015)**	**Species of *Citrullus* (classification after Fursa, 1972)[c]**	**Distribution**	**Use**	**Crossability**
Dessert watermelon	*Citrullus lanatus* [b]	*C. lanatus*, ssp. *vulgaris* (Schrad). Fursa: cordophanus Ter-Avan	World[d]	Fruit flesh[d]	*Citrullus mucosospermus* and *Citrullus amarus* (Levi et al., 2002, 2011)
Egusi watermelon	*Citrullus mucosospermus (Fursa) Fursa* [b]	C. lanatus ssp. *lanatus*	Regional[d]	Oily seeds[d]	*Citrullus amarus* and *Citrullus mucosospermus* (Levi et al., 2002, 2011)
Citron watermelon	*Citrullus amarus Schrad*[b]	*C. lanatus* var. *citroide*s by L.H. Bailey	Regional[d]	Fruit flesh[d]	*Citrullus lanatus* and *Citrullus mucosospermus* (Levi et al., 2002, 2011)
Colocynth	*Citrullus colocynthis* (L.) Schrad [b]		Regional[d]	Sparingly cultivated[d] Medicinal properties Oil seeds [e]	-
	C. rehmii De Winter [a]		-	Uncultivated[d]	-
	C. ecirrhosus Cogn.[a]		-	Uncultivated[d]	-
	*C.naudinianus (*Sond.)[a]		-	Uncultivated[d]	-

[a] Chomicki and Renner (2015)
[b] Renner et al. (2014)
[c] Fursa (1972)
[d] Paris (2015)
[e] Hussain et al. (2014)

1.1.2. Origin and distribution

Cucurbitaceae are believed to have originated in Asia during the Late Cretaceous period (ca 70 Myr ago). This is borne out by the observation that India contains more lineages of Cucurbitaceae than any other similar-sized geographical area (Schaefer et al., 2009). On the other hand, the genus *Citrullus* seems to be originated in Africa and belongs to the tribe Benincaseae, one of the youngest tribes of the family (Whitaker and Davis, 1962; Maynard and Maynard, 2000; Schaefer et al., 2009; Zohary et al., 2012).

The dessert watermelon (*C. lanatus*) is the result of many years of domestication and selection for desirable fruit quality traits. Previously, as mentioned above, the origins of this watermelon from within the African continent had not been clearly defined, due to the phenotypical variations between the citron, egusi and dessert watermelon varieties. The confusing taxonomy used in the past, together with the ease with which these varieties can cross, led to the proposal of the following four hypotheses as to the origin of this particular watermelon (Paris, 2015).

- Dessert watermelon descends from the colocynth (*Citrullus colocynthis*) of northern Africa (Singh, 1978; Sain et al., 2002; McCreight et al., 2013).
- Dessert watermelon descends from the citron watermelon (*Citrullus amarus*) of southern Africa (Robinson and Decker-Walters, 1997; Maynard and Maynard, 2000; Rubatsky, 2001).
- Dessert watermelon derives the egusi watermelon (*Citrullus mucosospermus*) of western Africa (Guo et al., 2013).
- Dessert watermelon originated and was first domesticated in northeastern Africa (Paris, 2015).

Paris (2015) reviewed genetic, archaeological and historical data, together with a phylogenetic analysis of *Citrullu*s (Chomicki and Renner, 2015), and rejected the first and second hypotheses, thus leaving west Africa and northeast Africa as possible centers of origin of dessert watermelon. Recently, Chomicki et al. (2019) and Guo et al. (2019) observed that *C. mucosospermus* and *C. lanatus* are closely related, having a shared ancestry. This suggests that they were derived from the same ancestral population and perhaps domesticated for different purposes, for example, for seed consumption and for fruit flesh.

The most plausible hypothesis is that the dessert watermelon, *C. lanatus*, is native to northeastern Africa. Wild *C. lanatus* populations in Sudan, reported as bearing small, round, inferior-quality fruits, are living representatives of the wild ancestor of today`s dessert watermelon (Paris, 2015). Hence, it was probably domesticated in northeastern Africa, Egypt, and Sudan over 4000 years ago. Later, it could have spread to the deserts of sub-Saharan Africa, the Middle East and Asia, as a source of water and nutrients, by nomadic peoples (Paris, 2015). African watermelons were introduced into China around 1100 A.D. (Anno Domini) and into India around 800 A.D. (Zhao et al., 2015). The Moorish inhabitants of what is now southern Spain are credited with their introduction into Europe, as there are records of this in Córdoba dated 961 A.D. Watermelons spread into other parts of Europe, but slowly; perhaps mainly because the summers are not generally hot enough for good yields. Spanish colonists introduced them into the New World; watermelons were being grown in Florida by 1576. By 1650, they were common in many parts of Latin America and also Brazil. They also spread throughout northeastern North America, where they were found abundantly prior to 1960 in British and Dutch colonies (Sauer, 1993). The many years of domestication and selection in order to obtain desirable fruit quality have led to a narrow genetic base among watermelon cultivars, probably initially due to the farming methods of the first agrarians and thereafter by the methods of breeders (Levi and Thomas, 2001; Levi et al., 2001; Guo et al., 2013). This narrow genetic base and the loss of resistance alleles may have contributed to the susceptibility of today's cultivated watermelon to a wide range of pests and diseases (Levi et al., 2001).

Phylogenomic data suggested that the egusi watermelon, *Citrullus mucosospermus*, was native to Western Africa, and was domesticated by its nutritious seeds (Achigan-Dako et al., 2015), which belonged to a different gene pool to that of the dessert watermelon (Chomicki et al., 2019) (Figure 1.1). The species *Citrullus amarus*, which originated in Southern Africa (Chomicki and Renner, 2015; Renner et al., 2017), is referred to as the 'citron melon' in its wild form in South Africa. It is also an important crop in the Mediterranean region (Chomicki et al., 2019) and is cultivated because of its edible but typically hard flesh, which is often cooked or pickled (Bush, 1978). The colocynth, *Citrullus colocynthis*, has not been domesticated (Figure 1.1) (Chomicki et al., 2019). It is native to the deserts and semi-arid regions of northern Africa and southwestern and central Asia, eastwards to Afghanistan, Pakistan, and India; it also native to some Mediterranean islands (Jeffrey, 1967; Burkill, 1978; Dane and Liu, 2007; Paris, 2015). *C. colocynthis* is known as the ´bitter apple` and is

cultivated for its numerous medicinal properties and the oil from its seeds (Hussain et al., 2014).

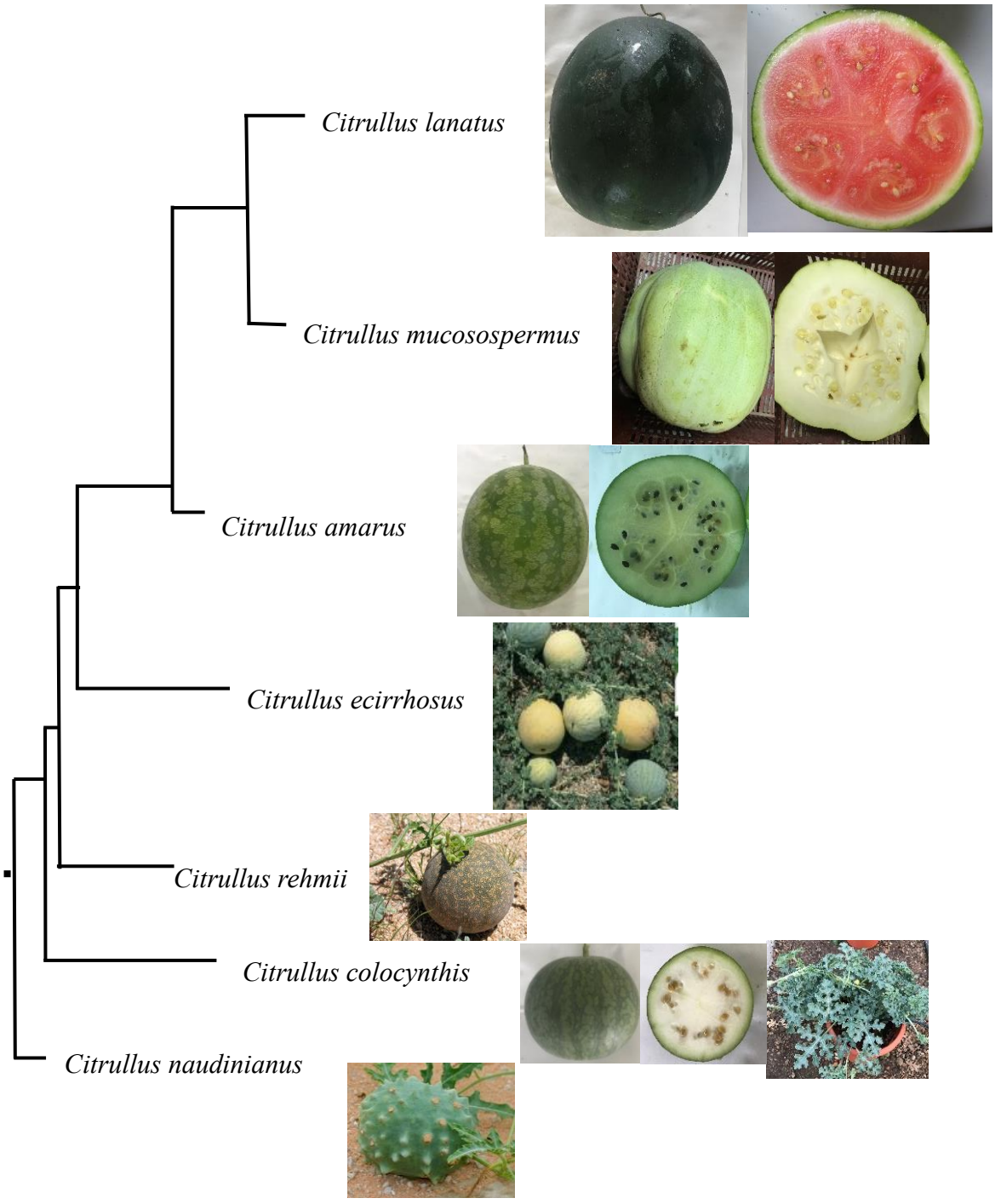


Figure 1.1. The phylogeny and domestication of the genus *Citrullus* using nuclear and chloroplastidial markers (Renner et al., 2019). Images of *Citrullus ecirrhosus*, *Citrullus rehmii* and *Citrullus naudinianus* are taken from Renner et al. (2017).

1.1.3. Phenotype of the watermelon plant and fruit

Watermelon plants require extensive space and a long sunny growing season in order for their fruits to mature. They are multiple-branched, tendril-bearing, and have extensive root systems. The leaf is pinnatifid and tendrils, flowers, and fruits develop from its leaf axils (Figure 1.2). Watermelons develop individual flowers which are 2–3 cm in diameter, with five pale yellow petals. The flowers open in the early morning but begin to wither in the mid-afternoon; subsequently, they do not re-open (Figure 1.2). The fruits are harvested at maturity, a month or more after anthesis.

Fruit size can vary, the weight of commercially grown fruit ranging from 3 to 13 kg. The fruit flesh has various textures, for instance; soft, crisp, or hard, and can be red, pink, orange, yellow, green, or white. Rind color ranges from light to dark green, and the fruit can be spherical or a long or short oval shape. There are many varieties of seeds, as black, brown, tan, white, yellow, red, purple, green and orange, and can be patterned with a second color. Throughout most of the world, watermelons are valued as a sweet, juicy, refreshing dessert fruit (Paris et al., 2013).

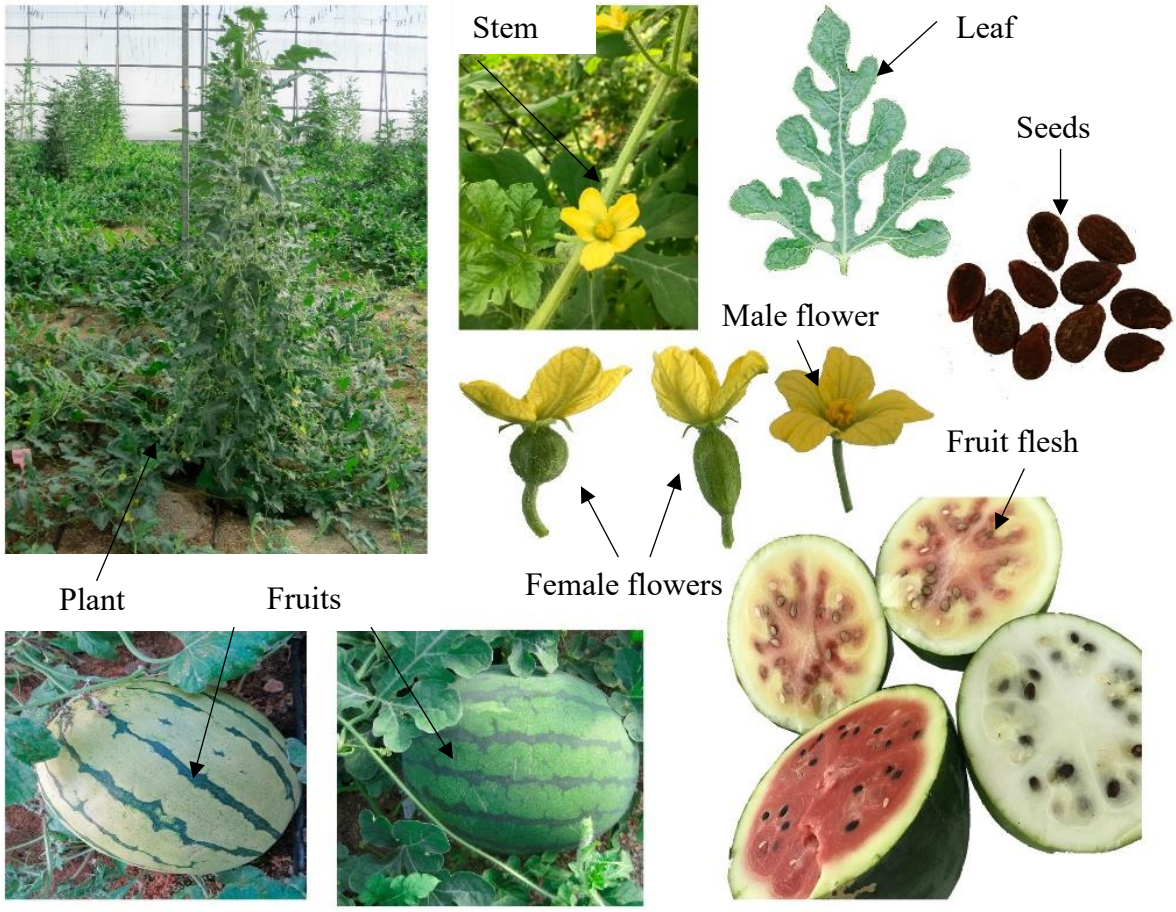


Figure 1.2. Phenotype of watermelon plant, leaf, flowers, seeds, and their fruits.

1.1.4. Genomic resources

Citrullus lanatus is a diploid species possessing eleven pairs of chromosomes (2n = 2x = 22). Recently, two genomes of watermelon were sequenced, one corresponding to the cultivar 'Charleston Gray', which is an American dessert watermelon, and the other to that of '97103', an East Asian cultivar (Guo et al., 2019; Wu et al., 2019). The recently sequenced and assembled 'Charleston Gray' genome extends to about 396.4 Mb, representing 94.6% of the complete genome of watermelon (Wu et al., 2019). On the other hand, the '97103' genome is composed of 31 scaffolds with a total size of 362.7 Mb (99.3% of the assemblage), and includes 22,596 high-confidence genes (Guo et al., 2019). The sequencing of these genomes provides a solid reference base for watermelon research and genetic improvement.

The watermelon genomic, transcriptomic and genetic data is available, together with those of melon, cucumber, pumpkin, squash and gourds in the Cucurbit Genomics Database (CuGenDB; http://cucurbitgenomics.org). This is therefore a platform for cucurbit research which will facilitate the storing, mining, analyzing, integration, and dissemination of these datasets. The database currently contains all available genomes and expressed sequence tag (EST) sequences, genetic maps, and transcriptome profiles for cultivated cucurbit species, and also includes sequence annotations, biochemical pathways and comparative genomic analysis results such as synteny blocks and homologous gene pairs between different cucurbit species (Zheng et al., 2019).

Recently, 1365 watermelon accessions (plant introductions, PIs), maintained at the U.S. National Plant Germplasm System, were genotyped using genotyping by sequencing (GBS). The sequencing data, derived from three watermelon species (*C. lanatus*, *C. mucosospermus* and *C. amarus*) collected from around the world, are also available in CuGenDB. Approximately 25.000 high quality single nucleotide polymorphisms (SNPs) were extracted from the GBS dataset using the genome of the cultivar 'Charleston Gray' as the reference genome. Phylogenetic analyses using these SNPs found that *Citrullus amarus* is separate from the closely-related species *C. lanatus* and *C. mucosospermus* (Wu et al., 2019). The identified SNPs are available for the identification and selection of genomic regions associated with important agronomic traits, including fruit quality attributes and genetic resistance to pests and diseases (Wu et al., 2019). Even more recently, 414 accessions from the seven extant species of *Citrullus* have been sequenced, and their polymorphic regions proved to be valuable for genome wide association analysis of key fruit traits (Guo et al.,

2019). The mentioned resources provide a genomic framework for future germplasm usage and watermelon improvement.

1.2. Sex expression and sex determination in cucurbits

Many of the species in the Cucurbitaceae family are able to develop unisexual female flowers (those with no stamen) and male ones (those with no carpel). They can also develop bisexual flowers (undeveloped stamens and fully developed carpels) and hermaphrodite ones (fully developed stamens and carpels) on the same plant or on different ones (Figure 1.3A).

The occurrence and distribution of these unisexual and hermaphrodite flowers on the plant results in the sex morphotypes found in this group of species: monoecy (separate male and female flowers on the same plant), andromonoecy (separate male and hermaphrodite flowers on the same plant), trimonoecy (male, female and hermaphrodite flowers on the same plant), subgynoecy (male flowers at early nodes, followed by female flowers only), gynoecy (only female flowers), gynomonoecy (female and hermaphroditic flowers), hermaphrodite (only hermaphrodite flowers) and androecy (only male flowers) (Figure 1.3B) (Rosa, 1928; Robinson and Decker-Walters, 1997; Trebitsh et al., 1997; Maynard, 2001; Ferreira et al., 2002). Despite this variety of sex phenotypes, the most common one in horticultural cucurbits, that is, *Cucumis sativus*, *Cucumis melo*, *Citrullus lanatus*, and *Cucurbita pepo*, is monoecy. In the basal nodes of monoecious cultivars, the plant produces male flowers, but after the transition to female flowering, the plant alternates the production of both female and male flowers. The distribution of male and female flowers varies between species or cultivars (Li et al., 2019). Slight variations on these sexual morphotypes have also been found in some species. Thus, the subgynoecy observed in some watermelon, zucchini and cucumber cultivars is characterized by plants showing an early transition to female flowering but no male flower production thereafter, while the trimonoecy or partially andromonoecy of different cucurbits is characterized by the production of male, female and bisexual flowers in response to high temperatures (Galun, 1961; Kubicki, 1969d, 1969e, 1969b, 1969a; Trebitsh et al., 1997; Martínez et al., 2014; Ji et al., 2015; Li et al., 2020b).

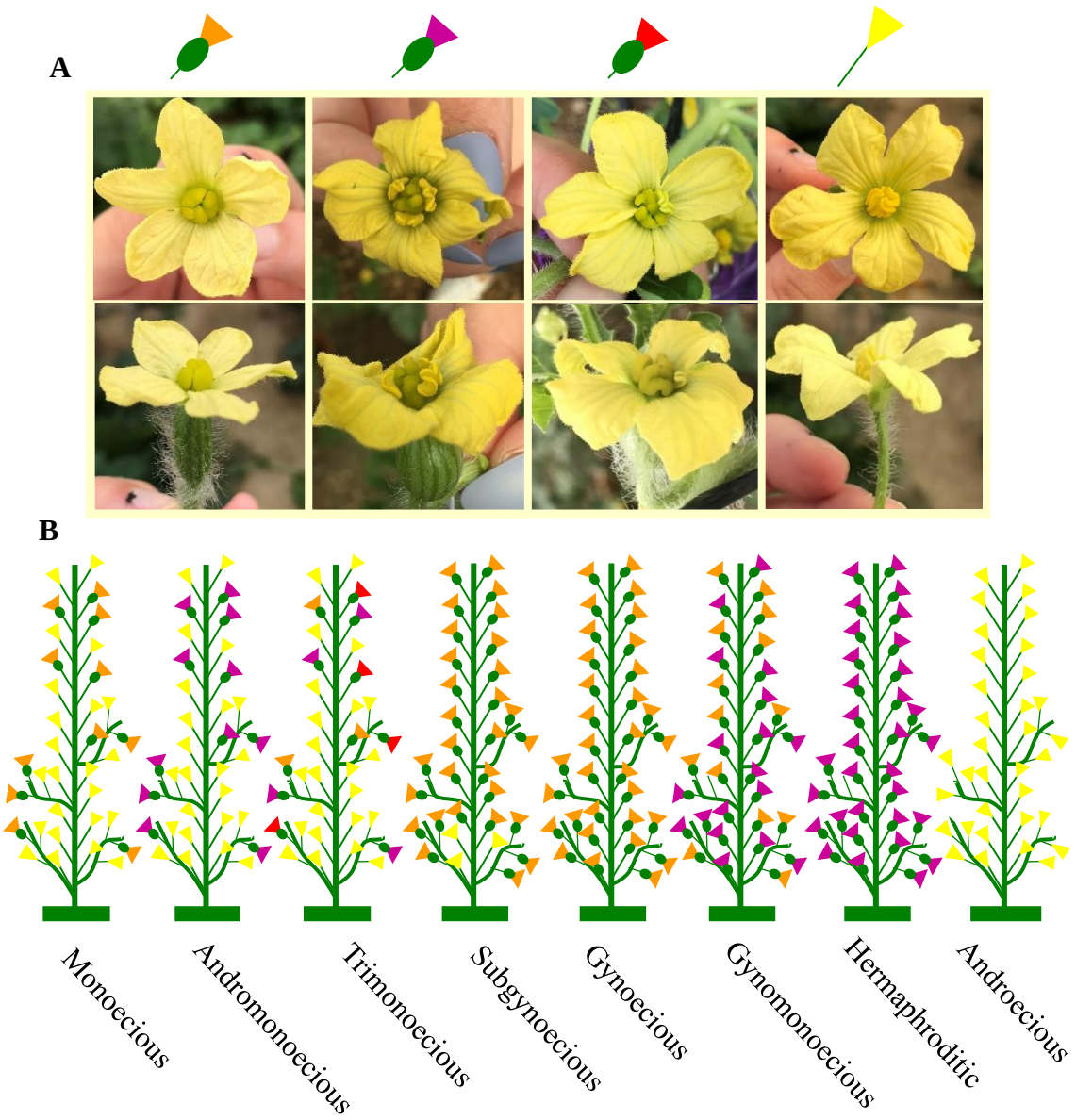


Figure 1.3. Sex expression in cucurbits. (**A**) Flower phenotypes in *C. lanatus*. From left to the right: (1) Female flower (green ovary and orange petals); (2) hermaphrodite flower (green ovary and purple petals); (3) bisexual flower (green ovary and red petals); (4) male (only yellow petals). (**B**) Schematic diagram of 8 different sex morphotypes found in the species of the Cucurbitaceae family, defined by the occurrence and distribution of each flower type on the plant.

This diversity of sex morphotypes makes the cucurbit an exceptional model to the study of the genetic and molecular mechanisms that regulate sex determination in plants. The original sex morphotype of cucurbits is thought to be the hermaphrodite type, and changes occurred later during evolution in response to environmental factors that caused the diversity of sex morphotypes observed nowadays (Megharaj et al., 2017). Starting from the primitive hermaphrodite type, different mutations have led to intermediate sex morphotypes such as

andromonoecious, trimonoecious and gynomonoecious, and later to monoecious, androecious and gynoecious (Robinson and Decker-Walters, 1997; Megharaj et al., 2017). Finally, the separation of male and female flowers in different individual plants led to dioecy. In the literature, there are cucurbits described as hermaphrodite species (i.e. Ridge gourd), as well as dioecious species (i.e. Pointed gourd, Ivy gourd). Most of the cultivated cucurbits, however, show an intermediate sex phenotype. Many of them are monoecious (cucumber, melon, pumpkin, summer squash, winter squash, watermelon, and gourds), but some mutant variants also resulted in types such as gynoecious (cucumber, melon, watermelon, bitter and ridge gourds), androecious (cucumber and melon), andromonoecious (cucumber, squash, melon and watermelon), gynomonoecious (cucumber, melon and ridge gourd), and trimonoecious (cucumber and watermelon) (Martínez et al., 2014; Ji et al., 2015; Megharaj et al., 2017).

1.2.1. Environmental and hormonal factors controlling sex determination

It is known that cultivated cucurbits are able to modify their sex expression in response to seasonal conditions such as temperature and photoperiod. Whereas winter conditions with short days, low light intensity and low night temperatures promote pistillate flower production, summer conditions increase staminate flowers production (Rudich et al., 1972; Rudich and Peles, 1976; Wien, 1997; Wien et al., 2002; Peñaranda et al., 2007; Li et al., 2018). Summer conditions delay pistillate flower transition in squash, and promote the conversion of female flowers into bisexual ones (Manzano et al., 2009, 2010; Martínez et al., 2014). High temperatures promote the production of bisexual flowers instead of female ones in watermelon (Manzano et al., 2014).

Besides environmental conditions, phytohormones also regulate sex determination in the Cucurbitaceae family. Gibberellins (GA_3), auxins, brassinosteroids and most especially ethylene are known to be essential regulators of flower development (Rudich et al., 1972; Trebitsh et al., 1987; Wien, 1997; Girek et al., 2013; Manzano et al., 2013, 2014; Zhang et al., 2017).

Gibberellins and auxins promote the production of male and female flowers respectively, although their function seems to be mediated by ethylene (Shannon and De La Guardia, 1969; Rudich et al., 1972; Trebitsh et al., 1987; Wien, 1997). The effect of the gibberellins (GA_3) on sex expression depends on the species and the sex morphotype. In the monecious

Momordica charantia, GA_3 favors pistillate flower production (Girek et al., 2013), while in monoecious and gynoecious lines of *C. sativus* and monoecious and andromonoecious lines of *C. melo*, the hormone induces plant masculinization, increasing the number of male flowers and delaying the occurrence of the first pistillate (female or bisexual) flower (Peterson and Anhder, 1960; Bukovac and Wittwer, 1961; Malepszy and Niemirowicz-Szczytt, 1991; Girek et al., 2013; Zhang et al., 2017). Auxins promote feminization in both monoecious and gynoecious lines of *C. sativus* (Shannon and De La Guardia, 1969; Trebitsh et al., 1987), as well as in monoecious lines of *C. melo* (Laibach and Kribben, 1950; Byers et al., 1972). In monoecious and andromonoecious watermelons GA_3 increases the percentage of male flowers, thereby delaying the occurrence of the first pistillate flower (female, bisexual or hermaphrodite flower) (Figure 1.4). However, the application of GA_3 does not promote masculinization in gynoecious and hermaphroditic watermelon plants, possibly due to differing levels of sensitivity to GA_3 (Zhang et al., 2017).

The role of brassinosteroids on cucurbit sex determination has not been thoroughly studied. The external application of brassinolide onto cucumber, melon and zucchini reduces the number of male flowers in the initial phase of development and promotes the female flowering transition on the main shoot (Papadopoulou and Grumet, 2005). Since brassinolide induces the production of ethylene, it has been proposed that the effect of external brassinosteroids application upon the sex expression of cucumber is mediated by ethylene (Papadopoulou and Grumet, 2005). The differential effect of the anti-brassinosteroid compound brasinazole in *C. pepo* genotypes contrasting for ethylene insensitivity, also indicates that ethylene production and perception could mediate the effects of brassinosteroids on cucurbit sex determination (Manzano et al., 2011).

Undoubtedly, ethylene is the key hormone regulating sex determination in the Cucurbitaceae family. Many works which refer to external applications, internal hormone content and genetic cues, point out the importance of ethylene in the sex determination of cucurbits. This hormone regulates the transition to pistillate flowering, the balance between male and female flowers in the plants, and the arrest of stamen primordia during the development of female flowers. Thus, applications of anti-ethylene biosynthesis agent (aminoethoxyvinylglycine, AVG) and the anti-ethylene perception agent (silver thiosulphate, STS) increase the number of male flowers per plant (Rudich et al., 1969; Byers et al., 1972; Den Nijs and Visser, 1980; Owens et al., 1980; Payán et al., 2006). Contrary to this, ethephon, an ethylene-releasing agent, has a feminizing effect upon differing sex morphotypes of melon, cucumber and

zucchini squash, reducing the initial male phase of development and increasing the female to male flower ratio (Rudich et al., 1969; Robinson et al., 1970; Manzano et al., 2011, 2014; Martínez et al., 2014; Zhang et al., 2017).

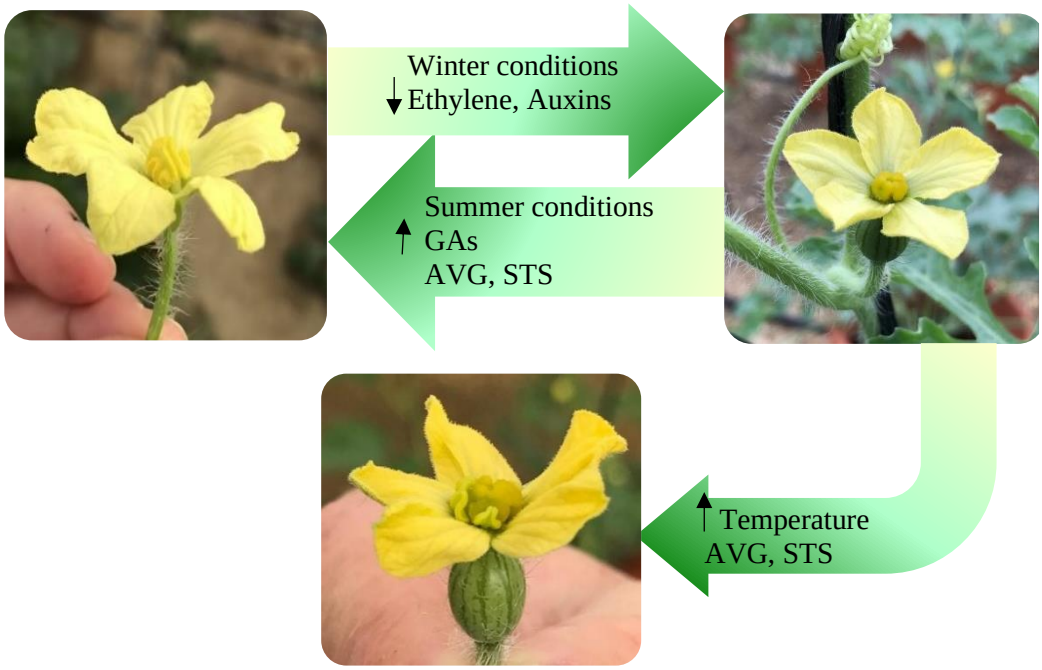
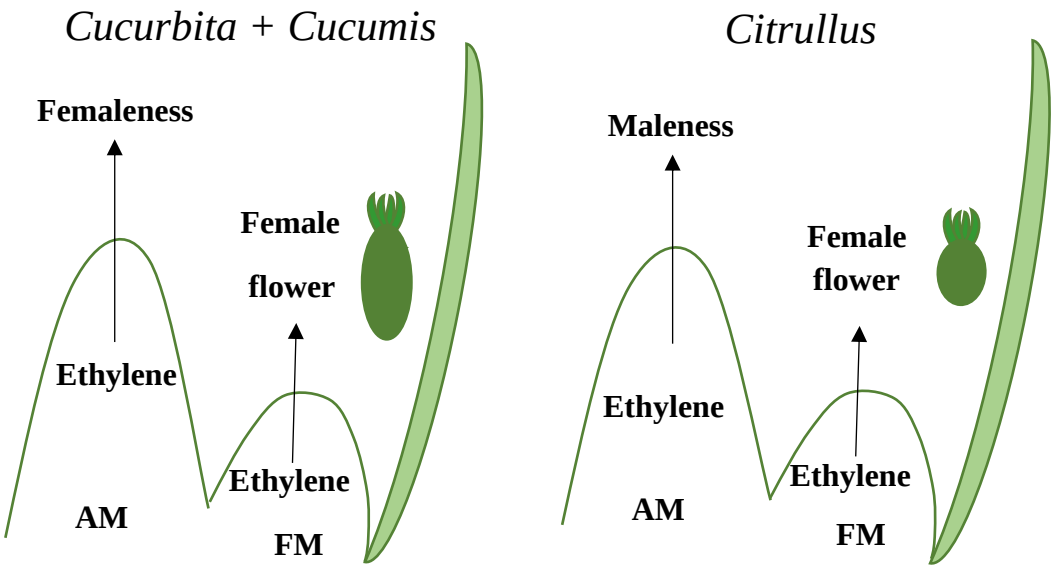


Figure 1.4. Environmental and hormonal factors regulating sex determination in watermelon.

Contrary to what is observed in other cucurbits of the genera *Cucurbita* and *Cucumis*, in watermelon ethylene regulates sex expression in a totally opposite way. Ethylene enhances maleness, delaying female flowering transition and reducing the number of pistillate (female/bisexual) flowers per plant (Manzano et al., 2014). In fact, a reduction in the endogenous level of ethylene, by inhibitors of both its biosynthesis and its perception, hastens the appearance of the first pistillate flower and increases the pistillate to male flower ratio (Rudich and Zamski, 1985; Sugimaya et al., 1998; Manzano et al., 2014; Zhang et al., 2017). However, as in other cucurbits also, the development of female watermelon flowers requires the presence of ethylene so as to arrest the growth of stamen primordia (Manzano et al., 2014). Blocking ethylene biosynthesis or perception using AVG or STS respectively, promotes the conversion of female into bisexual flowers in monoecious and gynoecious plants (Zhang et al., 2017). The model proposed by Manzano et al. (2014) to explain the role of ethylene in sex expression and sex determination in watermelon is shown in Figure 1.5.

Ethylene production (and action) in the apical meristem is regulated differently from that produced in the female floral buds. While the former controls sex expression, the latter regulates sex differentiation and female flower development.



Adapted from Manzano et al. 2014

Figure 1.5. Model explaining the key role of ethylene in the regulation of sex expression and sex determination in cucurbit species. In the apical meristem (AM) ethylene production regulates female flowering transition and the production of female flowers in *Cucurbita* and *Cucumis* but delays female flowering transition and induces the production of male flowers in *Citrullus*. On the other hand, in the floral meristem (FM) ethylene has the same role in all the cucurbits which have been studied. Its production and action determine the appropriate development of a female flower by arresting stamen development in either *Cucurbita*, *Cucumis* and *Citrullus*.

1.3. Ethylene: biosynthesis and signaling

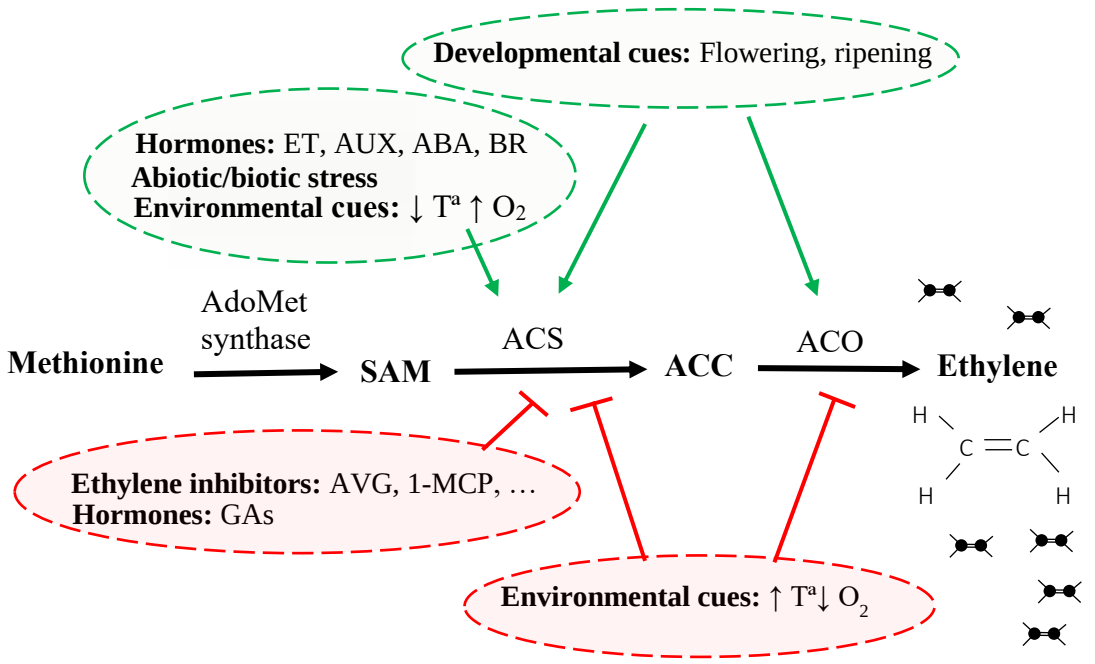
The ethylene is a gaseous hormone which regulates multiple processes during plant growth and development in response to internal and environmental cues. In many species, ethylene controls seed germination, dormancy rupture, plant growth, root nodulation, cellular respiration, sexual expression, fruit ripening and fruit maturation. It controls leaf epinasty, leaf and flower senescence, and it is activated in defense against abiotic and biotic stresses. Ethylene is activated during flower initiation in the meristem and during floral organs development, playing a key role in the regulation of floral sex expression in the cucurbits

species. For this reason, ethylene biosynthesis and its signaling pathways have been the subject of numerous studies that have proven the importance of this hormone in watermelon sex determination.

1.3.1. Ethylene biosynthesis

The route of ethylene biosynthesis is considered to be initiated with the S-adenosylmethionine (S-AdoMet), a precursor together with 1-aminocyclopropane-1-carboxylic acid (ACC) of ethylene (Figure 1.6) (Yang and Hoffman, 1984). Ethylene is synthesized, in two steps, from S-adenosyl methionine (SAM), which is derived from methionine by the action of AdoMet synthase (Johnson and Ecker, 1998). The first and most important step of ethylene biosynthesis is the conversion of SAM to ACC by the action of ACC synthase (ACS), an enzyme encoded by a multigene family in many species of plants (Kende, 1993; Johnson and Ecker, 1998). The second step consists of the oxidation of ACC by ACC oxidase (ACO), forming ethylene, CO2, and cyanide (Figure 1.6). ACO is also encoded by a multigene family in plants , but with a smaller number of genes than the ACS family (Pech et al., 2010).

Ethylene biosynthesis



Adapted from Argueso et al. 2007; Manzano 2009; Martínez 2013; Megías 2016

Figure 1.6. The two-step ethylene biosynthesis pathway in plants. In the first step, which is the rate-limiting step, ACC is synthesized from SAM by the enzyme ACS. SAM is produced from methionine in the "Yang cycle". In the second step, ACC is converted to ethylene by the enzyme ACO. The precursors and blocking agents of ACS and ACO enzymes are indicated in green and red, respectively. ET, ethylene; AUX, auxins; ABA, abscisic acid; BR, brassinosteroids; Tª, temperature; O_2, oxygen; AVG, aminoethoxyvinylglycine; GAs, gibberellins. Adapted from Argueso et al. 2007; Manzano 2009; Martínez 2013; Megías 2016.

ACS gene family in Cucurbits

Next Generation Sequencing technologies and the assembly of the whole genomes of different cultivated cucurbit species, together with the development of multiple bioinformatics analysis tools, allowed the identification and determining of the molecular structure of *ACS* genes in cucurbits. As observed in other plant species, such as Arabidopsis and tomato (Jakubowicz and Sadowski, 2002), *ACS* genes made up a large multigenic family (Nakajima et al., 1990; Nakagawa et al., 1991; Kamachi et al., 1997; Trebitsh et al., 1997; Ishiki et al., 2000; Kato et al., 2000; Watanabe et al., 2001; Boualem et al., 2008, 2009). Table 1.2 shows the structure and function of *ACS* genes within the most important cultivated cucurbits.

Table 1.2. List of *ACS* genes in different cultivated species of Cucurbitaceae.

Species	Genome	Gene ID	Gene name	Exon number	Watermelon ortholog (Charleston Gray)/	Known mutations	Paralogous gene ID	Sex determination function (known mutant phenotypes)	Reference
C. lanatus	Watermelon (Charleston Gray)/ Watermelon (97103)	ClCG01G003470 Cla011522 Cla006245	*CitACS10/ CitACS12*	4			ClCG05G007530		(Guo et al., 2013, 2015)
		ClCG05G007530 Cla011522 Cla006245	*CitACS10/ CitACS12*	4			ClCG01G003470		(Guo et al., 2013, 2015)
		ClCG03G015830 Cla011230	*CitACS4*	3				Sex determination (A/*a*, *Monoecy/ andromonoecy*)	(Salman-Minkov et al., 2008; Prothro et al., 2013; Guo et al., 2015)
		ClCG08G016770 Cla022653	*CitACS11*	4					(Guo et al., 2013, 2015)
		ClCG07G007900 Cla014057	*CitACS2*	5				Sex determination	(Salman-Minkov et al., 2008; Prothro et al., 2013; Guo et al., 2015)
		ClCG06G003420 Cla006634	*CitACS3*	5				Sex determination	(Guo et al., 2013, 2015)
		ClCG01G017870 Cla014652	*CitACS1*	4				Sex determination	(Guo et al., 2013, 2015)
		ClCG10G008330	*CitACS1*	4				Sex determination	(Salman-Minkov et al., 2008; Prothro et al., 2013; Guo et al., 2015)
C. sativus	Cucumber (Chinese Long) v2	Csa1G580750	*CsACS2*	3	ClCG03G015830	P209S		*M/m (Monoecious/* andromonoecious)	(Boualem et al., 2009)
		Csa2G353460	*CsACS11*	4	ClCG08G016770 ClCG01G017870	G39R or W58*	Csa6G496450	A (Androecious)	(Boualem et al., 2015)

		Csa6G496450	*CsACS1/ CsACS1G*	3	ClCG08G016770 ClCG01G017870	Duplication of *ACS1* (*CsACS1* + *CsACS1G*)	Csa2G353460	*F* (*Femaleness*). To copies of *ACS1* *Ff (Subgynoecy):* loss of the *CsACS1G* copy	(Trebitsh et al., 1997; Mibus and Tatlioglu, 2004; Knopf and Trebitsh, 2006; Li et al., 2020b)
		Csa5G157380	*CsACS10*	4	ClCG01G003470				NCBI
		Csa3G177920	*CsACS12*	4	ClCG01G003470		Csa5G157380		NCBI
		Csa4G049610	*CS-ACS2*	4	ClCG07G007900				NCBI
		Csa4G099220		4	Cla97C10G191820				
		Csa6G006800	*CS-ACS1*	5	ClCG06G003420				NCBI
C. melo	Melon (DHL92) v3.6.1	MELO3C015444	*CmACS7*	3	ClCG03G015830	A57V		*M/m* (*Monoecious*/*andromonoecious*)	(Boualem et al., 2008)
		MELO3C010779	*CmACS11*	4	ClCG08G016770 ClCG01G017870	L45F or S295F	MELO3C007662	*A/a* (*Monoecious*/*andrecious*)	(Boualem et al., 2015)
		MELO3C005597	*ACS10*	3	ClCG05G007530 ClCG01G003470		MELO3C006840		NCBI
		MELO3C006840	*ACS12*	4	ClCG05G007530 ClCG01G003470		MELO3C005597		NCBI
		MELO3C007662		3	ClCG08G016770 ClCG01G017870		MELO3C010779		
		MELO3C016340	*CMe-ACS2*	5	ClCG07G007900				NCBI
		MELO3C024891		1					
		MELO3C021182	*CMe-ACS1*	5	ClCG06G003420				NCBI
		MELO3C019008		2	Cla97C10G191820				
		MELO3C024893		1					

		Cp4.1LG18g03790	*CpACS27A*	3		S176A	*m* (andromonoecious)	(Martínez et al., 2014)
		Cp4.1LG04g10620	*CpACS27B*	19		Cp4.1LG18g03790		(Martínez et al., 2014)
		Cp4.1LG05g02460	*ACS10*	4	ClCG01G003470 ClCG05G007530	Cp4.1LG13g01900		NCBI
		Cp4.1LG13g01900	*ACS10*	4	ClCG01G003470 ClCG05G007530	Cp4.1LG05g02460		NCBI
		Cp4.1LG14g02300		5	ClCG07G007900	Cp4.1LG01g00210		
		Cp4.1LG11g01010		4	ClCG08G016770			
C. pepo	Cucurbita pepo (Zucchini) version	Cp4.1LG19g10460		4	ClCG01G017870 ClCG08G016770	Cp4.1LG10g01970		
		Cp4.1LG10g01970		3	ClCG01G017870 ClCG08G016770	Cp4.1LG19g10460		
		Cp4.1LG01g00210		10	ClCG07G007900	Cp4.1LG14g02300		
		Cp4.1LG01g09110		4	ClCG10G008330	Cp4.1LG14g00920		
		Cp4.1LG12g03970		5				
		Cp4.1LG12g03930		5	ClCG06G003420			
		Cp4.1LG14g00920		4	ClCG10G008330	Cp4.1LG01g09110		
		Cp4.1LG00g10840		4				

There are 8 *ACS* genes annotated in the cucumber genome, 10 in melon, 14 in zucchini and 9 in watermelon (CuGenDB; http://cucurbitgenomics.org). In all species, *ACS* genes vary in their structure, showing between 1 and 19 exons of different sizes (CuGenDB; http://cucurbitgenomics.org) (Table 1.2). The orthologous *ACS* genes involved in the sex determination of the different cucurbit species have a conserved molecular structure. The *CmACS7* of melon, the *CsACS2* of cucumber and the *CpACS27A* of squash are made up of only three exons, all three having a high level of homology (Table 1.2) (Boualem et al., 2008, 2009; Martínez et al., 2014). Similarly, *CmACS11* and *CsACS11* have 4 exons in both melon and cucumber, but with a long exon 4 (Boualem et al., 2015). Regarding cucumber gynoecy, a tandem duplication of 30.2 Kb has been found that comprises the genes *CsACS1* and *CsMYB* (Trebitsh et al., 1997; Mibus and Tatlioglu, 2004; Knopf and Trebitsh, 2006; Li et al., 2020b). Up to now, no sex determining *ACS* gene has been cloned and characterized in watermelon.

The ACS proteins belongs to PLP-dependent enzymes, that is, they require pyrodoxal-5´-phosphate (PLP) as cofactor (Argueso et al., 2007) and are classified based on the presence or absence of phosphorylation sites in the C-terminal sequences (Jakubowicz and Sadowski, 2002; Argueso et al., 2007; Xu and Zhang, 2014). Type I ACS isoforms contain an extended C-terminus domain with Serine residues target of phosphorylation by calcium-dependent protein kinase (CDPK) and mitogen-activated protein kinase (MAPK). Type II ACS isoforms have only one single potential CDPK phosphorylation site in their C-terminal, and type III ACS isoforms lack phosphorylation sites in their shortened C-terminus.

Different *ACS* genes found to have a central role in cucurbit sex determination (Kamachi et al. 1997; Yamasaki et al. 2003; Mibus and Tatlioglu 2004; Papadopoulou and Grumet 2005; Knopf and Trebitsh 2006; Salman-Minkov et al. 2008; Manzano et al. 2010; Martínez et al. 2014; Li et al. 2020), but in watermelon none of them was involved before the start of this thesis. Salman-Minkov et al. (2008) studied the expression of 4 *ACS* genes (*CitACS1*, *CitACS2*, *CitACS3* and *CitACS4*) during watermelon flower development. *CitACS1* is expressed in young floral buds of male, female, and bisexual flowers, but not at the stage of anthesis. The *CitACS3* (ClCG06G003420) gene is specifically expressed in male and bisexual flowers. *CitACS2* mRNA was not detectable in floral or vegetative tissue but became active in response to auxins. No *CitACS4* transcription was detected, and it had been speculated that it may be a non-functional *ACS* paralog. In fact, this gene shares 95% identity

with *CsACS2* (Kamachi et al., 1997), the latter being a cucumber gene responsible for the arrest of stamen primordia during the development of female flowers (Boualem et al., 2009).

***ACO* gene family in Cucurbits**

ACC oxidase (ACO) oxidizes the ACC compound into ethylene in the presence of oxygen. The *ACO* gene family contains a smaller number of genes compared with the *ACS* one. Five *ACO* genes having a molecular structure characterized by the occurrence of four exons interspersed with three introns, have been found in watermelon, although some *ACO* genes only have three exons (CuGenDB; http://cucurbitgenomics.org). Kahana et al. (1999) were the first to report on cloning and studying the expression *ACO* genes in flowers, shoot tips and leaves of different cucumber sex morphotypes. In female flowers *CS-ACO3* transcripts were accumulated in nectaries, pistil and arrested stamens, while *CS-ACO2* transcripts were localized to ovaries and staminoids. Male flowers accumulated both type of transcripts in petals and nectaries, and *CS-ACO2* also accumulated in the pollen of mature flowers (Kahana et al., 1999). In floral buds, high level of these two transcripts were observed in vascular bundles. In concordance with these results, a publication by Chen et al. (2016) has revealed that a mutation in *CsACO2* results in 50% loss of ethylene emission in shoot tips, and also that it confers androecy.

1.3.2. Ethylene biosynthesis regulation

ACS and *ACO* genes appear to be expressed in all plant tissues, but there are differences in the levels of accumulation of specific transcripts due to physiological and developmental changes (Ruduś et al., 2013). The level of expression of both gene families seems to be regulated at the transcriptional and post-transcriptional level (Argueso et al., 2007). Transcriptional regulation is produced in response to developmental, environmental, and internal factors (as in hormonal balance), which induce or repress gene expression.

Figure 1.6 shows ethylene precursors and blockers and their interaction with each other. Ethylene biosynthesis is promoted by auxins, abscisic acid (ABA), brassinosteroids and ethylene itself, this being considered one of the main precursors during petal senescence and fruit maturation (Jakubowicz and Sadowski, 2002; Argueso et al., 2007). Besides hormones, abiotic/biotic stress, and developmental cues (i.e. low temperature, flowering, and ripening)

cause an increase in the expression of *ACS*, as do *ACO* genes (i.e. ripening). Among the blockers, there are environmental cues (i.e. high temperature, together with low O_2, repress *ACO* expression) and hormones such as gibberellins, which represses *ACS* expression, while salicylic acid and ABA have been cited as down regulators of the *ACO* transcript level (Lee and Yoon, 2018). Chemicals, however, are the main blockers; vinylglycine analogues such as Aminoethoxyvinylglycine (AVG) inhibit PLP-dependent enzymes such as *ACS*, acting at a post-transcriptional level (Huai et al., 2001).

1.3.3. Ethylene signal transduction pathway

Key elements that are responsible for the primary response to ethylene have been studied and identified over the past couple of decades, using Arabidopsis as a model plant (Klee, 2004; Chen et al., 2005). The identification and characterization of these signaling elements have resulted in a model for ethylene signal transduction which follows a linear pathway (Figure 1.7). This pathway is initiated when ethylene binds to membrane-bound receptors and culminates in the transcriptional control of ethylene regulated genes at the nucleus.

The gaseous hormone ethylene is perceived in plants by a family of ethylene receptors (ETR1, ETR2, EIN4, ERS1 and ERS2 in Arabidopsis), which are histidine kinases predominantly localized at the endoplasmic reticulum (ER) membrane (Ju and Chang, 2012). For ethylene to bind to the receptor, it requires a copper cofactor provided by the action of the copper transporter RAN1 (Schaller and Bleecker, 1995; Hirayama et al., 1999; Rodríguez et al., 1999; Woeste and Kieber, 2000; Binder et al., 2012).

In the absence of ethylene, the receptors activate CTR1, which was the first ethylene signaling component to be cloned and encodes a Kinase with homology to Raf MAPK Kinase Kinase (MAPKKK). The latter suppresses the ethylene response (Kieber et al., 1993). CTR1 being a negative regulator of the downstream ethylene response pathway, the direct phosphorylation target of CTR1 is the ER-bound protein ETHYLENE INSENSITIVE2 (EIN2) (Stepanova and Alonso, 2009). In the absence of ethylene, the level of phosphorylated EIN2 is reduced by protein turnover involving the 26S proteasome degradation (Stepanova and Alonso, 2009) (Figure 1.7).

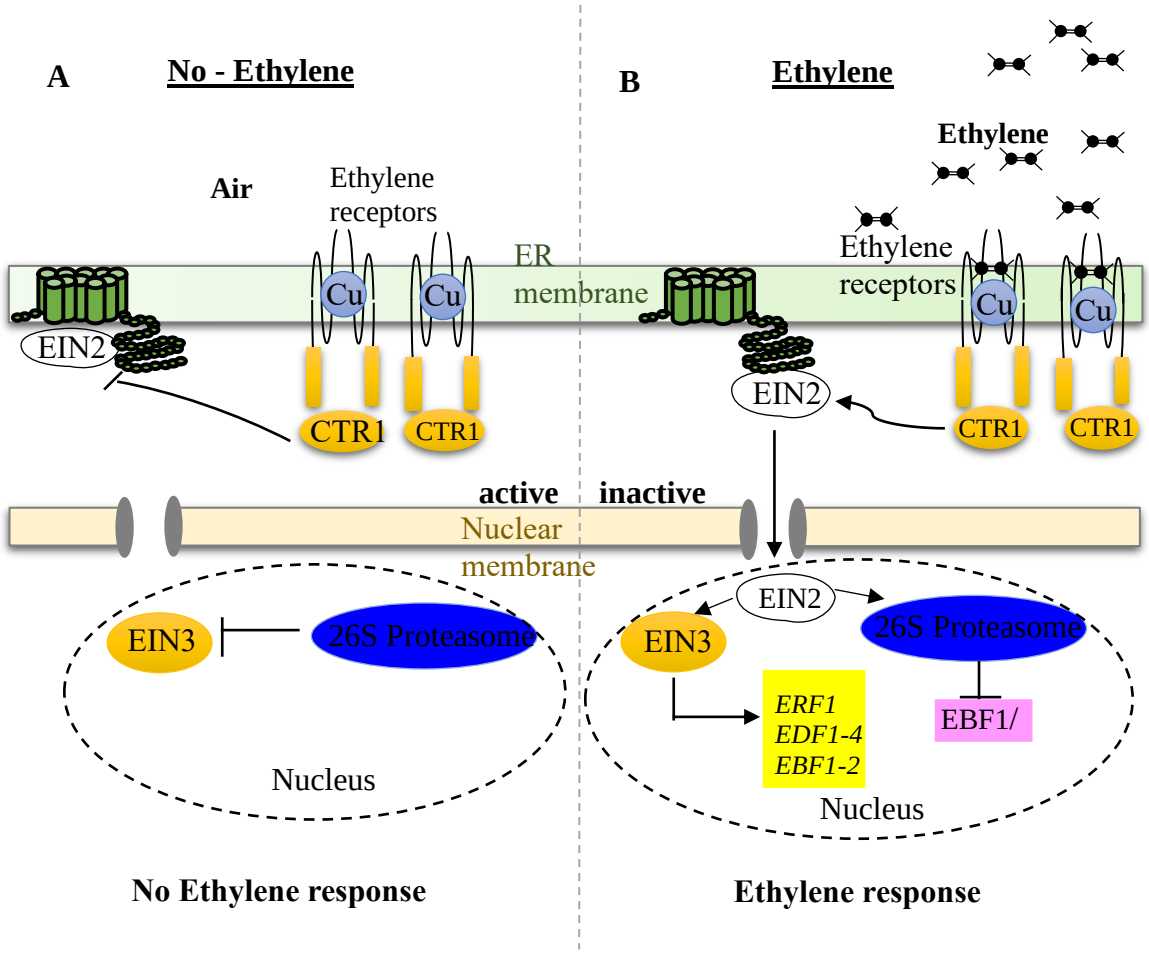


Figure 1.7. Ethylene signaling pathway. (**A**) In the absence of ethylene, the receptors stimulate the kinase activity of CTR1 (Kieber et al., 1993), which results in phosphorylation and in the inhibition of EIN2 activity, and also in the suppression of the ethylene response. (**B**) When ethylene is present, CTR1 is inactivated, which then leads to the activation of downstream components including EIN2, EIN3/ EILs, and ERF transcription factors (reviewed in Wang et al. 2002; Zhao and Guo 2011; Shakeel et al. 2013).

When ethylene is present however, CTR1 is inactivated and degraded by the proteasome. The dephosphorylated ER-membrane anchored EIN2 is proteolytically processed, such that its C-terminal domain is released to migrate to the nucleus (Ju and Chang, 2012; Qiao et al., 2012; Wen et al., 2012). Within the nucleus, via an unknown mechanism, EIN2 activates the transcription factors EIN3 and EIN3 like1 (EIL1) and ETHYLENE RESPONSE FACTORS (ERFs), so as to initiate the transcriptional response to ethylene (Chao et al., 1997; Solano et al., 1998; Alonso et al., 2003; Ju et al., 2012; Qiao et al., 2012; Wen et al., 2012) (Figure 1.7).

Five CTR-like genes have been cloned and characterized in *C. melo*, *C. sativus* and *C. pepo* (Alverson et al., 2010; Manzano et al., 2011). Recently, two semi-dominant mutations have been identified that affect two ethylene receptors of *C. pepo*, *CpETR1A* and *CpETR2B*, both of which confer ethylene insensitive and promote the conversion of monoecy into andromonoecy (García et al., 2020).

1.4. Genetic control of sex determination in cucurbits

1.1.2. Major genes controlling sex determination and flower development

In the second half of the twentieth century, Galun (1961) and Kubicki (1969a, b, c, d) established that three major independent loci could explain the flower morphotypes of cucumber. The locus *F/f* controls femaleness (gynoecy) and the *A/a* one controls maleness (androecy), while the locus *M/m* regulates monoecy (*M_)* and andromonoecy (*mm*). At the present time, cucumber, and melon, are the most studied species regarding sex determination. In cucumber, sex appears to be controlled by four major loci, *F*, *M*, *A*, and *Gy* (Kenigsbuch and Cohen, 1990; Malepszy and Niemirowicz-Szczytt, 1991; Boualem et al., 2009). These genes have been identified in cucumber and melon; most of them encode for ethylene biosynthesis enzymes (Table 1.3).

- Locus *F/f* (femaleness): In combination with the *M* allele, *FF* plants are gynoecious, *Ff*, subgynoecious, and *ff*, monoecious. The *F* allele consists of a tandem duplication of 30.2 kb that affects the *CsACS1* and *CsMYB* genes. The duplicated *ACS1* is known as *CsACS1G* with a distinctive promoter respect to *CsACS1*. Monoecious lines (*ff*) possess only one copy for *CsACS1* and one of *CsMYB* (Trebitsh et al., 1997; Mibus and Tatlioglu, 2004; Li et al., 2020b). Thus, the feminizing effect of *F* allele is conjectured as being a dose effect, which leads to an increase in ethylene production in gynoecious lines (Knopf and Trebitsh, 2006; Shiber et al., 2008; Li et al., 2020b). Gynoecy is unstable, since monoecious and subgynoecious plants can be found in the offspring of gynoecious lines. This gynoecy loss seems to be caused by unequal crossing over (UCO) within the tandem duplicated region in meiosis, which results in a single *CsACS1* gene (Li et al., 2020b).
- Locus *M/m* (monoecious): The dominant *M* allele confers monoecy while the recessive allele *m* confers andromonoecy. Plants with the genotype *mmff* are andromonoecious and those with *mmF_* one are hermaphrodite when combined with the *F/f* locus. *M_F_* plants

are female or gynoecious, whereas *M_ff* ones are monoecious with mostly male flowers. The *M* locus has been found to correspond to the *CmACS7* gene in melon, *CsACS2* in cucumber and *CpACS27A* in squash. The genes are expressed very early in the development of female flowers, and their loss of function mutations or their suppression causes a complete or partial conversion of female into bisexual or hermaphrodite flowers, and the conversion of monoecy into andromonoecy or partial andromonoecy (Boualem et al., 2008, 2009; Li et al., 2009b; Martínez et al., 2014).

- Locus *A/a* (androecious): The recessive *a* allele promote maleness (Galun, 1961). Plants with the genotypes *mmffaa* and *M_ffaa* are wholly male. In melon, androecy results from a loss of function of the *CsACS11* gene, and in cucumber from a loss of function of the *CmACS11* gene. Moreover, mutations in the ethylene biosynthesis gene *CsACO2* lead to androecy in cucumber, indicating that this gene is also involved in carpel development (Chen et al., 2016).
- Locus *Gy/gy* (gynoecious): The recessive allele *gy* is responsible for femaleness. In watermelon a gynoecious mutant is controlled by a single recessive gene *gy* (Schaefer et al. 2009; Zhang et al. 2019). It was discovered by Martin et al. (2009) that the *Gy/gy* locus of melon corresponded to *CmWIP1*, a gene encoding for a C2H2 zinc-finger-type transcription factor. *CmWIP1* regulates carpel abortion during the development of male flowers, and the mutations in this gene leads to gynoecy (Martin et al., 2009; Chen et al., 2016; Hu et al., 2017; Zhang et al., 2019b). The *WIP1* ortholog has recently been identified in watermelon, where it has the same function as in melon and cucumber (Zhang et al. 2019). *WIP1* expression is repressed by *ACS11*, allowing the coexistence of male and female flowers on the same plant (Boualem et al., 2015; Chen et al., 2016).

In watermelon, an additional locus, the *Tm/tm* (trimonoecious) one, has been reported as being responsible for trimonoecy (Ji et al., 2015; Li et al., 2019). Homozygous *tmtm* plants produce male, female and bisexual/hermaphrodite flowers, although the ratio of these three different types of flowers to each other could be influenced by environmental factors (Ji et al., 2015).

Table 1.3. List of genes involved in sex determination in different cucurbit species.

Gene	Protein	Function in sex determination	LOF mutations	References
CsACS1* + *CsACS1G	Ethylene biosynthesis enzyme	Femaleness, gynoecy	Loss of *CsACS1G* copy: subgynoecy, monoecy	(Trebitsh et al., 1997; Mibus and Tatlioglu, 2004; Knopf and Trebitsh, 2006; Li et al., 2020b)
CsACS2 ***CmACS7*** ***CpACS27A***	Ethylene biosynthesis enzyme	Arrest of stamen development in female flowers	Andromonoecy	(Boualem et al., 2008, 2009; Martínez et al., 2014)
CmACS11 ***CsACS11***	Ethylene biosynthesis enzyme	Promotion of carpel development in female flowers	Androecy	(Boualem et al., 2015)
CmWIP1 ***CsWIP1*** ***CsWIP1*** ***ClWIP1***	Transcription factor	Arrest of carpel development in male flowers	Gynoecy	(Martin et al., 2009; Chen et al., 2016; Hu et al., 2017; Zhang et al., 2019b)
CsACO2	Ethylene biosynthesis enzyme	Promotion of carpel development	Androecy	(Chen et al., 2016)

1.1.3. Spatio-temporal patterning of gene expression during flower formation

Plant sex determination is known to occur within the floral meristem, where the spatio-temporal pattern of gene expression throughout flower development will determine whether a floral bud will become a male, female, or bisexual flower. In cucumber, Bai et al. (2004) discerned 12 stages of development from primordia to anthesis in both male and pistillate flowers. During the first developmental phases (stages 1-5) - that is, up to ~ 0.5 mm - the development of male and pistillate floral buds are indistinguishable. The differences begin to be observed when the floral bud is 3-5 mm in size (stage 6), stage at which stamen and carpel primordia start to develop (Bai et al., 2004). In buds bound to be male flowers, the anther differentiates from stages 6 to 12, and buds committed to be female, carpels differentiate from stages 7 to 12.

This division of floral developmental into 12 stages has been useful in establishing the expression of each sex determining gene during the development of male, female, and bisexual flowers. During construction of the model published by Chen et al (2016) regarding

cucumber, they studied the spatio-temporal expression of a combination of sex-related genes in order to explain the different development paths of male, female and bisexual flowers. The model shown in Figure 1.8 is an adaptation of that proposed by Chen et al. (2016) and Boualem et al. (2015), incorporating the interacting functions of the *CsACS11*, *CsACS1/CsACS1G*, *CsWIP1*, *CsACO2* and *CsACS2* genes in sex determination (Che and Zhang 2019; Li et. 2019).

In situ hybridization studies in cucumber and melon found that *CsACS2* and *CmACS7* were expressed in carpel primordia between stage 4 to 7 of the development of female flowers (Boualem et al., 2008, 2009; Li et al., 2019). Due to the proposed interaction between *CsWIP1* and *CsACS2*, and between *CsACS11* and *CsWIP1* (Figure 1.8), it is expected that the expression of *CsACS11* and *CsWIP1* occurred before stage 6. In accordance with that, Boualem et al. (2015) found that the transcripts of the *CsACS11* cucumber gene and the *CmACS11* melon gene were located at stage 4, but not at stage 8 in female flowers. Similarly, *CsWIP1* and *CmWIP1* were found to be expressed in male flowers up to development stage 6 of monoecious plants, but not in female or bisexual forms of monoecious, andromonoecious or gynoecious plants (Boualem et al., 2008, 2009; Martin et al., 2009). Moreover, the *CsACO2* that was suggested as having a combined action with *CsACS11* and *CsACS1G*, was indeed found to be expressed in pistillate flowers up to stage 6 (Li et al., 2019). The interaction between *CsWIP1* and *CsACO2* was demonstrated by Chen et al. (2016). *CsWIP1* binds directly to the promoter of *CsACO2* to repress its expression. The role of the *CsACS1G/CsACS1* genes in the production of a complete female plant is still unknown. In conclusion, while the *CsACS1G*, *CsACS11*, *CsACS2* genes are expressed in the female flowers of cucumber, only the *CsACS1G*, *CsACS11* ones are expressed in bisexual flowers. *CsWIP1* is only expressed in male flowers and *CsACO2* has no sex specificity (Li et al., 2019).

The precise spatio-temporal and coordinated expression of ACS and ACO genes, together with the expression of transcription factor gene WIP1, explain the development of each flower type in cucurbits. Natural as well as induced mutations in these genes promote the conversion of one sex morphotype into another.

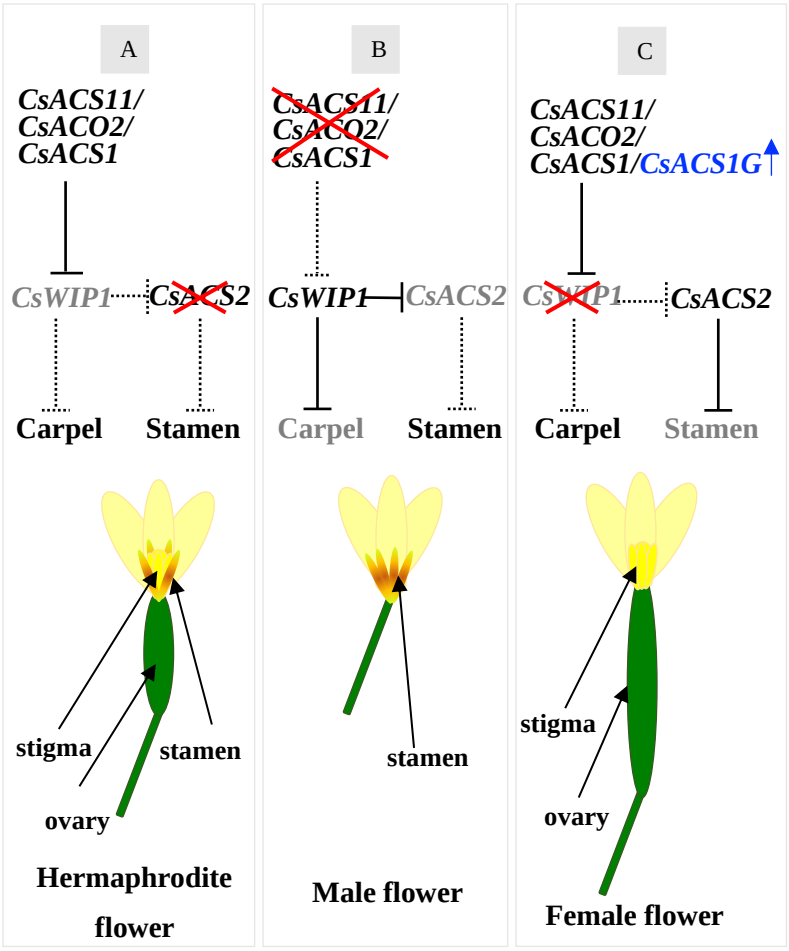


Figure 1.8. A simplified model outlining the interactions of major genes in the control of cucumber sex determination. (**A**) Mutations in *CsACS2* suppress the arrest of stamens. The production of ethylene by *CsACS11*, *CsACO2* and *CsACS*1 inhibit *CsWIP1*, which is responsible for the arrest of carpel development. The result is the production of a hermaphrodite flower (**B**) Mutations in *CsACS11* or *CsACO2* reduced the production of ethylene in the floral buds, thus activating the *CsWIP1* gene leading to the arrest of the carpel. The down regulation of *CsACS2* by *CsWIP1* inhibit the arrest of stamens and the production of a male flower. (**C**) The induction of ethylene production by *CsACS1G* or mutations in *CsWIP1* inhibit the arrest of carpels, but activate *CsACS2*, which is the responsible for stamen arrest. The result is the development of a female flower.

LOF mutations of the gene *ACS11* activate the expression of *WIP1*, favoring the development of male flowers (androecy). *CsACO2* LOF mutations also lead to androecy (Chen et al., 2016). The development of female flowers, however, requires the expression of *ACS11* and the repression of *WIP1*, leading to the expression of *ACS2* (Figure 1.8). On the other hand, *ACS2* mutations prevent the arrest of stamen development in pistillate flowers, resulting in the conversion of female into bisexual/hermaphrodite flowers, and monoecy into andromonoecy. In contrast, gynoecious plants are the result of the inactivation

of the *WIP1* function, while hermaphrodite plants are the consequence of the inactivation of both *WIP1* and *ACS7* (Martin et al., 2009). In cucumber, gynoecy can also be caused by the expression of an extra copy of *CsACS1* (*CsACS1G*); the lack of this additional *CsACS1G*, by unequal crossing-over, leads to subgynoecy or monoecy (Li et al., 2020b).

1.2. The involvement of ethylene in fruit set and development

The initiation of fruit growth and development, or fruit set, is a process controlled by many hormonal signals, induced by fertilized ovules (Ferrándiz et al., 1999). The hormones usually considered to be involved in this process are auxins, GAs and cytokinins (Ruan et al., 2012); however, the role of ethylene has been only partially investigated. Transcriptomic analysis and ethylene inhibitors applications performed in different species have shown that fruit set requires low levels of ethylene (Dussi et al. 2002; Ogata et al. 2002; Singh and Agrez 2002; Vriezen et al. 2008; Pascual et al. 2009; Carbonell-Bejerano et al. 2011; Martínez et al., 2013). In fact, ethylene genes are highly expressed in tomato ovary at anthesis, although a downregulation of these genes is observed after fertilization. This suggests that ethylene is an antagonist of auxins; consequently, it prevents carpel development before pollination and fertilization (Vriezen et al., 2008; Pascual et al., 2009). The *ethylene insensitive* (*ein*) mutants of Arabidopsis have a reduced response to ethylene and a long-lasting ovule lifespan. Also, they respond to GAs (Carbonell-Bejerano et al., 2011). In other words, ovules and ovary senescence is correlated with a loss of response to GAs, and an increase in ethylene production and sensitivity, suggesting that ethylene does indeed play a crucial role in ovule lifespan and ovary/fruit development.

With regard to cucurbits, the work by Martínez et al (2013) in *C. pepo* pointed out the critical role of ethylene in fruit set. Ethylene production during the first days after anthesis drives fruit to abortion; in contrast, a reduction in ethylene production during this developmental period favors fruit set and fruit development (Martínez et al., 2013). On the other hand, it should be noted that despite watermelon being classified as non-climacteric fruit (Paul et al., 2012), diverse studies indicate that ethylene production and/or sensitivity may well participate in physiological changes associated with fruit development and senescence. Mao et al. (2004) detected that exogenous ethylene causes accelerated softening and water-soaking of watermelon flesh. Wechter et al. (2008) found that more ethylene was produced

in watermelon during the green fruit stage than during the white and pink fruit stages. Taken together all of this data indicates that ethylene is not only a key hormone concerned with the control of flower development and sex determination, but is also a regulator of fruit set and development in cucurbit species.

2. Objectives

The phytohormone ethylene is the most important factor regulating sex expression and sex determination in the Cucurbitaceae family. Different ethylene biosynthesis and signaling genes have been cloned in the genera *Cucumis* (melon and cucumber) and *Cucurbita* (squash) that are involved in the arrest of stamens and carpels during the development of unisexual flowers. In watermelon (*Citrullus lanatus*), however, there is no known ethylene biosynthesis gene that is involved in sex determination.

The main objective of this thesis is to investigate the genetic and molecular regulation of sex determination and sex expression in watermelon, analyzing the role of specific ethylene biosynthesis genes in flower and fruit development, and identifying and characterizing the genomic regions involved in different floral morphotypes.

The specific objectives were as follows:

FIRST. To characterize at a molecular and functional level the ethylene biosynthesis gene *CitACS4* of watermelon, determining its role in arresting the stamen primordia during the development of female flowers. Gene expression will be studied in developing male, bisexual, and female flowers from andromonoecious and monoecious inbred lines, determining the possible cosegregation of monoecy/andromonoecy with the allelic variants that will be identified in the course of this research.

SECOND. To analyze the involvement of ethylene biosynthesis gene *CitACS4* in the control of sex expression traits such as the number of male and female flowers per plant, and the female flowering transitions, as well as its role in fruit and seed set, and in fruit development and fruit shape. A segregation study will be carried out between the alleles of the *CitACS4* gene and the traits of interest in segregating populations derived from biparental crosses between monoecious and andromonoecious inbred lines.

THIRD. To study the mode of inheritance of the partial andromonoecious (pa) phenotype, of watermelon, and to map the *pa* locus using combined strategies, the bulk segregant analysis sequencing (BSA-seq) and fine mapping, and genome-wide association analysis (GWAS) with a large panel of watermelon accessions from different parts of the world.

3. The ethylene biosynthesis gene *CitACS4* regulates monoecy/andromonoecy in watermelon (*Citrullus lanatus*)

3.1. Abstract

Monoecious and andromonoecious cultivars of watermelon are characterised by the production of male and female flower or male and hermaphrodite flowers, respectively. The segregation analysis in the offspring of crosses between monoecious and andromonoecious lines has demonstrated that this trait is controlled by a single gene pair, being the monoecious allele *M* semi-dominant to the andromonoecious allele *A.* The two studied F1 hybrids (*MA*) had a predominantly monoecious phenotype since both produced not only female flowers, but also bisexual flowers with incomplete stamens, and hermaphrodite flowers with pollen. Given that in other cucurbit species andromonoecy is conferred by mutations in the ethylene biosynthesis genes *CmACS7*, *CsACS2* and *CpACS27A* we have cloned and characterised *CitACS4*, the watermelon gene showing the highest similarity with the formers. *CitACS4* encoded for a type ACS type III enzyme that is predominantly expressed in pistillate flowers of watermelon. In the andromonoecious line we have detected a missense mutation in a very conserved residue of CitACS4 (C364W) that co-segregates with the andromonoecious phenotype in two independent F2 populations, concomitantly with a reduction in ethylene production in the floral buds that will develop as hermaphrodite flowers. The gene does not however cosegregates with other sex expression traits regulated by ethylene in this species, including pistillate flowering transition and the number of pistillate flowers per plant. These data indicate that *CitAC4* is likely to be involved in the biosynthesis of the ethylene required for stamen arrest during the development of female flowers. The C364W mutation would reduce the production of ethylene in pistillate floral buds, promoting the conversion of female into hermaphrodite flowers, and therefore of monoecy into andromonoecy.

3.2. Introduction

The cultivated species of the Cucurbitaceae family, including melon, cucumber, watermelon, squash, and gourds, are monoecious, developing unisexual male and female flowers on the same individual plant. Evolution has led, however, to a number of sex morphotypes in the species of this family, including andromonoecious (plant produces male and bisexual flowers), gynoecious (only female flowers), androecious (only male flowers) and hermaphrodite (only hermaphrodite flowers) lines. All these sex morphotypes have been detected in melon (Poole and Grimball, 1938; Kenigsbuch and Cohen, 1987, 1990),

cucumber (Kubicki, 1969b, 1969c) and watermelon (Rosa, 1928; Jiang and Lin, 2007; Ji et al., 2015). In squash, the predominant monoecious cultivars coexist with partially andromonoecious ones (Martínez et al., 2014), and some androecious mutants have been also recently described (Manzano et al., 2009; García et al., 2015), but no gynoecious squash have been identified so far.

Sex determination in this family is mainly controlled by the gaseous hormone ethylene. It has long been known that external treatment with ethylene favours the formation of female flowers in monoecious cultivars of melon, cucumber and squash, while the application of inhibitors of ethylene biosynthesis and response, including aminoethoxyvinylglycine (AVG) or silver thiosulphate (STS), favours the development of male flowers (Rudich et al., 1969; Byers et al., 1972; Den Nijs and Visser, 1980; Owens et al., 1980; Rudich, 1990; Manzano et al., 2011). Moreover, in melon and cucumber, the best characterised species of the family, the existence of several of the sexual morphotypes described is controlled by this hormone. Thus, the andromonoecious morphotype in cucumber, melon and zucchini squash, result from mutations in the three orthologous ethylene biosynthesis genes *CmACS7*, *CsACS2* and *CpACS27A*, respectively (Boualem et al., 2008, 2009; Martínez et al., 2014). These genes are expressed only in pistillate flower primordia and are responsible for the arrest of stamens during the development of unisexual female flowers. The gynoecy of cucumber also depends on an additional *ACS* gene which is only present in the gynoecious varieties (Trebitsh et al., 1997; Mibus and Tatlioglu, 2004; Knopf and Trebitsh, 2006). However, in melon gynoecy results from a transposon-mediated mutation in the promoter of the transcription factor *CmWIP1*, a negative regulator of *CmACS7*, responsible of the abortion of carpels and the promotion of stamen development (Martin et al., 2009). The genes responsible for androecy in melon and cucumber have been recently characterised. They correspond to *CmACS11* and *CsACS11*, both involved in the biosynthesis of ethylene in the phloem of flowers programmed to become females, and in melon this gene functions as a negative regulator of the male-promoting transcription factor gene *CmWIP1* (Boualem et al., 2015).

Sex determination mechanisms in watermelon have received little attention. Ethylene is also an important regulator of sex in this species, although external treatments with the hormone induce the production of male flowers, (Rudich and Zamski, 1985), while treatments with ethylene inhibitors hasten the appearance of the first female flower and increase the number of female flowers per plant (Rudich and Zamski, 1985; Sugimaya et al., 1998; Manzano et al., 2014), which it is contrary to what happens in the other cucurbit species. Recently we

have differentiated between two sex related processes: sex expression, i.e. the earliness and production of female flowers per plant, and sex determination, as the mechanism that leads to the proper development and differentiation of unisexual female and male flowers (Manzano et al., 2014). In contrast to what happens in other cucurbits, ethylene inhibits the transition from male to female flowering and reduces the number of female flowers per plant. Nevertheless, as in other cucurbit species, ethylene is necessary for the arrest of stamen development during the proper development of the female flower, and the reduction of ethylene production or action lead to the transformation of female into bisexual and hermaphrodite flowers (Manzano et al., 2014). In this paper it is shown that *CitACS4*, an homologous gene to *CmACS7*, *CsACS2* and *CpACS27A* of melon, cucumber and squash, is responsible for the arrest of stamens in female flower development, and that a recessive mutation in this gene reduces the production of ethylene in the floral bud, and leads to the conversion of female into bisexual or hermaphrodite flowers, and therefore monoecy into andromonoecy.

3.3. Materials and Methods

3.3.1. Plant material, growing conditions, and phenotyping

Three inbred lines of watermelon (*Citrullus lanatus*) two monoecious lines (P85 and P86) and one andromonoecious line (P87) were characterised in this paper. The F1 and F2 generations from two independent crosses (P85XP87 and P86XP87) were used to determine the inheritance of monoecy/andromonoecy in this species. The crosses were performed in spring-summer seasons of 2012 and 2013, and the final phenotyping carried out in plants grown under standard greenhouse conditions in the province of Almería (Spain) in the spring-summer of 2014 and 2015.

To evaluate monoecy in the different inbred lines and populations, the so-called Andromonoecy Index (AI, (Martínez et al., 2014)) were defined for each flower, plant and population. Pistillate flowers were scored from 1 to 3 according to their degree of stamen development. Female flowers with no stamen development were scored as AI = 1, while hermaphrodite flowers with complete stamens and anthers able to produce pollen were scored as AI = 3. A score of 2 was assigned to bisexual flowers not producing pollen with medium-sized stamens and anthers (Figure 3.1A). Based on the flower scores, the AI of each

plant in a population was calculated as the average score for at least five pistillate flowers. The average AI for inbred lines or F1 was then estimated from at least 10 plants with a minimum of 5 pistillate flowers evaluated per plant. Plants and genotypes with and AI = 1-1.2 were considered to be monoecious, while those with AI=1.2-2.7, partially andromonoecious, and those with AI ≥2.7 were phenotyped as andromonoecious.

Sex expression in each plant was assessed by both the number of initial nodes with male flowers before the production of the first pistillate flower in the main shoot (pistillate flowering transition), and the percentage of pistillate flowers per plant in the first 20 nodes of the main shoot. At least 10 plants were phenotyped to assess the sexual expression of each genotype.

3.3.2. Cloning and molecular characterization of *CitACS4*

To identify the watermelon ortholog for *CmACS7*, *CsACS2* and *CpACS27A* (Boualem et al., 2008, 2009; Martínez et al., 2014), we blasted the coding sequences of the known genes on the watermelon genome at Cucurbit Genome Database (http://www.icugi.org). Thereby a watermelon *ACS* gene having the highest homology with the formers was identified. The gene, called *CitACS4*, was cloned and characterised in monoecious and andromonoecious lines. Specific primers *CitACS4gen-F1/R1* and *Fw/Rw* (Supplementary Table 3.1) were designed to amplify a genomic region of 1232 bp, covering the complete sequence of *CitACS4* from P85, P86 and P87 genomes.

For the phylogenetic analysis, alignments were performed using Clustalw at GenomeNet Database Resources (http://www.genome.jp/tools/clustalw/), and the MEGA4 software (Tamura et al., 2007), which allowed the alignment of proteins and the construction of phylogenetic trees using the UPGMA method (Sneath and Sokal, 1973), with 2,000 replicates bootstrap (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 519 positions in the final dataset.

3.3.3. Genotyping *M* and *A* alleles of *CitACS4*

We have detected a single nucleotide polymorphism (SNP) between monoecious (P85 and P86) and andromonoecious (P87) lines that produce an amino acid substitution of a cysteine to a tryptophan in residue 364 of the CitACS4 protein (C364W). The respective alleles of for *CitACS4* in monoecious and andromonoecious lines were called *M* and *A*, respectively.

To genotype these two alleles in parental lines, and F1 and F2 generations, we used the specific primer pair *CitACS4MF/CitACS4gen-R1* or *CitACS4S-F/CitACS4M-R* (Supplementary Table 3.1), which were designed to specifically amplify the *M* allele, and primer pair *CitACS4A-F/CitACS4gen-R1* or *CitACS4S-F/CitACS4A-R*, that only amplified the allele *A*. DNA was isolated from frozen young leaves using the CTAB method (Levi and Thomas, 1999). 15–35 ng of purified DNA was used to amplify by PCR a 253 or 271 bp fragments of *CitACS4* gene. The amplifications were performed using the GeneAmp PCR System 2700 (Applied Biosystems) and PCR reactions consisted of 35 cycles of 30 s at 95°C, 30 s at 60°C and 90 s at 72°C. PCR fragments were resolved in agarose gels at 1.3 %.

3.3.4. Ethylene production and quantitative RT-PCR

The production of ethylene and the expression of *CitACS4* gene were studied in flower buds throughout four different stages of floral development (S0 to S3). The different developmental stages were separated on the basis of the corolla length: S0 = 4±1 mm, S1 = 8±2 mm, S2 = 12±2 mm, S3 = 15±2 mm (Manzano et al., 2010). Ethylene was determined in three biological replicates per sample, each one containing three female, hermaphrodite, or male flowers at the same stage of development. Floral buds were excised from the plant and incubated at room temperature for 6 h in hermetic glass containers in the dark. Ethylene production was determined by analysing 1 ml of gas from the *head*space on a Varian 3900 gas chromatograph apparatus, fitted with a flame ionization detector. The instrument was calibrated with standard ethylene gas. At least three technical replicates were made for each biological sample.

Gene expression analysis was performed on three biological replicates per sample. Each replication was the result of an independent extraction of total RNA from 3 different flowers at the same stage of development. RNA extractions were performed according to the protocol of the GeneJET Plant RNA Purification Kit (Thermo). The remaining DNA in RNA

samples was eliminated by digestion with RQ1 RNAse free DNAse (Promega). cDNA was then synthesized from 500 ng of total RNA using RevertAid RT Reverse Transcription Kit (Thermo). The expression of genes was evaluated through quantitative RT-PCR by using the Rotor gene thermocycler (Qiagen) and SYBR® Green Master Mix (BioRad). Supplementary Table 3.1 shows the different primers used. The q-PCR primers were designed from the 3′ non-coding regions of each gene by using the Primer Express v 2.0 (Applied Biosystem) software. To avoid possible cross-amplification, and before any q-PCR experiment, the size of the PCR products for each pair of primers was tested in agarose gels, and sequenced. Quantitative RT-PCR reactions consisted of 40 cycles of 20 s at 95°C, 15 s at 59°C and 20 s at 60°C.

Relative expression of each gene was determined by the comparative Ct (Cycle Threshold) method using *C. pepo* 18S ribosomal RNA and *ACTIN* genes as internal standards. To use this method, it was first demonstrated that the efficiency of amplification for each amplicon was roughly equivalent, regardless of the amount of template cDNA. The absolute value of the slope of ΔCt (Ct of the target gene-Ct of the reference gene) versus serial dilutions of cDNA for a given sample must be less than 0.1. The relative expression of each gene was then calculated relative to a calibrator sample using the formula $2^{-\Delta\Delta Ct}$, where ΔΔCt is the difference between the ΔCt of each sample and the ΔCt of the calibrator sample.

3.3.5. Statistical analysis

Simple and factorial analyses of variance (ANOVA) at $p < 0.05$ were performed by the STATISTIX 8.0 software package, and each two means were compared with the method of Fisher's least significant difference (LSD) or Tukey's multiple comparison test.

3.4. Results

3.4.1. Phenotypic and genetic characterisation of monoecious and andromonoecious lines of watermelon

The sexual phenotype of three watermelon inbred lines (P85, P86 and P87) were studied by phenotyping staminate and pistillate flowers in the first 20 nodes of the main shoot in at least 10 plants per genotype. Given that the development of stamens in pistillate flowers was

variable, these flowers were classified and scored according to their stamen development using the Andromonoecy Index (AI, (Martínez et al., 2014)). The female flowers with no stamen development were scored as AI = 1, while hermaphrodite flowers with complete stamens and pollen were scored as AI = 3. Ovary-bearing flowers with intermediate stamen development and no pollen production were classified as bisexual and scored as AI = 2 (Figure 3.1A). The AI of each plant, genotype and progeny was then calculated as the average score of a minimum of five pistillate flowers in each plant, and at least 10 plants for each genotype or progeny.

The distribution of staminate and pistillate flowers of the three inbred lines along the 20 first nodes of the plant are shown in Figure 3.1B. The sexual phenotype of line P87 was very stable for andromonoecy condition (AI = 3). Under our conditions P87 plants only produced staminate and hermaphrodite flowers with complete stamens and pollen (AI = 3). Lines P85 and P86 were monoecious, since the predominantly produced female flowers, but also produced bisexual flowers, which resulted in AI = 1.16 for both P85 and P86. On the basis of these results, plants and genotypes with AI = 1-1.19 were considered to be monoecious, those with AI = 1.2-2.69, partially andromonoecious, and those with AI ≥2.7 were considered andromonoecious.

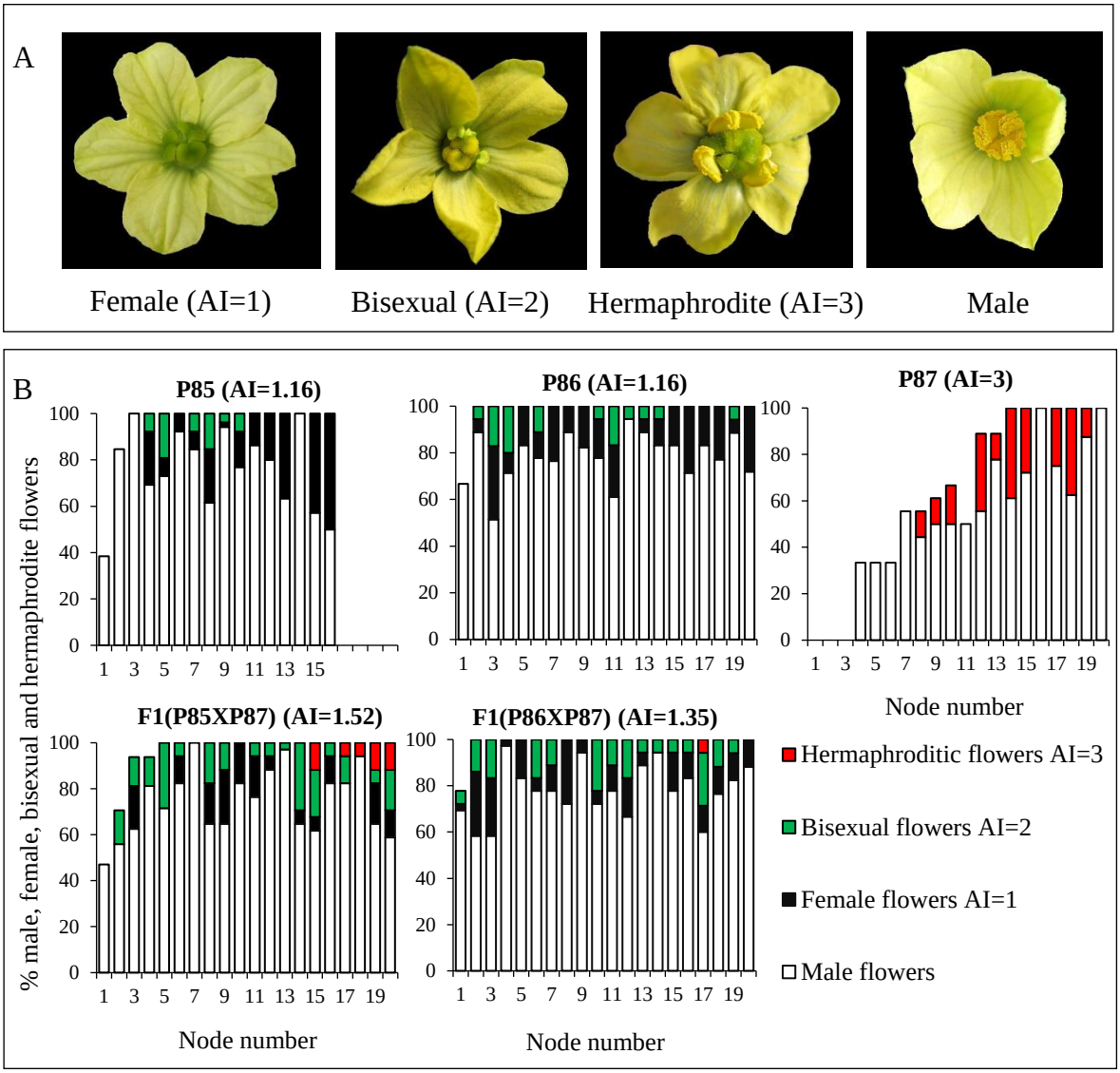


Figure 3.1. Sexual expression of watermelon lines P85, P86 and P87 and F1 hybrids derived from crosses P85XP87 and P86XP87. (**A**) Phenotype of watermelon hermaphrodite, bisexual, female, and male flowers. (**B**) Distribution of staminate and pistillate flowers in the 20 first nodes of the main shoot. In each node, white, black, green, and red bars represent the percentages of male, female, bisexual and hermaphrodite flowers in the total number of plants analysed (n ≥ 10 for each genotype). The lack of bar in a node indicates the absence of flower in that node for some of the analysed plants.

The sexual phenotype of the two F1 hybrids derived from crosses between monoecious and andromonoecious lines (P85XP87 and P86XP87) had an intermediate phenotype between monoecious and andromonoecious and were therefore classified as partially andromonoecious (Figure 3.1B). The two F1 populations had an intermediate AI (1.52 and 1.35), since both produced not only female, but also bisexual and hermaphrodite flowers (Figure 3.1B), suggesting that the monoecy allele in these two lines of watermelon is a semi-

dominant trait in respect of andromonoecy. The segregation of monoecious, andromonoecious and partial andromonoecious plants in the two F2 generations studied demonstrated that the trait is controlled by a single gene pair, being the monoecious allele (*M*) incompletely dominant over the andromonoecious allele (*A*). As expected, the segregation of monoecious, partially andromonoecious and andromonoecious plants in the two F2 populations fitted the 1:2:1 ratio, as expected if the homozygous plants *MM* and *AA* were monoecious and andromonoecious, respectively, while heterozygous plants *MA* had an intermediate phenotype between monoecy and andromonoecy (Table 3.1).

Table 3.1. Segregation ratio of monoecious, partially andromonoecious and andromonoecious plants in F2 populations derived from two crosses between monoecious and andromonoecious inbred lines.

Generation	No. of plants			Expected segregation	χ^2	*p-value*
	Monoecious	Partially andromonoecious	Andromonoecious			
Parental P87	0	0	10	-	-	-
Parental P85	13	0	0	-	-	-
Parental P86	18	0	0	-	-	-
F1 (P85XP87)	0	17	0	-	-	-
F1 (P86XP87)	3	15	0	-	-	-
F2 (P85XP87)	27	41	24	1:2:1	0.34	0.53
F2 (P86XP87)	24	34	13	1:2:1	1.02	0.17

The F2 plants were phenotyped on the basis of their average AI, scored from at least 5 flowers per plant. Monoecious ($1 \leq AI < 1.2$), partially andromonoecious ($1.2 \leq AI < 2.7$), andromonoecious ($2.7 \leq AI \leq 3$).

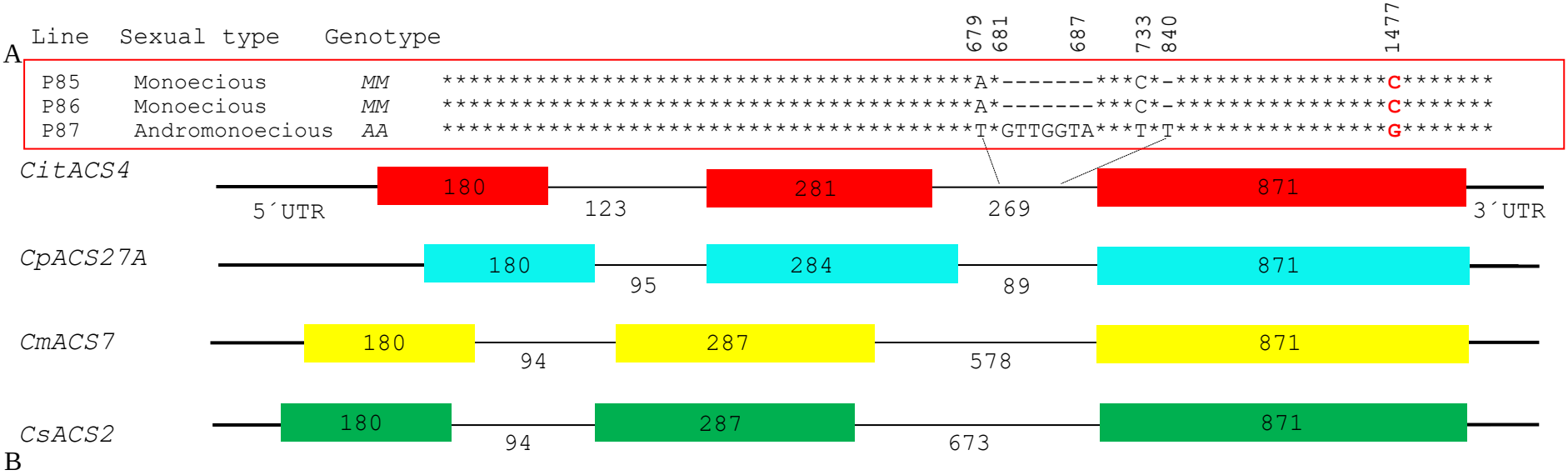
3.4.2. Cloning and characterization of *CitACS4*

Since in melon, cucumber and squash the andromonoecious phenotype is caused by mutations in the orthologs *CmACS7, CsACS2* and *CpACS27A* (Boualem et al., 2008, 2009; Martínez et al., 2014), a homology analysis was performed to identify the watermelon *ACS* gene showing the highest similarity with the former. The nucleotide sequences of these homologous genes were blasted on watermelon genome at Cucurbit Genome Database (http://www.icugi.org), and the highest homology (E-value = 0) was found with Cla011230 gene on chromosome 3, a partial sequence of which was previously reported as *CitACS4* by Salman-Minkov et al. (2008).

The coding sequence of *CitACS4* is 1332 bp, encoding for a protein of 444 amino acids. The gene consists of three exons of 180, 281 and 871 bp, and two introns of 123 and 269 bp, a genomic structure very similar to that found in the orthologs *CmACS7*, *CsACS2* and *CpACS27A* (Figure 3.2). The CitACS4 protein shares 91-93% similarity with CmACS7, CsACS2 and CpACS27A (Figure 3.3). These four enzymes are clustered together with the *Arabidopsis thaliana* AtACS7, in the branch corresponding to ACS type III (Figure 3.3), lacking the CDPK phosphorylation motif of type I, and the MAPK6 phosphorylation motif of type I and II ACS enzymes (Yoshida et al., 2005; Zhang et al., 2012).

Figure 3.2. (**A**) Gene structure of *CitACS4*, *CpACS27A*, *CmACS7* and *CsACS2* in watermelon, squash, melon, and cucumber, respectively. The numbers indicate the size of the three exons (filled boxes) and the two introns (black lines). The identified polymorphisms between DNA sequences in the monoecious and andromonoecious inbred lines are shown in above *CitACS4*. The missense mutation (C1477G) producing the amino acid substitution C364W in the protein is highlighted in red. (**B**) Alignment of watermelon *CitACS4* with *CpACS27A*, *CmACS7* and *CsACS2* in squash, melon, and cucumber. The amino acid changes between monoecious and andromonoecious lines in the different species are highlighted in red, blue, yellow, and green, respectively.

A



B

```
                                  G33C                A57V
CitACS4      MAIEIEIEQNPTVELSQIGTSETHGEDSPYFAGWKAYDEDPYNETTNPSGVIQMGLAENQVSFDLLEEYLEQNCEAEAN--CSGFRENALFQDYHGLLSFRTAMAGFMEEIRGGRAKFDP 118
CpACS27A     ******LD**PA*******M******************N************************V*******SDVLQAN-S*****D*******************Q******R*** 119
CmACS7       *****D**********R*************************S***********.**************E***G*G*YLN***************F***S**GS************** 120
CsACS2       ***********S*****R***************.***********S*************************E***G*G*YLN***************F***S**GS************** 120

                                                  S176A                P209S
CitACS4      NRVVLTAGATAANELLTFILANPGDALLVPTPYYPGFDRDLRWRTGVKIVPIHCDSSNNFQITPKALEEAYNTAMAMKIKVRGVLITNPSNPLGATIQRSTIEEILDFVTRKNIHLVSDE 238
CpACS27A     ************************************************************.********S**A***N*I***M*************************************** 239
CmACS7       ************************************************************************S**E*************************D*****************  240
CsACS2       ************************************************************************S*TE***********.**************D*****************  240

                                                                                                                          E355D
CitACS4      IYSGSVFSSAEFTSVAEVLESRGYKNAERVHIVYSLSKDLGLPGFRIGTIYSYNDKVVTTARRMSSFTLISSQTQRFLASMLSNRKFTEKYIKMNRDRLRKRYEMIIEGLRTAGIQCLEG 358
CpACS27A     ******************A*************************V*************************************************K***************D**K* 359
CmACS7       **********************S*************************************************************************K***************E**** 360
CsACS2       *********D*************************************V*****************************************N*********K***************E**** 360

             C364W                                S399L
CitACS4      NAGLFCWMNLSPLLKNKKLTREGEIEIWKRILKEVKLNISPGSSCHCSEPGWFRVCFANMSEKTLHVALERIHCFMERMRKEDEVN 444
CpACS27A     ***************D**TNTD****L*************************************************D*****V*KLK****A* 445
CmACS7       ***************D**-*K*************************************************D**RR*****K**N*A* 445
CsACS2       ***************D**-*K******************.*************************D**RR*****K**N*A* 445
```

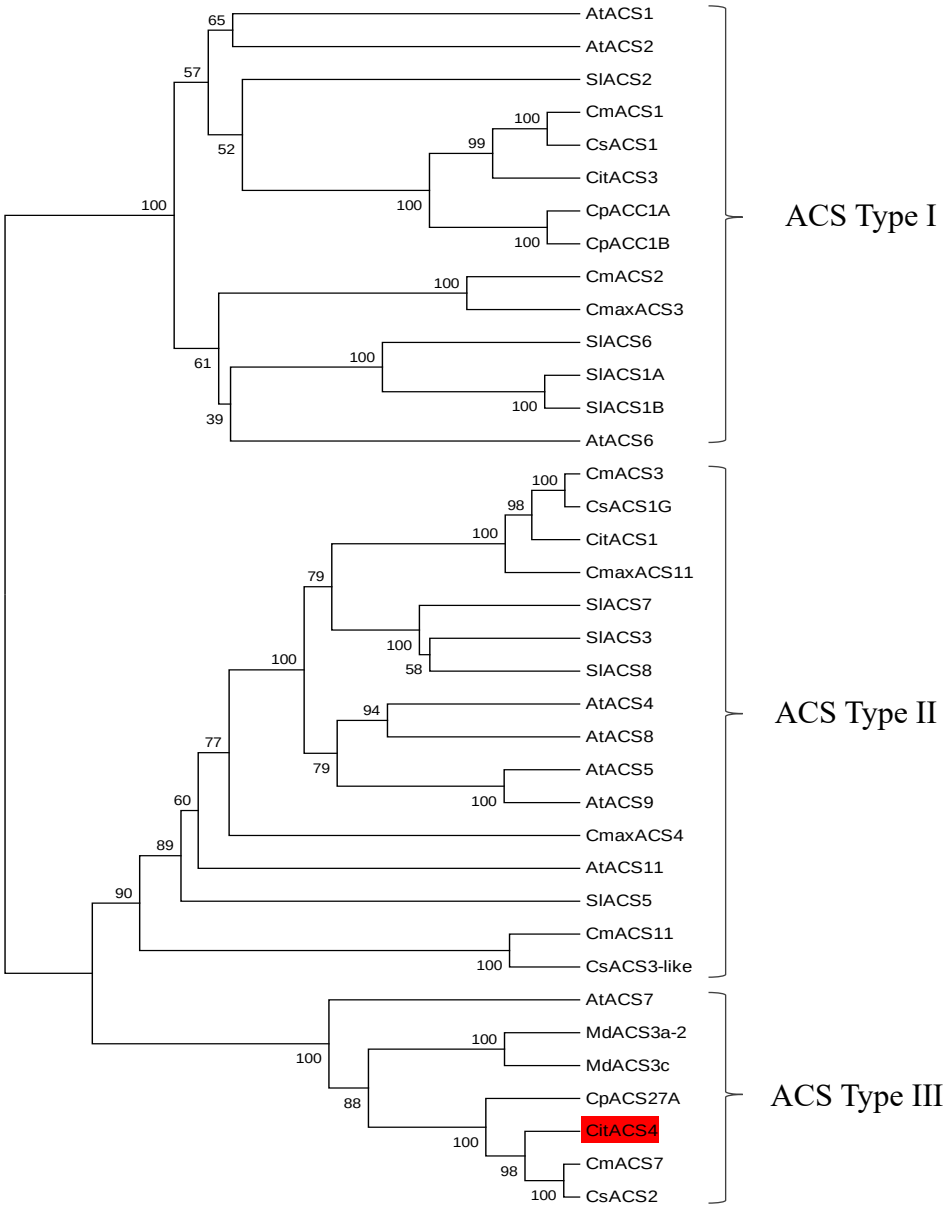


Figure 3.3. Phylogenetic analysis of CitACS4 protein. Evolutionary tree performed for 37 ACS proteins from different plants: *Arabidopsis thaliana* (AtACS1, AAM91649.1; AtACS2, AAG50097.1; AtACS4, Q43309.1; AtACS5, Q37001.1; AtACS6, Q9SAR0.2; AtACS7, AEE85169.1; AtACS8, Q9T065.1; AtACS9, Q9M2Y8.1; AtACS11, AEE82593.1), *Cucurbita maxima* (CmaxACS3, BAB47124.1; CmaxACS4, BAB47123.1; CmaxACS11, CBAA00839.1), *Cucurbita pepo* (CpACC1A, AAA33111.1; CpACC1B, AAA33112.1; CpACS27A, KF113530), *Cucumis melo* (CmACS1, BAA83618.1; CmACS2, BAB18464.1; CmACS3, ACO83163.1; CmACS7, ACG70849.1; CmACS11, XP_008445556.1), *Cucumis sativus* (CsACS1, BAA93714.1; CsACS1G, ABI33818.1; CsACS2, ACG70849.1; CsACS3-like, XP_004142909.2), *Citrullus lanatus* (CitACS1, AFI49625.1; CitACS3, ABO76787.1; CitACS4, EF154458.1), *Malus x domestica* (MdACS3a-2, AEP82201.1; MdACS3c, BAE94692.1) and *Solanum lycopersicon* (SlACS1A, AAF97614.1; SlACS1B, AAF97615.1; SlACS2, P18485.2; SlACS3, NP_001234026.1; SlACS5, NP_001234156.1; SlACS6, NP_001234164.1; SlACS7, AAK72432.1; SlACS8, AAK72431.1). The tree was inferred using the UPGMA method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) is shown next to the branches.

Expression of *CitACS4* was determined by quantitative RT-PCR in different plant organs. The gene was found to be specifically expressed in flowers, and predominantly in pistillate flowers (Figure 3.4). The expression in bisexual flowers was about half of that found in female flowers, and very low expression was detected in the male flowers. No *CitACS4* transcript was detected in the vegetative organs such as leaves or shoots (Figure 3.4).

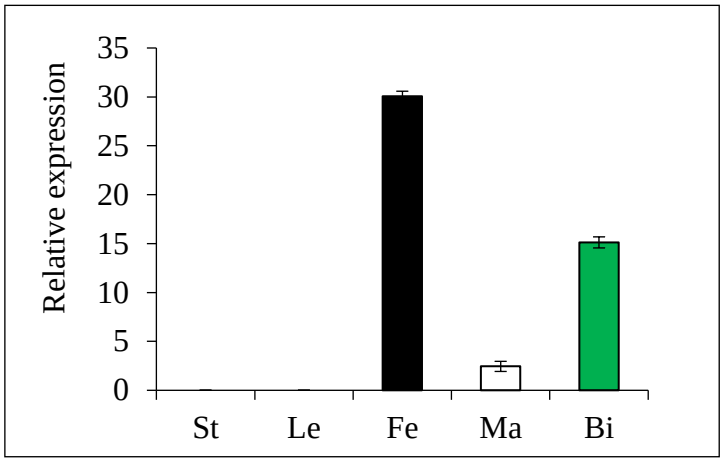


Figure 3.4. Relative expression of *CitACS4* in different tissues of watermelon cv. Premium. The values are the average and standard deviation of three biological replicates. St stem, Le leaves, Fe female flowers, Ma male flowers, Bi bisexual flowers. The utilized flowers were at early stages of development.

We have also compared the expression of *CitACS4* during the development of pistillate flowers in the monoecious (P85 and P86) and andromonoecious (P87) lines of watermelon (Figure 3.5). The maximum expression was found in the female flowers of the monoecious lines P85 and especially in the P86 at very early stages of development (stage S0, floral buds of about 4 mm). Subsequently gene expression decreased until cessation at stage S3 (floral buds of about 15 mm). In the hermaphrodite flowers of the andromonoecious line P87, *CitACS4* showed the same expression profile, although with a lower level (Figure 3.5B). No expression was detected in pistillate flowers at anthesis or post-anthesis stages of development (data not shown). In S1-S3 floral buds, where it was possible to separate the ovary from petals, style and stigma, it was found that the accumulation of *CitACS4* transcripts in the ovary was lower than that found in the other floral organs, including petals, style and stigma (Figure 3.5B). Ethylene production in flowers correlated to *CitACS4* expression. In comparison with female flowers of monoecious line P86, the hermaphrodite flowers of andromonoecious line P87 showed reduced ethylene production to a level that is similar to that produced by male flowers (Figure 3.5C).

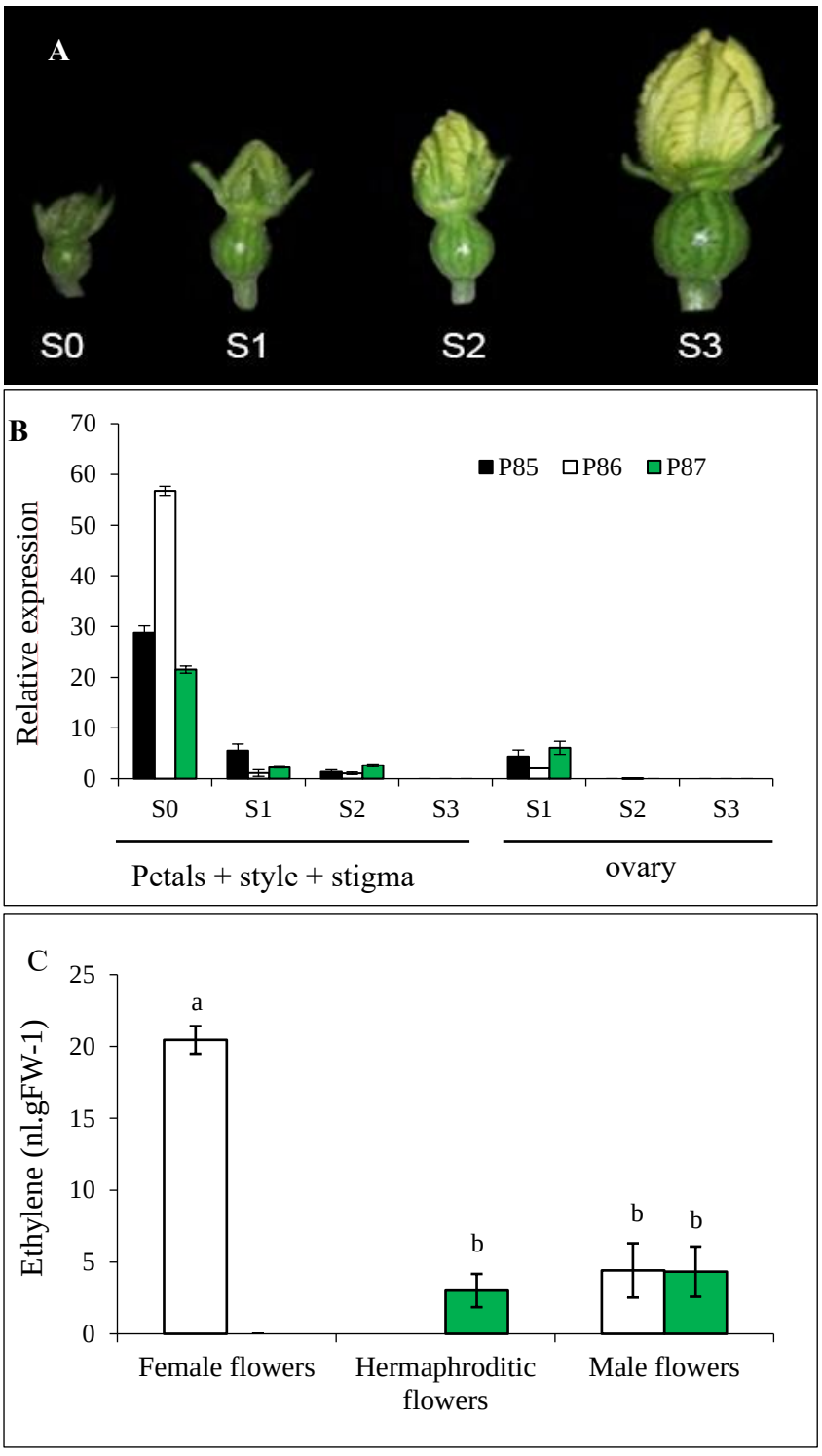


Figure 3.5. Expression of *CitACS4* and ethylene production during the development of pistillate flowers in monoecious and andromonoecious lines of watermelon. (**A**) Stages of development studied. (**B**) Relative expression of the gene in female flowers of monoecious (P85 and P86) and in the hermaphrodite flowers of andromonoecious (P87) lines. At S0, the expression corresponds to complete flowers, but in the other stages (S1 to S3), the expression in the ovary was separated to that in the rest of the floral organs (petals, style and stigma, and stamens). (**C**) Ethylene production in female, hermaphrodite, and male flowers of monoecious and andromonoecious lines. Each value is the average from at least three biological replicates. Error bars indicate standard deviation.

3.4.3. Cosegregation analysis of *CitACS4* with monoecious/andromonoecious phenotypes

Polymorphisms between the *CitACS4* gene in monoecious (P85 and P86) and andromonecious (P87) lines have been searched for, and the possible cosegregation of the alleles with the monoecious and andromonoecious phenotypes in segregating populations derived from crosses P85XP87 and P86XP87 have been analysed. In comparison with the monoecious lines, the andromonoecious line one displayed not only two SNPs and two insertions of 8 nucleotides in the second intron of the gene, but also a SNP in the third exon (C1477G) that produced an amino acid substitution of cysteine (C) by tryptophan (W) at the residue 364 of the protein (Figure 3.2A). The residue C^{364} in the monoecious lines was conserved not only in the orthologs CmACS7, CsACS2 and CpACS27A (Figure. 3.3), but also in other ACS enzymes from different plant species (data not shown), indicating that it is likely an essential residue for the enzymatic activity.

To study the possible involvement of *CitACS4* in the control of andromonoecy in watermelon, the C364W mutation in 163 plants from the F2 populations derived from two crosses P85XP87 and P86XP87 were genotyped (Table 3.2). All F2 plants homozygous for the mutated allele (genotype *AA*) were andromonoecious (average AI = 2.87±0.24 and 1.76±0.15 in each F2 population), while those homozygous for the WT allele (*MM*) were monoecious (average AI = 1.11±0.13 and 1.13±0.17 for each F2 population). The heterozygous plants (*MA*) showed a partially andromonoecious phenotype (average AI = 1.67±0.45 and 1.51±0.49 for each population), although some plants had a monoecious phenotype (Table 3.2). These data demonstrated that the andromonoecious phenotype in watermelon cosegregated with the mutated allele *A* of *CitACS4*, and therefore that the mutation C364W is likely the responsible for the andromonoecious phenotype in watermelon.

Table 3.2. Segregation of the *M* and *A* alleles of *CitACS4* with sex monoecy and andromonoecy phenotype in the two F2 populations derived from crosses monoecious x andromonoecious.

Generation	*CitACS4* genotype	Andromonoecious index (mean±sd)	No. of plants		
			Monoecious	Partially Andro-monoecious	Andro-monoecious
Parental P87	*AA*	3±0 a	0	0	9
F1 (P85XP87)	*MA*	1.52±0.19 b	0	17	0
F1 (P86XP87)	*MA*	1.35±0.21 b	3	15	0
Parental P85	*MM*	1.16±0.12 c	13	0	0
Parental P86	*MM*	1.16±0.12 c	18	0	0
F2 (P85XP87)	*AA*	2.87±0.24 a	0	0	24
	MA	1.67±0.45 b	5	41	0
	MM	1.11±0.23 c	22	0	0
F2 (P86XP87)	*AA*	2.76±0.15 a	0	0	13
	MA	1.51±0.49 b	7	34	0
	MM	1.13±0.17 c	17	0	0

a-c. Different letters indicate significant differences between genotypes.
Note that the *A* allele of *CitACS4* co-segregate with andromonoecy phenotype in the 169 F2 plants analysed.

A linkage analysis was also performed for two other sex expression traits that are also regulated by ethylene (Manzano et al., 2014): the number of nodes before the production of the first pistillate flower (pistillate flowering transition) and the number of pistillate flowers per plant (Table 3.3). The andromonoecious parental line P87 had a later pistillate flowering transition (average flowering node = 12.55) in comparison with the monoecious lines P85 (flowering node = 4.77) and P86 (flowering node 2.05) (Table 3.3). The two F1 generations had an early flowering phenotype (Table 3.3) but, in the F2 generations, the plants with the andromonoecious allele (genotype *AA*) did not flower later than those with the *M* allele (genotype *MM*). In fact, no significant differences were detected among F2 plants for three genotypes *MM*, *MA*, and *AA* (Table 3.3). For the number of pistillate flowers per plant, no significant differences were detected between andromonoecious (P87) and monoecious (P85 and P86) parental lines, nor between genotypes *MM*, *MA*, and *AA* in the F2 generation (Table 3.3). These data indicate that pistillate flowering transition and the percentage of female flowers, although controlled by ethylene, are not regulated the *CitACS4* gene.

Table 3.3. Evaluation of sex expression (transition to pistillate flowering and % pistillate flowers per plant) in F1 and F2 populations derived from crosses monoecious x andromonoecious.

Generation	CitACS4 genotype	Pistillate flowering transition	Percentage pistillate flowers
P87	*AA*	12.55±4.12 a	16.66±7.9 ab
P85	*MM*	4.77±1.92 b	13.84±5.46 b
P86	*MM*	2.05±0.72 cd	19.44±5.66 ab
F1(P85XP87)	*MA*	3.64±2.23 bc	22.05±5.15 a
F1(P86XP87)	*MA*	1.56±1.19 d	19.16±4.28 ab
F2(P85XP87)	*MM*	5.18±3.16 b	16.09±4.25 b
	MA	4.04±2.63 b	16.85±5.8 ab
	AA	4.26±2.54 b	20±7.07 ab
F2(P86XP87)	*MM*	4.65±2.54 b	15±4.3 b
	MA	3.83±2.61 b	16.78±5.27 ab
	AA	4.2±3.09 b	18.33±4.49 ab

a-d. For each trait, different letters indicate significant differences between genotypes.
No significant differences were detected among *MM*, *MA*, and *AA* genotypes for the two traits in the two F2 generations analysed, indicating that the gene *CitACS4* does not co-segregate with these two traits.

3.5. Discussion

Studies on the inheritance of watermelon sex morphotypes have indicated that monoecy is dominant to andromonoecy and controlled by a single gene with two alleles (Rosa, 1928; Poole and Grimball, 1945; Ji et al., 2015). The results from two crosses between monoecious and andromonoecious lines indicate that the F1 offspring has a predominantly monoecious phenotype. Nevertheless, the higher production of bisexual and hermaphrodite flowers in the F1 suggests that the monoecy of these two lines is not actually dominant but semi-dominant to andromonoecy. The contrasting data may reflect the existence of different monoecious or andromonoecious alleles in watermelon. Differences in the average AI between F1 offspring of the two crosses performed (monoecious x andromonoecious), should be caused by two distinct monoecious alleles in the parental lines P85 and P86, as the andromonoecious parental lines were the same in both cases. The two F1 generations produced female, bisexual and male flowers, but the F1 derived from the cross P85XP87 had a higher number of bisexual and hermaphrodite flowers and a higher AI value (AI = 1.52) than the F1 derived from the cross P86XP87 (AI = 1.35). This suggests that the monoecious allele derived from P85 is less dominant to andromonoecy than that derived from P86. Therefore, the existence

of completely dominant alleles for monoecy in other genotypes of watermelon is not excluded.

Different monoecious alleles may explain differences in the expression of *CitACS4* such as has been observed for P85 and P86 at the earliest stage of development. The higher expression of *CitACS4* in P86 and the higher production of ethylene in the pistillate floral bud can result in a higher monoecy stability and a higher dominance of the monoecious over the andromonoecious allele in the F1 generation. It is known that ethylene regulates sex determination in watermelon, not only controlling the number of floral buds that will be developed as male or pistillate flowers, but also the differentiation and development of individual floral buds as male or female flowers (Manzano et al., 2014). The arrest of stamens during the development of female flowers requires ethylene, since external treatments with ethylene inhibitors induce the transformation of female into bisexual flowers with variable stamen size and even into hermaphrodite flowers with viable pollen (Manzano et al., 2014). In this paper it is found that the ethylene required to arrest stamen development in pistillate flowers is likely to be produced by the action of *CitACS4*, a major ethylene biosynthesis gene, already proposed as a candidate for the control of monoecy/andromonoecy in watermelon (Salman-Minkov et al., 2008; Prothro et al., 2013). *CitACS4*, as other orthologs in melon, cucumber and squash (Boualem et al., 2008, 2009; Martínez et al., 2014), is mainly expressed in pistillate flowers. Moreover, the mutation C364W is a very conserved residue of CitACS4 that cosegregates with the andromonoecious phenotype in two independent F2 populations, concomitantly with a reduction in ethylene production in the floral buds that will develop as hermaphrodite flowers in andromonoecious plants of the F2 segregating populations. These data indicate therefore that the abortion of stamen during female flowers development in watermelon requires the production of ethylene mediated by *CitACS4*.

The genomic structure, nucleotide and protein sequence, and the expression profile of *CitACS4* also support that it is the orthologous gene to *CmACS7*, *CsACS2* and *CpACS27A*. Similar to the other three genes, *CitACS4* is composed of 3 exons and 2 introns of similar size, suggesting that the different genes have evolved from the same ancestral sequence. Moreover, the phylogenetic analysis carried out with different ACS enzymes on a variety of plant species has demonstrated that CitACS4 is a type-III ACC synthase with a short C-terminal tail, showing none of the identifiable phosphorylation sites in type-I and type-II ACS enzymes (Argueso et al., 2007). The expression pattern of these orthologous genes has

also been conserved through evolution. In melon, cucumber and squash the gene is specifically transcribed in the pistillate flowers, with a higher expression in female than in hermaphrodite flowers (Boualem et al., 2008, 2009; Martínez et al., 2014). In watermelon the expression of *CitACS4* is also higher in female than in hermaphrodite flowers, but a low level of transcripts were also detected in male flowers, indicating that the function of the gene is dosage-dependent. The differential expression of *CitACS4* gene in the two analysed monoecious lines, and the phenotype of F1 hybrids, also indicate that the level of *CitACS4* gene expression is essential to control the abortion of stamen development and monoecy stability through plant development.

Apart from andromonoecy, no cosegregation between *CitACS4* gene and other sex expression traits regulated by ethylene in this species, including pistillate flowering transition and the number of pistillate flowers per plant have been detected. These two sex expression traits should be regulated by other ethylene genes, which supporting previous data indicating that sexual expression of watermelon is an independent mechanism from sex determination of individual floral buds (Manzano et al., 2014). In fact, an increase of ethylene in the apical shoot does not induce the production of pistillate flowers, as occurs in melon, cucumber and squash, but on the contrary it reduces the number of pistillate flowers in the shoot (Manzano et al., 2014). This paper confirms therefore that there is a conserved molecular mechanism that makes use of the hormone ethylene for promoting the transformation of hermaphrodite to female flowers at the origin of monoecy in cucurbit species. The mechanisms that regulate the formation of male and female flower along main and lateral shoots, although still dependent on ethylene production and sensitivity, has diverged in watermelon (Manzano et al., 2014) from what occurs in other cucurbit cultivated species such as *Cucumis* (Yamasaki et al., 2001) and *Cucurbita* (Manzano et al., 2013).

4. The sex-determining gene *CitACS4* is a pleiotropic regulator of flower and fruit development in watermelon (*Citrullus lanatus*)

4.1. Abstract

In the species of the Cucurbitaceae family, the occurrence of separate male and female flowers in the same plant (monoecy) is controlled by an ethylene biosynthesis *ACS* gene, which specifically suppresses the development of stamen in the female flower. In watermelon, a mutation of loss of function in *CitACS4* promotes the conversion of female into hermaphrodite flowers, and of monoecious into andromonoecious plants. We have studied whether the ethylene produced by CitACS4 enzyme could also be involved in other ethylene regulated traits, including pistillate flowering transition and the number of female flowers per plant, the development of floral organs other than stamens, as well as fruit and seed set, and fruit development. A linkage analysis approach was performed in three independent F2 populations segregating for the two alleles of the gene (*M*, monoecious; *m*, andromonoecious), and the different traits under study. The *CitACS4 m* allele not only cosegregated with andromonoecy, but also with earlier pistillate transition, an increased number of pistillate flowers per plant, and a slower growth and maturation of petals and carpels, which delayed anthesis time in hermaphrodite flowers. The *m* allele was also found to be linked to a reduced fruit set, which was not caused by a deficiency in pollination or fertilization. The gene also affected the longitudinal and transverse growth rates of the ovary and fruit, which means that fruits from andromonoecious plants (*mm*) were rounder than those from monoecious (*MM*) ones. Taken together, these data indicate that the locus defined by the ethylene biosynthesis and sex determining gene *CitACS4* acts as a pleiotropic regulator of the complete development of the pistillate flower and the earlier development of the fruit.

4.2. Introduction

Watermelon (*Citrullus lanatus*) is a major horticultural crop worldwide, with a production of over 111 million tons in 2014 (FAOSTAT, 2018). Production related traits, including pollination efficiency and fruit set, are quite dependent on the sexual expression of the cultivar. The flowering pattern of watermelon *Citrullus* spp. is either monoecious (male and female flowers in the same plant), andromonoecious (male and hermaphrodite flowers in the same plant) or trimonoecious (female, hermaphrodite and male flowers in the same plant) (Rudich and Zamski 1985; Ji et al. 2015), Andromonoecy and trimonoecy are undesirable

traits in watermelon, since hermaphrodite flowers need to be emasculated when acting as female parents in the production of hybrid seed (Prothro et al., 2013), and also because the trait is usually associated with a reduction in fruit set and fruit quality (Monforte et al. 2005; Abdelmohsin and Pitrat 2008; Martínez et al. 2014).

Sex expression and flower development in watermelon is known to be regulated by several environmental factors and phytohormones such as ethylene and gibberellins. External treatments with ethylene and GA_3 inhibit the transition from male to female flowering and reduce the production of pistillate flowers, while treatments with the ethylene inhibitors AVG promote female flowering transition and increase the number of pistillate flowers per plant (Manzano et al., 2014; Zhang et al., 2017). High temperatures, and the concomitant reduction in ethylene production, are also responsible for the conversion of monoecious into partially andromonoecious plants. Treatments with silver sulphate, an inhibitor of ethylene action, also produce a total or partial transformation of female into hermaphrodite flowers (Zhang et al., 2017), indicating that ethylene, as occurs in other cucurbit species, is responsible for the arrest of stamen growth during female flower development (Manzano et al., 2014, 2016). Studies on the inheritance of watermelon sex morphotypes have indicated that monoecy is dominant to andromonoecy and controlled by a single gene with two alleles (Rosa, 1928; Poole and Grimball, 1945; Rudich and Zamski, 1985; Salman-Minkov et al., 2008). It has recently been demonstrated that monoecy is actually controlled by a single semi-dominant gene called *CitACS4* (Boualem et al. 2016; Manzano et al. 2016; Ji et al. 2016). The gene encodes for a flower specific ACS enzyme involved in the biosynthesis of the ethylene required for stamen arrest in the female flowers. A single missense mutation in the coding region of this gene produces an amino acid substitution of cysteine to tryptophan in residue 364 of the *CitACS4* protein (C364W), reducing the production of ethylene in pistillate floral buds, and promoting a complete conversion of female into hermaphrodite flowers, and therefore of monoecy into andromonoecy (Boualem et al. 2016; Manzano et al. 2016; Ji et al. 2016). The andromonoecious trait in other cucurbit species, including cucumber, melon and zucchini squash, also results from mutations in the orthologous ethylene biosynthesis genes *CmACS7, CsACS2,* and *CpACS27A*, respectively (Boualem et al. 2008, 2009; Martínez et al. 2014).

Besides sex determination, ethylene regulates several developmental processes associated with flower and fruit development. After pollination, the induction of ethylene production in the ovaries and petals appears to be responsible for coordinating ovary growth and petal senescence (Larsen et al., 1993; Balbi and Lomax, 2003; Wang et al., 2005; Stepanova et al., 2008). Recent studies have shown an interconnection between early ovule abortion and the size of the silique in Arabidopsis ethylene mutants (Carbonell-Bejerano et al., 2011). In squash, Martínez et al (2013) found that a reduction in ethylene production or signaling in the flower induces fruit set and early fruit development. Similarly, pollination and gibberellin treatments downregulate ethylene biosynthesis and signaling genes in tomato immediately after fruit set (Pandolfini et al., 2007; Stepanova et al., 2008). Fruit set in watermelon is unstable at low temperatures and under cloudy or rainy weather, as the activity of flower visiting insects is sluggish and the dehiscence of anthers is hindered (Tsukahara, 1988). Whether this fruit set is dependent on ethylene is unknown, but there are some data suggesting that fruit set improves in monoecious cultivars (those producing more ethylene in the female flower) in comparison to andromonoecious ones (those producing less ethylene in the female flower) (Wechter et al., 2008; Manzano et al., 2014).

Fruit shape is also related to sex expression in the species of Cucurbitaceae, which also suggests the potential involvement of ethylene in this developmental process. In cucumber and melon, the fruits developed from hermaphrodite flowers on andromonoecious plants are rounder than those derived from female flowers (Loy, 2006; Abdelmohsin and Pitrat, 2008; Sakata et al., 2013; Díaz et al., 2014). In watermelon, Rosa (1928) also reported that andromonoecious plants produced fruit that was rounder, and Poole and Grimball (1945) detected a genetic linkage between round fruits and andromonoecy, and between oval-shaped fruits and monoecy.

In the present study, we used watermelon populations that segregate for two alleles of the *CitACS4* gene, and therefore for monoecy and andromonoecy, to study whether *CitACS4*, and consequently the production of ethylene in the female flower, not only controls sex determination, but is also responsible for the regulation of the following traits: number of male and female flowers per plant, floral organ maturation, fruit and seed set, growth rate and shape of the watermelon ovary and fruit.

4.3. Materials and Methods

4.3.1. Plant material and growing conditions

Four inbred lines of watermelon (*C. lanatus*), three monoecious lines (P84, P85 and P86) and one andromonoecious line (P87), as well as the F2 generations derived from crosses between monoecious and andromonoecious lines (P84XP87, P85XP87 and P86XP87) were characterised. The number of phenotyped plants in the parent lines and in the plants genotyped as *MM*, *Mm,* and *mm* of the F2 generations is shown in Supplementary Table 4.1. Sex determination and sex expression in the crosses P85XP87 and P86XP87 were previously studied by Manzano et al. (2016). In this paper, we analysed the sex expression and sex determination in a new cross between the monoecy unstable line P84 and the andromonoecious line P87 and studied the floral and fruit traits detailed below in the three crosses.

Seeds of the different lines were simultaneously germinated in seed trays in both spring/summer and autumn/winter seasons, and seedlings transplanted in a greenhouse at the experimental station of the University of Almería (Spain), and grown under the same standard crop management of the region. Phenotypic evaluations were performed in the spring-summer and autumn-winter seasons of 2014, 2015 and 2016.

4.3.2. Genotyping for *CitACS4* alleles

The F2 seedlings from the three independent crosses (P84XP87, P85XP87 and P86XP87) were genotyped for *CitACS4* alleles before being transplanted to the greenhouse. The specific primer pairs *CitACS4MF/CitACS4gen-R1* or *CitACS4S-F/CitACS4M-R*, designed to specifically amplify the *M* allele, and *CitACS4A-F/CitACS4gen-R1* or *CitACS4S-F/CitACS4A-R*, which only amplified the *m* allele, were used for genotyping. These primers pair resulted in a 253 or 271 bp PCR fragment of the *CitACS4* gene, respectively. Plant DNA was extracted from frozen young leaves using the CTAB method (Manzano et al., 2016), and the PCR reactions were performed in the GeneAmp PCR System 2700 (Applied Biosystems). PCR reactions consisted of 35 cycles of 30 s at 95ºC, 30 s at 60ºC and 90 s at 72ºC. PCR fragments were resolved in agarose gels at 1% and plants classified in *MM*, *Mm*, and *mm*. At least 15 seedlings of each *CitACS4* genotype were transplanted to the greenhouse for phenotyping (Supplementary Table 4.1).

4.3.3. Phenotyping for monoecy stability and sex expression

To assess the level of monoecy in the different inbred lines and populations, the so-called Andromonoecy Index (AI; Martínez et al. 2014) was defined for each flower, plant and population. Pistillate flowers were scored from 1 to 3 according to their degree of stamen development. Female flowers with no stamen were scored as AI = 1, while hermaphrodite flowers with complete stamens and anthers able to produce pollen were scored as AI = 3. A score of 2 was assigned to bisexual flowers with medium-sized stamens and anthers (Figure 4.1). Based on the flower scores, the AI of each plant was assessed as the average AI of at least five pistillate flowers. The average AI in inbred lines was estimated from at least 15 plants. Plants with an AI = 1–1.2 were considered monoecious, those with AI = 1.2–2.7, partially andromonoecious, and those with AI ≥2.7 were phenotyped as andromonoecious (Manzano et al., 2016).

Sex expression in each plant was assessed by both the number of initial nodes with male flowers before the production of the first pistillate flower in the main shoot (female flowering transition), and the percentage of pistillate flowers per plant in the first 20 nodes of the main shoot. At least 15 plants were phenotyped to assess the AI and the sexual expression of both parental lines, and *MM*, *Mm*, and *mm* plants of the three F2 populations (Supplementary Table 4.1).

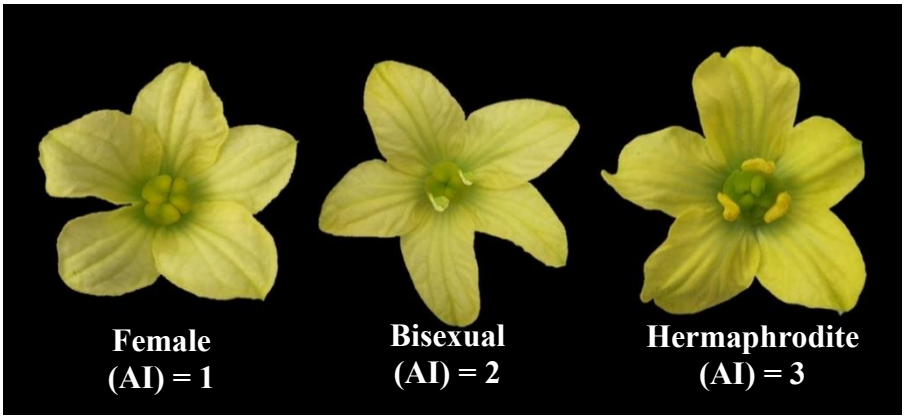


Figure 4.1. Phenotypes of watermelon hermaphrodite, bisexual and female flowers.

4.3.4. Phenotyping for floral and fruit traits

To assess floral organ development, the growth rates of ovaries and petals of each *CitACS4* genotype were determined by measuring the length of these floral organs every two days until anthesis, starting with flower buds of about 2 mm in length.

The evaluation of fruit set, and early fruit development was conducted in 15 pistillate flowers for each genotype (*MM, Mm,* or *mm*) in each of the three analysed F2 populations. Plants were hand pollinated with fresh pollen of the same plant for a total of 12 consecutive days, when the environmental conditions were similar, and always at the same time of the day (9:00-10:00 in spring, and 10:00-11:00 in fall). Pollination were done on day of anthesis for both pistillate and male flowers. To prevent flower damage and abortion, the hermaphrodite and bisexual flowers in *Mm* and mm plants were not emasculated before pollination. After hand pollination, the length and diameter of at least 15 ovaries/fruitlets were measured from anthesis to 14 days post anthesis (DPA). The ratio between the number of fruits that continued growing and the number of fruits whose growth aborted over this period of time was used to calculate the percentage of fruit set. When the number of abortions was very high, many more flowers were pollinated to reach a minimum of 10 fruits for seed set analyses (Supplementary Table 4.2).

The ovary/fruitlet shape (FS) throughout development was assessed by calculating the ratio of fruit length (FL) over maximum fruit diameter (FD) at anthesis, 14 DPA and in mature fruit (Díaz et al., 2014)**.**

In at least 10 fully mature fruits for each *CitACS4* genotype, harvested 60 days after pollination, the number of viable and non-viable seeds in 1/4 of each fruit was assessed, after which the number of seeds per kilogram of fresh fruit was calculated. The viability of seeds was determined using the floatation test. When newly extracted seeds were placed in a container with water, the submerged and floating seeds were classified as viable and non-viable, respectively. We verified that the floating seeds contained no embryo and did not germinate, while the submerged seeds contained embryos and most of them germinated under our conditions.

4.3.5. Evaluation of pollination and fertilization

Pollen-pistil interaction was analysed in female and hermaphrodite flowers of P84, P86 and P87 lines, determining the best fertility period for setting fruits in each *CitACS4* genotype, but also the possible failure associated with reduced fruit set in P87. Pistillate or hermaphrodite flower were hand pollinated with its own fresh pollen at anthesis and at -1, -2, +1 and +2 days from anthesis, always at the same time of day. Hermaphrodite flowers from P87 line were previously emasculated to avoid self-pollination before scheduled date. Flowers were fixed in FAE (Formaline: glacial acetic acid: ethanol 70%, in a ratio 1:2:17 v/v) 24 h after pollination. Fixed flowers were processed as explained in Cuevas et al. (1994) and stained using 0,1% aniline blue in phosphate buffer for observations under fluorescence microscopy (Martin, 1959) in a Nikon Labophot epifluorescence microscope. Pollen adhesion, germination, pollen tube growth, fertilization levels were determined in each flower and the results averaged for each pollination date. Pollen adhesion was estimated by counting the number of pollen grains in three different areas of the stigma (3,8 mm^2 each). Pollen germination was expressed as the ratio between pollen grains adhered and those germinated forming a pollen tube and penetrating the stigma. Pollen tubes in the style were observed and an approximate number range were indicated: 0-5 (very few pollen tubes), 5-25 (scarce number of pollen tubes) or >25 (high number of pollen tubes). In order to estimate fertilization, ovules were extracted from a section of the ovary of flowers pollinated at anthesis. Fertilization rates were calculated as the percentage of fertilized ovules. An ovule was considered fertilized if a pollen tube was present at the micropyle (Figure 4.5). The presence of pollen tubes in the ovary was also observed.

4.3.6. Linkage and statistical analysis

As the *CitACS4* gene is involved in ethylene production, we have studied whether this gene might also regulate other ethylene-regulated processes and traits in flowers and fruits. The expression of each trait was compared in *MM* and *mm* parental lines, as well as in the *MM*, *Mm*, and *mm* plants of the three segregating F2 generations. When differences between parental lines are maintained between *MM* and *mm* genotypes in the F2 generation, we conclude that *CitACS4* is cosegregating with the trait, and therefore that the gene is likely involved in its regulation.

For statistical comparison, simple and factorial analyses of variance (ANOVA) at $p < 0.05$ were performed by the STATISTIX 8.0 software package, and each two means were compared by Fisher´s Least Significant Difference (LSD) test. The Tukey's multiple comparison test was mainly used when the number of samples per comparing group was the same (only in the comparisons of pollination and fertilization events).

4.4. Results

4.4.1. Involvement of the *CitACS4* gene in sex determination and sex expression

Four types of flowers can be found in watermelon: female flowers, which develop carpels but no stamens, male flowers, developing stamens but no carpels, and hermaphrodite and bisexual flowers which are flowers producing both complete carpels and stamens, or complete carpels and partially developed stamens, respectively (Figure 4.1). To assess the sex phenotype of watermelon plants, we defined the andromonoecious index (AI). AI ranges from 1 to 3 and assesses the degree of development of stamens in each pistillate flower, and therefore the level of monoecy-andromonoecy per plant and population. Plants and lines with AI = 1-1.19 produced predominantly female flowers and were considered monoecious, those with AI = 1.2-2.69 produced female, bisexual and hermaphrodite flowers and were considered partially andromonoecious, and those with AI ≥2.7 were considered andromonoecious because they produced predominantly hermaphrodite flowers.

Table 4.1 summarizes AI and other sex related traits in the four parental lines, as well as in F2 generations derived from crosses between monoecious (P86, P85 and P84) and andromonoecious (P87) lines. The P87 line has a very stable andromonoecy, producing only male and hermaphrodite flowers with complete stamens and pollen (AI = 3). The monoecious P85 and P86 lines produced predominantly female flowers, although they also produced some bisexual flowers (AI = 1.16). The sexual phenotype of the two F1 hybrids P85XP87 and P86XP87, had an intermediate phenotype (AI = 1.52 and AI = 1.35) and were therefore classified as partially andromonoecious. The monoecious line P84 showed a higher andromonoecious index (AI = 1.26), suggesting that its monoecy is less stable than that of P85 and P86 lines. The AI of the F1 from P84XP87 (AI = 1.94) was also intermediate, although more biased to andromonoecy (Table 4.1). As previously demonstrated for P85 and P86 (Manzano et al., 2016), the monoecy of P84 was also controlled by a single semi-

dominant gene (Supplementary Table 4.3). Among the 137 phenotyped F2 plants (P84XP87), 31 were monoecious, 81 were partially andromonoecious and 25 were andromonoecious, which fits the segregation ratio 1:2:1 ($\chi^2 = 5.087$, $p = 0.078$), expected for a single semi-dominant gene controlling the trait (Supplementary Table 4.3).

Segregation data from P84XP87 confirmed that *CitACS4* regulates monoecy/andromonoecy in watermelon. Indeed, the *M* and *m* alleles of the gene cosegregated with either monoecious or andromonoecious phenotypes in all analysed F2 plants. Homozygous *MM* and *mm* plants were monoecious and andromonoecious, respectively, while heterozygous *Mm* plants had a partially andromonoecious phenotype (Table 4.1).

Table 4.1. Comparison of andromonoecious index (AI), pistillate flowering transition, and percentage of pistillate flowers per plant in monoecious and andromonoecious plants from parental, F1 and F2 generations.

Generation	*CitACS4* genotype	AI	Sex phenotype	Pistillate flowering transition	Percentage pistillate flowers/plant
P86	*MM*	1.16cd	Mono	4.70ab	20.00ab
P87	*mm*	3a	Andro	5.58a	25.42a
F1	*Mm*	1.35bc	PA	1.56c	19.16b
F2	*MM*	1.11d	Mono	4.65ab	15.00c
	Mm	1.51b	PA	3.83b	16.78bc
	mm	2.76a	Andro	4.20ab	18.33bc
P85	*MM*	1.16c	Mono	4.77ab	13.84c
P87	*mm*	3a	Andro	5.58a	25.42a
F1	*Mm*	1.52b	PA	3.64b	22.05ab
F2	*MM*	1.11c	Mono	5.18ab	16.09c
	Mm	1.67b	PA	4.04ab	16.85c
	mm	2.87a	Andro	4.26ab	20.00b
P84	*MM*	1.26c	Mono	6.78ab	22.78bc
P87	*mm*	3a	Andro	5.58b	25.42ab
F1	*Mm*	1.94b	PA	5.92b	22.92bc
F2	*MM*	1.22c	Mono	8.22a	21.22c
	Mm	1.8b	PA	7.46a	22.54bc
	mm	2.81a	Andro	6.52b	27.58a

The traits were assessed in monoecious (*MM*) and andromonoecious (*mm*) parental lines, and in *MM*, *Mm*, and *mm* F2 plants derived from monoecious x andromonoecious crosses. Mono, monoecious; PA, partially andromonoecious; Andro, andromonoecious. Statistical analysis was performed using LSD test ($p \leq 0.05$), and the different letters indicate significant differences between genotypes of the same cross.

A linkage analysis was also performed for two other sex expression traits that are known to be regulated by ethylene: the number of nodes before the production of the first pistillate flower (pistillate flowering transition) and the number of pistillate flowers per plant (Table 4.1). For pistillate flowering transition no difference was detected among the four parental lines. For the number of pistillate flowers per plant, however, monoecious lines P85 and P86 showed less number of female flowers than the andromonoecious P87, and in F2 populations derived from crosses between these lines, *MM* plants also produced fewer female flowers than *mm* plants (Table 4.1).

4.4.2. Involvement of the *CitACS4* in floral organ maturation

The anthesis time was measured as the number of days it takes a floral bud of 2 mm in length to reach anthesis in a minimum of 10 flowers for each *CitACS4* genotype. Female and male flowers in monoecious (*MM*) plants differed in the time to reach complete maturation at anthesis, but no statistical difference was found between hermaphrodite and male flowers in andromonoecious plants (Table 4.2), suggesting that the presence of stamens delayed the aperture of both male and hermaphrodite flowers.

Given that male and female flowers differ in their production of ethylene (Manzano et al., 2016), we studied whether the ethylene derived from *CitACS4* expression could also regulate anthesis time in pistillate and male flowers. The anthesis time of pistillate flowers in the andromonoecious line P87 was delayed (average = 9.3 days) in comparison with that in the monoecious lines P86 (average = 6.0 days), P85 (average = 6.01 days) and P84 (average = 6.0 days) (Table 4.2). In the F2 generation of P86XP87 and P85XP87, the hermaphrodite flowers of andromonoecious *mm* plants also delayed anthesis in comparison with female flowers of *MM* plants (Table 4.2).

Table 4.2. Comparison of anthesis time (days) in pistillate and male flowers of monoecious (*MM*) and andromonoecious (*mm*) plants from parental and F2 generations.

Generation	*CitACS4* genotype	Anthesis time (days)	
		Female flowers	Male flowers
P86	*MM*	6.0d	7.8c *
P87	*mm*	9.3ab	9.8b
F2(P86XP87)	*MM*	8.4c	11.0a *
	Mm	8.6bc	9.8b *
	mm	9.8a	11.0a
P85	*MM*	6.1c	7.8a *
P87	*mm*	9.3b	9.8b
F2(P85XP87)	*MM*	7.2c	10.0b *
	Mm	8.9b	10.0b *
	mm	11.7a	11.2b
P84	*MM*	6.0c	8.5d *
P87	*mm*	9.3a	9.8bc
F2(P84XP87)	*MM*	8.4b	10.4a *
	Mm	9.0ab	9.4c
	mm	9.0ab	10.0ab

The trait was assessed in monoecious (*MM*) and andromonoecious (*mm*) parental lines, and in F2 plants (*MM*, *Mm*, and *mm*) derived from monoecious x andromonoecious crosses. Statistical analysis was performed using the LSD test ($p \leq 0.05$). Different letters specify significant differences between genotypes within the same cross; *indicates significant differences between male and pistillate flowers of the same genotype and generation.

Differences in anthesis times were also found in male flowers of the monoecious (P84, P85 and P86) and andromonoecious (P87) lines, but those differences were not maintained among the *MM*, *Mm* and *mm* genotypes in the F2 generations (Table 4.2), suggesting that *CitACS4* and ethylene could control the maturation time of the pistillate flower, but not that of the male flower.

4.4.3. Involvement of the *CitACS4* gene in ovary and fruit development

A linkage analysis was performed between the *CitACS4* gene and floral organ size throughout development, including ovary and fruit. At earlier stages of pistillate flower development, ovary growth rate in *MM*, *Mm,* and *mm* flowers was very similar (Figure 4.2; Supplementary Table 4.4). Significant differences in the petal and ovary size were detected, however, after 6 days, when the ovary and the corolla of *MM* flowers in both parental lines and F2 plants were larger than those of *mm* flowers (Figure 4.2; Supplementary Table 4.4).

This higher growth rate in *MM* flowers was maintained up to anthesis, but given that *MM* flowers reached anthesis earlier than *mm* flowers, the size of the ovary at anthesis was smaller in the female flowers of monoecious *MM* plants than in the hermaphrodite flowers of *mm* plants (Figure 4.2, Supplementary Table 4.4). The ovary growth rate of heterozygous *Mm* flowers was intermediate with respect to that of the two homozygous genotypes in the three crosses (Figure 4.2). These data demonstrated that the larger ovary size of *mm* flowers at anthesis is not due to a higher growth rate of the organ throughout development, but rather because the full maturation of floral organs and anthesis is delayed in hermaphrodite flowers.

Immediately after anthesis, pollinated fruits of monoecious *MM* lines (P86, P85 and P84) also grew at a higher rate than those of the andromonoecious *mm* line, but these differences were not detected between *MM* and *mm* fruits in the F2 populations of the three crosses (Figure 4.3, Supplementary Table 4.4). At 14 DPA *MM* fruits were larger than *mm* fruits, but in the F2 generations significant differences between *MM* and *mm* fruits were only detected in the P84XP87 cross (Figure 4.3, Supplementary Table 4.4).

We also found a close linkage between the gene and fruit shape, in that *mm* fruits were rounder than *MM* ones. The fruit shape, estimated as the ratio between fruit length and width (FS), did not change between the two experimental seasons (Table 4.3). The P87 line produced round-shaped fruits at anthesis, 14 DPA and at the mature stage, while the monoecious lines P85 and P86 displayed oval-shaped fruits at anthesis (Figure 4.3; Table 4.3) which became rounded 14 DPA and at maturation (Table 4.3). In the F2 generations derived from P85XP87 and P86XP87, *MM* fruits at anthesis were also more elongated than *mm* fruits, but no significant difference was detected between *MM* and *mm* fruits at 14 DPA or at the mature stage. Fruits of the monoecious line P84, on the other hand, displayed a more elongated-shaped at anthesis, becoming oval-shaped at 14 DPA and at maturation, in both experimental seasons (Table 4.3), and in the F2 from cross P84XP87, *MM* plants produced a more elongated fruit than *mm* plants (Table 4.3), suggesting that the elongated fruit shape of P84 is linked to *CitACS4*.

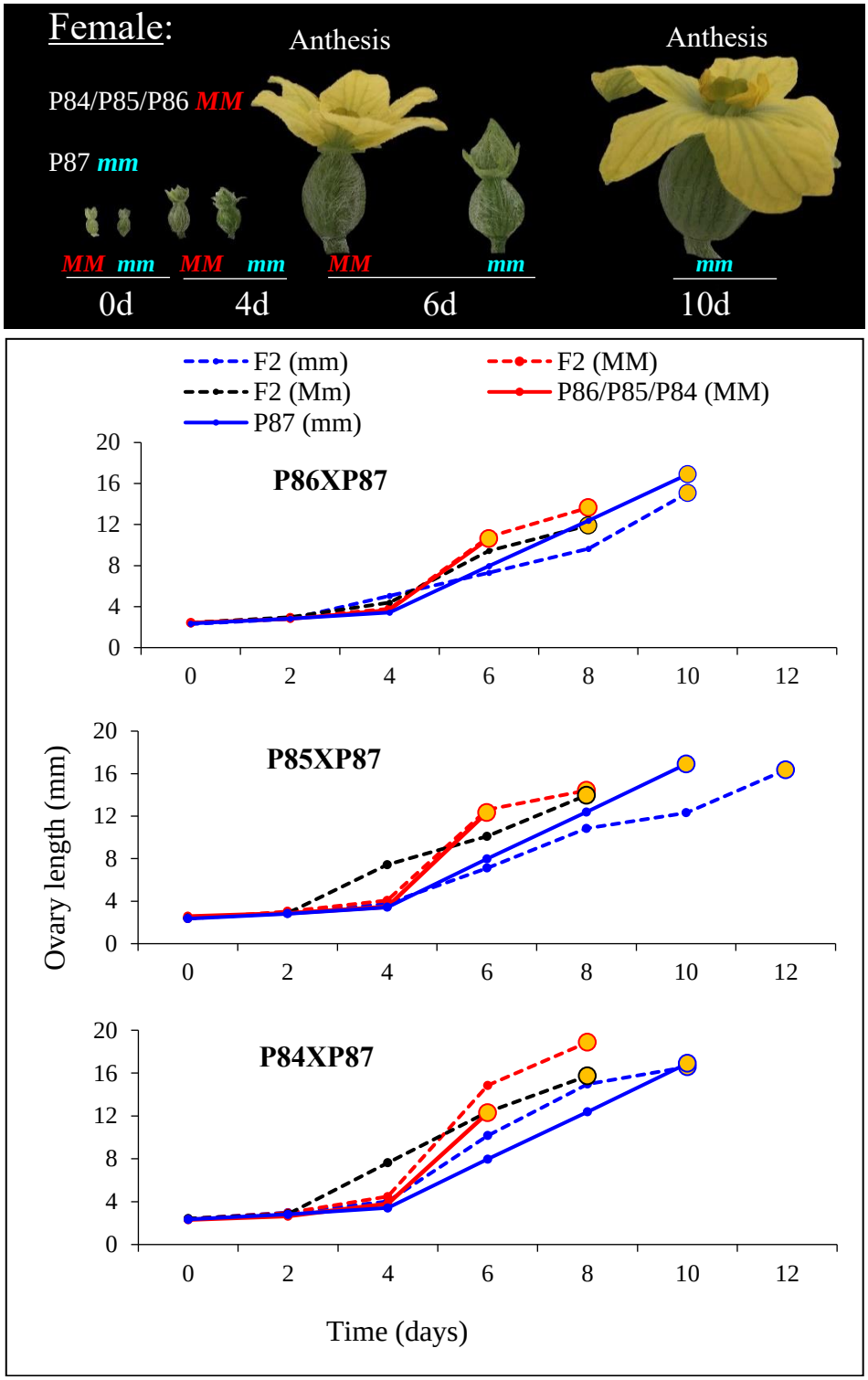


Figure 4.2. Ovary growth rate in monoecious *MM* (P84, P85, and P86) and andromonoecious mm (P87) lines, and in *MM*, *Mm*, and *mm* plants of three F2 generations derived from crosses between monoecious and andromonoecious lines. Flowers were labelled when they were 2mm in length and ovaries were measured every two days until anthesis. Average of 10-15 flowers and fruits for each *CitACS4* genotype. Yellow circles show the average anthesis day. Significant differences between genotypes on each sampled day are shown in Supplementary Table 4.4.

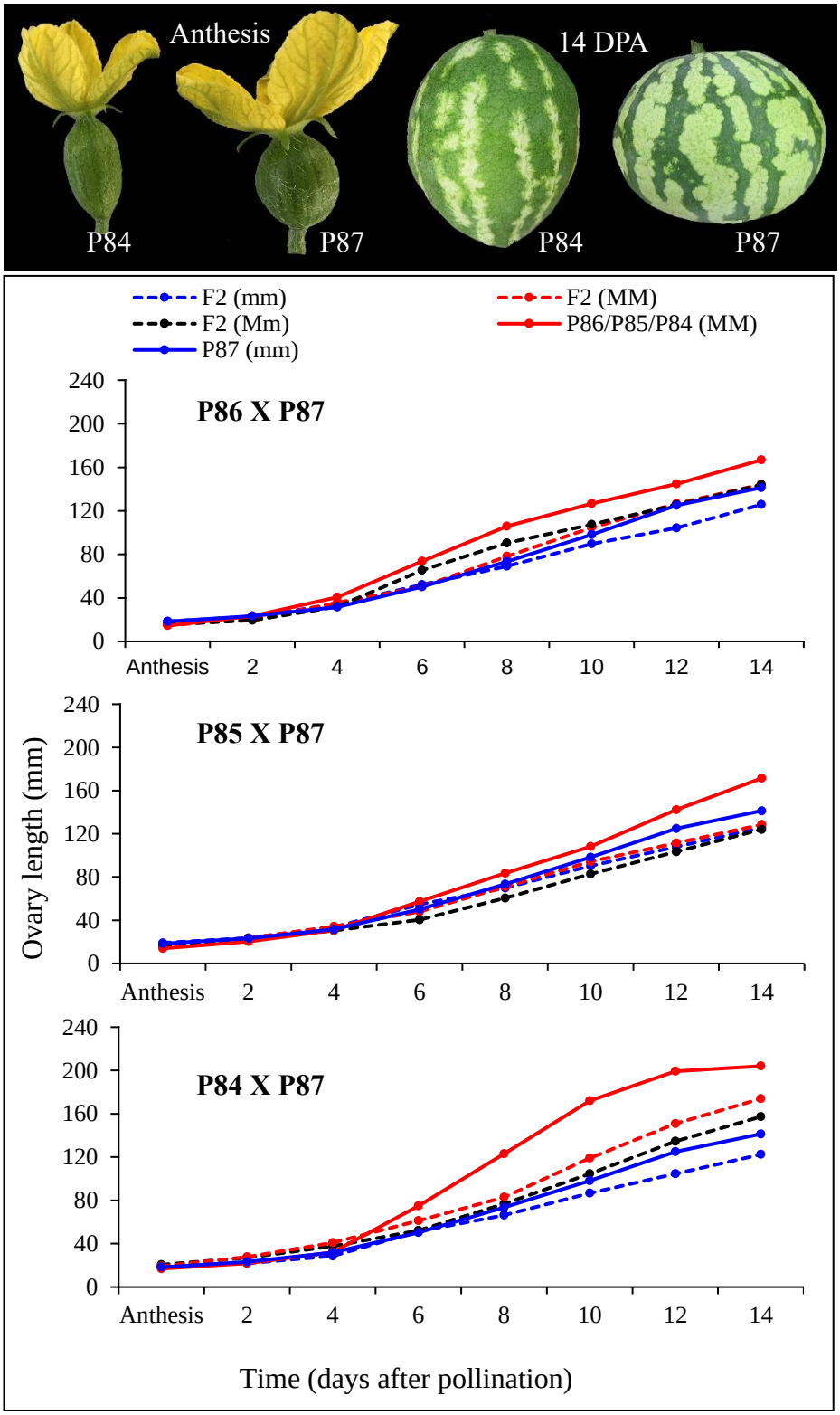


Figure. 4.3. Fruit growth rate in monoecious *MM* (P84, P85, and P86) and andromonoecious *mm* (P87) lines, and in *MM*, *Mm*, and *mm* plants of three F2 generations derived from crosses between monoecious and andromonoecious lines. Fruit size was recorded every two days from anthesis up to 14 DPA. Average of 10-15 flowers and fruits for each *MM*, *Mm*, and *mm* genotypes. Significant differences between genotypes on each sampled day are shown in Supplementary Table 4.4.

Table 4.3. Fruit shape index (FS) in *MM*, *Mm*, and *mm* plants of parental lines and F2 generation.

Generation	*CitACS4* genotype	FS (lenght/width ratio)					
		Spring/Summer			Autumn/Winter		
		Anthesis	14dpa	Mature	Anthesis	14dpa	Mature
P86	*MM*	1.46b	1.17a	1.15abc	1.33a	1.16a	1.06a
P87	*mm*	1.26c	1.12a	1.07c	1.22b	1.06b	0.99b
F2 (P86XP87)	*MM*	1.63a	1.15a	1.21a	1.28a	1.16a	1.09a
	Mm	1.41b	1.18a	1.17ab	1.30a	1.13ab	1.15a
	mm	1.26c	1.09a	1.11abc	1.23b	1.13ab	1.05ab
P85	*MM*	1.39a	1.15a	1.13ab	1.30a	1.14b	1.03bc
P87	*mm*	1.26b	1.12a	1.07b	1.22b	1.06c	0.99c
F2(P5XP87)	*MM*	1.42a	1.18a	1.14ab	1.31a	1.20ab	1.13a
	Mm	1.42a	1.14a	1.19a	1.26ab	1.22a	1.14a
	mm	1.22b	1.10a	1.09ab	1.24ab	1.15ab	1.07ab
P84	*MM*	1.48bc	1.21ab	ND	1.59b	1.35b	1.32b
P87	*mm*	1.26d	1.12c		1.22d	1.06c	0.99c
F2(P84XP87)	*MM*	1.82a	1.34a		1.82a	1.58a	1.69a
	Mm	1.50b	1.17bc		1.45c	1.45ab	1.31b
	mm	1.31cd	0.94c		1.17d	1.05c	0.96c

The trait was assessed in parental monoecious (*MM*) and andromonoecious (*mm*) lines, and in F2 plants (*MM*, *Mm*, and *mm*) derived from monoecious x andromonoecious crosses. FS was calculated as the ratio between fruit length and width. Data are the average of a minimum of 10 fruits per genotype. Statistical analysis was performed using the LSD method ($p \leq 0.05$) and the different letters indicate significant differences between *CitACS4* genotypes. ND, non-determined.

4.4.4. Involvement of *CitACS4* in fruit and seed set

Fruit and seed set were determined by hand-pollinating a minimum of 15 flowers for each *CitACS4* genotype, and then assessing the number of setting fruits and viable seeds in at least 10 mature fruits. Since emasculation could decrease fruit and seed set, none of the bisexual or hermaphrodite flowers were emasculated before hand pollination.

Fruit set varied between monoecious and andromonoecious lines under the two studied conditions (spring/summer and autumn/winter), with P85 and P86 showing significantly higher fruit set than P84 and P87. We have determined whether these differences cosegregated with *CitACS4* alleles in the F2 generations. As expected, no difference was detected between *MM* and *mm* plants in the F2 generation of the cross P84XP87 (Figure 4.4). In the P85XP87 and P86XP87 crosses, however, the higher fruit set of the monoecious parental lines was also observed in the monoecious *MM* plants of the segregating F2

generations (Figure 4.4). Heterozygous *Mm* plants showed an intermediate percentage of fruit set (Figure 4.4).

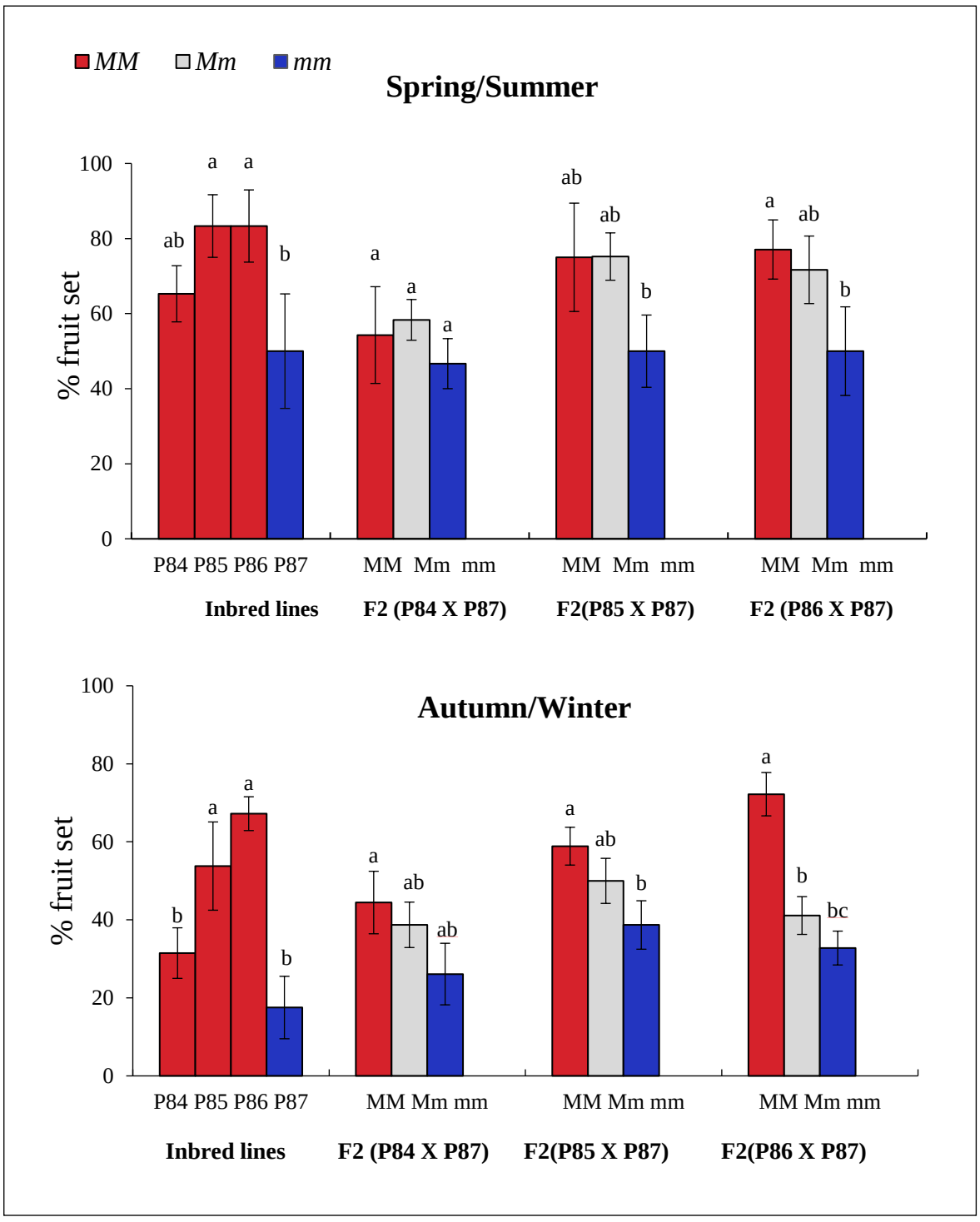


Figure 4.4. Percentage of fruit set in monoecious (P84, P85, P86) and andromonoecious (P87) lines, and in *MM*, *Mm*, and *mm* plants of three F2 generations derived from crosses P84XP87, P85XP87 and P86XP87. Bars represent SE of at least 15 fruits. Different letters indicate significant differences between genotypes ($p \leq 0.05$).

Table 4.4 compares the production of viable seeds in *MM*, *Mm*, and *mm* fruits in parental lines and F2 generations of plants growing under autumn/winter conditions. The higher number of seeds in the monoecious *MM* lines (P84, P85 and P86) cosegregated with *MM* plants only in the F2 population of the cross P86XP87 (Table 4.4). In the crosses P84XP87 and P85XP87, however, both *MM* and *mm* F2 fruits produced a very low number of seeds, displaying no significant difference in seed set (Table 4.4).

Table 4.4. Number of viable seeds per kg of fruit pulp in *MM*, *Mm*, and *mm* plants of parental lines and F2 generations growing under autumn/winter conditions.

Generation	*CitACS4* Genotype	Number of seeds
P84	*MM*	53.34a
P87	*Mm*	0b
F2 (P84XP87)	*MM*	0.69b
	Mm	0.59b
	mm	0b
P85	*MM*	50.04a
P87	*Mm*	0b
F2 (P85XP87)	*MM*	12.48b
	Mm	1.35b
	mm	4.12b
P86	*MM*	70.32a
P87	*Mm*	0 c
F2 (P86XP87)	*MM*	50.73ab
	Mm	27.6bc
	mm	10.61c

SE of at least 10 fruits per line and generation. Statistical analysis was performed using the LSD method ($p \leq 0.05$) and the different letters indicate statistical differences between genotypes within the same cross.

4.4.5. Pollination and fertilization in female and hermaphrodite flowers

To investigate the possible factors accounting for the differences in fruit and seed set between female and hermaphrodite flowers, we compared pollen-stigma interaction, pollen tube germination and growth, and ovule fertilization in pistillate flowers of monoecious (P86 and P84) and andromonoecious (P87) lines at anthesis, and at -2, -1, +1 and +2 days post anthesis (DPA). The results are shown in Table 4.5 and Figure 4.5. In female flowers of P84 and P86, pollen adhesion and germination occur similarly between flowers at different

phenological stages of the flowers, reaching a maximum around anthesis. In hermaphrodite flowers of P87, however, pollen adhesion was clearly reduced at -1 and -2 DPA, and pollen germination was almost nil in floral buds at -2 DPA (Table 4.5). The dynamic of pollen tube growth followed the same trend in female and hermaphrodite flowers of P86, P84 and P87, with a maximum number of pollen tubes in styles of flowers that were pollinated at anthesis (Table 4.5, Figure 4.5). Nor were any differences found in pollen tube penetration and fertilization between female and hermaphrodite flowers at any of the floral stages at which they were pollinated (Table 4.5, Figure 4.5). In the ovary of flowers pollinated at anthesis and +1 DPA, pollen tubes were frequently observed close to the ovules. When flowers were pollinated at anthesis, fertilization rates were similar in the three *CitACS4* genotypes, although slightly higher in P87 flowers (Table 4.5).

Table 4.5. Pollen -pistil interaction 24h after pollination at different dates in P84, P86 and P87 lines.

Line	Pollination date (days from anthesis)	Pollen adhesion (grains/mm2)	Pollen Germination (%)	Pollen tube number in style	Pollen tubes in ovary anf Fertilization 2 (%)
P84	-2	18.8b1	37.7b	5-25	Not observed
	-1	18.8ab	57.8ab	>25	Some pollen tubes
	0	22.5a	69.0a	>25	20±1
	+1	11.4b	69.1a	>25	Many pollen tubes
	+2	20.6ab	67.8a	>25	Many pollen tubes
P86	-2	10.6b	40.0b	5-25/>25	Very few pollen tubes
	-1	15.3b	42.8ab	>25	Few pollen tubes
	0	11.0b	69.5a	>25	27±1
	+1	29.8a	37.8ab	>25	Many tubes near ovules
	+2	17.1b	26.7b	<25	Some pollen tubes
P87	-2	7.1ab	0.76b	5-25/>25	Not observed
	-1	5.8b	60.3a	>25	Few pollen tubes
	0	25.4a	60.2a	>25	33±8
	+1	23.2ab	62.1a	>25	P.t. near ovules
	+2	16.9ab	49.8a	<25	P.t. near ovules

[1] Mean values (n = 10) followed by different letters in each line and column indicate means significant differences at $p < 0.05$ by Tukey test.

[2] Fertilization expressed as mean value± standard error (%).

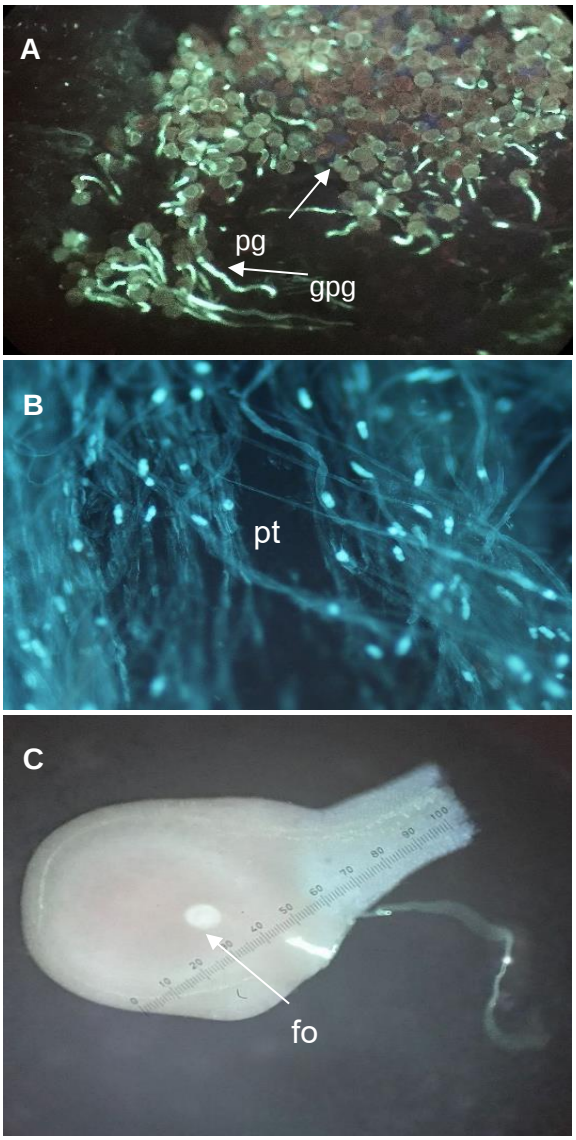


Figure 4.5. Pollination and fertilization in andromonoecious line P87. (**A**) Pollen adhesion and germination in the stigma. (**B**) Pollen tubes growing on the style. (**C**) Fertilized ovule. pg, pollen grain; gpg, germinated pollen grain; pt, pollen tubes; fo, fertilized ovule.

4.5. Discussion

In the monoecious species of the Cucurbitaceae family, sex determination, i.e. the conversion of a putative hermaphrodite floral meristem into a female or a male flower, is known to be regulated by ethylene (Manzano et al., 2014). The key regulator is an ACS enzyme encoded by the orthologous genes *CmACS7*, *CsACS2*, *CpACS2/7*, and *CitACS4* in melon, cucumber, zucchini, and watermelon, respectively. These genes are specifically expressed in female floral buds at early stages of development, and their function results in

the arrest of stamen development during the formation of the female flower (Boualem et al. 2008, 2009; Li et al. 2009; Martínez et al. 2014; Manzano et al. 2016; Ji et al. 2016). Although sex determination seems to be the main function of these genes, they could also control other developmental processes regulated by ethylene. In this paper we have studied whether the watermelon *CitACS4* gene could also be involved in sex expression, floral organ development, including petals, ovaries, and fruits, as well as in fruit and seed set. Results demonstrated that the ethylene produced in earlier pistillate floral buds is enough to control the entire development of pistillate flowers and fruits.

In *CitACS4* loss of function mutants (*mm*), female flowers are converted into hermaphrodite flowers, and monoecious into andromonoecious plants (Manzano et al. 2016; Boualem et al. 2016; Ji et al. 2016). Crosses indicated that the monoecious allele *M* of P84, P85 and P86 is semi-dominant to the andromonoecious allele *m* of P87. The stability of sex determination, however, varied among monoecious lines (*MM*), showing P85 and P86 a more stable monoecy than P84. Consequently, the P84 line not only produced male and female flowers, but also several bisexual flowers with stamens at different developmental stages.

4.5.1. Regulation of sex expression

Sex expression, i.e. the transition from male to female phases of development (pistillate flowering transition), and the female to male flower ratio are regulated by ethylene in watermelon and other cucurbit species, although in a completely opposite way. In *Cucumis* and *Cucurbita* species, ethylene promotes femaleness (Rudich, 1990; Perl-Treves, 1999; Manzano et al., 2011, 2013), while in watermelon it stimulates maleness, delaying female flowering transition, and reducing the number of pistillate flowers per plant (Sugimaya et al., 1998; Manzano et al., 2014; Zhang et al., 2017). Our data indicate that the reduced ethylene in *mm* pistillate floral buds could affect the number of pistillate flower per plant but not the pistillate flowering transition. In previous data we found that sex expression traits were not affected by *CitACS4* gene. The differences between the previous and current results are likely due to the influence of grafts. In the previous experiments, watermelon plants were grafted on *Cucurbita* rootstocks, which surely altered their sexual expression (Manzano et al., 2014). In the present study, however, none of the plants were grafted, and monoecious and andromonoecious plants showed more noticeable differences in the number of pistillate and male flowers per plant. Even so, P86 and P87 parental lines showed no difference in sex

expression, which suggests the influence of the genetic background or the existence of other genetic factors regulating this trait. Therefore, although sex determination and sex expression are different mechanisms (Manzano et al., 2014), our present data demonstrate that the ethylene required for sex determination in the pistillate flowers, converting the putative hermaphrodite floral buds into female flowers, can influence the number of female flower per plant. Given that sex expression mechanisms should be controlled from the apical meristem, it is likely that the ethylene produced in pistillate floral buds (encoded by *CitACS4*) can influence the production and/or the action of ethylene in the apical shoot, as proposed by Manzano et al. (2013) in zucchini squash.

4.5.2. Coordination of floral organ development

Our data also demonstrate that *CitACS4* is a coordinator of floral organ development, acting as a repressor of stamen development, but also as a promoter of ovary and corolla growth, and consequently of the maturation and aperture of the pistillate flower. The growth rates of both petals and ovaries were significantly higher in female flowers of *MM* plants than in hermaphrodite flowers of *mm* plants. Consequently, flower maturation and anthesis time are delayed in hermaphrodite flowers in comparison to female ones, which leads to larger *mm* ovaries than *MM* ones at anthesis, despite their reduced growth rates. The differences in flower maturation between *mm* and *MM* plants did not affect male flowers, which showed a significantly longer anthesis time than female flowers in *MM* but not in *mm* plants (Table 4.2). These data strongly suggest that the masculinization of the pistillate flower in the andromonoecious *mm* plants decreases the growth rate of the flower and delays its maturation and aperture for about two days, in a similar way as occurs in male flowers. Similar results have been found in zucchini, where the delayed anthesis of bisexual and hermaphrodite flowers resulted in ovaries much larger than those of female flowers (Martínez et al., 2013). Since female flowering occurs at later stages of the plant development, it is likely that the acceleration of the anthesis in female flowers was a coevolutionary mechanism that ensured pollination during the evolution of monoecy in the Cucurbitaceae family.

The role of ethylene as a promoter of carpel development, but also as a repressor of stamen development has been found not only in the unisexual flowers of cucurbit species (Boualem et al. 2008, 2009; Li et al. 2009; Martínez et al. 2014; Manzano et al. 2016), but also in the

hermaphrodite flowers of Arabidopsis and other species. Overexpression of the ethylene biosynthesis cucumber gene *CsACO2* represses stamen development in Arabidopsis (Duan et al., 2008), while downregulation of the ethylene receptor gene *ETR1* reduces the ethylene signaling repressor CTR1, resulting in the production of female flowers in Arabidopsis (Wang et al., 2010). Transgenic tobacco silencing an *ACO* gene has shown female sterility due to an arrest of megasporogenesis (De Martinis et al., 1999). Here we demonstrate that ethylene is also a positive regulator of petal development and maturation, and therefore of corolla aperture and anthesis time.

4.5.3. Regulation of fruit set and development

Data indicated that *CitACS4* affects developmental events occurring before or at anthesis, including fruit shape and fruit setting, but not those occurring after anthesis and pollination such as fruit growth rate and final size. In the P84XP87 cross, the monoecious *M* allele was linked to elongated fruits while the andromonoecious *m* allele was rather linked to round-shaped fruits. Since the final shape of a fruit depends on ovary shape at anthesis (Perin et al., 2002), it is not difficult to realize that a gene like *CitACS4*, which is specifically expressed in floral buds at earlier stages of development, can regulate the final shape of the fruit. The association of hermaphrodite flowers with round-shaped fruits was first reported by Rosa (1928), and later by Kubicki (1962) and Wall (1967) in melon. Among QTLs controlling fruit shape in melon, that in LGII seems to be a pleiotropic effect of the sex determining gene *CmACS7* (Perin et al., 2002; Díaz et al., 2014). The monoecious (*M*) and andromonoecious (*m*) alleles of the cucumber gene *CsACS2* also cosegregate with elongated and round fruits, respectively, although a novel allele of the gene (*m1*), encoding for a truncated protein, is responsible for elongated fruit shape and andromonoecy (Tan et al., 2015). In zucchini, hermaphrodite flowers of monoecious unstable cultivars produce larger ovaries and fruits, but fruit shape is not altered (Martínez et al. 2014).

The cosegregation between *m* allele and a reduced fruit set in P85XP87 and P86XP87 crosses could be the result of *CitACS4*, but the existence of other linked genes cannot be ruled out. Moreover, the fact that the monoecious line P84 does not differ in fruit set with respect to the andromonoecious one P87 also suggests that the trait is influenced by other unlinked loci. The role of ethylene in fruit setting has not been studied in depth. A downregulation of ethylene biosynthesis and signaling genes has been observed immediately

after anthesis in pollinated, GA_3 treated and parthenocarpic fruits of tomato (Vriezen et al., 2008) and zucchini (Martínez et al., 2013). Ethylene produced in the ovules appears to be responsible for both the ovule lifespan and the fate of the ovary/fruit in tomato (Olimpieri et al., 2007) and Arabidopsis (Carbonell-Bejerano et al., 2010, 2011), controlling fruit set in response to GA in Arabidopsis unfertilized ovaries. In zucchini, the inhibition of ethylene biosynthesis or response is sufficient to induce the set and early development of the fruit in absence of pollination, demonstrating a direct involvement of ethylene in fruit set (Martínez et al., 2013). This post-anthesis ethylene, which could be involved in ovule senescence and fruit abortion, does not appear to be the same as the one responsible for *mm* fruit abortion in watermelon. Although this interesting finding requires more research, it seems that fruit set requires higher ethylene production in the immature flower buds, probably for a coordinated development and maturation of floral organs at anthesis, but lower ethylene production in ovules and fruits immediately after anthesis, because at this later stage ethylene could trigger ovule senescence and consequently fruit abortion.

Under unfavourable environmental conditions of autumn/winter, the correct set of seeds in fruits was found to be linked to the *M/m* locus only in one of the analysed F2 populations (P86XP87), where monoecious *MM* fruits had higher seed yield than andromonoecious *mm* ones. This trait is very influenced by environmental conditions, especially temperature. The lack of linkage between seed set and *M/m* locus in the other two crosses could indicate the existence of other major loci regulating this trait in watermelon. Therefore, the role of the *M/m* locus in the regulation of watermelon seed set will require further experimental work. Comparison of pollen adhesion, pollen tube growth and ovule fertilization in monoecious and andromonoecious lines shedded also no light on seed abortive mechanisms in fruits derived from hermaphroditic *mm* flowers. The pollination window at which pollen adheres to the stigma is slightly delayed in *mm* flowers, but pollen adhesion and pollen tubes observed at anthesis and +1 and +2 DPA, ensured fertilization in both *MM* and *mm* flowers. Therefore, it is likely that the loss of seeds in the P87 line is not due to a lack of pollination or fertilization events, but rather to a premature abortion of fertilized ovules. Ethylene plays a significant role in ovule development and female gametophyte fertility (Tsai et al., 2008; Clark et al., 2010). In Arabidopsis the onset of ovule senescence and the time window for the pistil to respond to GA treatments is modulated by ethylene (Carbonell-Bejerano et al., 2011). Silencing of ethylene biosynthesis genes in transgenic tobacco plants results in a reversible inhibition of ovule development (De Martinis and Mariani, 1999). Moreover,

ethylene biosynthesis and signaling genes have been found to be hypo-methylated in the female sterile rice mutant *fsv*, in which the ethylene genes were up-regulated and then down-regulated during ovule development (Yang et al., 2016; Liu et al., 2017).

Taken together, the results presented in this paper indicate that, in addition to arresting stamen growth and development, and determining the sex identity of the female flower, the ethylene biosynthesis gene *CitACS4* is capable of regulating several developmental processes that occur in the pistillate flower and in the early development of the fruit (Figure 4.6). The decrease in the production of ethylene associated with the loss of function *m* allele, prevents stamen arrest, but inhibits the development of the petals and carpels, making the flower reach anthesis about 4 days later. The result at anthesis is a hermaphrodite flower with a rounder and larger ovary than that of the female flower. The lack of ethylene during the development of the hermaphrodite flower could also explain the reduced fruit set found in the andromonoecious *mm* plants (Figure 4.6).

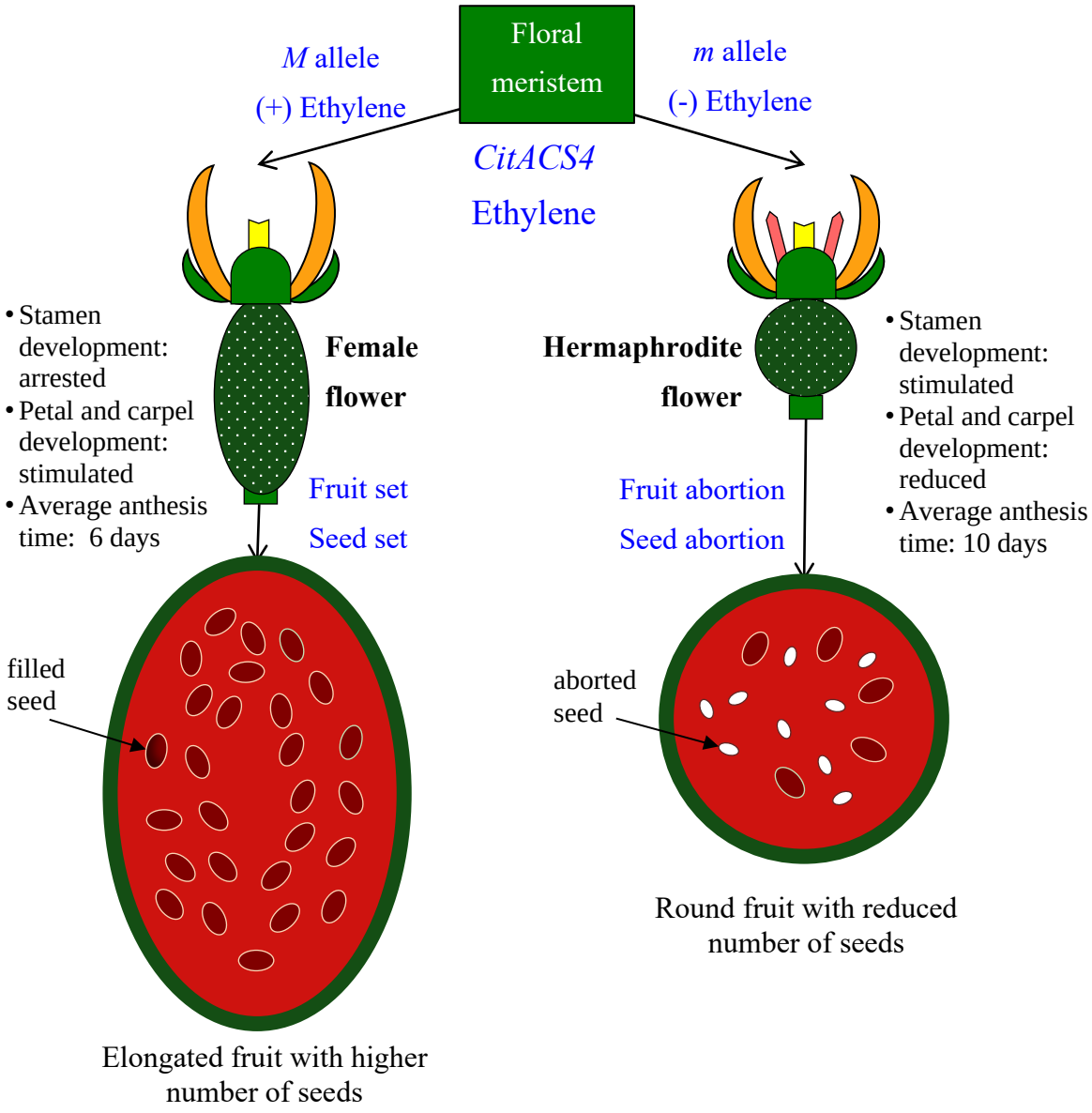


Figure 4.6. Involvement of the ethylene biosynthesis gene *CitACS4* in watermelon flower and fruit development. In the monoecious *MM* plants, the production of ethylene in the floral meristem arrests the development of stamens, but stimulates the growth of petals and carpels, which results in a female flower with an elongated ovary. In the *mm* plants, the lack of ethylene production prevents stamen arrest and reduces the growth rate of petals and carpels, which results in a hermaphrodite flower with a round-shaped ovary. After pollination, the reduced ethylene production in the *mm* flowers could be also responsible of fruit and seed abortion observed in the andromonoecious *mm* line.

5. Mapping a *partial andromonoecy* locus in *Citrullus lanatus* using BSA-seq and GWAS approaches

5.1. Abstract

The sexual expression of watermelon plants is the result of the distribution and occurrence of male, female, bisexual and hermaphrodite flowers on the main and secondary stems. Plants can be monoecious (producing male and female flowers), andromonoecious (producing male and hermaphrodite flowers), or partially andromonoecious (producing male, female, bisexual and hermaphrodite flowers) within the same plant. Sex determination of individual floral buds, that is, the mechanisms that determine whether the floral meristem will develop into a male, a female or a hermaphrodite flower, and the distribution of the different flower types on the plant, are both controlled by ethylene. A single missense mutation in the ethylene biosynthesis gene *CitACS4*, is able to promote the conversion of female into hermaphrodite flowers, and therefore of monoecy (genotype *MM*) into partial andromonoecy (genotype *Mm*) or andromonoecy (genotype *mm*). We phenotyped and genotyped, for the *M/m* locus, a panel of 207 *C. lanatus* accessions from all parts of the world, including 5 inbreds and hybrids, and found several accessions that were repeatedly phenotyped as PA (partially andromonoecious) in several locations and different years, despite being *MM*. A cosegregation analysis between a SNV in *CitACS4* and the PA phenotype, demonstrated that the occurrence of bisexual and hermaphrodite flowers in a PA line is not dependent on *CitACS4*, but conferred by an unlinked recessive gene which we called *pa*. Two different approaches were performed to map the *pa* gene in the genome of *C. lanatus*: bulk segregant analysis sequencing (BSA-seq) and genome wide association analysis studies (GWAS). The BSA-seq study was performed using two contrasting bulks, the monoecious M-bulk and the partially andromonoecious PA-bulk, each one generated by pooling DNA from 20 F2 plants possessing the most contrasting phenotypes for monoecy and PA, respectively. For GWAS, 122 accessions from USDA gene bank, already re-sequenced by genotyping by sequencing (GBS), were used. The combination of the two approaches indicates that *pa* maps onto a genomic region expanding across 32.24-36.44 Mb in chromosome 1 of watermelon. Fine mapping narrowed down the *pa* locus to an 867 Kb genomic region containing 101 genes. A number of candidate genes were selected, not only for their function in ethylene biosynthesis and signaling as well as in flower development and sex determination, but also by the impact of the SNPs and indels differentially detected in the two sequenced bulks.

5.2. Introduction

Watermelons, genus *Citrullus* (2n = 2x= 22) Schrad. ex Eckl. et Zeyh., are among the most grown vegetable fruit crops worldwide (Levi et al., 2017), representing a planted area of more than 3 million hectares, and production exceeding 100 million tons annually (http://faostat.fao.org/). The cultivated citron, egusi and dessert watermelons have been treated as subspecies of the single species *Citrullus lanatus* (Fursa, 1972), although there are some cross-ability as well as genome organization data that suggest separate them into three different species (Renner et al., 2014; Paris, 2015): *C. lanatus* (Thunb.) Matsum. & Nakai is the dessert watermelon (also known as *C. lanatus* subsp. *vulgaris*); *C. amarus* Schrad. is the citron watermelon (also known as *C. lanatus* subsp. *lanatus*), and *C. mucosospermus* (Fursa) Fursa is the egusi watermelon (also known as *C. lanatus* subsp. *mucosospermus*). The genus is also composed of four other species: the less cultivated *C. colocynth*is (L.) Schrad., (known as colocynth watermelon), and the non-cultivated species *C. ecirrhosus* Cogn., *C. rehmii* De Winter and *C. naudinianus* (Sond.) Hooker f. (Paris, 2015).

Watermelons, like as other members of the family Cucurbitaceae, produce unisexual male or female flowers as well as hermaphrodite or bisexual flowers on the same plant (Poole and Grimball, 1938). According to the manner in which these types of flowers are combined, the flowering pattern of watermelon is either monoecious (M, producing male and female flowers on the same plant), andromonoecious (A, producing male and hermaphrodite flowers on the same plant) or partially andromonoecious (PA, producing male, female, and hermaphrodite and bisexual flowers on the same plant) (Rudich and Zamski, 1985; Ji et al., 2015). The cucurbit fruit is developed from the ovaries of female or hermaphrodite flowers, and thus the initiation and the correct development of these flowers are crucial with regard to the level and quality of crop production (Latrasse et al., 2017). Andromonoecy and PA are undesirable traits in watermelon, since hermaphrodite flowers need to be emasculated when acting as female parents in the production of hybrid seed (Prothro et al., 2013). They are also undesirable because these traits are usually associated with a reduction in fruit set and fruit quality (Aguado et al., 2018).

In the Cucurbitaceae family, sex expression is controlled by environmental, hormonal, and genetic factors. With regard to the environmental factor, some changes of sex expression in

cucurbits have been attributed to variations in temperature and photoperiod regimes (Manzano et al. 2014; Li et al. 2018). Ethylene is the principal hormone regulating sex expression of cucurbits species. In watermelon, recent studies have demonstrated that ethylene inhibits the transition from male to female flowering and reduces the number of pistillate flowers per plant (Manzano et al. 2014; Zhang et al., 2017).

The combination of three pairs of genes explains the control of sex forms in watermelon: gynoecious (*gy*), andromonoecious (*a*), and trimonoecious (*tm*) (Ji et al., 2015). Andromonoecy has been considered recessive to monoecy, being determined by a single locus with two alleles (*M*, monoecious; *m*, andromonoecious) (Rosa 1928; Poole and Grimball 1945; Kubicki 1962; Rudich and Zamski 1985; Salman-Minkov et al. 2008). Recently, it was demonstrated that the *M/m* locus corresponds to the gene *CitACS4*, which encodes for a flower specific ACS enzyme. This enzyme is involved in the biosynthesis of ethylene, which is required for stamen arrest during the development of female flowers (Boualem et al., 2016; Ji et al., 2016; Manzano et al., 2016). As occurs in the orthologs *CmACS7*, *CsACS2* and *CpACS27A* of melon, cucumber and squash (Knopf and Trebitsh, 2006; Boualem et al., 2009; Martínez et al., 2014), a single missense mutation (*m*) in the coding region of *CitACS4*, promotes the conversion of monoecy (*MM*) either into andromonoecy (*mm*), or into PA (*Mm*) (Manzano et al., 2016). The andromonoecious allele (*m*) of *CitACS4* also has some pleiotropic effects on fruit development, for example, shortening and rounding the shape of ovaries and fruits, and reducing fruit set and fruit quality of watermelon (Kubicki, 1962; Aguado et al., 2018).

Other ethylene biosynthesis and perception genes are responsible for sex determination in cucurbit species. In cucumber, gynoecy is conferred by the dominant locus *F* (Female), which encodes for an additional copy of *CsACS1* (*CsACS1G*) (Trebitsh et al., 1997; Zhang et al., 2015; Li et al., 2020b). The recessive gynoecy of melon is, however, conferred by a loss of expression of the C_2H_2 zinc-finger-type transcription factor *CmWIP1*, which directly repress *CmACS7* (Martin et al., 2009). The CRISPR/Cas9 generated mutations of *WIP1* orthologs in cucumber and watermelon also lead to gynoecy, demonstrating that *WIP1* is required for carpel abortion during male flower development in different cucurbit species (Hu et al., 2017; Zhang et al., 2019b). On the other hand, mutations in the *CsACS11* and *CmACS11* orthologs of cucumber and melon, together with the *CsACO2* of cucumber, completely block female flower development, leading to androecy (Boualem et al., 2015; Chen et al., 2016). This androecious phenotype is the result of ethylene being produced by

ACS11 and *ACO2*, which downregulates the expression of the carpel-aborting gene *WIP1*. The ethylene receptors *ETR1* and *ETR2* also play an important role of sex determination of cucurbits. The ethylene insensitive mutants *etr1a* and *etr2b* of *Cucurbita pepo* both disrupt female flower development (converting monoecy into andromonoecy) and significantly increase the number of male flowers in the plant. This probably indicates that *ETR1* and *ETR2* are able to integrate the two ethylene biosynthesis pathways, perceiving and signaling the ethylene produced by *ACS2/7* as well as that produced by *ACS11* and *ACO2* (García et al., 2020).

We identified some inbred lines that are homozygous for the monoecious allele (*MM*), but exhibited a PA phenotype when grown under high temperature conditions (Manzano et al., 2014). This PA phenotype does not appear to be conferred by the *m* allele of *CitACS4* in heterozygous conditions (*Mm*), but by other allelic or non-allelic gene. The current selection of monoecy by the SNP marker linked to the *M* allele is therefore not enough to prevent the production of bisexual and hermaphrodite flowers. The PA phenotype is similar to that of the previously defined trimonoecious (plants producing male, female and bisexual flowers; Rosa 1928; Ji et al. 2015), however, we cannot ensure that it is actually the same trait. Previously trimonoecy was stated by the occurrence of the three types of flowers, but ignoring the ratio between them. The term PA is therefore more precise since it is defined on the basis of the andromonoecious index (AI), which is an index that score the level of andromonoecy in a plant or accession by considering the ratio between female, bisexual and hermaphrodite flowers per plant (Martínez et al., 2014).

In order to identify the gene controlling the PA phenotype, screened for this trait in a large number of watermelon accessions, determining the inheritance of it by means of biparental crossings between monoecious (*MM*) and partial andromonoecious (*MM*) inbred lines. Thereafter, we mapped the trait by using bulk segregant analysis combined with new generation sequencing (BSA-seq), and also by genome-wide association analysis (GWAS). These two methods allow the identifications of DNA markers closely linked to the causal gene for a given phenotype (Giovannoni et al., 1991; Michelmore et al., 1991; Mansfeld and Grumet, 2018). By combining these two approaches, we identified a genomic region on chromosome 1 that regulates the partial andromonoecy trait in watermelon.

5.3. Materials and Methods

5.3.1. Germplasm and growing conditions

Three inbred lines P84, P86 and P87 of *Citrullus lanatus* were used in this study: P84 (partially andromonoecious), P86 (monoecious) and P87 (andromonoecious). A total of 47 watermelon accessions coming from two Spanish gene banks, those of the Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) and of the Centro de Investigación y Tecnología Agroalimentaria de Aragón (BGH-CITA), and 155 watermelon PI accessions taken from the USDA National Plant Germplasm System (NPGS) were also used. All watermelon accessions belonged to the species *Citrullus lanatus* (Thunb.) Matsum. & Nakai. Geographically, 38 were from Africa, 70 from Asian, 76 were from Europe, 10 from North America, 7 from South America and 1 from Oceania. Two commercial cultivars of *C. lanatus*, Charleston Gray and Calhoun Gray, were also included (Supplementary Table 5.1).

Sex phenotyping of most of the PI accessions was performed in two locations and under two different environmental conditions in 2017 and 2018: open field conditions in Raleigh (North Carolina, USA) in the summer of 2017, and greenhouse conditions in Almería (Spain) in the summer of 2018 (Supplementary Table 5.1). Only the PI accessions showing a stable sex phenotype in each location and each year were used for GWAS analysis. The Spanish accessions were evaluated only during spring-summer 2018 in Almería. In the greenhouse, the plants were grown in 10 L pots under the standard agronomic conditions of Almería.

5.3.2. Evaluation of sex expression traits and inheritance of PA phenotype

The andromonoecious index (AI), defined by Martínez et al. (2014) and adapted for watermelon by Manzano et al. (2016), was used to evaluate the andromonoecy level of each plant, population and accession. Pistillate flowers were scored from 1 to 3 according to their degree of stamen development. Female flowers without stamens were scored as AI = 1, while pistillate flowers with medium size stamens and anthers were assigned as AI = 2 and pistillate flowers with primordial stamens with pollen were scored AI = 3. Based on these flower scores, the AI of each plant was calculated as the average score for at least 10 flowers of the first 30 nodes along the main stem and lateral branches. The average AI of each accession was calculated from at least 5 plants with a minimum of 10 pistillate flowers

evaluated per plant. Plants and genotypes with an AI of between 1 and 1.35 were considered to be monoecious, while those with AI scores between 1.36 and 2.7 were considered partially andromonoecious (PA) and those with AI higher than 2.7 were classified as andromonoecious.

The inheritance of the PA phenotype was determined by formal genetic analysis. The monoecious line P86 was crossed with the partially andromonoecious line P84, and the F1 was self-fertilized to produce F2 and thereafter subsequent selfing generations. From F2 and F3 generations, we selected the plants with extreme phenotypes which were then selfed so as to produce F3 and F4 offspring, respectively. Parental and offspring generations plants were all phenotyped under greenhouse conditions during the spring-summer season of 2017, 2018 and 2019. The χ^2 test for goodness-of-fit ($P < 0.05$) and homogeneity were used to examine segregation ratios (monoecious: partially andromonoecious) in the F2 and successive selfed generations.

5.3.3. Genotyping for *M* and *m* alleles of *CitACS4* gene

The collection of watermelon accessions were genotyped by real time PCR using Taq-Man probes for the *M* and *m* alleles of the *CitACS4* gene. Individual plants' leaves were sampled, and DNA extracted using the CTAB method (Murray and Thompson, 1980). The PCR reactions were done with Bioline SensiFAST™ Probe No-ROX Kit, a set of forward and reverse primers amplifying the polymorphic sequence, and two allele-specific probes descriptive of the SNP of interest (C-G). The monoecious *M*-allele probe was labelled with FAM dye and the andromonoecious *m*-allele probe was labelled with HEX reporter dye (metabiom). The BHQ1 quencher molecule was used in both probes (Supplementary Table 5.2). Reactions were performed in the Rotor-Gene Q thermocycler (Qiagen) by using green and yellow channels for FAM and HEX reporter dye, respectively. The annealing temperature in the reaction was 60ºC.

In a number of *MM*, *Mm* or *mm* accessions, a fragment of 1875 bp covering the complete coding region of the *CitACS4* was amplified by PCR using specific primers designed by Manzano et al. (2016), and sequenced for searching new polymorphisms that could be responsible of the PA phenotype.

5.3.4. Mapping PA respect to *CitACS4* gene

To be able to see whether the PA phenotype is linked to *CitACS4*, we searched for a DNA polymorphism in the gene between P84 (partially andromonoecious) and P86 (monoecious) inbred lines, and analysed the segregation of the phenotype and the gene in 256 F2 plants. Given that the coding region had no variation between the lines, we amplified and sequenced a 708 bp fragment of the promoter by using the specific primers CitACS4genF6/R6 and F4/R4 (Supplementary Table 5.2). The multiple sequence alignments were performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). A single nucleotide variant (Indel) at the position -369bp respect to the ATG start codon was detected between the two inbred lines.

To study the segregation between the PA phenotype and the *CitACS4* gene, the F2 population of the cross P84XP86 was phenotyped for the AI and some plants having extreme monoecious (AI = 1) and PA phenotypes (AI >1,6) were genotyped for the SNV which was detected in the promoter by using Sanger sequencing.

5.3.5. BSA-seq mapping approach

A total of 256 F2 plants from the cross P84XP86 (2016 assay) were phenotyped for AI. 40 plants having the most extreme phenotypes for both monoecy (plants with AI = 1) and partial andromonoecy (plants with AI >1.6) were selected. Two DNA bulks, the monoecious M-bulk and the partial andromonoecious PA-bulk, were constructed by mixing an equal amount of DNA from 20 monoecious and 20 PA F2 plants, respectively. Each bulked DNA was randomly sheared into short fragments of about 350 bp for library construction using the NEBNext® DNA Library Prep Kit. Following end repairing, dA-tailing, and further ligation with NEBNext adapter, the required fragments were briefly PCR enriched with indexed oligos. The qualified DNA libraries were pooled according to their effective concentration as well as the expected data production. Pair-end sequencing was performed on the Illumina® sequencing platform, with the read length of PE150 bp at each end. Raw data obtained from the sequencing contained adapter contamination and low-quality reads. Different quality control steps were used: (1) discard the paired reads when either read contains adapter contamination; (2) discard the paired reads when uncertain nucleotides (N) constitute more than 10 percent of either read; (3) discard the paired reads when low quality nucleotides (base quality less than 5, $Q \leq 5$) constitute more than 50 percent of either read.

The resulting sequencing data was aligned with the reference Charleston Gray genome v1 using BWA (Li and Durbin, 2009) software. The duplicates were removed by SAMTOOLS (Li et al., 2009a). Single nucleotide polymorphisms (SNPs) were detected by using GATK HaplotypeCaller (Depristo et al., 2011). ANNOVAR was used to annotate the detected SNPs (Zheng et al., 2019).

BSA-seq analysis was performed by using the R package QTLseqr designed by Mansfeld and Grumet (2018). M- and PA-bulk SNPs were filtered using the function "filterSNPs", thereby selecting SNPs with: a reference allele frequency between 0.3 and 0.7; total depth between 20 and 100; per sample read depth higher than 10; and genotype quality higher than 99.

Following the method developed by Takagi et al. (2013), the identification of QTLs was based on SNP-index and Δ(SNP-index) parameters, calculated by using the function “runQTLseqAnalysis”. SNP-index is the proportion of reads harbouring the variant that is different from the reference sequence. Δ(SNP-index) was calculated by subtracting the SNP-index of M-bulk from that of the PA-bulk. An average SNP-index of SNPs located in a given genomic interval was calculated using a sliding window analysis, with 1 Mb window size and 10 kb increment. The SNP-index graphs for M-pool and PA-pools, as well as corresponding Δ(SNP-index) graphs, were plotted. The Δ(SNP-index) value should not significantly differ from 0 in a genomic region with no major QTL of the target gene (Takagi et al., 2013). The 95% and 99% confidence intervals of the Δ(SNP-index) were calculated under the null hypothesis of no QTLs -using a simulation analysis of 10,000 replications for each bulk - and plotted alongside the Δ(SNP-index) (Takagi et al., 2013).

5.3.6. Fine mapping

The sequence derived from QTLseq was used to design 11 polymorphic SNP markers distributed equidistantly. The 11 SNPs were designed in KASP markers (Kompetitive Allele Specific PCR, www.lgcgroup.com) and were used for genotyping a subset of 220 individuals, following the protocol of LGC Genomics. A genetic map was constructed with the data of the 11 SNP markers, using JoinMap® 5 (Kyazma, B.V.). Linkage analysis and marker order were performed with the regression mapping algorithm, and genetic distance was calculated using the Kosambi mapping function (Crow and Dove, 1990). We performed QTL analysis using MapQTL6® (Kyazma B.V.) interval mapping analysis.

5.3.7. Genome-wide association (GWAS) mapping approach

The SNPs obtained from GBS analysis were used thereafter for GWAS analysis (Wu et al., 2019). The genome sequence of 'Charleston Gray' and the GBS SNPs are available at the Cucurbit Genomics Database (http://cucurbitgenomics.org; (Zheng et al., 2019). A diversity panel of 122 watermelon accessions of *Citrullus lanatus* was used for GWAS. The collection was phenotyped for AI in at least 10 pistillate flower per plant in each accession and their phenotypic data was used for GWAS. Only biallelic SNPs within *C. lanatus* accessions (a total of 12039 SNPs) were used for GWAS. The analysis was performed with TASSEL software (Bradbury et al., 2007; Gur et al., 2017), using the linear mixed model (MLM), which considers both population structure (Q matrix) and relatedness (kinship matrix), and the generalized linear model (GLM), which does not consider the population structure. Genome-wide significance thresholds of GWAS were determined using the Bonferroni correction at Pvalue = 0.05 (FDR 5%) and Pvalue = 0.01 (FDR 1%) for significant and extremely significant associations, respectively, as described in Li et al. (2012). Principal component analysis (PCA), and Quantile-quantile (Q-Q) plots, where distributions of P values expected the null hypothesis distribution, were performed by using TASSEL software (V 5.2.52) (Bradbury et al., 2007).

5.4. Results

5.4.1. Sex phenotyping a large germplasm panel of watermelon

Figure 5.1 shows the two main sex morphotypes of watermelon, monoecious (M) and andromonoecious (A), together with the intermediate sex phenotype of watermelon cultivars called partial andromonoecy (PA), the latter being characterized by the production of male, female, and bisexual and hermaphrodite flowers on the same plant. We studied the diversity of sex morphotypes among a large germplasm panel of watermelon, comprising 207 watermelon accessions: 155 from USDA-NPGS gene bank, 47 from the Spanish gene banks COMAV and BGH-CITA, plus 2 commercial cultivars and 3 inbred lines (Supplementary Table 5.1).

The sex phenotype of each accession was assessed by using the andromonoecious index (AI), using a minimum of 10 pistillate flowers per plant and 5 plants per accession (Figure 5.1). This index ranges from 1 to 3 and measures the degree of stamen development in

pistillate flowers, and therefore the level of andromonoecy of a genotype (Figure 5.1). An AI = 1 corresponds to complete monoecy (pistillate flowers are all female), and AI = 3 to complete andromonoecy (pistillate flowers are all hermaphrodite). AI between 1 and 3 is assigned to plants and accessions whose pistillate flowers can be either female, bisexual (partial arrest of stamen development) and hermaphrodite (complete development of stamens and pollen) (Figure 5.1). The inbred lines P86, P87 and P84, that were previously shown to have an average AI of 1.02, 3 and 1.55, respectively, were used as control indices of M, A and PA, respectively.

Taking into account the sex phenotype of these three control genotypes under our environmental conditions, we stablished that accessions with $1 \leq AI \leq 1.35$ were monoecious, those with $2.7 \leq AI \leq 3$ were andromonoecious, and those with $1.35 < AI < 2.7$ were partially andromonoecious. In accordance with these criteria, 18% of the phenotyped accessions were classified as monoecious, 43% as andromonoecious, and 39% as partially andromonoecious (Supplementary Table 5.1).

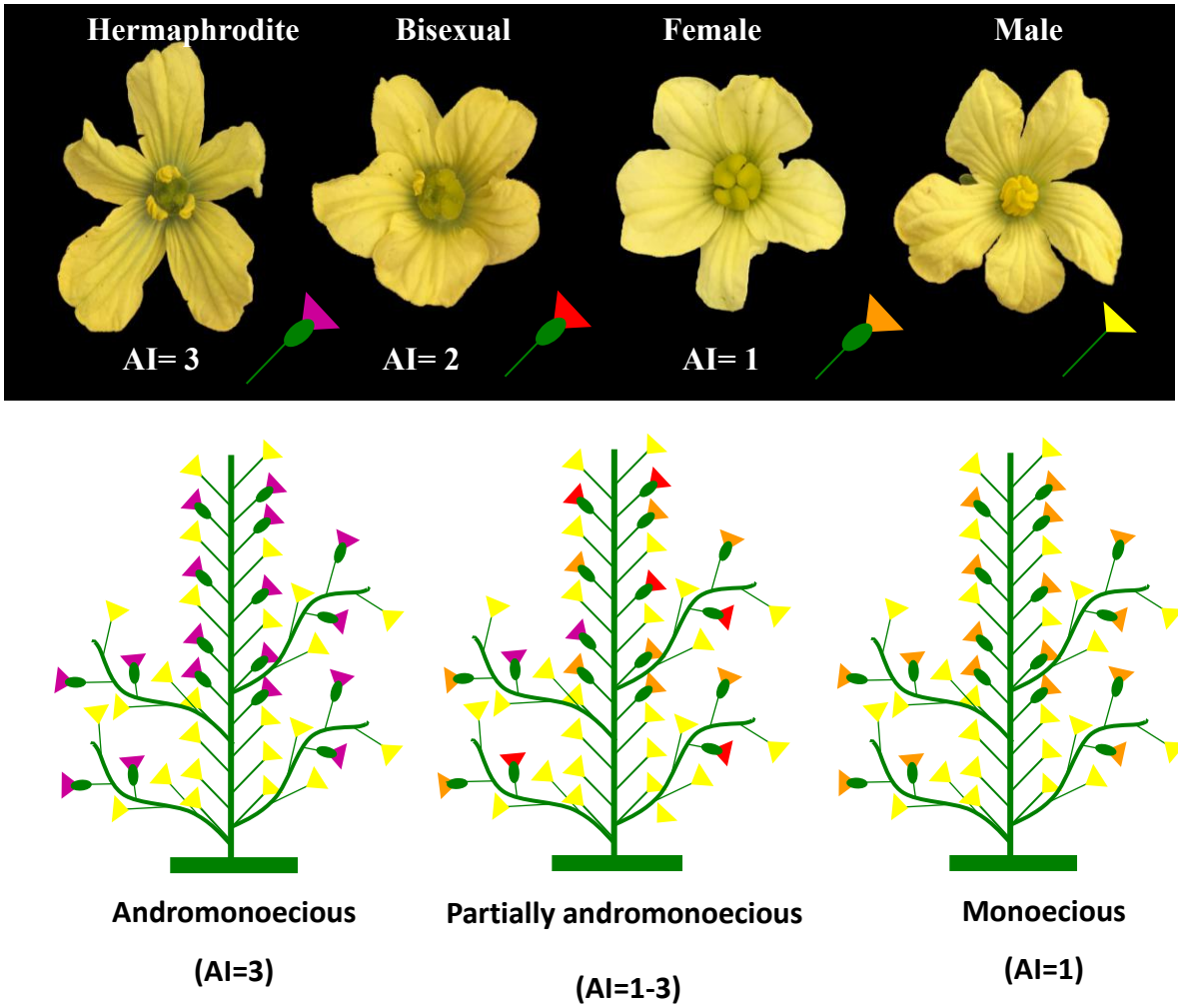


Figure 5.1. Types of flowers and sex morphotypes in watermelon: monoecious, andromonoecious and partially andromonoecious. Flowers are classified as male (only stamens), female (only carpels), bisexual (producing carpels and some undeveloped stamens) and hermaphrodite (producing both mature carpels and mature stamens with pollen). The Andromonoecious Index (AI) was used to estimate the level of andromonoecy of each plant and accession. For this, female, bisexual, and hermaphroditic flowers were assigned the values of AI = 1, 2 and 3, respectively. The AI of each plant was calculated as the average of the AI of a minimum of 10 pistillate flowers per plant. An AI value of 1 corresponds to plants producing only female flowers, and a value of 3 indicates that plants produce only hermaphrodite flowers.

The locus *M/m*, which underlies the ethylene biosynthesis gene *CitACS4*, is the main regulator of monoecy. Homozygous *MM* and *mm* plants are monoecious and andromonoecious respectively, while heterozygous *Mm* are partially andromonoecious (Manzano et al., 2016).

So as to investigate whether the variability found in sex determination can be explained solely on the basis of this gene, we genotyped all the accessions for the *M* and *m* alleles of *CitACS4* by using a bulked DNA from 5 plants of each accession. Some accessions were segregating for the two alleles and exhibited the three sex phenotypes: monoecious, andromonoecious and partially andromonoecious (Supplementary Table 5.1). All the accessions that were genotyped as *mm* were andromonoecious. The majority of the *MM* accessions exhibited a stable monoecy and produced male and female flowers (Figure 5.2). However, a number of accessions were found (BGHZ4849, BGHZ5441, BGHZ5442, BGHZ5993, BGV002674, PI 164992, PI 379237 and Calhoun Gray) that were *MM* but showed a partially andromonoecious phenotype, characterized by the production of male, female, bisexual and hermaphrodite flowers (Supplementary Table 5.1; Figure 5.2). The PA phenotype of these *MM* accessions could possibly be conferred by an andromonoecious allele on the *CitACS4* gene other than *m*, or by a different gene.

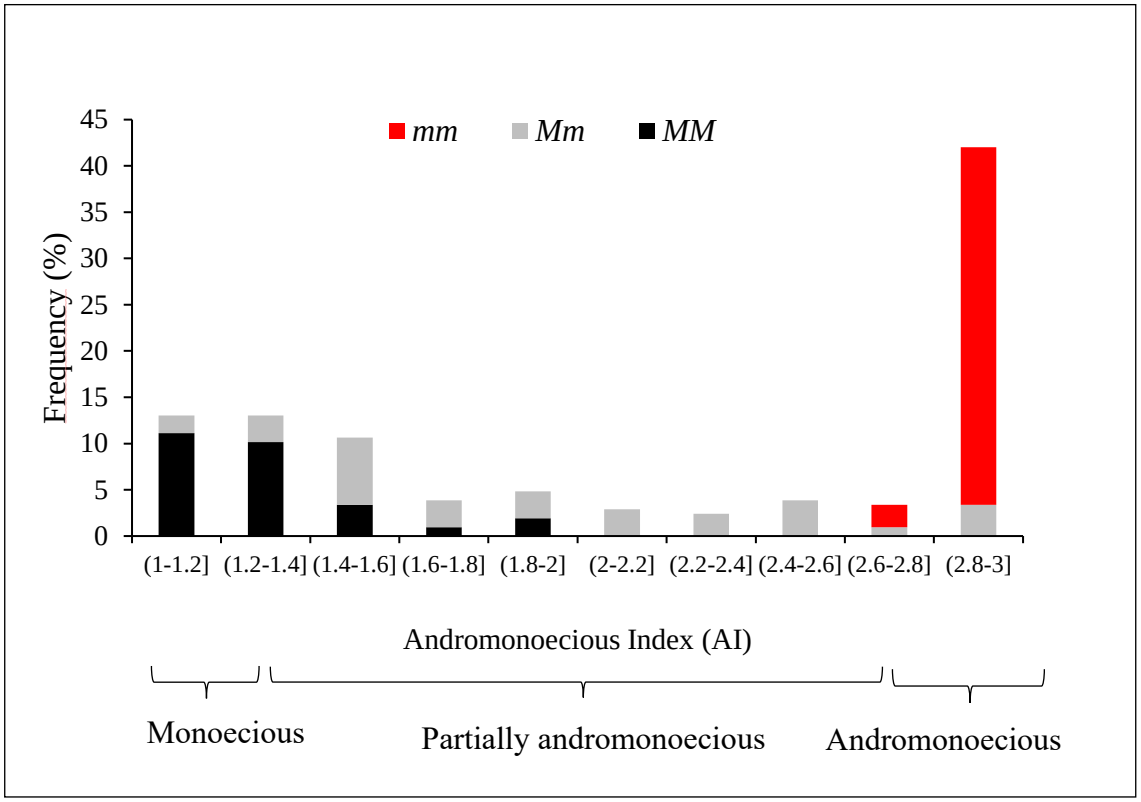
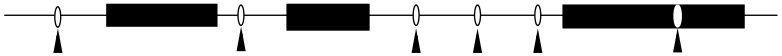


Figure 5.2. Frequency distribution of Andromonoecious Index (AI) in 207 accessions of watermelon. Bar colors indicate the genotype of accessions for the *M* (*monoecious*) and *m* (*andromonoecious*) alleles of *CitACS4* gene. Given accessions were initially genotyped as a bulk, the *Mm* accessions are those having both *M* and *m* alleles, but plants of those accessions were proved to be *MM*, *Mm*, or *mm*.

To investigate the occurrence of new andromonoecious alleles in *CitACS4*, the gene was amplified and sequenced in 16 *MM* accessions showing either monoecious or partial andromonoecious phenotypes, along with 4 *mm* andromonoecious accessions (Table 5.1). The control lines P84, P86 and P87 were included in the sequencing analysis. In the coding region of *CitACS4*, we found no variation other than the mutation C to G in the third exon, which is responsible for the andromonoecious phenotype in *mm* plants (Manzano et al., 2016). One single nucleotide variant was found in the first intron, and three in the second intron of the gene, but none of them was linked to the PA phenotype (Table 5.1). Based on this data, we concluded that the partial andromonoecy of *MM* plants is not likely to be conferred by a *CitACS4* allele. It seems to be caused by a different gene.

Table 5.1. Polymorphic sites of *CitACS4* gene in 20 watermelon accessions. The top shows the structure of the gene, and the identified polymorphic sites. The *M* and *m* alleles are defined by the polymorphism at position 1,477 in the third exon. The sex phenotypes of each accession are also indicated. Pink, green, and yellow boxes indicate the nucleotide variants in the different accessions. The polymorphic site (indel) in the promoter was only analysed in the P84 and P86 parental inbred lines. M, monoecious; A, andromonoecious; PA, partially andromonecious.



				Genotype						Phenotype			
				Promoter	Intron 1	Intron 2			Exon 3				
Accession	Species	Origin	*CitACS4*	-239	196	678	740	840	1477	Andromo-noecious index (AI)	Sex Morpho-type	Percentage pistillate flowers/plant	Pistillate flowering transition
P86	*C. lanatus*	Inbred line	***MM***	T	G	CA-------	C	----	C	1.02±0.03	M	18.33±3.08	5.13±1.89
BGV002712	*C. lanatus*	Spain, Islas Canarias	***MM***		G	CA-------	C	----	C	1±0.00	M	9.44±4.91	12.00±3.16
PI 169299	*C. lanatus*	Turkey, Hatay	***MM***		A	CA-------	C	----	C	1±0.00	M	13.33±0.00	12.00±0.00
PI 482328	*C. lanatus*	Zimbabwe	***MM***		G	CTAGTTGGT	C	TTT-	C	1±0.00	M	6.67±4.71	17.50±7.78
PI 271774	*C. lanatus*	South Africa, Transvaal	***MM***		A	CA-------	C	----	C	1±0.00	M	5.00±2.36	24.50±6.36
BGHZ4848	*C. lanatus*	Spain, Andalucía	***MM***		G	CA-------	C	----	C	1±0.00	M	14.00±2.79	9.60±1.67
BGV001875	*C. lanatus*	Spain, Cataluña	***MM***		A	CA-------	C	----	C	1±0.00	M	8.89±3.44	10.50±3.73
P84	*C. lanatus*	Inbred line	***MM***	-	G	CA-------	C	----	C	1.55±0.23	PA	15.19±2.94	8.22±2.17
BGHZ5993	*C. lanatus*	Spain, Huelva	***MM***		G	CA-------	C	----	C	1.47±0.37	PA	9.33±4.35	16±4.06
Calhoum	*C. lanatus*	Commercial	***MM***		A	CA-------	C	----	C	1.58±0.62	PA	10.67±5.48	12.20±6.06
BGV002674	*C. lanatus*	Spain, Las Palmas	***MM***		G	CA-------	C	----	C	1.65±0.32	PA	8.67±2.98	10.4±4.39
BGHZ4849	*C. lanatus*	Spain, Andalucía	***MM***		G	CA-------	C	----	C	1.79±0.43	PA	10±2.98	13.33±3.01
PI 164992	*C. lanatus*	Turkey, Ankara	***MM***		G	CA-------	C	----	C	1.86±0.12	PA	16.00±2.79	7.20±0.45
BGHZ5441	*C. lanatus*	Spain, Andalucía	***MM***		A	CA-------	C	----	C	1.96±0.38	PA	8.00±1.83	17.60±4.45
PI 379237	*C. lanatus*	Macedonia	***MM***		A	CA-------	C	----	C	1.98±0.81	PA	14.17±3.19	7.50±4.80
BGHZ5442	*C. lanatus*	Spain, Andalucía	***MM***		A	CA-------	C	----	C	2±0.60	PA	11.33±1.83	13.20±2.17
P87	*C. lanatus*	Inbred line	***mm***		G	CTAGTTGGT	T	TT--	G	3±0.00	A	22.00±2.98	9.40±1.34
PI 368501	*C. lanatus*	Macedonia	***mm***		G	CTAGTTGGT	T	TTT-	G	3±0.00	A	13.33±2.36	12.00±3.00
PI 537274	*C. lanatus*	Pakistan, Punjab	***mm***		G	CTAGTTGGT	T	TT--	G	2.60±0.30	A	15.33±1.83	10.00±3.32
BGHZ6020	*C. lanatus*	Spain, Castilla-La Mancha	***mm***		G	CTAGTTGGT	T	TTT-	G	3±0.00	A	12.67±5.48	13.60±3.65

5.4.2. Partial andromonoecy (PA) phenotype of P84 is conferred by a single recessive gene other than *CitACS4*

Inheritance of the PA phenotype was studied in the P84 (PA) X P86 (M) cross. P86 is a monoecious *MM* line producing only male and female flowers (average AI = 1.02), while P84 is a partially andromonoecious *MM* line producing male, female, bisexual and hermaphrodite flowers (average AI = 1.55) (Table 5.1). The average AI of the F1 progeny plants was 1.21, while that of the F2 plants ranged between 1 and 1.9 (Figure 5.3). As stablished above, the plants with $1 \leq AI \leq 1.35$ were considered monoecious, and those with $1.35 < AI < 2.7$ partially andromonoecious. The F1 was monoecious, and the F2 progeny segregates for 181 M:75 PA, which fit the expected 3:1 ratio for a single gene ($\chi^2 = 2.52$, P value = 0.11; Table 5.2). F3 and F4 progenies of PA plants were always PA, while those of M plants were either M or segregated for M and PA (Table 5.2). This data indicates that the PA phenotype of P84 is conferred by a single recessive gene, the one we have called *pa*.

A segregation analysis in the F2 generation of P84XP86 confirmed that *pa* is not an allele of *CitACS4*, but a newly found gene controlling sex determination in *C. lanatus*. The gene *CitACS4*, including exons, introns, and the 708 bp of the promoter, was sequenced in P84 and P86 lines. No nucleotide variation was found in the coding region of *CitACS4* between P84 and P86, but a single nucleotide deletion was detected in the promoter region of P84, at nucleotide position -369 respect to the ATG start codon (Table 5.1). 25 F2 plants were genotyped for this indel, but the deletion did not cosegregated with the PA phenotype in the F2 population of P84XP86. This demonstrated that the PA phenotype of P84 is not conferred by *CitACS4* but by a different, unlinked gene.

Table 5.2. Segregation ratio of monoecious and partially andromonoecious in F2, F3 and F4 populations derived from crosses between monoecious and partially andromonoecious inbred lines (P86XP84).

Generation	No. of plants		Expected segregation	χ^2	p-value
	Monoecious	Partially andromonoecious			
Parental *P84*		12			
Parental *P86*	12				
F1 (*P84XP86*)	12				
F2 (F1 ⊗)	181	75	3:1	2.52	0.1124
F3					
(*Papa* ⊗)		60			
(*PAPA* ⊗)	22				
(*Papa* ⊗)	45	14			
F4					
(*papa* ⊗)		50			
(*PaPA* ⊗)	30				
(*Papa* ⊗)	36	13			

The F2, F3 and F4 plants were phenotyped on the basis of their average AI, scored from at least 10 flowers per plant. Monoecious ($1 \leq AI \leq 1.35$) and partially andromonoecious ($1.35 < AI < 2.7$). The genotypes for the partial andromonoecious locus (*PA/pa*) in F3 and F4 generations were assigned on the basis of the segregation ratio. ⊗, self-pollination.

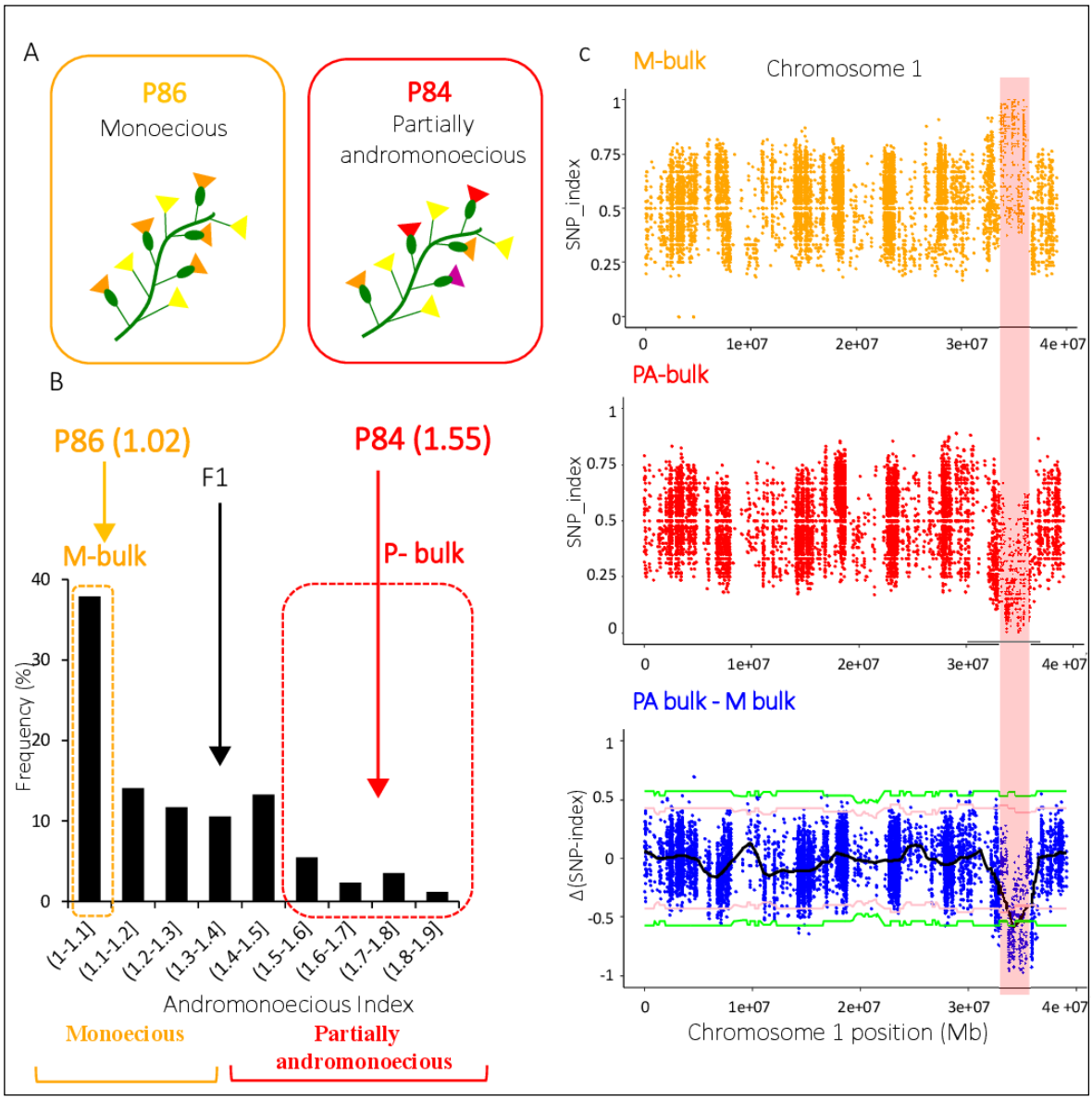


Figure 5.3. QTL-seq applied to watermelon F2 progeny identified the locus conferring partial andromonoecy (*pa*) on chromosome 1. (**A**) Phenotype of two watermelon inbred lines P84 and P86 for andromonoecy index. P84 is partially andromonoecious whereas P86 is monoecious. (**B**) Frequency distribution of andromonoecious index in 256 F2 progenies. The average andromonoecy index (AI) for P86 and P84 is indicated. The dashed orange and red boxes indicate the 20 F2 plants with AI = 1 that were used to make the M-bulk, and the 20 F2 / AI >1.6 plants used to make the PA-bulk. (**C**) SNP-index plots of M-bulk and PA-bulk, and the Δ(SNP-index) plot on chromosome 1. SNP-index was calculated based on a 1 Mb interval with a 10 kb sliding window. A candidate QTL on chromosome 1 (32.2 - 36.4 Mb interval) with the criteria that the SNP-index in M-pool was near 1, the SNP-index in PA-pool was near 0, and the Δ(SNP-index) was above the confidence value (P <0.001). The Δ (SNP-index) plot shows the tricube smoothed Δ(SNP-index) and the statistical confidence intervals under the null hypothesis of no QTLs (pink, P < 0.05; green, P < 0.01).

5.4.3. QTL-seq identified a major locus controlling PA on chromosome 1

Bulked-segregant analysis coupled with whole genome sequencing (BSA-seq) was used to map the *pa* locus, which is responsible for the partial andromonoecy of the P84 line. 20 plants with extreme phenotypes in the F2 from the cross P84 (PA) X P86 (M) were selected and their DNA pooled together to construct two DNA bulks: the M-bulk and the PA-bulk (Figure 5.3). The M and PA parental lines, along with the M- and PA-bulks were resequenced by WGS (Whole Genome Sequencing). Illumina high-throughput sequencing resulted in 55,875,160 short reads (350 bp) from M-bulk (average 18.3x depth and 99.05 % coverage), and 55,743,486 short reads (350pb) from PA-bulk (18.11x depth and 99.06% coverage) (Table 5.3).The alignment of the reads to the Charleston Gray reference genome resulted in more than 370,000 SNPs in each of the bulked DNA samples (Table 5.3).

Table 5.3. Summary of sequencing data and alignment result of BSA-Seq.

Sample	Total reads	Mapped reads	Clean data (G)	Q30(%)	% Mapping reads	Average depth (X)	Coverage at least 4X (%)	SNP number
M_bulk	55,875,160	55,202,891	8.40	93.02	98.80	18.30	97.79	370,332
PA_bulk	55,743,486	54,904,988	8.40	92.48	98.50	18.11	97.76	374,678
P84	51,301,618	50,549,756	7.70	92.65	98.53	16.52	97.39	193,904
P86	56,832,990	55,956,004	8.50	92.73	98.46	17.28	97.18	356,384

SNPs from the two bulked datasets were used to run the QTLseqr R package, calculating SNP-index for M- and PA-bulks, and Δ(SNP-index). An average SNP-index was computed in a 1 Mb interval using a 10 kb sliding window. SNP-index was calculated for the M- and PA-bulk and plotted against the positions along each chromosome of the Charleston Gray genome (Supplementary Figure 5.1). By combining the SNP-index information in M- and PA-bulks, the Δ(SNP-index) was also computed and plotted against the genome positions (Supplementary Figure 5.1). As expected, most of the genomic regions were not relevant to the phenotypic variation (andromonoecious index) and showed identical SNP-index graphs for the M- and PA-bulks. However, a single region on chromosome 1 is the most probable *pa* locus, since it exhibits unequal contributions from P84 and P86 parental genomes (Figure 5.3). The region on chromosome 1 comprised from 32,237,329 to 36,436,269 bp, had an average SNP-index higher than 0.63 in M-bulk, while the SNP-index in the corresponding region of PA-bulk was lower than 0.33. After examining SNP haplotypes in the M- and PA-

bulks, it was found that most haplotypes in the 32.24–36.44 Mb region of chromosome 1 corresponded to P86 (monoecious) in the M-bulk and to P84 (PA) in the PA-bulk (Figure 5.3). The Δ(SNP-index) value should be significantly different from 0 if a genomic region harbours a major QTL of the target gene. At 95 % and 99% significance level, only the genomic region on chromosome 1 from 32.24 to 36.44 Mb had an Δ(SNP-index) value that differed significantly different from 0 (Figure 5.3). This major QTL on chromosome 1 (QTL1) is therefore most likely to underlie the watermelon PA phenotype.

5.4.4. Genome-wide association studies (GWAS) with regard to andromonoecy and partial andromonoecy

Genome-wide association studies (GWAS) has proven to be a powerful tool in dissecting the genetic basis of variation for both simple and complex traits. This results in the identification of nucleotide variants that are associated with the trait being sought. Using the nucleotide variation approach, that is, the variation reported in the USDA watermelon collection after genotyping by sequencing (GBS) (Wu et al., 2019), we performed a GWAS analysis of the Andromonoecious Index (AI) trait. We filtered out - from all the phenotyped accessions - those traits with no GBS data, and also those showing a high inter-replication variance for the trait in question. Accessions that clearly segregated for monoecy and andromonoecy (accessions with plants having AI = 1 or AI = 3) were also discarded. A panel of 122 PI accessions was finally selected (Supplementary Table 5.1). A total 15,681 GBS-SNPs with a minor allele frequency of 0.01 or greater and a missing data rate of 5% or less in the panel were used for GWAS. The density of SNPs per chromosome ranged between 1425.54 and 39.55 SNPs per Mb, with an average of one SNP for each 25.3 Kb.

Principal component analysis (PCA) was performed in order to ensure that any possible associations between SNPs and the trait cannot be attributed to population structure, that is, any kind of relatedness between genotypes in the sample. The PCA plots of accessions panels are shown in Supplementary Figure 5.2. For association mapping, we used both General Linear Model (GLM) and Mixed Linear Model (MLM). The former includes a correction for population structure as covariate (GLM+Q), while the latter also incorporating a correction for kinship (MLM+Q+K), which is used for association mapping. Using the panel of 122 accessions, and using both GLM+Q (P-value = 1.74E-19) and MLM+Q+K (P-value = 1.96E-7) statistical models, two SNPs at positions 30,740,460 and 30,740,501 on

chromosome 3 were found to be significantly associated with monoecy/andromonoecy traits (Figure 5.4). The two SNPs lie approximately 80 kb upstream of the locus *M*/*m* (*CitACS4*), the major causative gene regarding monoecy/andromonoecy in watermelon (Manzano et al., 2016).

Given that the PCA plot for the panel of 122 accessions showed a cluster of andromonoecious accessions that were clearly related, all of them were derived from Africa (Supplementary Figure 5.2), we discarded them and generated a new panel comprising 96 accessions (Supplementary Table 5.1 and Supplementary Figure 5.2). In the new panel, GLM+Q resulted in an additional SNP at position 31,376,952 on chromosome 1 that was found to be significantly associated with andromonoecy and PA (P-value = 4.16E-7). Moreover, in an attempt to map the locus controlling partial andromonoecy (pa), we also discarded all the andromonoecious accessions (mm) of the panel. The remaining 47 accessions (Supplementary Figure 5.2), 17 monoecious and 30 partially andromonoecious, were subjected to the same association analysis (Supplementary Table 5.1 and Figure 5.2; Figure 5.4). The homogeneous distribution of accessions in the PCA plot with the two first principal component indicated there is no reason to suspect that population structure could influence the GWAS so as to give false positive associations (Supplementary Figure 5.2). Although the panel size was small, and MLM+Q+K detected no associated SNPs, the generalized linear model (GLM+Q) resulted in only one single SNP of significance on chromosome 1 (position 37,154,162) that was associated with partial andromonoecy (P-value = 2.94E−6). This was situated very close to the locus evidenced by BSA-seq analysis (Figure 5.4).

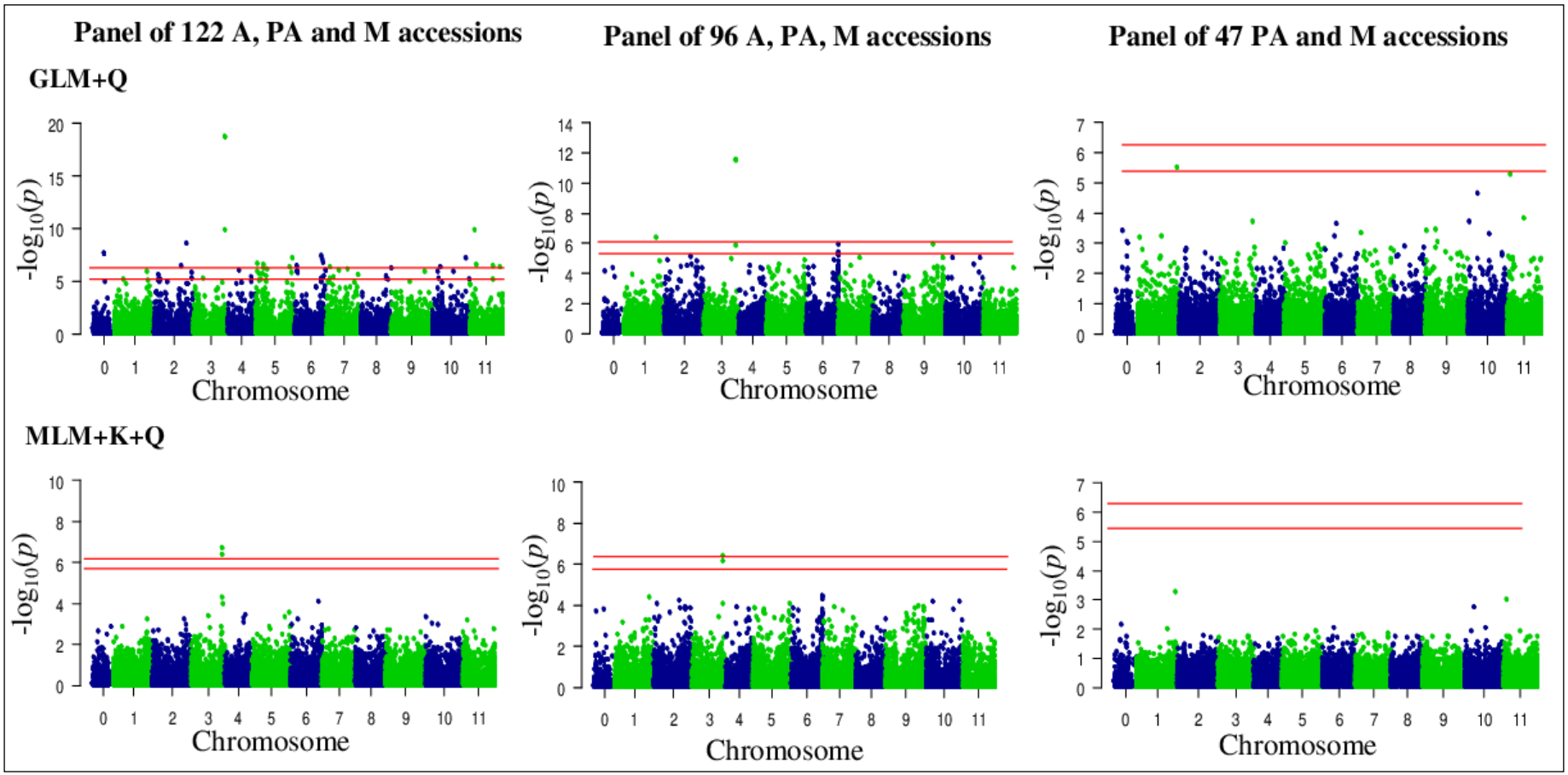


Figure 5.4. Manhattan plots of GWAS in watermelon for andromonoecious index (AI). Two statistical models have been applied: GLM+Q (General linear model, corrected for population structure), and MLM+K+Q (Mixed linear model, corrected for both population structure and relatedness). The GWAS was performed using populations of differing sizes: 122 accessions, comprising the sex morphotypes A, PA, M; 96 accessions, comprising the morphotypes A, PA and M, and 47 accessions, comprising the morphotypes PA and M accessions. Red horizontal lines indicate the Bonferroni corrected significance thresholds of GWAS at p = 0.05 and p = 0.01, respectively.

5.4.5. Fine Mapping of QTL1

Since QTL1 has a total size of 4.2 Mb (position 32.24 Mb to 36.44 Mb on chromosome 1), a fine-mapping study was designed to narrow down the genomic interval at the *pa* locus. A total of 220 F2 plants were genotyped for 11 SNPs (identified by the BSA-seq approach) that were about 0.5 Mb distant (Supplementary Table 5.3). 25 informative plants showing recombination events in the region were used to delimit QTL1 within an interval of 867 kb, between markers 1_33997548 and 1_ 34864233 (Figure 5.5). The genotypes and phenotypes of F2 plants (Figure 5.5) clearly indicate that this 867 Kb region cosegregates with the monoecious/partial andromonoecious trait.

The 867 Kb genomic region contains 101 annotated genes (Supplementary Table 5.4). We investigated whether the function of these genes was related to ethylene biosynthesis and signaling, or with flower development and sex determination (Table 5.4). The narrowest interval contains an F-box gene (ClCG01G020700); it has been reported some of the members of this family are known to regulate ethylene response pathway (Qiao et al., 2009), as well as four linked chitinase-like genes (ClCG01G020770, ClCG01G020780, ClCG01G020790 and ClCG01G020800), some of these are known to control developmental processes by regulating ethylene biosynthesis (Zhong et al., 2002; Gu et al., 2019). The 867 Kb genomic region also contains several genes known to be involved in flower development (Table 5.4). The identified genes encode for: an FG-GAP repeat-containing family protein (ClCG01G020030); a lateral root primordium family protein (ClCG01G020040); a Glutaredoxin family protein (GRXs) (ClCG01G020060); GATA transcription factor (ClCG01G020080); a WUSCHEL related homeobox 1 transcription factor (ClCG01G020260) and lastly for an MS5 male sterility family protein (ClCG01G020430) (Table 5.4).

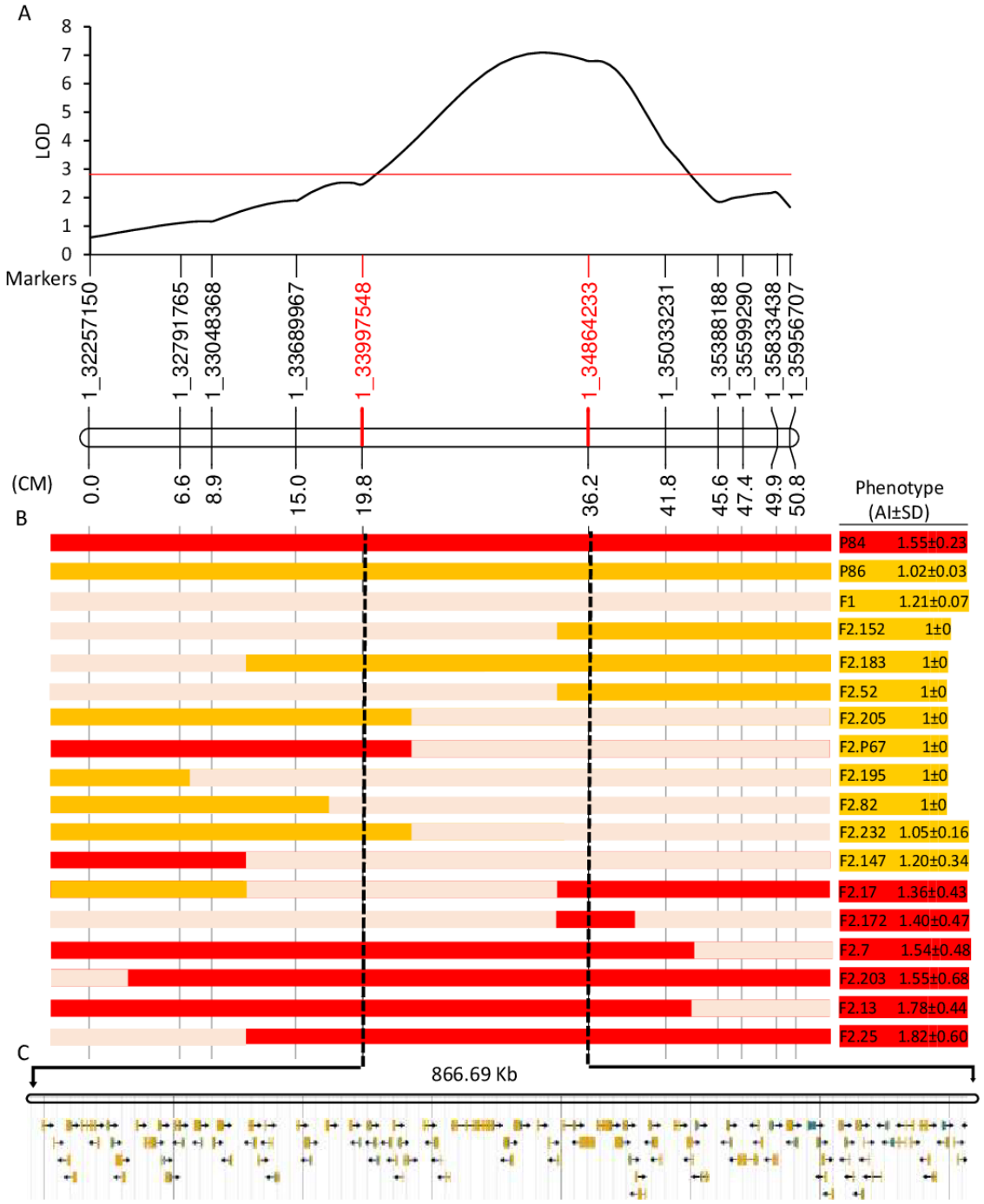


Figure 5.5. Fine mapping of partial andromonoecy (*pa*) locus. (**A**) Interval mapping result on chromosome 1, showing the logarithm of the odds (LOD) score (y-axis), and the genetic distances for the 11 SNPs (x-axis) that were used. The numbers of the markers correspond to their genomic position in bp. (**B**) Genotypes and Phenotypes for parents, F1, and 15 F2 informative individuals showing recombination in genomic interval. The color bars show the genotypes along chromosome 1: red, homozygous for P84 allele (partially andromonoecious); orange, homozygous for the P86 allele (monoecious); light pink, heterozygous. The phenotypes for the andromonoecy index (AI) are shown on the right: red; partial andromonoecious (1.35<AI<2.7); orange, monoecious (1≤AI≤1.35). (**C**) The *pa* locus maps on an interval of 866.69 Kb between markers 1_3399758 and 1_34864233, which contains 101 annotated genes.

Searching for candidate genes for Pa/Pa locus we also investigated the impact of 901 SNPs and 169 INDELs in the mapped region. 41 of the SNPs were exonic, and 21 caused non-synonymous changes located in the coding sequences of 17 genes (Supplementary Table 5.4). Of the exonic and non-synonymous SNPs, only two genes - ClCG01G019850 which encodes for a Cytochrome P450, and ClCG01G020620, which encodes for an unknown protein - were found to be distributed in M- and PA-bulk as expected according to the parental phenotypes and the inheritance mode of the trait (Supplementary Table 5.4). Among a total of 169 INDELs, only one was found to be located within a coding region. It corresponds to a frameshift deletion of 38 bp that was found in gene ClCG01G020800, which was annotated in the watermelon genome as a Chitinase-like protein (Table 5.4). The deletion was found solely in the PA parental line (P84) and in the PA-bulk.

Table 5.4. Annotated genes in the mapping region of the watermelon *pa* locus that are related to ethylene biosynthesis and response, or to flower development.

Gene_ID	Annotation	Function in flower development	References
ClCG01G020030	FG-GAP repeat-containing family protein	Some family members regulate stamen and pollen development.	Zou et al. 2020
ClCG01G020040	Lateral root primordium family protein	The SHI/STY family of transcription factors regulate the development of different plant organs, including stamens and carpels.	Kuusk et al. 2006; Estornell et al. 2018; Singh et al. 2020
ClCG01G020060	Glutaredoxin family protein (GRXs)	ROXY1 and ROXY2 of Arabidopsis are involved in anther development. The double *roxy1 roxy2* mutants does not have anthers and is consequently male sterile.	Reichheld et al. 2010; Xing and Zachgo 2008
ClCG01G020080	GATA transcription factor	Floral organ development in Arabidopsis. Mutants for the *HANABA TARANU* gene in Arabidopsis have fused sepals and reduced organ number in the four whorls, but especially in petals and stamens.	Zhao et al. 2004; Mara and Irish 2008
ClCG01G020260	WUSCHEL related homeobox	Lateral organ outgrowth and floral organ fusion in Arabidopsis, Petunia and Medicago.	Matsumoto and Okada 2001; Yu et al. 2003; Vandenbussche et al. 2009
ClCG01G020430	Male sterility MS5 family protein	Microsporogenesis.	Sanders et al. 1999
ClCG01G020700	F-box family protein	Components of the Ubiquitin-proteosome system for protein degradation. Regulation of EIN2 and EIN3 ethylene signaling elements, and flower developmental genes such as UFO. DDF1-1 of rice control B-class floral homeotic genes.	Chae et al. 2008; Qiao et al. 2009; Duan et al. 2012; Xing et al. 2012; Lin et al. 2016
ClCG01G020770 ClCG01G020780 ClCG01G020790 ClCG01G020800	Chitinase	The Arabidopsis CHITINASE LIKE 1 (CTL1) controls root elongation by negatively regulating ethylene biosynthesis and response. Floral organ development in Arabidopsis and rice.	Gu et al. 2019; Takakura et al. 2000

5.5. Discussion

5.5.1. Partial andromonoecious phenotype is conferred by a single recessive gene other than *CitACS4*

As monoecy is an important trait in watermelon, the occurrence of partial or complete andromonoecy not only implies the use of manual emasculation for the production of hybrid seed (Wehner, 2008), but also a reduced fruit set and quality (Aguado et al., 2018). Three watermelon inbred lines, P86, P84 and P87, have been previously identified and characterised; they are representative of monoecy, partially andromonoecy and andromonoecy respectively (Manzano et al., 2016). The P86 and P87 are very stable lines under different environmental conditions, producing either male and female flowers (P86) and male and hermaphrodite flowers (P87), respectively. The P84 line was more unstable, that is, besides male and female, also developed bisexual and hermaphrodite flowers when plants were grown under a high-temperature regime (Aguado et al., 2018). By taking into account the andromonoecious index (AI) (Martínez et al., 2014), as well as studying the sex phenotypes of these control lines, more than 200 *C. lanatus* accessions were separated into three sex phenotypic classes: monoecious (plants with AI = 1-1,35), partially andromonoecious (plants with AI = 1,35-2,7), and andromonoecious (plants with AI = 2,7-3). It has been reported that andromonoecy phenotype is caused by the *CitACS4* mutant allele *m* (Boualem et al., 2016; Ji et al., 2016; Manzano et al., 2016), although is also responsible for PA under heterozygous conditions. So, *MM* and *mm* plants are monoecious and andromonoecious, respectively, while *Mm* are PA, with AI ranging from 1,3 to 2,7 (Manzano et al., 2016).

During this study, we discovered that the partial andromonoecy of differing watermelon accessions, including that of P84, is not solely dependent upon the action of the *CitACS4* gene, but also upon that of a different recessive gene. We came to this conclusion in the following way: First, a number of accessions exhibited a PA phenotype, but all were homozygous for the WT allele of *CitACS4* (*MM*). Second, no variation was found in the coding sequence of monoecious and PA *MM* plants. Third, no genetic linkage was found between PA and *CitACS4* in a segregating F2 population derived from the cross P86 (*MM*, monoecious) X P84 (*MM*, PA) lines. The phenotypic segregations in the F2 (3:1 for M:PA), F3 and F4 generations (Table 5.2) confirmed that the PA phenotype is conferred by a single recessive locus called *pa*. Therefore, the three most important sex morphotypes of

watermelon can be conferred by the combination of the loci *M/m* and *Pa/pa*: i) monoecious, *MM Pa_*; ii) partial andromonoecious, *MM papa* and *Mm _ _*; and iii) andromonoecious, *mm _ _*. It is possible that the *pa* locus could be equivalent to the trimonoecious (*tm*) locus proposed by Rosa (1928) and Ji et al. (2015), but these authors considered trimonoecious to the plants having the three flower types, without defining the level of andromonoecy that we calculated in this paper on the basis of the Andromonoecy Index (AI). For this reason, it is probable that the trimonoecious trait is not exactly what we have called here partial andromonoecy (PA).

According to our data, breeding for monoecy in watermelon will not only require the selection of the monoecious *M* allele of the major gene *CitACS4*, but also a counter-selection of the *pa* allele associated with partial andromonoecy. The selection of these two loci would improve the stability of monoecy in watermelon lines and commercial hybrids, thus increasing fruit set and fruit quality associated with female flowers in monoecious genotypes (Aguado et al., 2018).

5.5.2. Location of the pa locus on a 867 kb genomic region of chromosome 1 using a combination of BSA-seq, GWAS and fine mapping

BSA-seq (Takagi et al., 2013) is a helpful alternative to the conventional gene mapping approach that has been used in plant studies. It is cheaper and less time consuming than conventional mapping as it requires less phenotyping and genotyping work. The combination of bulked segregant analysis of a pair of bulked DNA from plants with extreme phenotypes, together with whole genome re-sequencing, proved to be effective for mapping agronomic traits in differing plant species, including soybean (Song et al., 2017), Brassica (Zhu et al., 2019a; Li et al., 2020a), rice (Tao et al., 2018; Gao et al., 2019), sunflower (Imerovski et al., 2019), barley (Xu et al., 2019), pea (Zheng et al., 2018), maize (Klein et al., 2018), and cucurbit species such as cucumber and melon (Pujol et al., 2019; Zhang et al., 2019c), amongst others. The approach has been very successfully used for mapping several important agronomic traits in watermelon, including short internode length (Gebremeskel et al., 2020), dwarf phenotype (Zhu et al., 2019b), fruit colour (Dou et al., 2018a; Oren et al., 2019), and fruit shape (Dou et al., 2018b). Utilising an F2 segregating population and a BSA-seq approach, we were also able to quickly map the locus *Pa*/*pa* on a single genomic region of chromosome 1. The same region was also revealed by Genome

Wide association analysis (GWAS) in different alternative panels of 122, 96 and 47 accessions of dessert watermelon. GWAS easily detected several SNPs on chromosome 3, which are tightly linked to major andromonoecious gene *CitACS4* (Boualem et al., 2016; Ji et al., 2016; Manzano et al., 2016). It also identified two SNPs on chromosome 1 as being responsible for the PA phenotype. GWAS was used to good effect to map other important agronomic traits in watermelon, melon and cucumber (Nimmakayala et al., 2016; Yagcioglu et al., 2016; Dou et al., 2018b, 2018a; Hou et al., 2018; Wang et al., 2018; Bo et al., 2019b, 2019a; Oren et al., 2019). The major sex-determining gene for monoecy/andromonoecy in melon (*CmACS7*) was also successfully detected using the GWAS approach (Gur et al., 2017; Zhang et al., 2019a). No other minor andromonoecy loci were already have yet been found in cucurbits. The future use of panels with a greater number of watermelon accessions will surely increase the powerful of GWAS for finding all the natural genetic variability involved in the control of this trait. This will undoubtedly favor the identification of new molecular markers and genes for the selection of materials with an increasingly stable monoecy.

Fine mapping narrowed down the *pa* locus to a genomic region of 867 kb containing 101 annotated genes. The region was firstly scanned for ethylene related genes, since this hormone is the main regulator of sex determination in watermelon and other cucurbits (Manzano et al., 2014, 2016; Ji et al., 2015). The genomic interval has no ethylene biosynthesis or response genes but contains Chitinase- and F-box-like genes that could modulate ethylene biosynthesis and signaling pathways (Table 5.4). Arabidopsis *CHITINASE LIKE1* (*CTL1*) controls root development in etiolated seedlings by negatively regulating ethylene biosynthesis genes in response to perturbed cell wall integrity (Zhong et al., 2002; Hermans et al., 2010; Gu et al., 2019), and mutations in *CTL1* and other *Chitinase-like* genes overproduce ethylene and enhance responsiveness to ethylene (Gu et al., 2019). It has long been proposed that Chitinases enhance a plant's defense against pathogens by hydrolyzing chitin, which is the main component of many fungal cell walls. However, there is increasing evidence demonstrating the role of these proteins in plant development also, including that of flowers (Grover, 2012). Certain Chitinases are predominantly expressed in the floral organs of some plant species, irrespective of a pathogenesis response (Lotan et al., 1989; Takakura et al., 2000). In tobacco, certain Chitinases that accumulate in normal stamens were found to be absent in a cytoplasmic male sterile mutant with reduced and malformed stamens (Lotan et al., 1989). Chitinases are also induced in abscission zones of

floral organs (McKim et al., 2008), this being associated with the lack of abscission in the flowers of the Arabidopsis *ida* (*inflorescence deficient in abscission*) mutants (Butenko and Simon, 2015). Given that the Chitinase-like ClCG01G020800 watermelon gene in the PA line (P84) and the PA-bulk contain a frameshift deletion of 38 bp, it is possible to speculate that this gene could be responsible of the PA phenotype.

On the other hand, F-box proteins are important components of the ubiquitinin proteosome pathway, a regulatory system of many plant hormone receptors and signaling proteins (Yu et al., 2007), including the ethylene signaling factor EIN2 and the ethylene transcriptional activator EIN3 (Guo and Ecker, 2003; Potuschak et al., 2003; Gagne et al., 2004; Qiao et al., 2009). Moreover, the mutation *dwarf and deformed flowers 1-1* (*ddf1-1*) of rice, which alters the identity of whorls 2 an 3 floral organs, is affected in a gene that encodes for an F-box protein regulating B-class homeotic genes (Duan et al., 2012). The auxin receptor F-box protein TRANSPORT INHIBITOR RESPONSE 1 (TIR1) has recently been reported to be associated with pistil development in the unisexual flowers of bitter gourd *Momordica charantia* (Lin et al., 2016), which encourages the role of this F-box protein in sex determination within the Cucurbitaceae family.

Other flower developmental genes that map in the *pa* locus, but are not related with ethylene (Table 5.4), include: a Wuschel-related homeobox (WOX) transcription factor, which has been associated with lateral organ outgrowth and floral organ fusion in different plant species (Matsumoto and Okada, 2001; Yu et al., 2003; Vandenbussche et al., 2009); a glutaredoxin family protein, some of its members are required for stamen development in Arabidopsis (Xing and Zachgo, 2008; Reichheld et al., 2010); - a GATA-like transcription factor, GATA being a protein family that plays a vital role in the functioning of various physiological and developmental processes including the development and the identity of floral organs (Zhao et al., 2004; Mara and Irish, 2008); a lateral root primordium protein, from a gene family that includes SHY/STY transcription factors involved in the promotion of the development of carpels and stamens, among other organs (Kuusk et al., 2006; Estornell et al., 2018); an FG-GAP protein and a male sterility MS5 protein, some of whose members are known to participate in anther and pollen development (Sanders et al., 1999; Zou et al., 2020). Identifying the gene that is actually involved in the partial suppression of stamen arrest in partially andromonoecious watermelon plants will, however, require a more detailed study of this genomic region on watermelon chromosome 1. Identifying this will not only benefit watermelon breeding, but will also improve the stability of monoecy in other crops such as

melon, cucumber and zucchini, where instability also leads to a decrease in production and quality (Boualem et al., 2009; Díaz et al., 2014; Martínez et al., 2014; Tan et al., 2015; Manzano et al., 2016; Martos-Fuentes et al., 2017; Aguado et al., 2018).

6. General conclusions

1. The segregation ratio 1:2:1 (andromonoecious: partially andromonoecious: monoecious) in two F2 populations derived from crosses between monoecious and andromonoecious lines demonstrated that monoecy/andromonoecy trait is conferred by a single gene. Monoecy (*MM*) is semi-dominant to andromonoecy (*mm*), with heterozygous plants (*Mm*) being partially andromonoecious.

2. The *C. lanatus CitACS4* gene encodes for a protein showing high similarity with the ethylene biosynthesis enzymes CmACS7, CsACS2, and CpACS27, involved in the control of monoecy in melon, cucumber and squash, respectively. According to this function, the gene is specifically expressed in carpel primordia of female flowers during the earliest stages of development, but not in hermaphroditic and male flowers.

3. The abortion of stamen during female flower development requires the production of ethylene mediated by *CitACS4*. The andromonoecious sex morphotype in watermelon is therefore caused by a *CitACS4* mutation (*m*) that reduce the production of ethylene in the floral buds will be developed as hermaphrodite flowers.

4. *CitACS4* is a pleiotropic gene that exhibits multiple phenotypic expression in flower and fruit development. Regarding sex expression, the mutant *m* allele of the gene is responsible of an increased number of pistillate flowers per plant, but the gene does not seem to control female flowering transition.

5. The mutant *m* allele of *CitACS4* gene was found associated with reduced growth rate and maturation of petals and carpels, which results in hermaphrodite flowers with delayed anthesis and larger ovaries.

6. The ethylene biosynthesis gene *CitACS4* regulates fruit development in watermelon. Fruit and seed set in monoecious *MM* plants was higher than in andromonoecious *mm* plants, and the ovary and fruit of hermaphrodite *mm* flowers are rounder than those of female flowers.

7.A number of accessions of *Citrullus lanatus* have been identified as having a partially andromonoecious phenotype but they are homozygous for the monoecious *M* allele of *CitACS4*. The partially andromonoecious phenotype of these accessions is not conferred by this ethylene biosynthesis gene, but for a different recessive locus which we call *partial andromonoecy (pa)*.

8. GWAS analysis with a panel of more than 100 accessions of *C. lanatus* identified a number of SNP markers associated with andromonoecy in the genomic region linked to

CitACS4 gene on chromosome 3, but also a region expanding across 32.24-36.44 Mb in chromosome 1 that was associated with the partial andromonoecy *pa* locus.

9. Bulk segregant-analysis sequencing (BSA-seq) and fine mapping of a segregating populations derived from a biparental cross between monoecious and partially andromonoecious lines, mapped the *pa* locus to a 887 kb region in chromosome 1 of watermelon. This genomic region contains 101 genes, but 8 candidate genes have been selected not only by their roles in flower development and ethylene biosynthesis and signaling, but also by the impact of their SNPs and indels in the two sequenced bulks.

7. Supplementary material

Supplementary Table 3.1. Primers used in quantitative real time RT-PCR reactions and to amplify a full sequence of *CitACS4* gene.

Gene	Primer name	Sequence	Use
CitACS4	*CitACS4gen-F1*	GGCTACATTCAACAGTCTTCCA	PCR, S
	CitACS4gen-R1	TTCATCTTCCTTCCTCATCCTC	PCR, S, G
	Fw	AGTTCAAATTTTTTAAACTGGGTTG	S
	Rv	AGTGCATGAAATTAAATCAACTTACA	S
	CitACS4A-F	GAATGCCGGTTTATTTTGG	PCR, G
	CitACS4M-F	GAATGCCGGTTTATTTTGC	PCR, G
	CitACS4A-R	CGGGCTTAAATTCATCCAC	PCR, G
	CitACS4M-R	CGGGCTTAAATTCATCCAG	PCR, G
	CitACS4S-F	TCCCGGGTTTAGAATCG	PCR, G
	CitACS4-FWD1	CTGCAGCCAATGAGCTTC	qPCR
	CitACS4-FWD2	CACTCCTTACTATCCTGGATTTG	qPCR
	CitACS4-REV3	GGTTCCGTTTTCTCCTCG	qPCR
	CitACS4-REV4	CGGTCCACAATTGAGGAG	qPCR
	Cit ACS-REV5	CCGCGGCTATAAAAACG	qPCR
ACTINE	*CitActin-F*	TGCCATTCTCCGTTTGGACC	qPCR
	CitActin-R	GCAACGGAATCTCTAGCTCC	qPCR

S: Sequencing; G: Genotyping *CitACS4* gene

Supplementary Table 4.1. Number of phenotyped plants in the four parental lines (P84, P85, P86 and P87), and the three F2 populations derived from crosses between monoecious and andromonoecious lines.

Generation	Spring/Summer				Autumn/Winter			
	Genotyped plants	Selected plants for phenotyping			Genotyped plant	Selected plants for phenotyping		
		MM	***Mm***	***mm***		***MM***	***Mm***	***mm***
P84	20	15			20	15		
P85	20	15			20	15		
P86	20	15			20	15		
P87	20			15	20			15
F2 (P84XP87)	137	15	29	15	134	18	30	15
F2 (P85XP87)	100	15	19	17	126	15	21	15
F2 (P86XP87)	100	15	17	15	112	19	23	19

After sowing more than 100 seeds from each F2 population, plants were genotyped for *CitACS4* gene as *MM*, *Mm* and *mm*, and a minimum of 15 plants of each genotype were transplanted to greenhouse for phenotyping.

Supplementary Table 4.3. Segregation ratio of monoecious, partially andromonoecious and andromonoecious plants from the F2 population derived from the cross P84XP87.

Genotype	Mono	PA	Andro	Expected segregation	$\chi 2$	p-value
P84	3	6				
P87			12			
F1(P84XP87)		12				
F2(P84XP87)	31	79	27	1:2:1	3.5	0.177

The F2 plants were phenotyped on the basic of their average AI in monoecious, (Mono, 1≤AI≤1.2), partially andromonoecious (PA, 1.2≤AI≤2.7), and andromonoecious (Andro, 2.7≤AI≤3).

Supplementary Table 4.2. Number of flowers and fruits used for assessing fruit set and the number of viable seeds in parental lines (P84, P85, P86 and P87), and the three F2 populations derived from crosses between monoecious and andromonoecious lines.

Generation	*CitACS4* genotype	Number pollinated flowers used for fruit set		Number of mature fruits used for seed set	
		Spring /Summer	Autumn/ Winter	Spring /Summer	Autumn/ Winter
P84	*MM*	15	32	10	10
P85	*MM*	15	27	13	15
P86	*MM*	15	27	13	18
P87	*mm*	19	54	10	10
F2 (P84XP87)	*MM*	19	21	10	10
	Mm	29	25	17	10
	mm	20	37	10	10
F2 (P85XP87)	*MM*	15	17	11	10
	Mm	21	20	16	10
	mm	20	26	10	10
F2 (P86XP87)	*MM*	15	18	12	13
	Mm	21	25	15	10
	mm	20	30	10	10

The flowers and the fruit derived from the plants shown in Supplementary Table 4.1. When fruit set was reduced, the number of pollinated flowers was increased to obtain a minimum of mature 10 fruits.

Supplementary Table 4.4. Comparison of floral organ size in monoecious (*MM*) and andromonoecious (*mm*) lines, and their respective F2 offspring.

		P84	P87	F2 (P84 X P87)			P85	P87	F2 (P85 X P87)			P86	P87	F2 (P86 X P87)		
				MM	*Mm*	*mm*			*MM*	*Mm*	*mm*			*MM*	*Mm*	*mm*
Ovary length (mm)	**t0**	2.3a	2.4a	2.4a	2.4a	2.4a	2.5a	2.4a	2.3a	2.4a	2.4a	2.4a	2.4a	2.4a	2.4a	2.4a
	t2	2.7d	2.8c	2.9ab	2.9bc	3.0a	2.9ab	2.8b	3.0a	2.9ab	2.9ab	2.8b	2.8b	2.9ab	3.0a	2.8b
	t4	3.8bc	3.4c	4.5b	7.6a	4.1bc	3.5c	3.4c	4.1b	7.4a	3.8bc	3.7c	3.4c	3.8c	4.4b	5.1a
	t6	12.3ab	8.0c	14.9a	12.4ab	10.6bc	12.3a	8.0c	13.0a	10.1b	7.1c	10.6a	8.0bc	10.8a	9.4ab	7.3c
	t8		12.4c	18.9a	15.7b	15.0bc		12.4ab	14.4a	14.0a	10.8b		12.4a	13.6a	11.9ab	9.6b
	t10		16.9a			16.6a		16.9a			12.3b		16.9a			15.1a
	t12										16.4					
Female Petal length (mm)	**t0**	2.4a	2.5a	2.4a	2.5a	2.5a	2.3a	2.5a	2.4a	2.5a	2.5a	2.4b	2.5a	2.5a	2.4a	2.6a
	t2	2.6b	2.9b	2.8b	3.3a	2.9b	2.5b	2.9a	3.0a	2.9a	2.9a	2.7c	2.9b	3.0ab	3.0ab	3.2a
	t4	3.1b	3.4b	3.5b	5.9a	3.7b	2.9c	3.4bc	3.7b	5.0a	3.3bc	3.0c	3.4bc	3.6bc	3.8b	4.7a
	t6	12.8a	7.2d	9.3bc	10.3b	7.6cd	13.4a	7.2c	10.9b	7.1c	5.3c	13.2a	7.2cd	8.8bc	9.2b	6.8d
	t8		13.1ab	15.3ab	16.7a	10.7b		13.1a	10.9ab	11.9ab	8.4b		13.1a	11.8ab	13.0ab	8.8b
	t10		17.2a			16.2a		17.2a			10.2b		17.2a			16.4a
	t12										17.4					
Ovary length of after pollination (mm)	**Anthesis**	16.9b	18.3b	19.2ab	20.6a	18.5ab	13.9c	18.3a	15.8b	16.2b	19.1a	14.5b	18.32a	14.9b	15.6b	18.36a
	4 dpa	32.12a	31.7a	40.9a	37.6a	28.6a	30.8a	31.7a	34.4a	30.5a	33.5a	40.6a	31.7b	35.6ab	31.8ab	34.3ab
	10 dpa	171.9a	98.2b	119.1b	104.5b	86.8b	108.2a	98.2a	94.4a	90.3a	82.9a	126.6a	98.2b	104.5ab	107.5ab	89.6b
	14 dpa	199.3a	124.9cd	150.8b	134.5bc	104.6d	171.5a	141.3ab	128.52b	124.14b	125.8b	166.74a	141.3ab	144.2ab	143.8ab	125.7b

Supplementary Table 5.1. List of *Citrullus lanatus* accessions used in the present study and their origin, genotype for the *CitACS4* gene and phenotypic data of the Andromonoecious Index (AI) in two locations.

Accession	Origin	Genotype for gene *CitACS4*	Andromonoecious index (AI)				Panel of 122 accessions for GWAS	Panel of 96 accessions for GWAS	Panel of 47 accessions for GWAS
			Open field evaluation (Raleigh, North Carolina) Summer 2017	Greenhouse evaluation (Almería, Spain) Summer 2018	Range	Sex morphotype			
P84	Inbred line	*MM*		1.55±0.23	1.43-1.80	PA			
P86	Inbred line	*MM*		1.02±0.03	1.00-1.07	M			
P87	Inbred line	*mm*		3.00±0.00	3.00-3.00	A			
Calhoun Gray	Commercial	*MM*		1.58±0.62	1.00-2.40	PA			
Charleston Gray	Commercial	*MM*	1±0.00	1.32±0.25	1.00-1.66	M			
BGHZ0674	Spain, Badajoz	*MM, Mm, mm*		2.45±0.64	1.43-3.00	PA			
BGHZ0675	Spain, Badajoz	*MM, Mm, mm*		1.56±0.83	1.00-3.00	PA			
BGHZ0984	Spain, Badajoz	*MM, Mm, mm*		1.79±0.78	1.00-3.00	PA			
BGHZ1254	Spain, Badajoz	*MM, Mm, mm*		1.87±0.71	1.10-2.88	PA			
BGHZ1414	Spain, Badajoz	*mm*		3.00±0.00	3.00-3.00	A			
BGHZ3027		*MM*		1.14±0.26	1.00-1.59	M			
BGHZ4155	Spain, Badajoz	*mm*		3.00±0.00	3.00-3.00	A			
BGHZ4156	Spain, Sevilla	*MM, Mm, mm*		1.59±0.47	1.00-2.38	PA			
BGHZ4157	Spain, Córdoba	*mm*		2.98±0.06	2.88-3.00	A			
BGHZ4636	Spain, Córdoba	*MM, Mm, mm*		2.29±0.70	1.50-3.00	PA			

BGHZ4637	Spain, Córdoba	*MM*	1.15±0.18	1.00-1.50	M
BGHZ4847	Spain, Granada	*MM*	1.28±0.17	1.11-1.50	M
BGHZ4848	Spain, Granada	*MM*	1.00±0.00	1.00-1.13	M
BGHZ4849	Spain, Granada	*MM*	1.79±0.43	1.17-2.36	PA
BGHZ5064	Spain, Cádiz	*MM*	1.22±0.22	1.00-1.50	M
BGHZ5132	Spain, Cáceres	*MM, Mm, mm*	2.05±0.63	1.00-2.67	PA
BGHZ5248	Spain, Islas Baleares	*MM, Mm, mm*	2.98±0.04	2.91-3.00	A
BGHZ5249	Spain, Islas Baleares	*MM, Mm, mm*	2.60±0.27	2.17-2.90	PA
BGHZ5250	Spain, Islas Baleares	*MM, Mm, mm*	2.31±0.49	1.70-3.00	PA
BGHZ5251	Spain, Jaén	*MM, Mm, mm*	1.49±0.33	1.14-1.88	PA
BGHZ5252	Spain, Jaén	*MM*	1.37±0.23	1.15-1.88	PA
BGHZ5253	Spain, Jaén	*MM, Mm, mm*	1.41±0.58	1.00-2.20	PA
BGHZ5428	Spain, Huelva	*MM, Mm, mm*	1.34±0.27	1.00-1.60	M
BGHZ5441	Spain, Málaga	*MM*	1.96±0.38	1.44-2.50	PA
BGHZ5442	Spain, Málaga	*MM*	2.00±0.60	1.40-2.83	PA
BGHZ5443	Spain, Málaga	*MM, Mm, mm*	2.37±0.55	1.71-3.00	PA
BGHZ5622	Spain, Orense	*MM, Mm, mm*	1.19±0.19	1.00-1.46	M
BGHZ5623	Spain, Orense	*MM*	1.28±0.24	1.00-1.50	M
BGHZ5746	Spain, Granada	*MM, Mm, mm*	2.95±0.11	2.75-3.00	A
BGHZ5854	Spain, Huelva	*MM*	1.34±0.24	1.11-1.71	M
BGHZ5993	Spain, Huelva	*MM*	1.47±0.37	1.11-2.00	PA
BGHZ6020	Spain, Toledo	*mm*	3.00±0.00	3.00-3.00	A
BGV000743	Spain, Granada	*MM, Mm, mm*	2.60±0.89	1.00-3.00	PA
BGV000747	Spain, Cádiz	*MM, Mm, mm*	1.49±0.39	1.14-2.10	PA
BGV000754	Spain, Jaén	*MM, Mm, mm*	1.47±0.32	1.10-2.92	PA

BGV000762	Spain, Granada	*mm*		2.83±0.11	2.71-3.00	A			
BGV000765	Spain, Huelva	*MM, Mm, mm*		3.00±0.00	3.00-3.00	A			
BGV001875	Spain, Tarragona	*MM*		1.00±0.00	1.00-1.00	M			
BGV001880	Spain, Girona	*MM, Mm, mm*		2.74±0.58	1.71-3.00	A			
BGV001887	Spain, Girona	*MM*		1.47±0.38	1.00-2.00	PA			
BGV002674	Spain, Las Palmas	*MM*		1.65±0.32	1.25-2.00	PA			
BGV002712	Spain, Las Palmas	*MM*		1.00±0.00	1.00-1.22	M			
BGV004081	Spain, Cáceres	*MM, Mm, mm*		2.93±0.11	2.75-3.00	A			
BGV005163	Spain, Castellón	*MM*		1.19±0.13	1.00-1.33	M			
BGV005264	Spain, Valencia	*MM*		1.11±0.25	1.00-1.67	M			
BGV010408	Spain, Ciudad Real	*mm*		3.00±0.00	3.00-3.00	A			
BGV016388		*MM*		1.11±0.24	1.00-1.55	M			
Grif 5595	India	*MM*	1.00±0.00	1.32±0.04	1.00-1.00	M	XXXX	XXXX	XXXX
PI 163204	India, Punjab	*MM, Mm, mm*	2.00±1.00	2.43±0.62	1.59-2.71	PA	XXXX	XXXX	XXXX
PI 164992	Turkey, Ankara	*MM*	1.60±0.89	1.86±0.12	1.61-2.00	PA	XXXX	XXXX	XXXX
PI 164998	Turkey, Ankara	*MM*	1.40±0.89	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 166993	Turkey, Hatay	*mm*	2.80±0.45	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 169235	Turkey, Manisa	*mm*	2.63±0.74	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 169243	Turkey, Antalya	*mm*	3.00±0.00	2.98±0.06	2.88-3.00	A	XXXX	XXXX	
PI 169254	Turkey, Izmir	*MM, Mm, mm*	3.00±0.00	2.25±0.87	1.00-3.00	PA			
PI 169270	Turkey, Kirklareli	*mm*	3.00±0.00	2.95±0.11	2.75-3.00	A	XXXX	XXXX	
PI 169281	Turkey	*mm*	2.43±0.98	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 169287	Turkey, Bursa	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 169300	Turkey, Hatay	*mm*	2.00±1.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 172803	Turkey, Maras	*mm*	2.75±0.71	3.00±0.00	3.00-3.00	A	XXXX	XXXX	

PI 174099	Turkey, Elazig	*mm*	2.71±0.76	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 174109	Turkey, Elazig	*mm*	2.20±1.10	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 175665	Turkey, Kayseri	*mm*	2.67±0.82	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 175651	Turkey, Balikesir	*mm*	3.00±0.00	2.98±0.04	2.91-3.00	A	XXXX	XXXX	
PI 176487	Turkey, Tunceli	*mm*	2.60±0.55	2.85±0.29	2.33-3.00	A	XXXX	XXXX	
PI 176490	Turkey, Sivas	*MM, Mm, mm*	1.40±0.89	2.62±0.58	2.29-3.00	PA			
PI 176495	Turkey, Konya	*mm*	3.00±0.00	2.81±0.36	2.18-3.00	A			
PI 176915	Turkey, Konya	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 177325	Turkey, Hakkari	*MM*	2.14±1.07	1.39±0.10	1.00-1.58	PA	XXXX	XXXX	XXXX
PI 178872	Turkey, Kutahya	*mm*	2.20±1.10	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 178876	Turkey, Bursa	*MM, Mm, mm*	3.00±0.00	2.52±0.96	1.38-3.00	PA			
PI 179234	Turkey, Bursa	*mm*	2.33±1.03	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 179242	Iraq	*MM*	2.17±0.98	1.36±0.10	1.00-1.55	PA	XXXX	XXXX	XXXX
PI 182177	Turkey, Kirklareli	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 180426	India, Madhya Pradesh	*mm*	3.00±0.00	2.87±0.18	2.65-3.00	A	XXXX	XXXX	
PI 183673	India, Maharashtra	*MM, Mm, mm*	1.00±0.00	1.88±0.79	1.00-2.91	PA			
PI 185636	Ghana	*MM, Mm, mm*	2.23±1.01	1.87±1.04	1.00-3.00	PA			
PI 190050	Serbia	*MM, Mm, mm*	1.80±0.84	2.05±1.10	1.00-3.00	PA			
PI 193490	Ethiopia	*MM, Mm, mm*	1.00±0.00	1.70±0.78	1.40-3.00	PA			
PI 195928	Ethiopia	*mm*	2.36±0.84	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 195562	Ethiopia	*MM, Mm, mm*	3.00±0.00	2.20±0.76	1.00-3.00	PA			
PI 203551	United States, New Mexico	*MM, Mm, mm*	3.00±0.00	2.58±0.88	1.00-3.00	PA			
PI 220779	Afghanistan	*MM*	1.50±0.84	1.40±0.14	1.00-1.72	PA	XXXX	XXXX	XXXX
PI 222713	Iran, Bakhtaran	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 228238	Israel	*mm*	2.60±0.89	3.00±0.00	3.00-3.00	A	XXXX	XXXX	

PI 234603	New Zealand, Auckland Island	*MM, Mm, mm*	1.33±0.82	1.50±1.00	1.00-3.00	PA			
PI 249559	Thailand	*MM, Mm, mm*	1.40±0.89	1.90±0.60	1.36-2.46	PA	XXXX	XXXX	XXXX
PI 254622	Sudan, Khartoum	*mm*	2.17±0.98	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 255139	South Africa	*MM, Mm, mm*	2.80±0.45	2.04±0.39	1.00-2.63	PA	XXXX	XXXX	XXXX
PI 269679	Belize	*MM, Mm, mm*	1.00±0.00	1.80±0.53	1.00-2.17	PA			
PI 260733	Sudan, Khartoum	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 266027	Venezuela, Aragua	*MM*	1.00±0.00	1.00±0.00	1.00-1.00	M			
PI 274561	Portugal	*mm*	3.00±0.00	2.94±0.10	2.78-3.00	A	XXXX	XXXX	
PI 276659	Russian Federation	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 277989	Turkey, Aydin	*mm*	3.00±0.00	2.61±0.52	1.80-3.00	PA			
PI 277988	Turkey, Aydin	*mm*	2.60±0.89	2.99±0.03	2.94-3.00	A	XXXX	XXXX	
PI 277990	Turkey, Aydin	*MM, Mm, mm*	3.00±0.00	2.35±0.67	1.00-2.33	PA	XXXX	XXXX	XXXX
PI 277993	Turkey, Bilecik	*mm*	2.83±0.41	3.00±0.00	3.00-3.00	A			
PI 278002	Turkey, Bursa	*mm*	2.50±0.84	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 306365	Gabon	*MM, Mm, mm*	1.00±0.00	1.00±0.00	1.00-1.00	M			
PI 278006	Turkey, Gaziantep	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 278008	Turkey, Gaziantep	*mm*	3.00±0.00	2.97±0.06	2.87-3.00	A	XXXX	XXXX	
PI 278023	Turkey, Kars	*MM, Mm, mm*	1.40±0.89	1.55±0.27	1.00-1.99	PA	XXXX	XXXX	XXXX
PI 278030	Turkey, Kirsehir	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 357746	Former Serbia and Montenegro	*MM*	1.00±0.00	1.02±0.04	1.00-1.07	M	XXXX	XXXX	XXXX
PI 278055	Turkey, Tunceli	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 279460	Japan	*MM, Mm, mm*	2.43±0.98	1.92±0.67	1.00-2.67	PA	XXXX	XXXX	XXXX
PI 319236	Japan	*MM, Mm, mm*	1.20±0.45	1.47±0.20	1.00-1.86	PA	XXXX	XXXX	XXXX
PI 344066	Turkey, Gaziantep	*mm*	2.50±0.84	2.99±0.03	2.94-3.00	A	XXXX	XXXX	
PI 357696	Macedonia	*mm*	1.20±0.45	3.00±0.00	3.00-3.00	A	XXXX	XXXX	

PI 357678	Macedonia	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 357718	Macedonia	*MM, Mm, mm*	3.00±0.00	2.86±0.22	2.50-3.00	A	XXXX	XXXX	
PI 357731	Macedonia	*MM, Mm, mm*	1.80±1.03	1.33±0.64	1.00-2.47	M			
PI 368493	Macedonia	*mm*	2.60±0.89	2.88±0.20	2.58-3.00	A	XXXX	XXXX	
PI 368496	Macedonia	*MM, Mm, mm*	3.00±0.00	1.73±0.77	1.38-3.00	PA			
PI 368501	Macedonia	*MM, Mm, mm*	2.60±0.89	1.49±0.86	1.00-3.00	PA			
PI 369220	Soviet Union, Former	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 183125	India, Maharashtra	*mm*	2.40±0.55	3.00±0.00	3.00-3.00	A			
PI 494820	Zambia	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 368512	Former Serbia and Montenegro	*MM, Mm, mm*	3.00±0.00	2.60±0.89	1.00-3.00	PA			
PI 370424	Macedonia	*MM, Mm, mm*	1.40±0.89	1.36±0.08	1.00-1.51	PA	XXXX	XXXX	XXXX
PI 379237	Macedonia	*MM*	1.50±0.93	1.98±0.48	1.42-2.75	PA	XXXX	XXXX	XXXX
PI 379249	Macedonia	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 379257	Former Serbia and Montenegro	*MM, Mm, mm*	1.80±1.10	2.50±0.80	1.51-3.00	PA			
PI 381705	India	*MM, Mm, mm*	1.00±0.00	1.60±0.35	1.00-1.83	PA	XXXX	XXXX	XXXX
PI 381711	India	*MM, Mm, mm*	1.00±0.00	1.60±0.39	1.00-1.89	PA	XXXX	XXXX	XXXX
PI 381715	India	*MM, Mm, mm*	1.00±0.00	1.75±1.09	1.00-3.00	PA			
PI 435282	Iraq	*MM*	1.00±0.00	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 381733	India	*mm*	3.00±0.00	2.97±0.07	2.85-3.00	A	XXXX	XXXX	
PI 381737	India	*mm*	2.60±0.89	2.98±0.06	2.88-3.00	A	XXXX	XXXX	
PI 438675	Mexico, Chiapas	*MM*	1.00±0.00	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 458738	Paraguay, Chaco	*MM, Mm, mm*	1.00±0.00	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 459075	Botswana	*mm*	2.60±0.52	3.00±0.00	3.00-3.00	A	XXXX		
PI 470249	Indonesia, Kalimantan	*MM, Mm, mm*	1.40±0.55	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 476325	Ukraine	*mm*	2.43±0.98	2.87±0.19	2.55-3.00	A	XXXX	XXXX	

PI 482248	Zimbabwe	*MM, Mm, mm*	1.88±0.99	1.76±0.84	1.00-3.00	PA			
PI 482269	Zimbabwe	*mm*	2.00±1.10	3.00±0.00	3.00-3.00	A	XXXX		
PI 482275	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 482281	Zimbabwe	*mm*	1.00±0.00	2.80±0.45	2.00-3.00	A	XXXX		
PI 482295	Zimbabwe	*mm*	2.43±0.53	2.93±0.15	2.67-3.00	A	XXXX		
PI 482305	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 482320	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 482330	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 482337	Zimbabwe	*mm*	3.00±0.00	2.98±0.04	2.92-3.00	A	XXXX		
PI 482346	Zimbabwe	*mm*	2.88±0.35	3.00±0.00	3.00-3.00	A	XXXX		
PI 482349	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 482350	Zimbabwe	*mm*	2.80±0.45	3.00±0.00	3.00-3.00	A	XXXX		
PI 482359	Zimbabwe	*mm*	1.86±1.07	2.95±0.10	2.80-3.00	A	XXXX		
PI 482364	Zimbabwe	*mm*	2.60±0.89	3.00±0.00	3.00-3.00	A	XXXX		
PI 482365	Zimbabwe	*mm*	2.91±0.30	3.00±0.00	3.00-3.00	A	XXXX		
PI 482371	Zimbabwe	*mm*	3.00±0.00	2.98±0.03	2.92-3.00	A	XXXX		
PI 482376	Zimbabwe	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 500301	Zambia	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 500305	Zambia	*mm*	3.00±0.00	3.00±0.00	3.00-3.00	A	XXXX		
PI 500311	Zambia	*mm*	3.00±0.00	2.97±0.06	2.86-3.00	A	XXXX		
PI 500314	Zambia	*MM, Mm, mm*	1.40±0.89	1.40±0.09	1.00-1.53	PA	XXXX	XXXX	XXXX
PI 500345	Zambia	*MM, Mm, mm*	3.00±0.00	2.83±0.24	2.67-3.00	A	XXXX		
PI 500349	Zambia	*MM, Mm, mm*	3.00±0.00	2.93±0.14	2.81-3.00	A	XXXX		
PI 505587	Zambia	*mm*	2.78±0.67	2.95±0.11	2.75-3.00	A	XXXX		
PI 505590	Zambia	*mm*	3.00±0.00	2.94±0.13	2.71-3.00	A	XXXX		

PI 506439	Moldova	*mm*	2.20±1.10	3.00±0.00	3.00-3.00	A	XXXX		
PI 507864	Hungary, Szabolcs-Szatmar	*MM, Mm, mm*	1.00±0.00	1.88±0.82	1.10-2.80	PA			
PI 512342	Spain, Zaragoza	*MM*	1.29±0.76	1.36±0.06	1.00-1.45	PA	XXXX	XXXX	XXXX
PI 512349	Spain, Tarragona	*MM*	1.67±1.03	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 512352	Spain, Toledo	*MM*	2.67±0.82	1.47±0.27	2.00-1.00	PA	XXXX	XXXX	XXXX
PI 512358	Spain, Caceres	*MM, Mm, mm*	1.80±1.10	1.41±0.77	1.00-2.78	PA			
PI 512373	Spain, Alicante	*mm*	2.00±1.00	2.88±0.21	2.64-3.00	A	XXXX	XXXX	
PI 512401	Spain, Cadiz	*MM*	1.00±0.00	1.41±0.15	1.00-1.64	PA	XXXX	XXXX	XXXX
PI 512392	Spain, Castellon de Plana	*MM, Mm, mm*	3.00±0.00	2.06±0.89	1.44-3.00	PA			
PI 512395	Spain, Valencia	*MM*	2.00±1.04	1.38±0.06	1.00-1.48	PA	XXXX	XXXX	XXXX
PI 512398	Spain, Granada	*MM, Mm, mm*	3.00±0.00	2.11±0.89	1.00-3.00	PA			
PI 512403	Spain, Cadiz	*mm*	2.20±0.84	2.99±0.03	2.94-3.00	A	XXXX	XXXX	
PI 525099	Egypt, Matruh	*mm*	3.00±0.00	2.62±0.53	1.90-3.00	PA			
PI 532816	China	*MM, Mm, mm*	3.00±0.00	1.40±0.81	1.00-2.83	PA			
PI 534532	Syria	*mm*	3.00±0.00	2.94±0.11	2.78-3.00	A	XXXX	XXXX	
PI 534589	Syria	*mm*	3.00±0.00	2.98±0.04	2.90-3.00	A	XXXX	XXXX	
PI 534593	Syria	*mm*	2.33±1.03	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 537274	Pakistan, Punjab	*mm*	2.40±0.89	2.70±0.30	2.28-3.00	A	XXXX	XXXX	
PI 537472	Spain, Alicante	*MM*	1.50±0.84	1.39±0.10	1.00-1.53	PA	XXXX	XXXX	XXXX
PI 556994	United States, Alabama	*MM*	2.20±1.10	1.02±0.03	1.00-1.06	M	XXXX	XXXX	XXXX
PI 593347	China, Henan	*MM*	2.60±0.89	1.37±0.03	1.33-1.44	PA	XXXX	XXXX	XXXX
PI 593350	China, Henan	*MM*	2.33±1.03	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 542118	Botswana	*MM*	2.60±0.89	1.00±0.00	1.00-1.00	M			
PI 593351	China, Henan	*MM, Mm, mm*	1.00±0.00	1.40±0.89	1.00-3.00	PA			
PI 593377	China, Xinjiang	*MM*	1.14±0.38	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX

PI 482343	Zimbabwe	*mm*	2.14±0.69	2.62±0.52	2.00-3.00	PA			
PI 593384	China, Xinjiang	*MM*	2.20±1.10	1.01±0.03	1.00-1.06	M	XXXX	XXXX	XXXX
PI 595200	United States, Georgia	*MM*	1.00±0.00	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 595202	United States, Georgia	*MM*	1.00±0.00	1.32±0.05	1.00-1.41	M	XXXX	XXXX	XXXX
PI 600951	United States	*MM*	1.00±0.00	1.36±0.04	1.00-1.42	PA	XXXX	XXXX	XXXX
PI 601228	United States	*MM*	1.86±1.07	1.39±0.06	1.00-1.51	PA	XXXX	XXXX	XXXX
PI 612464	South Korea, Pusan	*MM*	1.14±0.38	1.37±0.04	1.00-1.44	PA	XXXX	XXXX	XXXX
PI 612468	South Korea, Pusan	*MM, Mm, mm*	1.00±0.00	1.43±0.09	1.00-1.55	PA	XXXX	XXXX	XXXX
PI 629109	United States, Colorado	*MM*	1.80±1.10	1.32±0.03	1.00-1.37	M	XXXX	XXXX	XXXX
PI 632754	Bulgaria	*MM*	1.00±0.00	1.37±0.07	1.00-1.51	PA	XXXX	XXXX	XXXX
PI 635610	United States, California	*MM*	1.00±0.00	1.32±0.06	1.00-1.45	M	XXXX	XXXX	XXXX
PI 635626	United States, California	*MM*	1.67±1.03	1.41±0.09	1.00-1.58	PA	XXXX	XXXX	XXXX
PI 635682	Argentina	*MM, Mm, mm*	1.00±0.00	1.47±0.20	1.00-1.89	PA	XXXX	XXXX	XXXX
PI 635712	United States, Mississippi	*MM*	1.00±0.00	1.00±0.00	1.00-1.00	M	XXXX	XXXX	XXXX
PI 658554	Turkmenistan	*mm*	2.80±0.63	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 169260	United States, Mississippi	*mm*	2.40±0.89	3.00±0.00	3.00-3.00	A	XXXX	XXXX	
PI 600962	Turkmenistan	*MM*	1.33±0.52	1.38±0.14	1.00-1.64	PA	XXXX	XXXX	XXXX

Supplementary Table 5.2. Primers used for genotyping, for PCR and for sequencing the full coding region and also for the promoter of the *CitACS4* gene.

Gene	Primer name	Sequence	Use
CitACS4 coding region	*Forward*	CGGCTCAGGAAACGATA	Genotyping
	Reverse	GCTTCCAAATCTCAATTTCAC	Genotyping
	Probe C	FAM—TTGCTGGATGAATTTAAGCCCGTTAT-BHQ1	Genotyping
	Probe G	HEX—TTGGTGGATGAATTTAAGCCCGTTAT-BHQ1	Genotyping
	CitACS4gen-F1	GGCTACATTCAACAGTCTTCCA	PCR, sequencing
	CitACS4gen-R1	TTCATCTTCCTTCCTCATCCTC	PCR, sequencing
	Fw	AGTTCAAATTTTTTAAACTGGGTTG	Sequencing
	CitACS4-RTR1	TCGTGCCGATTCTAAACCCG	Sequencing
CitACS4 Promoter	*CitACSF6*	GTGGCTATTTATAGCTTTACGGG	PCR, sequencing
	CitACSR6	GGAACGTCGGAAACTCAC	PCR, sequencing

Supplementary Table 5.3. List of SNPs marker used for fine mapping.

Marker_ID	Genome	Chromosome	Position (bp)	Flanking sequence (100_SNP_100)
1_32257150	Charleston Gray	Chr1	32257150	AAGTGGAGAAATAAATATATTAGGCCATTATTCTTTAGTTTAGTTATATT[T/C]AGTTATGAT TAGTAAAAGTAAGTAAAATTTTTAATCCCTATATATCCAAA
1_32791765	Charleston Gray	Chr1	32791765	ACTTGATCAGACTTTCTCCAAATGCCAATGAGGAGATACAGAATCTAACC[T/C]AAGATCCC AACCCAACTATTTCCATCTCCTCCCACCTCAAAAACAGGTTA
1_33048368	Charleston Gray	Chr1	33048368	CACTTCCATTAACGTCCCATTTGAAGAACACCCTTCTTCTGGAATTTGTA[C/T]TTTGGCCGTG GCTTGAAAAATCGCCATTGCAGAAACTGAATCCAAAGAAA
1_33689967	Charleston Gray	Chr1	33689967	CATTTTAAAGTAAAATTGCCAGCATAACATGAGCATAGTTCGATCGTAAA[A/G]TATGCCTT ATCGATTATGAAAACCACACATGTTTAGCAAATCAGTTGCTT
1_33997548	Charleston Gray	Chr1	33997548	GCAAAAGAGAATTTAATTATGTTCATTTTATAGTATTTGTAGCTCATCAT[T/G]TTCTTCAATC TTTTTGATGTGGAACAATTACAATTATGATATTTTTGCCA
1_34864233	Charleston Gray	Chr1	34864233	CAATCCAAAACAGGCTTATTTTCAGCTTATATATTGTCAAAGAAAATTGA[G/A]CAACTGCG TCTTTCCATTAGTATTTCCATCAATAATTTCACATTTTAATA
1_35033231	Charleston Gray	Chr1	35033231	TTTTAAGTCTATCGGCCATGAGTTTAACCCATCCAACCCCATCTGTCTCA[C/A]GTAGGTTAT GGAGCTGCATAGAAAGAAAGGGTTAGGTATGCAGGCCAACC
1_35388188	Charleston Gray	Chr1	35388188	GGAAAGTGGGTTCAAACTTTCATCTAACAAACTCCAACTGCTACCTTGAT[G/A]TTAGCCCTT TCTTTCTTGTTGTTTGTTTTACCATATACTTCTTTTCACCT
1_35599290	Charleston Gray	Chr1	35599290	TTGATAAAATCATTCCTGAAGCTCAAGGAATTGGCAGTAGAAGCAGACCC[G/A]GAATCAGA AACAGCCACCAACAGAGAAGAAACTCGAGCACAACCCTGGAA
1_35833438	Charleston Gray	Chr1	35833438	CATATATGAGTTAGGCTATTATTGGAAATGTATTTTTCAAAAGTTGTGTT[T/C]AGTAAAGTA TTTGTGTATCGTCTATTATTGGTTTCTTAGATTTGATTTTG
1_35956707	Charleston Gray	Chr1	35956707	AAAAAGCTTTTTTCTTTTTCACAGACGGAAAAGGAAATTTCAACTGAAAC[G/T]ACGCTGTCT CGGCTCCCATTAATTTAAAACGATGGAACCCCATTCCAGTT

Supplementary Table 5.4. List of annotation genes.

Gene_ID	Annotation	Uniprot
ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	https://www.uniprot.org/uniprot/D7KUA7
ClCG01G019820	Heavy metal transport/detoxification superfamily protein LENGTH=126	https://www.uniprot.org/uniprot/A0A1P8B9B2
ClCG01G019830	UPF0261 protein	https://www.uniprot.org/uniprot/Q8FUQ2
ClCG01G019840	Unknown protein	
ClCG01G019850	Cytochrome P450, putative	https://www.uniprot.org/uniprot/Q8LB43
ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	https://www.uniprot.org/uniprot/F4J8L2
ClCG01G019870	ATP-dependent zinc metalloprotease FtsH	https://www.uniprot.org/uniprot/P0AAI3
ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	https://www.uniprot.org/uniprot/Q0WVF7
ClCG01G019890	Thioredoxin domain-containing protein, putative	https://www.uniprot.org/uniprot/B7P4N2
ClCG01G019900	Phosphatidylinositol transfer-like protein II	https://www.uniprot.org/uniprot/F4I1W0
ClCG01G019910	basic helix-loop-helix (bHLH) DNA-binding superfamily protein LENGTH=453	https://www.uniprot.org/uniprot/A0A1P8B885
ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	https://www.uniprot.org/uniprot/A0A1P8AYT6
ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	
ClCG01G019950	Peroxidase 49 family protein	https://www.uniprot.org/uniprot/O23237
ClCG01G019960	50S ribosomal protein L7/L12, putative	https://www.uniprot.org/uniprot/C5LDW4
ClCG01G019970	Mitochondrial carrier protein, expressed	https://www.uniprot.org/uniprot/Q6AVS9
ClCG01G019980	Cytochrome P450	https://www.uniprot.org/uniprot/Q8VZC2
ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	https://www.uniprot.org/uniprot/Q9FK02
ClCG01G020000	Beta-galactosidase	https://www.uniprot.org/uniprot/Q9SCV1
ClCG01G020010	Thioredoxin	https://www.uniprot.org/uniprot/Q39241
ClCG01G020020	Purple acid phosphatase, putative	https://www.uniprot.org/uniprot/Q9LMG7
ClCG01G020030	**FG-GAP repeat-containing family protein**	https://www.uniprot.org/uniprot/F4J381
ClCG01G020040	**Lateral root primordium family protein**	https://www.uniprot.org/uniprot/Q94CK9
ClCG01G020050	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein LENGTH=895	https://www.uniprot.org/uniprot/Q0WL44
ClCG01G020060	**Glutaredoxin family protein (GRXs)**	https://www.uniprot.org/uniprot/B3H604
ClCG01G020070	Leucine-rich receptor-like protein kinase family protein LENGTH=1072	https://www.uniprot.org/uniprot/Q8VYG7
ClCG01G020080	**GATA transcription factor**	https://www.uniprot.org/uniprot/Q8LAU9
ClCG01G020090	Glycine-rich protein	https://www.uniprot.org/uniprot/Q94FQ3
ClCG01G020100	Zinc finger CCCH domain-containing protein	https://www.uniprot.org/uniprot/Q8GX84
ClCG01G020110	Cytochrome P450	https://www.uniprot.org/uniprot/Q8VZC2
ClCG01G020120	UPF0183 protein	https://www.uniprot.org/uniprot/Q9SD33
ClCG01G020130	Zinc finger family protein	https://www.uniprot.org/uniprot/Q9LZD5
ClCG01G020140	Leucine-rich repeat receptor-like protein kinase family protein LENGTH=1141	https://www.uniprot.org/uniprot/Q9SSL9
ClCG01G020150	NAC domain protein	https://www.uniprot.org/uniprot/Q0WV96
ClCG01G020160	Histone h1/h5, putative	https://www.uniprot.org/uniprot/P94109
ClCG01G020170	Zinc finger CCCH domain-containing protein	https://www.uniprot.org/uniprot/Q9SY74
ClCG01G020180	Peptide methionine sulfoxide reductase msrA	https://www.uniprot.org/uniprot/Q9FKF7
ClCG01G020190	50S ribosomal protein L2	https://www.uniprot.org/uniprot/P56791
ClCG01G020200	YEATS domain-containing protein, putative	https://www.uniprot.org/uniprot/Q9LK65
ClCG01G020210	Pentatricopeptide repeat-containing protein, putative	https://www.uniprot.org/uniprot/P0C7Q7
ClCG01G020220	Unknown protein	
ClCG01G020230	Zinc finger (Ran-binding) family protein	https://www.uniprot.org/uniprot/Q9C7M2
ClCG01G020240	WGS project CAID00000000 data, contig chromosome 18	
ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	https://www.uniprot.org/uniprot/Q9C6F0
ClCG01G020260	**WUSCHEL related homeobox 1 LENGTH=350**	https://www.uniprot.org/uniprot/Q6X7K0-1
ClCG01G020270	WD repeat protein-like	https://www.uniprot.org/uniprot/Q9FGX4
ClCG01G020280	Unknown protein	

ClCG01G020290	WEB family protein	https://www.uniprot.org/uniprot/Q5XVC7
ClCG01G020300	Inositol or phosphatidylinositol kinase, putative	https://www.uniprot.org/uniprot/Q9XID0
ClCG01G020310	Calcium-dependent protein kinase	https://www.uniprot.org/uniprot/Q06850
ClCG01G020320	Basic helix-loop-helix transcription factor	https://www.uniprot.org/uniprot/F4IQH8
ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	https://www.uniprot.org/uniprot/Q5XF07
ClCG01G020340	Chaperone protein dnaJ	https://www.uniprot.org/uniprot/Q39079
ClCG01G020350	Armadillo repeat protein 1	https://www.uniprot.org/uniprot/Q9SW41
ClCG01G020360	Cysteine-rich receptor-like protein kinase	https://www.uniprot.org/uniprot/Q9M0X5
ClCG01G020370	Thaumatin	https://www.uniprot.org/uniprot/A0A1P8B554
ClCG01G020380	Oligopeptide transporter 7	https://www.uniprot.org/uniprot/O82485
ClCG01G020390	Proteasome 54kD subunit	https://www.uniprot.org/uniprot/Q42134
ClCG01G020400	Lysosomal alpha-mannosidase	
ClCG01G020410	Proteasome subunit alpha type	https://www.uniprot.org/uniprot/O81149
ClCG01G020420	Unknown protein	
ClCG01G020430	**Male sterility MS5 family protein**	https://www.uniprot.org/uniprot/Q9SUC3
ClCG01G020440	Ascorbate peroxidase	https://www.uniprot.org/uniprot/Q1PER6
ClCG01G020450	Proline-rich family protein	https://www.uniprot.org/uniprot/Q9LYH4
ClCG01G020460	Cell cycle checkpoint protein RAD17	https://www.uniprot.org/uniprot/Q9MBA3
ClCG01G020470	Heavy metal transport/detoxification superfamily protein LENGTH=147	https://www.uniprot.org/uniprot/A0A178UU72
ClCG01G020480	GDSL esterase/lipase	https://www.uniprot.org/uniprot/Q9FJ25
ClCG01G020490	Mid1-complementing activity 2	https://www.uniprot.org/uniprot/Q3EBY6
ClCG01G020500	Pentatricopeptide repeat-containing family protein	https://www.uniprot.org/uniprot/Q8RWS8
ClCG01G020510	Protein preY, mitochondrial, putative	https://www.uniprot.org/uniprot/Q9LSQ2
ClCG01G020520	La domain-containing family protein	https://www.uniprot.org/uniprot/Q0WV96
ClCG01G020530	Methyltransferase-like protein 23	https://www.uniprot.org/uniprot/Q9SIZ3
ClCG01G020540	Peroxidase	https://www.uniprot.org/uniprot/Q96512
ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	https://www.uniprot.org/uniprot/F4J3C0
ClCG01G020560	Unknown protein	
ClCG01G020570	Early-responsive to dehydration stress-related protein	https://www.uniprot.org/uniprot/O48832
ClCG01G020580	Phosphatase 2C family protein	https://www.uniprot.org/uniprot/A0A1R7T3I5
ClCG01G020590	Protein kinase 2A	https://www.uniprot.org/uniprot/Q08466
ClCG01G020600	Aconitate hydratase 1	https://www.uniprot.org/uniprot/Q42560
ClCG01G020610	Prolyl 4-hydroxylase alpha subunit, putative	https://www.uniprot.org/uniprot/Q8GXT7
ClCG01G020620	Unknown protein	
ClCG01G020630	RNA binding protein, putative	https://www.uniprot.org/uniprot/O22791
ClCG01G020640	ankyrin repeat protein LENGTH=435	https://www.uniprot.org/uniprot/Q05753
ClCG01G020650	Ammonium transporter, putative	https://www.uniprot.org/uniprot/Q9LK16
ClCG01G020660	Senescence-associated protein DIN1, putative	https://www.uniprot.org/uniprot/P27626
ClCG01G020670	NAD(P)H dehydrogenase (Quinone)s	https://www.uniprot.org/uniprot/Q9LSQ5
ClCG01G020680	UPF0548 protein	https://www.uniprot.org/uniprot/Q8GXB1
ClCG01G020690	Serine-aspartate repeat-containing protein F	https://www.uniprot.org/uniprot/O86489
ClCG01G020700	**F-box family protein**	https://www.uniprot.org/uniprot/Q9LPL4
ClCG01G020710	S-acyltransferase	https://www.uniprot.org/uniprot/Q3EC11
ClCG01G020720	Regulator of Vps4 activity in the MVB pathway protein	https://www.uniprot.org/uniprot/F4JNS8
ClCG01G020730	Unknown protein	
ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	https://www.uniprot.org/uniprot/O81803
ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	https://www.uniprot.org/uniprot/O81803
ClCG01G020760	Pentatricopeptide repeat-containing protein, putative	https://www.uniprot.org/uniprot/Q9FNG8
ClCG01G020770	**Chitinase**	https://www.uniprot.org/uniprot/Q9MA41
ClCG01G020780	**Chitinase**	https://www.uniprot.org/uniprot/Q9MA41
ClCG01G020790	**Chitinase Ib**	https://www.uniprot.org/uniprot/Q9MA41
ClCG01G020800	**Chitinase**	https://www.uniprot.org/uniprot/Q9MA41
ClCG01G020810	2-succinylbenzoate-CoA ligase	https://www.uniprot.org/uniprot/Q8VYJ1
ClCG01G020820	3-isopropylmalate dehydrogenase	https://www.uniprot.org/uniprot/Q9FMT1

List of SNPs

Chromosome	Gene_ID	Annotation	Position	Reference	Alternative	ANNO_REGION	Mutation_impact
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33997548	T	G	upstream	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33997704	A	T	upstream	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33997729	T	C	upstream	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33997973	A	G	upstream	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33998215	C	T	exonic	synonymous SNV
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33999696	A	G	intronic	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	34000809	A	G	downstream	
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	34001751	G	A	downstream	
CG_Chr01	ClCG01G019820	Heavy metal transport/detoxification superfamily protein LENGTH=126	34008935	C	T	upstream	
CG_Chr01	ClCG01G019820	Heavy metal transport/detoxification superfamily protein LENGTH=126	34009324	G	T	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34020688	A	G	upstream	
CG_Chr01	ClCG01G019830	UPF0261 protein	34021621	T	C	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34022383	A	G	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34023469	A	G	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34023825	A	G	exonic	synonymous SNV
CG_Chr01	ClCG01G019830	UPF0261 protein	34024263	A	G	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34025276	G	C	intronic	
CG_Chr01	ClCG01G019830	UPF0261 protein	34025869	G	C	downstream	
CG_Chr01	ClCG01G019830	UPF0261 protein	34026328	G	C	downstream	
CG_Chr01	ClCG01G019850	Cytochrome P450, putative	34027093	T	A	downstream	
CG_Chr01	ClCG01G019850	Cytochrome P450, putative	34027382	G	A	downstream	
CG_Chr01	ClCG01G019850	Cytochrome P450, putative	34027726	A	T	downstream	
CG_Chr01	ClCG01G019850	Cytochrome P450, putative	34028012	C	G	downstream	
CG_Chr01	ClCG01G019850	Cytochrome P450, putative	34028849	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34038924	T	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34039144	G	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34039681	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34040792	C	G	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34041258	C	T	exonic	synonymous SNV
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34043237	T	G	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34043288	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34044795	G	T	intronic	

CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34044933	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34046881	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34046927	C	G	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34046942	G	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047087	T	C	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047088	G	T	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047090	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047335	G	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047409	C	T	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047495	G	T	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047549	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34047562	C	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34048095	T	C	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34048180	T	C	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34048914	A	T	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34050096	A	T	exonic	synonymous SNV
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34050564	C	T	downstream	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34050943	G	A	downstream	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34050988	T	C	downstream	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34051045	G	A	downstream	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34051230	C	T	downstream	
CG_Chr01	ClCG01G019870	ATP-dependent zinc metalloprotease FtsH	34051895	T	G	downstream	
CG_Chr01	ClCG01G019870	ATP-dependent zinc metalloprotease FtsH	34053699	T	A	upstream	
CG_Chr01	ClCG01G019870	ATP-dependent zinc metalloprotease FtsH	34053933	G	A	upstream	
CG_Chr01	ClCG01G019870	ATP-dependent zinc metalloprotease FtsH	34054237	G	A	upstream	
CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34062444	T	C	exonic	nonsynonymous SNV

CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34063036	C	A	downstream	
CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34063039	T	A	downstream	
CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34063260	C	T	downstream	
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34064001	A	G	upstream	
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34064363	G	A	upstream	
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34065054	T	C	intronic	
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34065863	A	G	exonic	synonymous SNV
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34066232	A	C	UTR3	
CG_Chr01	ClCG01G019890	Thioredoxin domain-containing protein, putative	34067157	G	A	downstream	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34069607	C	G	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34070083	C	G	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34070845	G	C	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34072596	G	T	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34072956	G	A	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34074080	A	C	intronic	
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34074329	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G019900	Phosphatidylinositol transfer-like protein II	34075010	G	A	downstream	
CG_Chr01	ClCG01G019910	basic helix-loop-helix (bHLH) DNA-binding superfamily protein LENGTH=453	34077196	A	G	intronic	
CG_Chr01	ClCG01G019910	basic helix-loop-helix (bHLH) DNA-binding superfamily protein LENGTH=453	34078774	G	A	upstream	
CG_Chr01	ClCG01G019910	basic helix-loop-helix (bHLH) DNA-binding superfamily protein LENGTH=453	34078984	G	T	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34088775	T	C	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089009	G	T	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089020	A	G	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089134	T	C	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089170	T	A	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089171	T	G	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089173	T	A	upstream	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34089951	A	T	intronic	
CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34090798	T	C	downstream	

CG_Chr01	ClCG01G019930	Basic-leucine zipper (bZIP) transcription factor family protein LENGTH=262	34091379	G	A	downstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34095816	T	G	upstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34095899	G	A	upstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34096285	T	C	upstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34096327	C	A	upstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34096925	G	T	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34100153	G	A	intronic	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34102519	G	A	intronic	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34102617	A	G	intronic	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34103085	T	C	intronic	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34105643	G	A	exonic	synonymous SNV
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34108515	C	T	downstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34108921	C	T	downstream	
CG_Chr01	ClCG01G019950	Peroxidase 49 family protein	34109695	A	T	upstream	
CG_Chr01	ClCG01G019960	50S ribosomal protein L7/L12, putative	34112757	A	C	upstream	
CG_Chr01	ClCG01G019960	50S ribosomal protein L7/L12, putative	34112890	T	C	upstream	
CG_Chr01	ClCG01G019960	50S ribosomal protein L7/L12, putative	34115681	A	C	intronic	
CG_Chr01	ClCG01G019970	Mitochondrial carrier protein, expressed	34116783	G	A	upstream	
CG_Chr01	ClCG01G019970	Mitochondrial carrier protein, expressed	34116806	G	A	upstream	
CG_Chr01	ClCG01G019970	Mitochondrial carrier protein, expressed	34119331	G	A	intronic	
CG_Chr01	ClCG01G019970	Mitochondrial carrier protein, expressed	34120090	A	G	intronic	
CG_Chr01	ClCG01G019970	Mitochondrial carrier protein, expressed	34120321	A	G	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34126517	A	G	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34126518	A	G	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34127064	T	A	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34127706	G	A	exonic	synonymous SNV
CG_Chr01	ClCG01G019980	Cytochrome P450	34128490	T	A	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34129641	T	G	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34129923	C	T	intronic	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34132813	A	G	downstream	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34134393	A	G	intronic	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34134928	C	T	upstream	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34135022	C	T	upstream	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34135783	T	C	upstream	
CG_Chr01	ClCG01G019990	Ribosomal RNA small subunit methyltransferase A	34135904	T	C	upstream	
CG_Chr01	ClCG01G020000	Beta-galactosidase	34144026	A	G	upstream	
CG_Chr01	ClCG01G020000	Beta-galactosidase	34144210	G	A	upstream	
CG_Chr01	ClCG01G020000	Beta-galactosidase	34144235	A	G	upstream	
CG_Chr01	ClCG01G020000	Beta-galactosidase	34149152	G	T	intronic	

CG_Chr01	ClCG01G020000	Beta-galactosidase	34150784	A	G	exonic	synonymous SNV
CG_Chr01	ClCG01G020000	Beta-galactosidase	34151552	G	A	downstream	
CG_Chr01	ClCG01G020010	Thioredoxin	34153298	T	G	upstream	
CG_Chr01	ClCG01G020010	Thioredoxin	34153393	T	C	upstream	
CG_Chr01	ClCG01G020010	Thioredoxin	34153648	T	C	upstream	
CG_Chr01	ClCG01G020010	Thioredoxin	34153791	A	T	upstream	
CG_Chr01	ClCG01G020040	Lateral root primordium family protein	34181560	G	T	upstream	
CG_Chr01	ClCG01G020050	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein LENGTH=895	34198749	G	A	exonic	synonymous SNV
CG_Chr01	ClCG01G020050	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein LENGTH=895	34199059	A	C	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020050	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein LENGTH=895	34199744	T	C	intronic	
CG_Chr01	ClCG01G020060	Glutaredoxin family protein	34205600	G	A	upstream	
CG_Chr01	ClCG01G020070	Leucine-rich receptor-like protein kinase family protein LENGTH=1072	34207317	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020080	GATA transcription factor	34212852	C	T	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213142	T	C	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213180	C	T	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213469	A	G	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213487	G	A,*	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213487	G	A,*	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34215144	T	A	downstream	
CG_Chr01	ClCG01G020090	Glycine-rich protein	34216992	T	G	downstream	
CG_Chr01	ClCG01G020090	Glycine-rich protein	34217100	C	T	downstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34243566	A	C	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34243610	T	A	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34243798	T	C	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34244035	C	T	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34244094	A	T	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34244258	T	C	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020110	Cytochrome P450	34244442	A	G	exonic	synonymous SNV
CG_Chr01	ClCG01G020110	Cytochrome P450	34244619	T	C	exonic	synonymous SNV
CG_Chr01	ClCG01G020110	Cytochrome P450	34244775	A	G	exonic	synonymous SNV
CG_Chr01	ClCG01G020110	Cytochrome P450	34245390	C	T	intronic	
CG_Chr01	ClCG01G020110	Cytochrome P450	34245657	T	G	intronic	
CG_Chr01	ClCG01G020110	Cytochrome P450	34246176	G	A	exonic	nonsynonymous SNV

CG_Chr01	ClCG01G020110	Cytochrome P450	34247245	G	T	downstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251038	G	T	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251074	G	A	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251200	T	C	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251381	A	G	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251517	G	T	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251805	C	T	upstream	
CG_Chr01	ClCG01G020120	UPF0183 protein	34251911	A	G	UTR5	
CG_Chr01	ClCG01G020120	UPF0183 protein	34252140	T	C	intronic	
CG_Chr01	ClCG01G020120	UPF0183 protein	34252807	T	A	intronic	
CG_Chr01	ClCG01G020120	UPF0183 protein	34253044	A	G	intronic	
CG_Chr01	ClCG01G020120	UPF0183 protein	34255485	C	T	exonic	synonymous SNV
CG_Chr01	ClCG01G020120	UPF0183 protein	34257192	T	C	downstream	
CG_Chr01	ClCG01G020150	NAC domain protein	34296891	C	T	intronic	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34303836	T	G	downstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34304329	T	C	downstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306470	A	T	upstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306473	A	T,C	upstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306473	A	T,C	upstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306474	A	C	upstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306477	A	C	upstream	
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306634	A	G	upstream	
CG_Chr01	ClCG01G020170	Zinc finger CCCH domain-containing protein	34315156	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020170	Zinc finger CCCH domain-containing protein	34315565	T	G	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020170	Zinc finger CCCH domain-containing protein	34316906	G	C	downstream	
CG_Chr01	ClCG01G020170	Zinc finger CCCH domain-containing protein	34317088	G	T	downstream	
CG_Chr01	ClCG01G020180	Peptide methionine sulfoxide reductase msrA	34318731	C	A	downstream	
CG_Chr01	ClCG01G020180	Peptide methionine sulfoxide reductase msrA	34318739	A	T	downstream	
CG_Chr01	ClCG01G020180	Peptide methionine sulfoxide reductase msrA	34319213	G	A	UTR3	
CG_Chr01	ClCG01G020190	50S ribosomal protein L2	34323252	T	C	intronic	
CG_Chr01	ClCG01G020190	50S ribosomal protein L2	34323895	G	A	exonic	synonymous SNV
CG_Chr01	ClCG01G020200	YEATS domain-containing protein, putative	34336042	A	C	upstream	
CG_Chr01	ClCG01G020200	YEATS domain-containing protein, putative	34336924	C	T	upstream	
CG_Chr01	ClCG01G020220	Unknown protein	34341501	G	C	upstream	
CG_Chr01	ClCG01G020230	Zinc finger (Ran-binding) family protein	34343814	A	C	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020230	Zinc finger (Ran-binding) family protein	34346345	C	T	downstream	
CG_Chr01	ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	34348650	C	G	intronic	
CG_Chr01	ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	34350147	C	T	intronic	

CG_Chr01	ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	34350896	A	T	intronic	
CG_Chr01	ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	34351564	T	C	frame	synonymous SNV
CG_Chr01	ClCG01G020250	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein	34353993	T	A	downstream	
CG_Chr01	ClCG01G020260	WUSCHEL related homeobox 1 LENGTH=350	34366151	G	A	upstream	
CG_Chr01	ClCG01G020290	WEB family protein	34391532	T	A	intronic	
CG_Chr01	ClCG01G020300	Inositol or phosphatidylinositol kinase, putative	34401811	G	T	intronic	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34443351	G	A	upstream	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34443482	T	A	upstream	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34443599	G	C	upstream	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34444095	A	T	upstream	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34445200	T	G	intronic	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34449305	A	G	UTR3	
CG_Chr01	ClCG01G020310	Calcium-dependent protein kinase	34449567	A	T	downstream	
CG_Chr01	ClCG01G020320	Basic helix-loop-helix transcription factor	34452599	T	A	exonic	synonymous SNV
CG_Chr01	ClCG01G020320	Basic helix-loop-helix transcription factor	34453060	T	C	intronic	
CG_Chr01	ClCG01G020320	Basic helix-loop-helix transcription factor	34454459	A	G	upstream	
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34455067	C	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34457672	T	A	intronic	
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34461654	T	G	UTR3	
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34461973	T	C	downstream	
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34462529	A	G	downstream	
CG_Chr01	ClCG01G020330	DNA-(Apurinic or apyrimidinic site) lyase	34462734	T	C	downstream	
CG_Chr01	ClCG01G020350	Armadillo repeat protein 1	34489157	C	T	exonic	synonymous SNV
CG_Chr01	ClCG01G020350	Armadillo repeat protein 1	34489873	A	C	downstream	
CG_Chr01	ClCG01G020350	Armadillo repeat protein 1	34490269	A	G	downstream	
CG_Chr01	ClCG01G020360	Cysteine-rich receptor-like protein kinase	34497378	A	G	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34502881	A	G	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34503572	T	C	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34506740	T	A	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34507187	T	A	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34508507	A	T	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34509184	G	A	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34509280	G	A	intronic	
CG_Chr01	ClCG01G020370	Thaumatin	34510439	A	C	intronic	
CG_Chr01	ClCG01G020380	Oligopeptide transporter 7	34519191	T	C	intronic	
CG_Chr01	ClCG01G020380	Oligopeptide transporter 7	34520193	G	A	exonic	synonymous SNV
CG_Chr01	ClCG01G020380	Oligopeptide transporter 7	34524772	T	G	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34536340	A	T	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34536629	G	A	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34536682	T	G	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34537795	A	G	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34537929	A	G	intronic	
CG_Chr01	ClCG01G020390	Proteasome 54kD subunit	34539088	C	T	intronic	
CG_Chr01	ClCG01G020430	Male sterility MS5 family protein	34571640	A	T	upstream	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34612232	T	C	intronic	

CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34612301	A	T	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34612459	T	C	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34612673	G	A	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34612987	C	T	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34613342	A	G	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34613354	G	A	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34614083	T	C	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34614201	G	A	intronic	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34614567	G	A	downstream	
CG_Chr01	ClCG01G020480	GDSL esterase/lipase	34614963	G	T	downstream	
CG_Chr01	ClCG01G020490	Mid1-complementing activity 2	34621271	A	C	exonic	synonymous SNV
CG_Chr01	ClCG01G020490	Mid1-complementing activity 2	34624884	T	C	intronic	
CG_Chr01	ClCG01G020490	Mid1-complementing activity 2	34625051	C	T	intronic	
CG_Chr01	ClCG01G020500	Pentatricopeptide repeat-containing family protein	34629344	T	C	upstream	
CG_Chr01	ClCG01G020510	Protein preY, mitochondrial, putative	34629496	C	A	upstream	
CG_Chr01	ClCG01G020510	Protein preY, mitochondrial, putative	34629826	A	T	upstream	
CG_Chr01	ClCG01G020520	La domain-containing family protein	34634903	C	T	downstream	
CG_Chr01	ClCG01G020520	La domain-containing family protein	34637443	A	G	intronic	
CG_Chr01	ClCG01G020520	La domain-containing family protein	34639273	G	A	intronic	
CG_Chr01	ClCG01G020530	Methyltransferase-like protein 23	34661628	T	C	intronic	
CG_Chr01	ClCG01G020540	Peroxidase	34664819	A	T	upstream	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34670435	T	C	downstream	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34674779	C	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34675084	T	C	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34675409	G	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34675616	A	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34675638	T	C	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34679221	T	C	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34679508	G	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34679566	A	G	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34679631	A	G	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34679685	T	G	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34680676	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34680762	C	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681045	T	C	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681168	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681597	C	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681646	C	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681648	C	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34682124	C	T	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34682878	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34684083	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34684138	C	T	intronic	

CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34685689	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34685806	G	A	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34688135	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34688775	A	G	intronic	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689594	C	T	upstream	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689752	T	A	upstream	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689913	C	T	upstream	
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34690366	G	T	upstream	
CG_Chr01	ClCG01G020560	Unknown protein	34690888	C	A	intronic	
CG_Chr01	ClCG01G020560	Unknown protein	34691232	T	C	upstream	
CG_Chr01	ClCG01G020560	Unknown protein	34692067	A	G	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34695214	T	C	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696375	G	A	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696423	A	T	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696439	C	T	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696603	C	T	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696616	T	C	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696739	C	T	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696750	C	T	upstream	
CG_Chr01	ClCG01G020570	Early-responsive to dehydration stress-related protein	34696984	C	T	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700601	T	C	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700718	C	A	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700765	G	A	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700785	T	C	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700791	A	G	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700818	A	G	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700826	C	T	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700964	C	T	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34701014	T	C	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34701024	A	T	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34701081	G	A	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34701219	T	A	upstream	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34701625	T	C	UTR5	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34702900	C	T	exonic	synonymous SNV
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34704309	C	G	UTR3	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34704375	A	G	UTR3	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34705391	C	A	UTR3	
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34705592	A	C	UTR3	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34707616	G	A	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34707786	A	T	intronic	

CG_Chr01	ClCG01G020590	Protein kinase 2A	34708298	G	A	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34708320	A	G	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34708329	G	C	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34708474	A	G	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34708605	G	T	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34708946	T	C	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34709059	A	T	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34709088	C	T	intronic	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34710513	C	T	upstream	
CG_Chr01	ClCG01G020590	Protein kinase 2A	34710524	T	C	upstream	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34721100	A	T	intronic	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34723093	C	G	intronic	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724024	G	C	UTR3	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724322	A	T	downstream	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724552	A	G	downstream	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724622	T	C	downstream	
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724729	A	T	downstream	
CG_Chr01	ClCG01G020610	Prolyl 4-hydroxylase alpha subunit, putative	34726935	G	C	intronic	
CG_Chr01	ClCG01G020610	Prolyl 4-hydroxylase alpha subunit, putative	34729008	T	C	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34737355	T	C	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34737411	T	C	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34737659	C	A	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34737701	C	T	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34737856	T	A	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34738117	A	T	upstream	
CG_Chr01	ClCG01G020620	Unknown protein	34738656	G	A	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34738694	A	C	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34738937	T	G	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020620	Unknown protein	34741471	A	G	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34741493	A	C	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34741695	C	G	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34741744	T	A	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34742203	C	T	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34743113	A	T	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34743360	C	T	UTR3	
CG_Chr01	ClCG01G020620	Unknown protein	34743411	A	G	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34743582	T	C	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34743673	C	T	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34744798	C	T	intronic	
CG_Chr01	ClCG01G020620	Unknown protein	34745475	C	G	intronic	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34746346	C	T	upstream	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34746348	A	C	upstream	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34746765	G	C	upstream	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34746773	G	A	upstream	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34749056	G	T	intronic	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34749079	T	C	intronic	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34750728	T	C	downstream	
CG_Chr01	ClCG01G020640	ankyrin repeat protein LENGTH=435	34753281	T	A	intronic	
CG_Chr01	ClCG01G020640	ankyrin repeat protein LENGTH=435	34753819	A	G	intronic	
CG_Chr01	ClCG01G020650	Ammonium transporter, putative	34755618	A	G	intronic	
CG_Chr01	ClCG01G020650	Ammonium transporter, putative	34756742	A	G	upstream	

CG_Chr01	ClCG01G020650	Ammonium transporter, putative	34756925	T	A	upstrea m	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34761100	T	A	downstr eam	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34761108	A	G	downstr eam	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34761568	C	A	downstr eam	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34761629	C	A	downstr eam	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34762125	G	C	intronic	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34763811	T	C	upstrea m	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34763823	T	A	upstrea m	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34763854	C	T	upstrea m	
CG_Chr01	ClCG01G020720	Regulator of Vps4 activity in the MVB pathway protein	34793170	A	T	upstrea m	
CG_Chr01	ClCG01G020720	Regulator of Vps4 activity in the MVB pathway protein	34793173	A	C	upstrea m	
CG_Chr01	ClCG01G020720	Regulator of Vps4 activity in the MVB pathway protein	34793482	T	C	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34796499	A	G	downstr eam	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34796611	C	G	downstr eam	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34796775	A	C	downstr eam	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34797527	G	A	exonic	nonsyno nymous SNV
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798182	A	G	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798453	C	T	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798495	T	C	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798496	T	C	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798741	T	C	upstrea m	
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34799104	A	C	upstrea m	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34801895	T	G	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34802364	T	C	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34802382	T	A	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34802787	G	A	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34803408	A	G	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34804343	C	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34804846	G	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34804979	T	A	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34804993	A	G	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34805012	T	C	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34805402	A	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34806711	C	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34806927	A	C	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34806999	A	G	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34807047	T	G	intronic	

CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34807146	T	G	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34807233	A	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34808826	G	C	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34809049	G	A	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34809935	T	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34810640	A	T	upstream	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34811267	C	T	upstream	
CG_Chr01	ClCG01G020760	Pentatricopeptide repeat-containing protein, putative	34815284	C	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34819111	T	C	UTR3	
CG_Chr01	ClCG01G020770	Chitinase	34820229	C	T	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34821269	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822848	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822867	A	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822897	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822906	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822959	G	T	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822983	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34822989	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823026	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823043	C	T	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823049	A	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823094	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823097	C	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823137	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823146	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823177	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823180	T	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823228	T	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823237	G	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34823243	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34824246	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34825179	C	T	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34825285	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34828282	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34831112	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34831289	T	C	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34831487	G	A	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34832694	A	G	intronic	
CG_Chr01	ClCG01G020770	Chitinase	34832723	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34835065	A	T	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34835566	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34835607	G	A	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34835856	A	G	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34836181	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34836715	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34837031	T	A	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34837072	G	T	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34837228	A	G	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34837325	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34837408	T	A	intronic	
CG_Chr01	ClCG01G020790	Chitinase Ib	34838102	G	A	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020780	Chitinase	34838206	G	T	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34838846	T	G	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34840536	T	C	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34840676	A	G	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34845426	T	C	upstream	
CG_Chr01	ClCG01G020800	Chitinase	34846850	C	T	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34847481	G	A	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34847641	T	C	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34847753	A	G	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34847896	T	C	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34847900	T	C	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34849585	T	G	intronic	

CG_Chr01	ClCG01G020800	Chitinase	34849974	A	G	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34850194	A	T	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34850401	A	T	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34850462	A	G	intronic	
CG_Chr01	ClCG01G020810	2-succinylbenzoate-CoA ligase	34855158	A	C	exonic	nonsynonymous SNV
CG_Chr01	ClCG01G020810	2-succinylbenzoate-CoA ligase	34856721	T	A	intronic	
CG_Chr01	ClCG01G020810	2-succinylbenzoate-CoA ligase	34856745	G	A	intronic	
CG_Chr01	ClCG01G020810	2-succinylbenzoate-CoA ligase	34857581	G	T	upstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34858558	A	G	downstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34858594	G	A	downstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34858638	T	C	downstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34860101	T	G	intronic	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34862860	C	G	intronic	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34863808	G	T	upstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34864098	C	T	upstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34864106	C	A	upstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34864122	A	C	upstream	
CG_Chr01	ClCG01G020820	3-isopropylmalate dehydrogenase	34864233	G	A	upstream	

ANNO_REGION: Corresponding annotation data for each genomic region. The Anno_region corresponding to intronic data was removed for a better understanding.

List of INDELs

Chromosome	Gene_ID	Annotation	Position	Reference	Alternative	ANNO_REGION	Mutation_ impact
CG_Chr01	ClCG01G019810	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	33998660	T	TAAA	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34042685	A	ATATATATATATATATATATAT	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34044677	ATTTTTT	A	intronic	
CG_Chr01	ClCG01G019860	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=451	34048366	G	GAATAATAAATACATATATGGAGTAGAAGGATTCAAATCCATATTTAATATAATAAC	intronic	
CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34063012	C	CTATATATA,CTATATATATATA	downstream	
CG_Chr01	ClCG01G019880	P-loop containing nucleoside triphosphate hydrolases superfamily protein LENGTH=495	34063012	C	CTATATATA,CTATATATATATA	downstream	
CG_Chr01	ClCG01G019940	Genomic DNA, chromosome 3, P1 clone: MDJ14	34100128	TA	T	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34126483	C	CT	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34126509	GA	G	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34126631	CT	C	intronic	
CG_Chr01	ClCG01G019980	Cytochrome P450	34130860	TTCTCTC	T	intronic	
CG_Chr01	ClCG01G020000	Beta-galactosidase	34149354	GTTT	G	intronic	
CG_Chr01	ClCG01G020010	Thioredoxin	34153557	A	AT	upstream	
CG_Chr01	ClCG01G020010	Thioredoxin	34153640	C	CT	upstream	
CG_Chr01	ClCG01G020060	Glutaredoxin family protein	34202305	G	GA	UTR3	
CG_Chr01	ClCG01G020060	Glutaredoxin family protein	34206001	C	CATTCCGAACACAACTGACTTTTGGCGATGCACCGATTTTGCTTTCAAAACAGTAATTTCCTA,CATTCCGAACACAATTGACTTTTGGCGATGCACCGATTTTGCTTTCAAAGCAGTAATTTCCAA	upstream	
CG_Chr01	ClCG01G020060	Glutaredoxin family protein	34206001	C	CATTCCGAACACAACTGACTTTTGGCGATGCACCGATTTTGCTTTCAAAACAGTAATTTCCTA,CATTCCGAACACAATTGACTTTTGGCGATGCACCGATTTTGCTTTCAAAGCAGTAATTTCCAA	upstream	
CG_Chr01	ClCG01G020080	GATA transcription factor	34213482	AAAAAG	A	upstream	
CG_Chr01	ClCG01G020110	Cytochrome P450	34245475	AT	A	intronic	
CG_Chr01	ClCG01G020130	Zinc finger family protein	34259344	C	CTTT	upstream	

CG_Chr01	ClCG01G020150	NAC domain protein	34294590	TC	T	upstream
CG_Chr01	ClCG01G020150	NAC domain protein	34298114	TA	T	downstream
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34303793	TA	T	downstream
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306465	TAAA	T	upstream
CG_Chr01	ClCG01G020160	Histone h1/h5, putative	34306469	A	AT	upstream
CG_Chr01	ClCG01G020180	Peptide methionine sulfoxide reductase msrA	34318911	C	CT	downstream
CG_Chr01	ClCG01G020200	YEATS domain-containing protein, putative	34336414	TTATATATATA	T	upstream
CG_Chr01	ClCG01G020270	WD repeat protein-like	34379469	CAT	C	intronic
CG_Chr01	ClCG01G020270	WD repeat protein-like	34381390	C	CG	upstream
CG_Chr01	ClCG01G020320	Basic helix-loop-helix transcription factor	34453850	A	AT	upstream
CG_Chr01	ClCG01G020350	Armadillo repeat protein 1	34487592	T	TTTCTTC	upstream
CG_Chr01	ClCG01G020370	Thaumatin	34506252	C	CT	intronic
CG_Chr01	ClCG01G020370	Thaumatin	34506727	ATTT	A	intronic
CG_Chr01	ClCG01G020370	Thaumatin	34506730	T	TAA	intronic
CG_Chr01	ClCG01G020370	Thaumatin	34509016	A	AATATATATAT	intronic
CG_Chr01	ClCG01G020430	Male sterility MS5 family protein	34571464	GA	G	upstream
CG_Chr01	ClCG01G020490	Mid1-complementing activity 2	34624512	C	CT	intronic
CG_Chr01	ClCG01G020520	La domain-containing family protein	34640509	G	GA	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34671036	AT	A	downstream
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34672791	G	GTATATA	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34672793	G	GTATGTATGTA	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34675553	C	CA	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34680581	A	AT	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34681748	AAG	A	intronic
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689771	T	TC	upstream
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689827	A	ATC	upstream
CG_Chr01	ClCG01G020550	Eukaryotic aspartyl protease family protein LENGTH=513	34689828	A	ACTAAAATCAGGT	upstream
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700606	T	TCTCTCTCTC,TTTTCTCTCTCTC	upstream
CG_Chr01	ClCG01G020580	Phosphatase 2C family protein	34700606	T	TCTCTCTCTC,TTTTCTCTCTCTC	upstream
CG_Chr01	ClCG01G020600	Aconitate hydratase 1	34724630	A	AATGGGCTCTATATTGTTGATGATAGATTCTTAAAGAGCTGGAATTAAGCTTACAACATGTGCAAAGTTTTTTTCGTTGT	downstream
CG_Chr01	ClCG01G020620	Unknown protein	34742458	TTCAAGGACTAAAATAGAC	T	intronic

CG_Chr01	ClCG01G020620	Unknown protein	34742638	A	AT	intronic	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34748164	T	TGAACATTTA	intronic	
CG_Chr01	ClCG01G020630	RNA binding protein, putative	34749069	CT	C	intronic	
CG_Chr01	ClCG01G020640	ankyrin repeat protein LENGTH=435	34751278	AAT	A	upstream	
CG_Chr01	ClCG01G020640	ankyrin repeat protein LENGTH=435	34753597	TTCTCTCTCTCTC TCTCTCTC	T	intronic	
CG_Chr01	ClCG01G020640	ankyrin repeat protein LENGTH=435	34753637	C	CTCT	intronic	
CG_Chr01	ClCG01G020650	Ammonium transporter, putative	34756914	G	GA	upstream	
CG_Chr01	ClCG01G020660	Senescence-associated protein DIN1, putative	34760970	AT	A	downstream	
CG_Chr01	ClCG01G020680	UPF0548 protein	34773257	C	CT	upstream	
CG_Chr01	ClCG01G020700	F-box family protein	34783567	CAAG	C	exonic	nonframeshift deletion
CG_Chr01	ClCG01G020740	Arabidopsis protein of unknown function (DUF241) LENGTH=325	34798258	T	TA	upstream	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34800466	TTATATATATATA TATATA	T	downstream	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34802614	G	GT	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34803335	A	AT	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34806642	TA	T	intronic	
CG_Chr01	ClCG01G020750	Arabidopsis protein of unknown function (DUF241) LENGTH=503	34809115	AAAAG	A	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34840135	A	AACAC	intronic	
CG_Chr01	ClCG01G020780	Chitinase	34840140	A	ACAC	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34850261	C	CTTTT	intronic	
CG_Chr01	ClCG01G020800	Chitinase	34854371	A	ATGGATGCAGTGATAGCATTCAAGACGGCCAT TTGGTTT	exonic	frameshift insertion
CG_Chr01	ClCG01G020810	2-succinylbenzoate-CoA ligase	34857872	AAT	A	upstream	

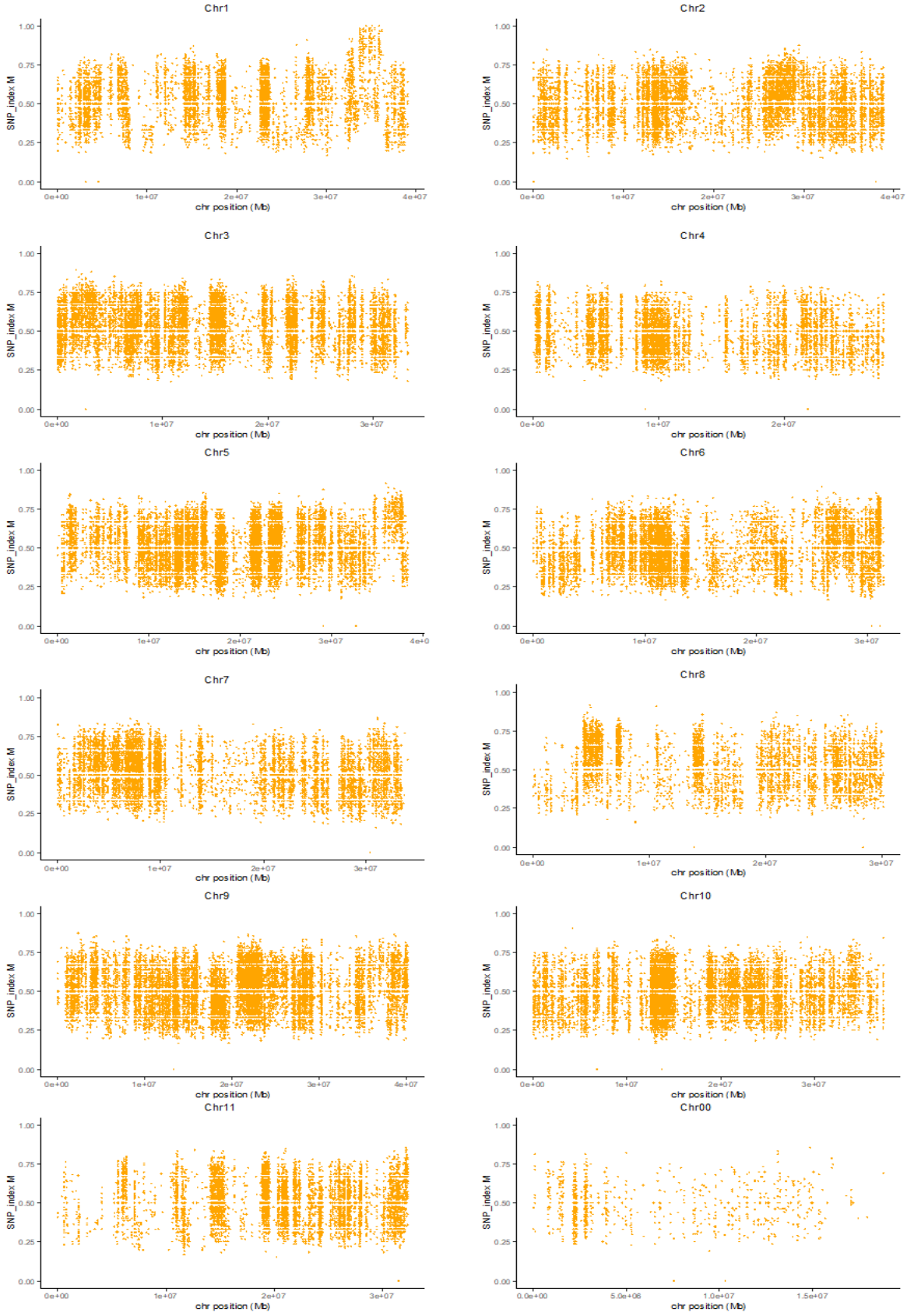
ANNO_REGION: Corresponding annotation data for each genomic region. The Anno_region corresponding to intronic data was removed for a better understanding.

Supplementary Figure 5.1. SNP-index and Δ(SNP-index) plots for 11 chromosomes of watermelon bulked DNA.

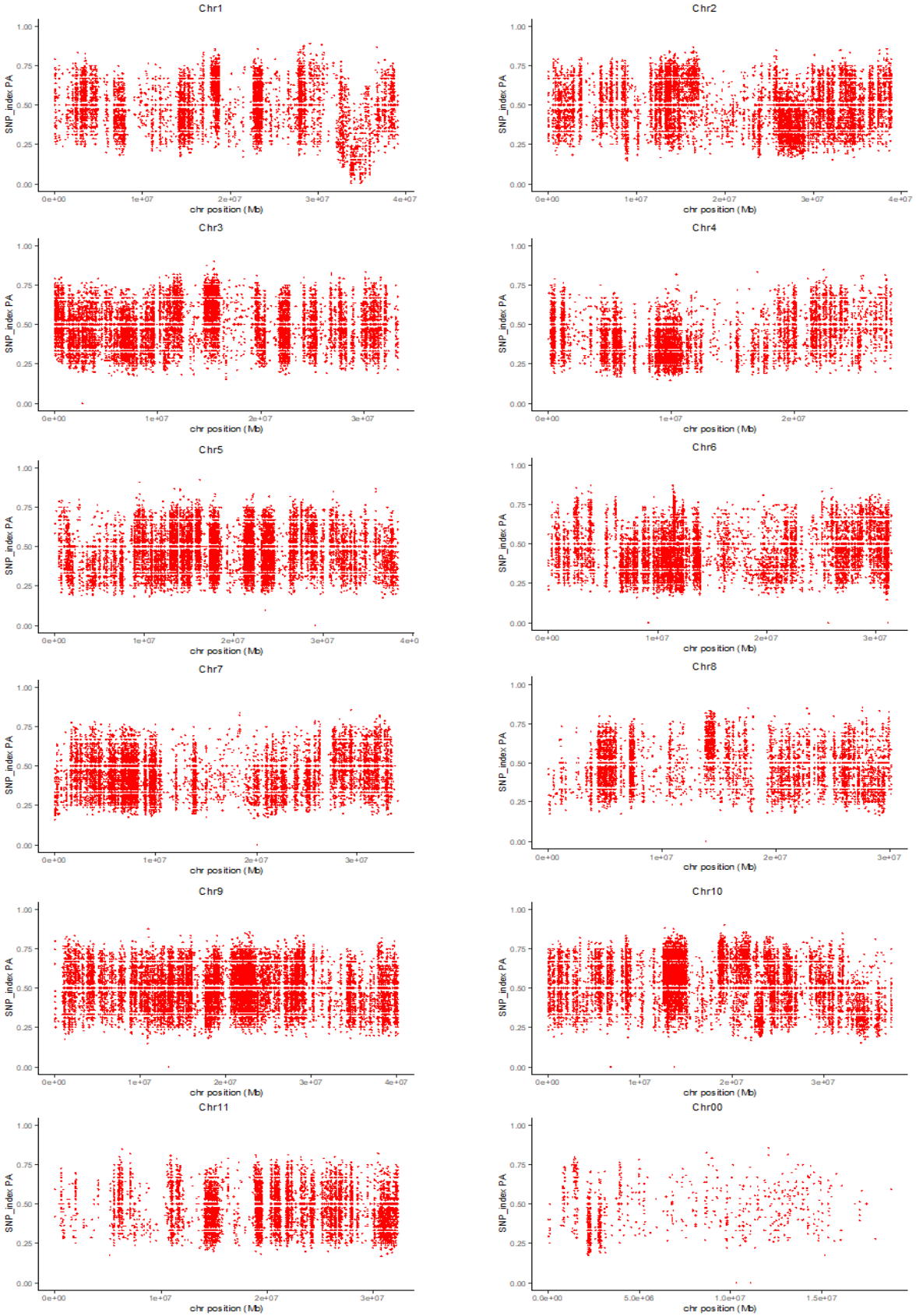
(**A**-**C**) The result of F2 progeny derived from a cross between P84 and P86. SNP-index graphs of M-bulk (**A**), PA-bulk (**B**) and Δ (SNP-index) graph (**C**) from QTL-seq analysis. X-axis represents the position of the eleven chromosomes and Y-axis represents the SNP-index.

Black lines indicate the sliding window average of 1Mb interval with 10 kb increment for Δ(SNP-index).

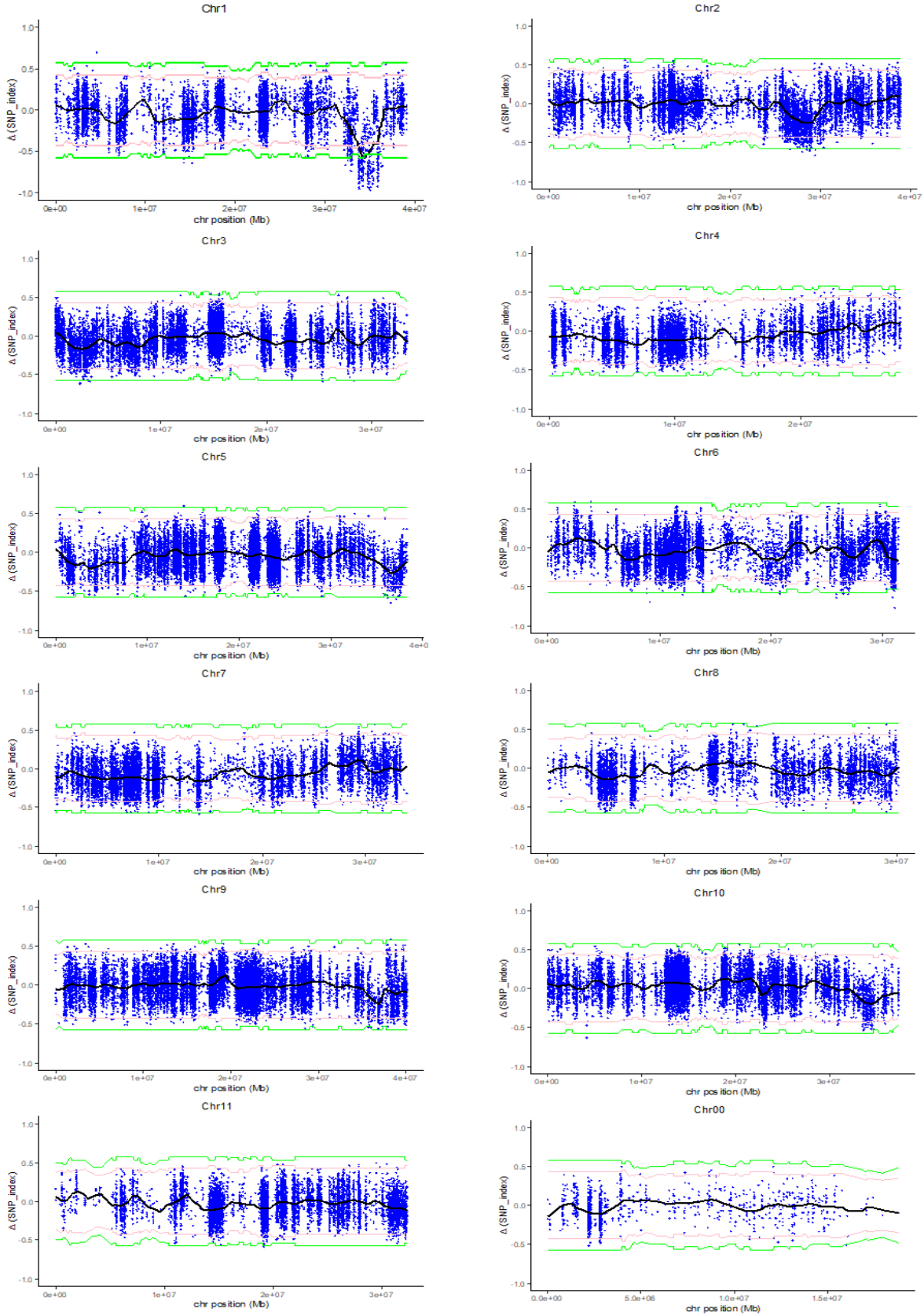
Reference allele frequency was defined in 0.3 (0.3 ≤ reference allele frequency ≤ 0.7), minimum total sample read depth ≥20, maximum total sample read depth ≤100, minimum per sample read depth ≥10, minimum genotype quality ≥99. A candidate QTL (QTL1) location was identified on watermelon chromosome 1 (32.2 - 36.4 Mb interval) with the criteria that the SNP-index in M-bulk (**A**) was near 1, SNP-index in PA-bulk (**B**) was near 0 and the Δ(SNP-index) (**C**) was above the confidence value.



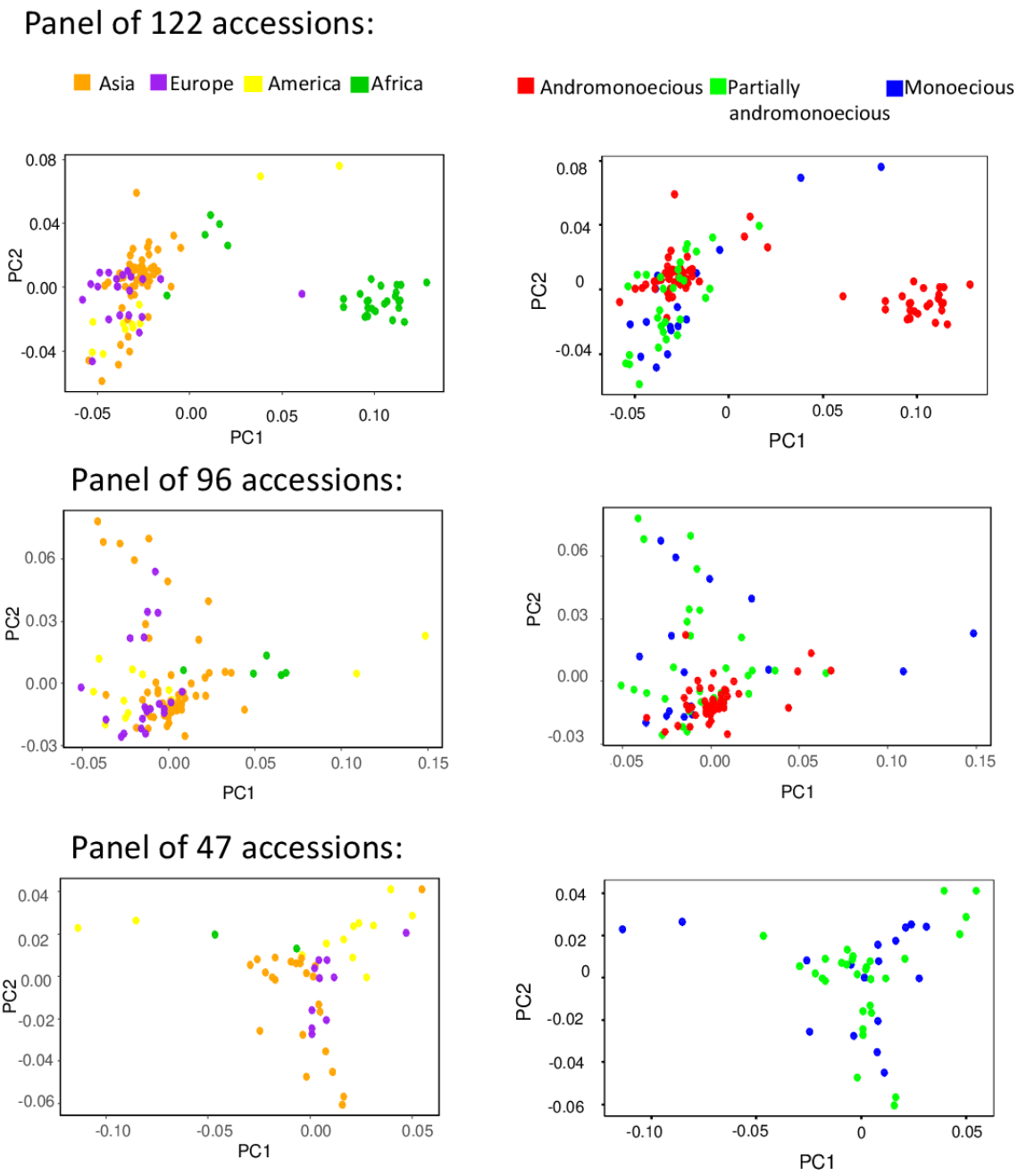
(**A**). SNP-index plot for bulked DNA of F2 showing the monoecious phenotype (M-bulk) derived from a cross between P84 and P86.



(**B**). SNP-index plot for bulked DNA of F2 showing the partially andromonoecious phenotype (PA-bulk) derived from a cross between P84 and P86.



(**C**). The Δ(SNP-index) plot obtained by subtraction of M-bulk SNP-index from PA-bulk SNP-index for F2 obtained from a cross between P84 and P86. Statistical confidence intervals under the null hypothesis of no QTL are shown as pink (P <0.05) or green (P <0.01).



Supplementary Figure 5.2. PCA plot based on GBS SNPs for the three panels used for GWAS. PC1 and PC2 indicate the scores of principal components 1 and 2, respectively. On the left, accessions were colored according to their geographical origin, while on the right they were colored based on their sex morphotype.

8. References

Abdelmohsin, M., and Pitrat, M. (2008). Pleiotropic effect of sex expression on fruit shape in melon. in *Cucurbitaceae 2008, Proceedings of the IXth EUCARPIA meeting on genetics and breeding of Cucurbitaceae (Pitrat M, ed), INRA, Avignon (France)*, 551–556.

Achigan-Dako, E. G., Avohou, E. S., Linsoussi, C., Ahanchede, A., Vodouhe, R. S., and Blattner, F. R. (2015). Phenetic characterization of *Citrullus* spp. (Cucurbitaceae) and differentiation of egusi-type (*C. mucosospermus*). *Genetic Resources and Crop Evolution* 62, 1159–1179. doi:10.1007/s10722-015-0220-z.

Aguado, E., García, A., Manzano, S., Valenzuela, J. L., Cuevas, J., Pinillos, V., et al. (2018). The sex-determining gene *CitACS4* is a pleiotropic regulator of flower and fruit development in watermelon (*Citrullus lanatus*). *Plant Reproduction* 31, 411–426. doi:10.1007/s00497-018-0346-1.

Alonso, J. M., Stepanova, A. N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F. M., et al. (2003). Five components of the ethylene-response pathway identified in a screen for weak ethylene-insensitive mutants in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* 100, 2992–2997. doi:10.1073/pnas.0438070100.

Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., and Palmer, J. D. (2010). Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution* 27, 1436–1448. doi:10.1093/molbev/msq029.

Argueso, C. T., Hansen, M., and Kieber, J. J. (2007). Regulation of ethylene biosynthesis. *Journal of Plant Growth Regulation* 26, 92–105. doi:10.1007/s00344-007-0013-5.

Bai, S. L., Peng, Y. Ben, Cui, J. X., Gu, H. T., Xu, L. Y., Li, Y. Q., et al. (2004). Developmental analyses reveal early arrests of the spore-bearing parts of reproductive organs in unisexual flowers of cucumber (*Cucumis sativus* L.). *Planta* 220, 230–240. doi:10.1007/s00425-004-1342-2.

Balbi, V., and Lomax, T. (2003). Regulation of early tomato fruit development by the *diageotropica* gene. *Plant Physiology* 131, 186–197. doi:10.1104/pp.010132.auxin.

Bo, K., Miao, H., Wang, M., Xie, X., Song, Z., Xie, Q., et al. (2019a). Novel loci *fsd6.1* and *Csgl3* regulate ultra-high fruit spine density in cucumber. *Theoretical and Applied Genetics* 132, 27–40. doi:10.1007/s00122-018-3191-6.

Bo, K., Wei, S., Wang, W., Miao, H., Dong, S., Zhang, S., et al. (2019b). QTL mapping and genome-wide association study reveal two novel loci associated with green flesh color in cucumber. *BMC Plant Biology* 19, 1–13. doi:10.1186/s12870-019-1835-6.

Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A., Morin, H., et al. (2008). A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science* 321, 836–838. doi:10.1126/science.1159023.

Boualem, A., Lemhemdi, A., Sari, M. A., Pignoly, S., Troadec, C., Choucha, F. A., et al. (2016). The andromonoecious sex determination gene predates the separation of *Cucumis* and *Citrullus* genera. *PLoS ONE* 11, 1–13. doi:10.1371/journal.pone.0155444.

Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M.-A., et al. (2015). A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science* 350, 688–691. doi:10.1126/science.aac8370.

Boualem, A., Troadec, C., Kovalski, I., Sari, M. A., Perl-Treves, R., and Bendahmane, A. (2009). A conserved ethylene biosynthesis enzyme leads to andromonoecy in two Cucumis species. *PLoS ONE* 4(7), e6144. doi:10.1371/journal.pone.0006144.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi:10.1093/bioinformatics/btm308.

Bukovac, M. J., and Wittwer, S. H. (1961). Gibberellin modification of flower sex expression in *Cucumis sativus* L. *Gibberellins. Advances in Chemistry 28*, 80–88. doi: 10.1021/ba-1961-0028.ch009

Burkill, H. M. (1978). The useful plants of west tropical Africa. In: Bush, A. (ed). Kew: Royal Botanic Gardens. Citron melon for cash and condiment. Econ Bot, 182–4.

Bush, A. (1978). Citron melon for cash and condiment. *Economic Botany* 32, 182–184. doi:10.1007/BF02866871.

Butenko, M. A., and Simon, R. (2015). Beyond the meristems: Similarities in the CLAVATA3 and INFLORESCENCE DEFICIENT IN ABSCISSION peptide mediated signalling pathways. *Journal of Experimental Botany* 66, 5195–5203. doi:10.1093/jxb/erv310.

Byers, R. E., Baker, L. R., Sell, H. M., Herner, R. C., and Dilley, D. R. (1972). Ethylene: a natural regulator of sex expression of *Cucumis melo* L. *Proceedings of the National Academy of Sciences USA* 69, 717–720. doi:10.1073/pnas.69.3.717.

Carbonell-Bejerano, P., Urbez, C., Carbonell, J., Granell, A., and Perez-Amador, M. A. (2010). A Fertilization-Independent Developmental Program Triggers Partial Fruit Development and Senescence Processes in Pistils of Arabidopsis. *Plant Physiology* 154, 163–172. doi:10.1104/pp.110.160044.

Carbonell-Bejerano, P., Urbez, C., Granell, A., Carbonell, J., and Perez-Amador, M. A. (2011).

Ethylene is involved in pistil fate by modulating the onset of ovule senescence and the GA-mediated fruit set in Arabidopsis. *BMC Plant Biology* 11:84. doi:10.1186/1471-2229-11-84.

Chae, E., Tan, Q. K. G., Hill, T. A., and Irish, V. F. (2008). An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* 135, 1235–1245. doi:10.1242/dev.015842.

Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., and Ecker, J. R. (1997). Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89, 1133–1144. doi:10.1016/S0092-8674(00)80300-1.

Che, G., and Zhang, X. (2019). Molecular basis of cucumber fruit domestication. *Current Opinion in Plant Biology* 47, 38–46. doi:10.1016/j.pbi.2018.08.006.

Chen, H., Sun, J., Li, S., Cui, Q., Zhang, H., Xin, F., et al. (2016). An ACC Oxidase Gene Essential for Cucumber Carpel Development. *Molecular Plant* 9, 1315–1327. doi:10.1016/j.molp.2016.06.018.

Chen, Y. F., Etheridge, N., and Schaller, G. E. (2005). Ethylene signal transduction. *Annals of Botany* 95, 901–915. doi:10.1093/aob/mci100.

Chomicki, G., and Renner, S. S. (2015). Watermelon origin solved with molecular phylogenetics including Linnaean material: another example of museomics. *New Phytologist* 205, 526–532. doi:10.1111/nph.13163.

Chomicki, G., Schaefer, H., and Renner, S. S. (2019). Origin and domestication of Cucurbitaceae crops: insights from phylogenies, genomics and archaeology. *New Phytologist*. doi:10.1111/nph.16015.

Clark, G., Torres, J., Finlayson, S., Guan, X., Handley, C., Lee, J., et al. (2010). Apyrase (Nucleoside Triphosphate-Diphosphohydrolase) and Extracellular Nucleotides Regulate Cotton Fiber Elongation in Cultured Ovules. *Plant Physiology* 152, 1073–1083. doi:10.1104/pp.109.147637.

Crow, J. F., and Dove, W. F. (1990). Mapping Functions. *Genetics Society of America*, 669–671. doi:10.1007/978-1-4684-2982-4_10.

Cuevas, J., Rallo, L., and Rapoport, H. F. (1994). Procedure to study ovule senescence in olive. *Acta Horticulturae* 356, 252–256.

Dane, F., and Liu, J. (2007). Diversity and origin of cultivated and citron type watermelon (*Citrullus lanatus*). *Genetic Resources and Crop Evolution* 54, 1255–1265. doi:10.1007/s10722-006-9107-3.

De Martinis, D., Haenen, I., Pezzotti, M., Benvenuto, E., and Mariani, C. (1999). Ethylene and Flower Development in Tobacco Plants. In: Kanellis AK, Chang C, Klee H, Bleecker AB, Pech JC, Grierson D (eds). Biology and Biotechnology of the Plant Hormone Ethylene II *Springer, Dordrecht*, 157–164.

De Martinis, D., and Mariani, C. (1999). Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic Tobacco plants. *Plant Cell* 11, 1061–1072. doi:10.1105/tpc.11.6.1061.

Den Nijs, A., and Visser, D. (1980). Induction of male flowering in (*Cucumis sativus* L.) by silver ions. *Euphytica* 29, 273–280. doi:https://doi.org/10.1007/BF00025124.

Depristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43, 491–501. doi:10.1038/ng.806.

Díaz, A., Zarouri, B., Fergany, M., Eduardo, I., Álvarez, J. M., Picó, B., et al. (2014). Mapping and introgression of QTL involved in fruit shape transgressive segregation into "Piel de Sapo" melon (*Cucucumis melo* L.). *PLoS ONE* 9. doi:10.1371/journal.pone.0104188.

Dou, J., Lu, X., Ali, A., Zhao, S., Zhang, L., He, N., et al. (2018a). Genetic mapping reveals a marker for yellow skin in watermelon (*Citrullus lanatus* L.). *PLoS ONE* 13, 1–15. doi:10.1371/journal.pone.0200617.

Dou, J., Zhao, S., Lu, X., He, N., Zhang, L., Ali, A., et al. (2018b). Genetic mapping reveals a candidate gene (*ClFS1*) for fruit shape in watermelon (*Citrullus lanatus* L.). *Theoretical and Applied Genetics* 131, 947–958. doi:10.1007/s00122-018-3050-5.

Duan, Q. H., Wang, D. H., Xu, Z. H., and Bai, S. N. (2008). Stamen development in Arabidopsis is arrested by organ-specific overexpression of a cucumber ethylene synthesis gene *CsACO2*. *Planta* 228, 537–543. doi:10.1007/s00425-008-0756-7.

Duan, Y., Li, S., Chen, Z., Zheng, L., Diao, Z., Zhou, Y., et al. (2012). *Dwarf and deformed flower 1*, encoding an F-box protein, is critical for vegetative and floral development in rice (*Oryza sativa* L.). *Plant Journal* 72, 829–842. doi:10.1111/j.1365-313X.2012.05126.x.

Dussi, M. C., Sosa, D., and Calvo, G. S. (2002). Effects of Retain™ on fruit maturity and fruit set of Pear cultivars Williams and Packham´s Triumph. *Acta Horticulturae* 596, 767–771. doi:10.17660/actahortic.2002.596.132.

Estornell, L. H., Landberg, K., Cierlik, I., and Sundberg, E. (2018). *SHI/STY* genes affect pre- and post-meiotic anther processes in auxin sensing domains in Arabidopsis. *Frontiers in Plant Science* 9, 1–16. doi:10.3389/fpls.2018.00150.

FAOSTAT (2018). http://www.fao.org/faostat/es/#data/QC. Accessed 1 Jun 2019. Accessed 1 Dec 2018.).

Felsenstein, J. (1985). Phylogenies and the comparative method. *American Naturalist* 125, 1–15. doi:10.1086/284325.

Ferrándiz, C., Pelaez, S., and Yanofsky, M. F. (1999). Control of carpel and fruit development in Arabidopsis. *Annu Rev Biochem* 68, 321–354.

Ferreira, M. A. J. F., Queiróz, M. A., Vencovsky, R., Braz, L. T., Vieira, M. L. C., and Borges, R. M. E. (2002). Sexual expression and mating system in watermelon: implications for breeding programs. *Crop Breeding and Applied Biotechnology* 2, 39–48. doi:10.12702/1984-7033.v02n01a06.

Fursa, T. B. (1972). On the taxonomy of the genus *Citrullus* Schad. *Bot Zh* 57, 31–4.

Gagne, J. M., Smalle, J., Gingerich, D. J., Walker, J. M., Yoo, S. D., Yanagisawa, S., et al. (2004). *Arabidopsis* EIN3-binding F-box 1 and 2 form ubiquitin-protein ligases that repress ethylene action and promote growth by directing EIN3 degradation. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6803–6808. doi:10.1073/pnas.0401698101.

Galun, E. (1961). Study of the inheritance of sex expression in the cucumber. The interaction of major genes with modifying genetic and non-genetic factors. *Genetica* 32, 134–163. doi:10.1007/BF01816091.

Gao, J., Dai, G., Zhou, W., Liang, H., Huang, J., Qing, D., et al. (2019). Mapping and identifying a candidate gene *Plr4*, a recessive gene regulating purple leaf in rice, by using bulked segregant and transcriptome analysis with next-generation sequencing. *International Journal of Molecular Sciences* 20, 1–15. doi:10.3390/ijms20184335.

García, A., Aguado, E., Martínez, C., Loska, D., Beltrán, S., Valenzuela, J. L., et al. (2020). The ethylene receptors *CpETR1A* and *CpETR2B* cooperate in the control of sex determination in *Cucurbita pepo*. *Journal of experimental botany* 71, 154–167. doi:10.1093/jxb/erz417.

García, A., Manzano, S., Martínez, C., Megías, Z., Gázquez, J. C., and Jamilena, M. (2015). Isolation and characterisation of ethylene insensitive mutants from a collection of *Cucurbita pepo* L. *Acta Horticulturae* 1151, 151–156. doi:10.17660/ActaHortic.2017.1151.24.

Gebremeskel, H., Dou, J., Li, B., Zhao, S., Muhammad, U., Lu, X., et al. (2020). Molecular mapping and candidate gene analysis for GA3 responsive short internode in watermelon (*Citrullus lanatus*). *International Journal of Molecular Sciences* 21. doi:10.3390/ijms21010290.

Giovannoni, J. J., Wing, R. A., Ganal, M. W., and Tanksley, S. D. (1991). Isolation of molecular

markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research* 19, 6553–6568. doi:10.1093/nar/19.23.6553.

Girek, Z., Prodanovic, S., Zdravkovic, J., Zivanovic, T., Ugrinovic, M., and Zdravkovic, M. (2013). The effect of growth regulators on sex expression in melon (*Cucumis melo* L.). *Crop Breeding and Applied Biotechnology* 13, 165–171. doi:10.1590/S1984-70332013000300003.

Grover, A. (2012). Plant Chitinases: Genetic Diversity and Physiological Roles. *Critical Reviews in Plant Sciences* 31, 57–73. doi:10.1080/07352689.2011.616043.

Gu, S. Y., Wang, L. C., Cheuh, C. M., and Lo, W. S. (2019). *CHITINASE LIKE1* regulates root development of dark-grown seedlings by modulating ethylene biosynthesis in *Arabidopsis thaliana*. *Frontiers in Plant Science* 10, 1–17. doi:10.3389/fpls.2019.00600.

Guo, H., and Ecker, J. R. (2003). Plant responses to ethylene gas are mediated by $SCF^{EBF1/EBF2}$-dependent proteolysis of EIN3 transcription factor. *Cell* 115, 667–677. doi:10.1016/S0092-8674(03)00969-3.

Guo, S., Sun, H., Zhang, H., Liu, J., Ren, Y., Gong, G., et al. (2015). Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS ONE* 10, 1–21. doi:10.1371/journal.pone.0130267.

Guo, S., Zhang, J., Sun, H., Salse, J., Lucas, W. J., Zhang, H., et al. (2013). The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genetics* 45, 51–58. doi:10.1038/ng.2470.

Guo, S., Zhao, S., Sun, H., Wang, X., Wu, S., Lin, T., et al. (2019). Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nature Genetics* 51, 1616–1623. doi:10.1038/s41588-019-0518-4.

Gur, A., Tzuri, G., Meir, A., Sa'ar, U., Portnoy, V., Katzir, N., et al. (2017). Genome-Wide Linkage-Disequilibrium Mapping to the Candidate Gene Level in Melon (*Cucumis melo*). *Scientific Reports* 7, 9770. doi:10.1038/s41598-017-09987-4.

Hermans, C., Porco, S., Verbruggen, N., and Bush, D. R. (2010). Chitinase-like protein CTL1 plays a role in altering root system architecture in response to multiple environmental conditions. *Plant Physiology* 152, 904–917. doi:10.1104/pp.109.149849.

Hou, J., Zhou, Y. F., Gao, L. Y., Wang, Y. L., Yang, L. M., Zhu, H. Y., et al. (2018). Dissecting the genetic architecture of melon chilling tolerance at the seedling stage by association mapping and identification of the elite alleles. *Frontiers in Plant Science* 871, 1–12. doi:10.3389/fpls.2018.01577.

Hu, B., Li, D., Liu, X., Qi, J., Gao, D., Zhao, S., et al. (2017). Engineering Non-transgenic

Gynoecious Cucumber Using an Improved Transformation Protocol and Optimized CRISPR/Cas9 System. *Molecular Plant* 10, 1575–1578. doi:10.1016/j.molp.2017.09.005.

Huai, Q., Xia, Y., Chen, Y., Callahan, B., Li, N., and Ke, H. (2001). Crystal Structures of 1-Aminocyclopropane-1-carboxylate (ACC) Synthase in Complex with Aminoethoxyvinylglycine and Pyridoxal-5′-Phosphate Provide New Insight into Catalytic Mechanisms. *Journal of Biological Chemistry* 276, 38210–38216. doi:10.1074/jbc.M103840200.

Hussain, A. I., Rathore, H. A., Sattar, M. Z. A., Chatha, S. A. S., Sarker, S. D., and Gilani, A. H. (2014). *Citrullus colocynthis* (L.) Schrad (bitter apple fruit): A review of its phytochemistry, pharmacology, traditional uses and nutritional potential. *Journal of Ethnopharmacology* 155, 54–66. doi:10.1016/j.jep.2014.06.011.

Imerovski, I., Dedić, B., Cvejić, S., Miladinović, D., Jocić, S., Owens, G. L., et al. (2019). BSA-seq mapping reveals major QTL for broomrape resistance in four sunflower lines. *Molecular Breeding* 39. doi:10.1007/s11032-019-0948-9.

Ishiki, Y., Oda, A., Yaegashi, Y., Orihara, Y., Arai, T., Hirabayashi, T., et al. (2000). Cloning of an auxin-responsive 1-aminocyclopropane-1-carboxylate synthase gene (*CMe-ACS2*) from melon and the expression of *ACS* genes in etiolated melon seedlings and melon fruits. *Plant Science* 159, 173–181. doi:10.1016/S0168-9452(00)00298-3.

Jakubowicz, M., and Sadowski, J. (2002). 1-Aminocyclopropane-1-carboxylate synthase - Genes and expression. *Acta Physiologiae Plantarum* 24, 459–478. doi:10.1007/s11738-002-0043-3.

Jeffrey, C. (1967). Cucurbitacea*e*. In: Milne-Redhead E, Polhill RM (eds). Flora of Tropical East Africa. London: Published under the Authority of the Minister for Overseas Development by the Crown Agents for Oversea Governments and Administration.

Ji, G., Zhang, J., Gong, G., Shi, J., Zhang, H., Ren, Y., et al. (2015). Inheritance of sex forms in watermelon (*Citrullus lanatus*). *Scientia Horticulturae* 193, 367–373. doi:10.1016/j.scienta.2015.07.039.

Ji, G., Zhang, J., Zhang, H., Sun, H., Gong, G., Shi, J., et al. (2016). Mutation in the gene encoding (*CitACS4*) led to andromonoecy in watermelon. *Journal of integrative plant biology*, 4–7. doi:10.1111/jipb.12466.

Jiang, X., and Lin, D. (2007). Discovery of watermelon gynoecious gene *gy*. *Acta Horticulturae Sinica* 34, 141–142.

Johnson, P. R., and Ecker, J. R. (1998). The ethylene GAS signal transduction pathway: A Molecular Perspective. *Annual Review of Genetics* 32, 227–254. doi:10.1146/annurev.genet.32.1.227.

Ju, C., and Chang, C. (2012). Advances in ethylene signalling: protein complexes at the endoplasmic reticulum membrane. *AoB Plants* 2012, pls031–pls031. doi:10.1093/aobpla/pls031.

Ju, C., Yoon, G. M., Shemansky, J. M., Lin, D. Y., Ying, Z. I., Chang, J., et al. (2012). CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 109, 19486–19491. doi:10.1073/pnas.1214848109.

Kahana, A., Silberstein, L., Kessler, N., Goldstein, R. S., and Perl-Treves, R. (1999). Expression of ACC oxidase genes differs among sex genotypes and sex phases in cucumber. *Plant Molecular Biology* 41, 517–528. doi:10.1023/A:1006343707567.

Kamachi, S., Sekimoto, H., Kondo, N., and Sakai, S. (1997). Cloning of a cDNA for a 1-Aminocyclopropane-1-Carboxylate Synthase That is expressed during development of female flowers at the apices of *Cucumis sativus* L. *Plant and Cell Physiology* 38, 1197–1206. doi:10.1093/oxfordjournals.pcp.a029106.

Kato, M., Hayakawa, Y., Hyodo, H., Ikoma, Y., and Yano, M. (2000). Wound-induced ethylene synthesis and expression and formation of 1-aminocyclopropane-1-carboxylate (ACC) synthase, ACC oxidase, phenylalanine ammonia-lyase, and peroxidase in wounded mesocarp tissue of *Cucurbita maxima*. *Plant and Cell Physiology* 41, 440–447. doi:10.1093/pcp/41.4.440.

Kende, H. (1993). Ethylene biosynthesis. *Annual Review of Plant Biology 44* 44, 283–307. doi:10.1146/annurev.pp.44.060193.001435.

Kenigsbuch, D., and Cohen, Y. (1987). Inheritance of gynoecious Sex type in Muskmelon. *Cucurbit Genetics Cooperative Report* 10, 47–48.

Kenigsbuch, D., and Cohen, Y. (1990). The inheritance of gynoecy in muskmelon. *Genome* 33, 317–320. doi:https://doi.org/10.1139/g90-049.

Kieber, J. J., Rothenberg, M., Roman, G., Feldmann, K. A., and Ecker, J. R. (1993). CTR1, a negative regulator of the ethylene response pathway in arabidopsis, encodes a member of the Raf family of protein kinases. *Cell* 72, 427–441. doi:10.1016/0092-8674(93)90119-B.

Klee, H. J. (2004). Ethylene signal transduction. Moving beyond Arabidopsis. *Plant Physiology* 135, 660–667. doi:10.1104/pp.104.040998.

Klein, H., Xiao, Y., Conklin, P. A., Govindarajulu, R., Kelly, J. A., Scanlon, M. J., et al. (2018). Bulked-segregant analysis coupled to whole genome sequencing (BSA-Seq) for rapid gene cloning in maize. *G3: Genes, Genomes, Genetics* 8, 3583–3592. doi:10.1534/g3.118.200499.

Knopf, R. R., and Trebitsh, T. (2006). The female-specific *Cs-ACS1G* gene of cucumber. A case of

gene duplication and recombination between the non-sex-specific 1-aminocyclopropane-1-carboxylate synthase gene and a branched-chain amino acid transaminase gene. *Plant and Cell Physiology* 47, 1217–1228. doi:10.1093/pcp/pcj092.

Kubicki, B. (1962). Inheritance of some characters in muskmelons (*Cucumis melo*). *Genetica Polonica* 3, 265–274.

Kubicki, B. (1969a). Comparative studies on sex determination in cucumber (*Cucumis sativus* L.) and muskmelon (Cucumis melo L.). *Genet Pol* 10, 167–183.

Kubicki, B. (1969b). Investigation on sex determination in cucumber (*Cucumis sativus* L.). III Variability of sex expression in the monoecious and gynoecious lines. *Genet Pol* 10, 5–22.

Kubicki, B. (1969c). Investigation on sex determination in cucumber (*Cucumis sativus* L.). V genes controlling intensity of femaleness. *Genet Pol* 10, 69–85.

Kubicki, B. (1969d). Investigations on sex determination in cucumber (*Cucumis sativus* L.). VII. Andromonoecism and hermaphroditism. *Genet Pol* 10, 101–120.

Kubicki, B. (1969e). Investigations on sex determination in cucumber (*Cucumis sativus* L.). VIII. Trimonoecism. *Genet Pol*, 123–142.

Kuusk, S., Sohlberg, J. J., Magnus Eklund, D., and Sundberg, E. (2006). Functionally redundant *SHI* family genes regulate Arabidopsis gynoecium development in a dose-dependent manner. *Plant Journal* 47, 99–111. doi:10.1111/j.1365-313X.2006.02774.x.

Laibach, F., and Kribben, F. J. (1950). Der Einfluss von Wuchstoff auf die bildung mannlicher und weiblicher Bluten bei einer monoezischen Pflanze (*Cucumis sativus* L.). *Ber. deut. bot. Ges* 62: 53-55.

Larsen, P.B., Woltering, E.J., and Woodson, W.R. (1993). Ethylene and interorgan signaling in flowers following pollination. In: Raskin, I. and Schultz, J. (ed) Plant Signals in Interactions with Other Organisms. *American Society of Plant Physiologist*s, 171-181.

Latrasse, D., Rodriguez-Granados, N. Y., Veluchamy, A., Mariappan, K. G., Bevilacqua, C., Crapart, N., et al. (2017). The quest for epigenetic regulation underlying unisexual flower development in *Cucumis melo*. *Epigenetics and Chromatin* 10, 1–17. doi:10.1186/s13072-017-0132-6.

Lee, H. Y., and Yoon, G. M. (2018). Regulation of ethylene biosynthesis by phytohormones in etiolated rice (*Oryza sativa* L.) seedlings. *Molecules and Cells* 41, 311–319. doi:10.14348/molcells.2018.2224.

Levi, A., Jarret, R., Kousik, S., Wechter, W. P., Nimmakayala, P., and Reddy, U. (2017). Genetic Resources of Watermelon. In: Grumet, R., Katzir, N., García-Mas, J. (eds) Genetics and

Genomics of Cucurbitaceae. Plant Genetics and Genomics: Crops and Models. *Springer, Cham*. doi: 87–110. doi:10.1007/7397_2016_34.

Levi, A., Thies, J. A., Wechter, W. P., Harrison, H. F., Simmons, A. M., Reddy, U. K., et al. (2013). High frequency oligonucleotides: Targeting active gene (HFO-TAG) markers revealed wide genetic diversity among *Citrullus* spp. accessions useful for enhancing disease or pest resistance in watermelon cultivars. *Genetic Resources and Crop Evolution* 60, 427–440. doi:10.1007/s10722-012-9845-3.

Levi, A., and Thomas, C. (1999). An improved procedure for isolation of high quality DNA from watermelon and melon leaves. *Cucurbit Genet Coop 22*, 41–42.

Levi, A., and Thomas, C. E. (2001). Low genetic diversity indicates the need to broaden the genetic base of cultivated watermelon. *HortScience* 36, 1096–1101. doi:10.21273/hortsci.36.6.1096.

Levi, A., Thomas, C. E., Joobeur, T., Zhang, X., and Davis, A. (2002). A genetic linkage map for watermelon derived from a testcross population: (*Citrullus lanatus* var. *citroides* x *C. lanatus* var. *lanatus*) x *Citrullus colocynthis*. *Theoretical and Applied Genetics* 105, 555–563. doi:10.1007/s00122-001-0860-6.

Levi, A., Thomas, C. E., Keinath, A. P., and Wehner, T. C. (2001). Genetic diversity among watermelon (*Citrullus lanatus* and *Citrullus colocynthis*) accessions. *Genetic Resources and Crop Evolution* 48, 559–566. doi:10.1023/A:1013888418442.

Levi, A., Wechter, W. P., Massey, L., Carter, L., and Hopkins, D. (2011). Genetic Linkage Map of *Citrullus lanatus* var. *citroides* Chromosomal Segments Introgressed into the watermelon cultivar Crimson Sweet (*Citrullus lanatus* var. *lanatus*) Genome. *American Journal of Plant Sciences* 2, 93–110. doi:10.1017/CBO9781107415324.004.

Li, D., Sheng, Y., Niu, H., and Li, Z. (2019). Gene interactions regulating sex determination in Cucurbits. *Frontiers in Plant Science* 10, 1–12. doi:10.3389/fpls.2019.01231.

Li, H., and Durbin, R. (2009). Making the Leap: Maq to BWA. *Mass Genomics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009a). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

Li, M. X., Yeung, J. M. Y., Cherny, S. S., and Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant *p*-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics* 131, 747–756. doi:10.1007/s00439-011-1118-2.

Li, P., Su, T., Zhang, B., Li, P., Xin, X., Yue, X., et al. (2020a). Identification and fine mapping of *qSB.A09*, a major QTL that controls shoot branching in *Brassica rapa* ssp. *chinensis* Makino. *Theoretical and Applied Genetics* 133, 1055–1068. doi:10.1007/s00122-020-03531-1.

Li, X., Zhang, X., Shen, D., Wu, Y., Wang, H., Song, J., et al. (2018). Temperature and photoperiod changes affect cucumber sex expression by different epigenetic regulations. *BMC Plant Biology* 18, 1–13. doi:10.1186/s12870-018-1490-3.

Li, Z., Han, Y., Niu, H., Wang, Y., Jiang, B., and Weng, Y. (2020b). Gynoecy instability in cucumber (*Cucumis sativus* L.) is due to unequal crossover at the copy number variation-dependent *Femaleness* (*F*) locus. *Horticulture Research* 7. doi:10.1038/s41438-020-0251-2.

Li, Z., Huang, S., Liu, S., Pan, J., Zhang, Z., Tao, Q., et al. (2009b). Molecular isolation of the *M* gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. *Genetics* 182, 1381–1385. doi:10.1534/genetics.109.104737.

Lin, C. K., Lee, N. Y., Huang, P. L., and Do, Y. Y. (2016). Gene structure and expression characteristics of the auxin receptor TIR1 ortholog in *Momordica charantia* and developmental analysis of its promoter in transgenic plants. *Journal of Plant Biochemistry and Biotechnology* 25, 253–262. doi:10.1007/s13562-015-0336-4.

Liu, H., Wu, Y., Cao, A., Mao, B., Zhao, B., and Wang, J. (2017). Genome-Wide Analysis of DNA methylation during ovule development of Female-Sterile Rice *fsv1*. *G3 (Genes|Genomes|Genetics)* 7, 3621–3635. doi:10.1534/g3.117.300243.

Lotan, T., Ori, N., and Fluhr, R. (1989). Pathogenesis-related proteins are developmentally regulated in Tobacco flowers. *The Plant cell* 1, 881–887. doi:10.1105/tpc.1.9.881.

Loy, J. B. (2006). Fruit size in melon in monoecious and andromonoecious Isolines J. Brent Loy. *Cucurbits Genetics Cooperative Report* 28–29, 12–13.

Malepszy, S., and Niemirowicz-Szczytt, K. (1991). Sex determination in cucumber (*Cucumis sativus*) as a model system for molecular biology. *Plant Science* 80, 39–47. doi:10.1016/0168-9452(91)90271-9.

Mansfeld, B. N., and Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* 11, 1–5. doi:10.3835/plantgenome2018.01.0006.

Manzano, S. (2009). Regulación genética de la determinación sexual en *Cucurbita pepo*: clonación, caracterización y análisis funcional de genes implicados en la biosíntesis, percepción y respuesta a etileno. Thesis, University of Almería.

Manzano, S., Aguado, E., Martínez, C., Megías, Z., García, A., and Jamilena, M. (2016). The Ethylene Biosynthesis Gene *CitACS4* Regulates Monoecy/Andromonoecy in Watermelon

(*Citrullus lanatus*). *PloS one* 11, e0154362. doi:10.1371/journal.pone.0154362.

Manzano, S., Martínez, C., Domínguez, V., Avalos, E., Garrido, D., Gómez, P., et al. (2009). A major gene conferring reduced ethylene sensitivity and maleness in *Cucurbita pepo*. *Journal of Plant Growth Regulation* 29, 73–80. doi:10.1007/s00344-009-9116-5.

Manzano, S., Martínez, C., García, J. M., Megías, Z., and Jamilena, M. (2014). Involvement of ethylene in sex expression and female flower development in watermelon (*Citrullus lanatus*). *Plant Physiology and Biochemistry* 85, 96–104. doi:10.1016/j.plaphy.2014.11.004.

Manzano, S., Martínez, C., Gómez, P., Garrido, D., and Jamilena, M. (2010). Cloning and characterisation of two *CTR1*-like genes in *Cucurbita pepo*: Regulation of their expression during male and female flower development. *Sexual Plant Reproduction* 23, 301–313. doi:10.1007/s00497-010-0140-1.

Manzano, S., Martínez, C., Megías, Z., Garrido, D., and Jamilena, M. (2013). Involvement of ethylene biosynthesis and signalling in the transition from male to female flowering in the monoecious *Cucurbita pepo*. *Journal of Plant Growth Regulation* 32, 789–798. doi:10.1007/s00344-013-9344-6.

Manzano, S., Martínez, C., Megías, Z., Gómez, P., Garrido, D., and Jamilena, M. (2011). The role of ethylene and brassinosteroids in the control of sex expression and flower development in Cucurbita pepo. *Plant Growth Regulation* 65, 213–221. doi:10.1007/s10725-011-9589-7.

Mao, L., Karakurt, Y., and Huber, D. J. (2004). Incidence of water-soaking and phospholipid catabolism in ripe watermelon (*Citrullus lanatus*) fruit: Induction by ethylene and prophylactic effects of 1-methylcyclopropene. *Postharvest Biology and Technology* 33, 1–9. doi:10.1016/j.postharvbio.2003.12.007.

Mara, C. D., and Irish, V. F. (2008). Two GATA transcription factors are downstream effectors of floral homeotic gene action in Arabidopsis. *Plant Physiology* 147, 707–718. doi:10.1104/pp.107.115634.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., et al. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature* 461, 1135–1138. doi:10.1038/nature08498.

Martin, F. W. (1959). Staining and observing pollen tubes in the style by means of fluorescence. *Stain Technology* 34, 125–128. doi:10.3109/10520295909114663.

Martínez, C. (2013). Involvement of ethylene biosynthesis and signaling genes in determination and parthenocarpic fruit set and development in Zucchini (*Cucurbita pepo* L.). Thesis, University of Almería.

Martínez, C., Manzano, S., Megías, Z., Barrera, A., Boualem, A., Garrido, D., et al. (2014). Molecular and functional characterization of CpACS27A gene reveals its involvement in monoecy instability and other associated traits in squash (*Cucurbita pepo* L.). *Planta* 239, 1201–1215. doi:10.1007/s00425-014-2043-0.

Martínez, C., Manzano, S., Megías, Z., Garrido, D., Picó, B., and Jamilena, M. (2013). Involvement of ethylene biosynthesis and signalling in fruit set and early fruit development in zucchini squash (*Cucurbita pepo* L.). *BMC plant biology* 13, 139. doi:10.1186/1471-2229-13-139.

Martos-Fuentes, M., Lizarzaburu, J. A., Aguayo, E., Martínez, C., and Jamilena, M. (2017). Pleiotropic effects of *CmACS7* on fruit growth and quality parameters in melon (*Cucumis melo*). in *Acta Horticulturae* (International Society for Horticultural Science (ISHS), Leuven, Belgium), 115–120. doi:10.17660/ActaHortic.2017.1151.19.

Matsumoto, N., and Okada, K. (2001). regulates lateral axis-dependent development of Arabidopsis flowers. *Genes & Development*, 3355–3364. doi:10.1101/gad.931001.from.

Maynard, D. (2001). An introduction to the watermelon. In: Maynard DN, (ed)Watermelons: characteristics, production, and marketing. Alexandria, VA: *ASHS Press*, 9–20.

Maynard, D., and Maynard, D. (2000). Cucumbers, melons, and watermelons. In: Kiple KF, Ornelas KC, (eds) The Cambridge world history of food. *Cambridge: Cambridge University Press*, 298–313

McCreight, J. D., Staub, J. E., Wehner, T. C., and Dhillon, N. P. S. (2013). Gone global: Familiar and exotic cucurbits have Asian Origins. *HortScience* 48, 1078–1089. doi:10.21273/hortsci.48.9.1078.

McKim, S. M., Stenvik, G. E., Butenko, M. A., Kristiansen, W., Cho, S. K., Hepworth, S. R., et al. (2008). The *BLADE-ON-PETIOLE* genes are essential for abscission zone formation in *Arabidopsis*. *Development* 135, 1537–1546. doi:10.1242/dev.012807.

Megharaj, K. C., Ajjappalavara, P. S., appa, R., Manjunathagowda, D. C., and Bommesh, J. C. (2017). Sex Manipulation in Cucurbitaceous Vegetables. *International Journal of Current Microbiology and Applied Sciences* 6, 1839–1851. doi:10.20546/ijcmas.2017.609.227.

Megías, Z. (2016). Genetic and molecular analysis of cold tolerance and postharvest fruit quality in zucchini (*Cucurbita pepo* L.). Thesis, University of Almería.

Mibus, H., and Tatlioglu, T. (2004). Molecular characterization and isolation of the *F*/*f* gene for femaleness in cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics* 109, 1669–1676. doi:10.1007/s00122-004-1793-7.

Michelmore, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-

resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci U S A* 88, 9828–9832. doi:10.1073/pnas.88.21.9828.

Monforte, A. J., Eduardo, I., Abad, S., and Arús, P. (2005). Inheritance mode of fruit traits in melon: Heterosis for fruit shape and its correlation with genetic distance. *Euphytica* 144, 31–38. doi:10.1007/s10681-005-0201-y.

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8, 4321–4326. doi:10.1093/nar/8.19.4321.

Nakagawa, N., Mori, H., Yamazaki, K., and Imaseki, H. (1991). Cloning of a complementary DNA for Auxin-Induced 1-Aminocyclopropane-1-carboxylate Synthase and differential expression of the gene by auxin and wounding. *Plant and Cell Physiology* 32, 1153–1163. doi:10.1093/oxfordjournals.pcp.a078192.

Nakajima, N., Mori, H., Yamazaki, K., and Imaseki, H. (1990). Molecular cloning and sequence of a complementary DNA encoding 1-Aminocyclopropane-l-carboxylate Synthase Induced by Tissue Wounding. *Plant and Cell Physiology* 31, 1021–1029. doi:10.1093/oxfordjournals.pcp.a077998.

Nimmakayala, P., Tomason, Y. R., Abburi, V. L., Alvarado, A., Saminathan, T., Vajja, V. G., et al. (2016). Genome-wide differentiation of various melon horticultural groups for use in GWAS for fruit firmness and construction of a high resolution genetic map. *Frontiers in Plant Science* 7, 1–15. doi:10.3389/fpls.2016.01437.

Ogata, T., Hirota, T., Shiozaki, S., Horiuchi, S., Kawase, K., and Ohashi, M. (2002). Effects of 2-aminoethoxyvinylglycine and high temperatures on fruit set and fruit characteristics of heat-cultured Satsuma Mandarin. *J Japan Soc Hortic Sci* 71, 348–354.

Olimpieri, I., Siligato, F., Caccia, R., Soressi, G. P., Mazzucato, A., Mariotti, L., et al. (2007). Tomato fruit set driven by pollination or by the parthenocarpic fruit allele are mediated by transcriptionally regulated gibberellin biosynthesis. *Planta* 226, 877–888. doi:10.1007/s00425-007-0533-z.

Oren, E., Tzuri, G., Vexler, L., Dafna, A., Meir, A., Faigenboim, A., et al. (2019). The multi-allelic *APRR2* gene is associated with fruit pigment accumulation in melon and watermelon. *Journal of experimental botany* 70, 3781–3794. doi:10.1093/jxb/erz182.

Owens, K. W., Peterson, C. E., and Tolla, G. E. (1980). Production of hermaphrodite flowers on gynoecious muskmelon by silver nitrate and aminoethoxyvinylglycine. *HortScience* 15, 654–655.

Pandolfini, T., Molesini, B., and Spena, A. (2007). Molecular dissection of the role of auxin in fruit initiation. *Trends in Plant Science* 12, 327–329. doi:10.1016/j.tplants.2007.06.011.

Papadopoulou, E., and Grumet, R. (2005). Brassinosteriod-induced femaleness in cucumber and relationship to ethylene production. *HortScience* 40, 1763–1767. doi:10.21273/hortsci.40.6.1763.

Paris, H. S. (2015). Origin and emergence of the sweet dessert watermelon, *Citrullus lanatus*. *Annals of Botany* 116, 133–148. doi:10.1093/aob/mcv077.

Paris, H. S., Daunay, M. C., and Janick, J. (2013). Medieval iconography of watermelons in Mediterranean Europe. *Annals of Botany* 112, 867–879. doi:10.1093/aob/mct151.

Pascual, L., Blanca, J. M., Cãizares, J., and Nuez, F. (2009). Transcriptomic analysis of tomato carpel development reveals alterations in ethylene and gibberellin synthesis during *pat3/pat4* parthenocarpic fruit set. *BMC Plant Biology* 9, 1–18. doi:10.1186/1471-2229-9-67.

Paul, V., Pande, R., and Srivastava, G. C. (2012). The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene-An overview. *Journal of Food Science and Technology* 49, 1–21. doi:10.1007/s13197-011-0293-4.

Payán, M. C., Peñaranda, A., Rosales, R., Garrido, D., Gómez, P., Jamilena, M., et al. (2006). Ethylene mediates the induction of fruits with attached flower in zucchini squash. In: Holmes, G.J. *Proceedings of Cucurbitaceae*. 171–179.

Pech, J.-C., Latché, A., and Bouzayen, M. (2010). Ethylene biosynthesis. In: Davies. *Springer, Dordrecht*. doi:10.1146/annurev.arplant.44.1.283.

Peñaranda, A., Payan, M. C., Garrido, D., Gómez, P., and Jamilena, M. (2007). Production of fruits with attached flowers in zucchini squash is correlated with the arrest of maturation of female flowers. *Journal of Horticultural Science and Biotechnology* 82, 579–584. doi:10.1080/14620316.2007.11512276.

Perin, C., Gomez-Jimenez, M. C., Hagen, L., Dogimont, C., Pech, J. C., Latché, A., et al. (2002). Molecular and genetic characterization of a non-climacteric phenotype in melon reveals two loci conferring altered ethylene response in fruit. *Plant Physiology* 129, 300–309. doi:10.1104/pp.010613.

Perl-Treves, R. (1999). Male to female conversion along the cucumber shoot: approaches to studying sex genes and floral development in *Cucumis sativu*s. In: Ainsworth, C.C. (ed) Sex determination in plants. *BIOS Scientifi*c, *Oxford*, 193–221.

Peterson, C. E., and Anhder, L. (1960). Induction of staminate flowers on gynoecious cucumbers with gibberellic A3. *Science*, 1673–1674.

Poole, C. F., and Grimball, P. C. (1945). Interaction of sex, shape, and weight genes in watermelon. *Journal of Agricultural Research, Washington, D C* 71, 533–552.

Poole, C., and Grimball, P. (1938). Inheritance of new sex forms in *Cucumis melo* L. *Journal of Heredity* 30, 21–25. doi:10.1093/oxfordjournals.jhered.a104626.

Potuschak, T., Lechner, E., Parmentier, Y., Yanagisawa, S., Grava, S., Koncz, C., et al. (2003). EIN3-dependent regulation of plant ethylene hormone signaling by two *Arabidopsis* F box proteins: EBF1 and EBF2. *Cell* 115, 679–689. doi:10.1016/S0092-8674(03)00968-1.

Prothro, J., Abdel-Haleem, H., Bachlava, E., White, V., Knapp, S., and McGregor, C. (2013). Quantitative Trait Loci Associated with Sex Expression in an Inter-subspecific Watermelon Population. *Journal of the American Society for Horticultural Science* 138, 125–130. doi:10.21273/jashs.138.2.125.

Pujol, M., Alexiou, K. G., Fontaine, A. S., Mayor, P., Miras, M., Jahrmann, T., et al. (2019). Mapping Cucumber Vein Yellowing Virus Resistance in Cucumber (*Cucumis sativus* L.) by Using BSA-seq Analysis. *Frontiers in Plant Science* 10, 1–10. doi:10.3389/fpls.2019.01583.

Qiao, H., Chang, K. N., Yazaki, J., and Ecker, J. R. (2009). Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in *Arabidopsis. Genes and Development* 23, 512–521. doi:10.1101/gad.1765709.

Qiao, H., Shen, Z., Huang, S. S. C., Schmitz, R. J., Urich, M. A., Briggs, S. P., et al. (2012). Processing and subcellular trafficking of ER-tethered EIN2 control response to ethylene gas. *Science* 338, 390–393. doi:10.1126/science.1225974.

Reddy, U. K., Abburi, L., Abburi, V. L., Saminathan, T., Cantrell, R., Vajja, V. G., et al. (2015). A genome-wide scan of selective sweeps and association mapping of fruit traits using microsatellite markers in watermelon. *Journal of Heredity* 106, 166–176. doi:10.1093/jhered/esu077.

Reddy, U. K., Aryal, N., Islam-Faridi, N., Tomason, Y. R., Levi, A., and Nimmakayala, P. (2013). Cytomolecular characterization of rDNA distribution in various *Citrullus* species using fluorescent in situ hybridization. *Genetic Resources and Crop Evolution* 60, 2091–2100. doi:10.1007/s10722-013-9976-1.

Reddy, U. K., Nimmakayala, P., Levi, A., Abburi, V. L., Saminathan, T., Tomason, Y. R., et al. (2014). High-resolution genetic map for understanding the effect of genome-wide recombination rate on nucleotide diversity in watermelon. *G3(Genes|Genomes|Genetics)* 4, 2219–2230. doi:10.1534/g3.114.012815.

Reichheld, J. P., Riondet, C., Delorme, V., Vignols, F., and Meyer, Y. (2010). Thioredoxins and

glutaredoxins in development. *Plant Science* 178, 420–423. doi:10.1016/j.plantsci.2010.03.001.

Renner, S. S., Chomicki, G., and Greuter, W. (2014). (2313) Proposal to conserve the name *Momordica lanata* (*Citrullus lanatus*) (watermelon, *Cucurbitaceae*), with a conserved type, against *Citrullus battich*. *Taxon* 63, 941–942. doi:10.12705/634.29.

Renner, S. S., Pérez-Escobar, O. A., Silber, M. V., Nesbitt, M., Preick, M., Hofreiter, M., et al. (2019). A 3500-year-old leaf from a Pharaonic tomb reveals that New Kingdom Egyptians were cultivating domesticated watermelon. *bioRxiv*, 642785. doi:10.1101/642785.

Renner, S. S., and Schaefer, H. (2016). Phylogeny and Evolution of the Cucurbitaceae. In: Grumet, R., Katzir, N., Garcia-Mas, J. (eds) Genetics and Genomics of Cucurbitaceae. Plant Genetics and Genomics: Crops and Models, vol 20. *Springer, Cham, 13-23.* doi:https://doi.org/10.1007/7397_2016_14.

Renner, S. S., Sousa, A., and Chomicki, G. (2017). Chromosome numbers, sudanese wild forms, and classification of the watermelon genus *Citrullus*, with 50 names allocated to seven biological species. *Taxon* 66, 1393–1405. doi:10.12705/666.7.

Robinson, R. W., and Decker-Walters, D. S. (1997). Cucurbits. CAB International, Wallingford, Oxon, U.K.

Robinson, R. W., Whitaker, T. W., and Bohn, G. W. (1970). Promotion of pistillate flowering in *Cucurbita* by 2-chloroethylphosphonic acid. *Euphytica* 19, 180–183. doi:10.1007/BF01902942.

Rosa, J. T. (1928). The inheritance of flower types in *Cucumis* and *Citrullus*. *Hilgardia* 3, 233–250. doi:10.3733/hilg.v03n09p233.

Ruan, Y. L., Patrick, J. W., Bouzayen, M., Osorio, S., and Fernie, A. R. (2012). Molecular regulation of seed and fruit set. *Trends in Plant Science* 17, 656–665. doi:10.1016/j.tplants.2012.06.005.

Rubatsky, V. (2001). Origin, distribution, and uses. In: Maynard DN (ed). Watermelons: characteristics, production, and marketing. Alexandria, VA: ASHS Press, 21–26.

Rudich, J. (1990). Biochemical aspects of hormonal regulation of sex expression in cucurbits. , eds. Bates, D., Robinson, R., and Jeffrey, C (eds). Cornell University Press, Ithaca, NY.

Rudich, J., Halevy, A. H., and Kedar, N. (1969). Increase in femaleness of three cucurbits by treatment with Ethrel, an ethylene releasing compound. *Planta* 86, 69–76. doi:10.1007/BF00385305.

Rudich, J., Halevy, A. H., and Kedar, N. (1972). The level of phytohormones in monoecious and

gynoecious cucumbers as affected by photoperiod and ethephon. *Plant Physiology* 50, 585–590. doi:10.1104/pp.50.5.585.

Rudich, J., and Peles, A. (1976). Sex expression in watermelon as affected by photoperiod and temperature. *Scientia Horticulturae* 5, 339–344. doi:10.1016/0304-4238(76)90129-1.

Rudich, J., and Zamski, E. (1985). *Citrullus lanatus. In: Halevy, A (ed) Handbook of flowering CRC Press, Boca Raton, FL*, 272–274.

Ruduś, I., Sasiak, M., and Kepczyński, J. (2013). Regulation of ethylene biosynthesis at the level of 1-aminocyclopropane-1-carboxylate oxidase (*ACO*) gene. *Acta Physiologiae Plantarum* 35, 295–307. doi:10.1007/s11738-012-1096-6.

Sain, R. S., Joshi, P., and Divakara Sastry, E. V. (2002). Cytogenetic analysis of interspecific hybrids in genus *Citrullus* (Cucurbitaceae). *Euphytica* 128, 205–210. doi:10.1023/A:1020800113252.

Sakata, Y., Fukino, N., Ohara, T., Sugiyama, M., and Yoshioka, Y. (2013). Effect of monoecious trait on fruit shape and total soluble solid contents of Melon. *Hort Res (Japan)* 12, 15–22.

Salman-Minkov, A., Levi, A., Wolf, S., and Trebitsh, T. (2008). ACC synthase genes are polymorphic in watermelon (*Citrullus* spp.) and differentially expressed in flowers and in response to auxin and gibberellin. *Plant and Cell Physiology* 49, 740–750. doi:10.1093/pcp/pcn045.

Sanders, P. M., Bui, A. Q., Weterings, K., McIntire, K. N., Hsu, Y. C., Lee, P. Y., et al. (1999). Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sexual Plant Reproduction* 11, 297–322. doi:10.1007/s004970050158.

Sauer, J. D. (1993). *Historical geography of crop plants: a select roster*. CRC Press.

Schaefer, H., Heibl, C., and Renner, S. S. (2009). Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society B: Biological Sciences* 276, 843–851. doi:10.1098/rspb.2008.1447.

Schaefer, H., and Renner, S. S. (2011). Cucurbitaceae, in The Families and Genera of Vascular Plants. In: Kubitzki K. (ed). Berlin: Springer-Verlag, 112–174.

Shakeel, S. N., Wang, X., Binder, B. M., and Schaller, G. E. (2013). Mechanisms of signal transduction by ethylene: Overlapping and non-overlapping signalling roles in a receptor family. *AoB PLANTS* 5, 1–16. doi:10.1093/aobpla/plt010.

Shannon, S., and De La Guardia, M. (1969). Sex expression and the production of ethylene induced by auxin in the cucumber (*Cucumis sativus* L.). *Nature*, 186–223.

Shiber, A., Gaur, R. K., Rimon-Knopf, R., Zelcer, A., Trebitsh, T., and Pitrat, M. (2008). The origin

and mode of function of the *Female* locus in cucumber. *Proceedings of the IXth EUCARPIA meeting on genetics and breeding of Cucurbitaceae*, 263-270.

Singh, A. K. (1978). Cytogenetics of Semi Arid Plants III . A natural interspecific hybrid of Cucurbitacae (*Citrullus colocynthis* x *vulgaris* Schrad). *Cytologia* 43, 569–574.

Singh, S., Yadav, S., Singh, A., Mahima, M., Singh, A., Gautam, V., et al. (2020). Auxin signaling modulates *LATERAL ROOT PRIMORDIUM1* (*LRP1*) expression during lateral root development in Arabidopsis. *Plant Journal* 101, 87–100. doi:10.1111/tpj.14520.

Singh, Z., and Agrez, V. (2002). Fruit set, retention and yield of mango in relation to ethylene. *Acta Horticulturae* 575, 805–811. doi:10.17660/ActaHortic.2002.575.95.

Sneath, P., and Sokal, R. (1973). Numerical Taxonomy: The Principles and Practice of Numerical Classification. WF Freeman & Co., San Francisco, 573.

Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998). Nuclear events in ethylene signaling: A transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes and Development* 12, 3703–3714. doi:10.1101/gad.12.23.3703.

Song, J., Li, Z., Liu, Z., Guo, Y., and Qiu, L. J. (2017). Next-generation sequencing from bulked-segregant analysis accelerates the simultaneous identification of two qualitative genes in soybean. *Frontiers in Plant Science* 8, 1–11. doi:10.3389/fpls.2017.00919.

Stepanova, A. N., Robertson-Hoyt, J., Yun, J., Benavente, L. M., Xie, D. Y., Doležal, K., et al. (2008). *TAA1*-Mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* 133, 177–191. doi:10.1016/j.cell.2008.01.047.

Sugimaya, K., Kanno, T., and Morishita, M. (1998). Evaluation method of female flower bearing ability in watermelon using silver thiosulfate (STS). *J Japan Soc Hortic Sci* 38, 226–226. doi:10.1248/cpb.37.3229.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., et al. (2013). QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant Journal* 74, 174–183. doi:10.1111/tpj.12105.

Takakura, Y., Ito, T., Saito, H., Inoue, T., Komari, T., and Kuwata, S. (2000). Flower-predominant expression of a gene encoding a novel class I chitinase in rice (*Oryza sativa* L.). *Plant Molecular Biology* 42, 883–897. doi:10.1023/A:1006401816145.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24, 1596–1599. doi:10.1093/molbev/msm092.

Tan, J., Tao, Q., Niu, H., Zhang, Z., Li, D., Gong, Z., et al. (2015). A novel allele of *monoecious* (*m*) locus is responsible for elongated fruit shape and perfect flowers in cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics* 128. doi:10.1007/s00122-015-2603-0.

Tao, Y., Niu, Y., Wang, Y., Chen, T., Naveed, S. A., Zhang, J., et al. (2018). Genome-wide association mapping of aluminum toxicity tolerance and fine mapping of a candidate gene for *Nrat1* in rice. *PLoS ONE* 13, 1–21. doi:10.1371/journal.pone.0198589.

Trebitsh, T., Rudich, J., and Riov, J. (1987). Auxin, biosynthesis of ethylene and sex expression in cucumber (*Cucumis sativus*). *Plant Growth Regulation*, 105–113.

Trebitsh, T., Staub, J. E., and O'Neill, S. D. (1997). Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the *female* (*F*) locus that enhances female sex expression in cucumber. *Plant Physiology* 113, 987–995. doi:10.1104/pp.113.3.987.

Tsai, W. C., Hsiao, Y. Y., Pan, Z. J., Kuoh, C. S., Chen, W. H., and Chen, H. H. (2008). The role of ethylene in orchid ovule development. *Plant Science* 175, 98–105. doi:10.1016/j.plantsci.2008.02.011.

Tsukahara, A. (1988). Semi-forcing culture in watermelon. *In: The system of agricultural technology, melon and watermelon Soc Cult Agr Vill Press, Tokyo*, 194–196.

Vandenbussche, M., Horstman, A., Zethof, J., Koes, R., Rijpkema, A. S., and Gerats, T. (2009). Differential recruitment of *WOX* transcription factors for lateral development and organ fusion in Petunia and *Arabidopsis*. *Plant Cell* 21, 2269–2283. doi:10.1105/tpc.109.065862.

Vriezen, W. H., Feron, R., Maretto, F., Keijman, J., and Mariani, C. (2008). Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. *New Phytologist* 177, 60–76. doi:10.1111/j.1469-8137.2007.02254.x.

Wall, J. R. (1967). Correlated inheritance of sex expression and fruit shape in *Cucumis*. *Euphytica*, 199–208.

Wang, D. H., Li, F., Duan, Q. H., Han, T., Xu, Z. H., and Bai, S. N. (2010). Ethylene perception is involved in female cucumber flower development. *Plant Journal* 61, 862–872. doi:10.1111/j.1365-313X.2009.04114.x.

Wang, K. L., Li, H., and Ecker, J. R. (2002). Ethylene Biosynthesis and Signaling Networks. *Plant Cell* 14, 131–152. doi:10.1105/tpc.001768.S-.

Wang, X., Bao, K., Reddy, U. K., Bai, Y., Hammar, S. A., Jiao, C., et al. (2018). The USDA cucumber (*Cucumis sativus* L.) collection: genetic diversity, population structure, genome-wide association studies, and core collection development. *Horticulture Research* 5. doi:10.1038/s41438-018-0080-8.

Wang, X., Goshe, M. B., Soderblom, E. J., Phinney, B. S., Kuchar, J. A., Li, J., et al. (2005). Identification and functional analysis of in vivo phosphorylation sites of the Arabidopsis BRASSINOSTEROID-INSENSITIVE1 receptor kinase. *Plant Cell* 17, 1685–1703. doi:10.1105/tpc.105.031393.tion.

Watanabe, T., Seo, S., and Sakai, S. (2001). Wound-induced expression of a gene for 1-aminocyclopropane-1-carboxylate synthase and ethylene production are regulated by both reactive oxygen species and jasmonic acid in *Cucurbita maxima*. *Plant Physiology and Biochemistry* 39, 121–127. doi:10.1016/S0981-9428(00)01224-9.

Wechter, W. P., Levi, A., Harris, K. R., Davis, A. R., Fei, Z., Katzir, N., et al. (2008). Gene expression in developing watermelon fruit. *BMC Genomics* 9, 1–13. doi:10.1186/1471-2164-9-275.

Wehner, T. C. (2008). Watermelon. In: J. Prohens, J., Nuez, F., (eds). Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae, and Cucurbitaceae. *Springer*. doi:10.1007/978-0-387-30443-4.

Wen, X., Zhang, C., Ji, Y., Zhao, Q., He, W., An, F., et al. (2012). Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. *Cell Research* 22, 1613–1616. doi:10.1038/cr.2012.145.

Whitaker, T. W., and Davis, G. N. (1962). Cucurbits: Botany, cultivation, and utilization. Leonard Hill, London (United Kingdom), 249

Wien, H. (1997). The physiology of vegetable crops. *Wallingford: CABI*.

Wien, H., Stapleton, S., Maynard, D., McClurg, C., Nyankanga, R., and Riggs, D. (2002). Regulation of female flower development in pumpkin (*Cucurbita* spp.) by temperature and light. In: Maynard, D. N. (ed). *Cucurbitaceae.* Alexandria: ASHS Press, 307–15.

Wu, S., Wang, X., Reddy, U., Sun, H., Bao, K., Gao, L., et al. (2019). Genome of 'Charleston Gray', the principal American watermelon cultivar, and genetic characterization of 1,365 accessions in the U.S. National Plant Germplasm System watermelon collection. *Plant Biotechnology Journal*, 1–13. doi:10.1111/pbi.13136.

Xing, L., Li, Z., Khalil, R., Ren, Z., and Yang, Y. (2012). Functional identification of a novel F-box/FBA gene in tomato. *Physiologia Plantarum* 144, 161–168. doi:10.1111/j.1399-3054.2011.01543.x.

Xing, S., and Zachgo, S. (2008). *ROXY1* and *ROXY2*, two Arabidopsis glutaredoxin genes, are required for anther development. *Plant Journal* 53, 790–801. doi:10.1111/j.1365-313X.2007.03375.x.

Xu, D., Sun, D., Diao, Y., Liu, M., Gao, J., Wu, B., et al. (2019). Fast mapping of a chlorophyll b synthesis-deficiency gene in barley (*Hordeum vulgare* L.) via bulked-segregant analysis with reduced-representation sequencing. *Crop Journal* 7, 58–64. doi:10.1016/j.cj.2018.07.002.

Xu, J., and Zhang, S. (2014). Regulation of ethylene biosynthesis and signaling by protein kinases and phosphatases. *Molecular Plant* 7, 939–942. doi:10.1093/mp/ssu059.

Yagcioglu, M., Gulsen, O., Yetisir, H., Solmaz, I., and Sari, N. (2016). Preliminary studies of genom-wide association mapping for some selected morphological characters of watermelons. *Scientia Horticulturae* 210, 277–284. doi:10.1016/j.scienta.2016.08.001.

Yamasaki, S., Fujii, N., Matsuura, S., Mizusawa, H., and Takahashi, H. (2001). The *M* locus and ethylene-controlled sex determination in andromonoecious cucumber plants. *Plant and Cell Physiology* 42, 608–619. doi:10.1093/pcp/pce076.

Yamasaki, S., Fujii, N., and Takahashi, H. (2003). Photoperiodic regulation of *CS-ACS2*, *CS-ACS4* and *CS-ERS* gene expression contributes to the femaleness of cucumber flowers through diurnal ethylene production under short-day conditions. *Plant, Cell and Environment* 26, 537–546. doi:10.1046/j.1365-3040.2003.00984.x.

Yang, L., Wu, Y., Yu, M., Mao, B., Zhao, B., and Wang, J. (2016). Genome-wide transcriptome analysis of female-sterile rice ovule shed light on its abortive mechanism. *Planta* 244(5), 1011–1028. doi:10.1007/s00425-016-2563-x.

Yang, S. F., and Hoffman, N. E. (1984). Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology* 35, 155–89. doi:https://doi.org/10.1146/annurev.pp.35.060184.001103.

Yoshida, H., Nagata, M., Saito, K., Kevin, W. L. C., and Ecker, J. R. (2005). *Arabidopsis* ETO1 specifically interacts with and negatively regulates type 2 1-aminocyclopropane-1-carboxylate synthases. *BMC Plant Biology* 5, 1–13. doi:10.1186/1471-2229-5-14.

Yu, H., Wu, J., Xu, N., and Peng, M. (2007). Roles of F-box proteins in plant hormone responses. *Acta Biochimica et Biophysica Sinica* 39, 915–922. doi:10.1111/j.1745-7270.2007.00358.x.

Yu, L. P., Miller, A. K., and Clark, S. E. (2003). *POLTERGEIST* encodes a protein phosphatase 2C that regulates CLAVATA pathways controlling stem cell identity at *Arabidopsis* shoot and flower meristems. *Current Biology* 13, 179–188. doi:10.1016/S0960-9822(03)00042-3.

Zhang, H., Li, X., Yu, H., Zhang, Y., Li, M., Wang, H., et al. (2019a). A High-Quality Melon Genome Assembly Provides Insights into Genetic Basis of Fruit Trait Improvement. *iScience* 22, 16–27. doi:10.1016/j.isci.2019.10.049.

Zhang, J., Guo, S., Ji, G., Zhao, H., Sun, H., Ren, Y., et al. (2019b). A unique chromosome

translocation disrupting *ClWIP1* leads to gynoecy in watermelon. *The Plant Journal* 2007, 1–13. doi:10.1111/tpj.14537.

Zhang, J., Shi, J., Ji, G., Zhang, H., Gong, G., Guo, S., et al. (2017). Modulation of sex expression in four forms of watermelon by gibberellin , ethephone and silver nitrate. *Horticultural Plant Journal* 3, 91–100. doi:10.1016/j.hpj.2017.07.010.

Zhang, T. C., Qiao, Q., and Zhong, Y. (2012). Detecting adaptive evolution and functional divergence in aminocyclopropane-1-carboxylate synthase (ACS) gene family. *Computational Biology and Chemistry* 38, 10–16. doi:10.1016/j.compbiolchem.2012.04.001.

Zhang, T., Liu, J., Liu, S., Ding, Z., Luan, F., and Gao, P. (2019c). Bulked-segregant analysis identified a putative region related to short internode length in melon. *HortScience* 54, 1293–1298. doi:10.21273/HORTSCI14052-19.

Zhang, Z., Mao, L., Chen, H., Bu, F., Li, G., Sun, J., et al. (2015). Genome-wide mapping of structural variations reveals a copy number variant that determines reproductive morphology in cucumber. *Plant Cell* 27, 1595–1604. doi:10.1105/tpc.114.135848.

Zhao, Q., and Guo, H. W. (2011). Paradigms and paradox in the ethylene signaling pathway and interaction network. *Molecular Plant* 4, 626–634. doi:10.1093/mp/ssr042.

Zhao, X., Niu, X., and Fan, M. (2015). Adventitious shoots induction and plant regeneration from cotyledons of watermelon (*Citrullus lanatus* L.). *African Journal of Biotechnology* 14, 2202–2210. doi:10.5897/ajb2015.14741.

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J. C., Yu, H., Sakai, H., et al. (2004). Hanaba taranu is a GATA transcription factor that regulates shoot apical meristem and flower development in Arabidopsis. *Plant Cell* 16, 2586–2600. doi:10.1105/tpc.104.024869.

Zheng, Y., Wu, S., Bai, Y., Sun, H., Jiao, C., Guo, S., et al. (2019). Cucurbit Genomics Database (CuGenDB): A central portal for comparative and functional genomics of cucurbit crops. *Nucleic Acids Research* 47, D1128–D1136. doi:10.1093/nar/gky944.

Zheng, Y., Xu, F., Li, Q., Wang, G., Liu, N., Gong, Y., et al. (2018). QTL Mapping combined with Bulked Segregant Analysis identify SNP markers linked to leaf shape traits in *Pisum sativum* using SLAF sequencing. *Frontiers in Genetics* 9, 1–11. doi:10.3389/fgene.2018.00615.

Zhong, R., Kays, S. J., Schroeder, B. P., and Ye, Z.-H. (2002). Mutation of a Chitinase-Like Gene Causes Ectopic Deposition of Lignin, Aberrant Cell Shapes, and Overproduction of Ethylene. *The Plant Cel* 14, 165–179. doi:10.1105/tpc.010278.

Zhu, H., Zhai, W., Li, X., and Zhu, Y. (2019a). Two QTLs controlling Clubroot resistance identified from Bulked Segregant Sequencing in Pakchoi (*Brassica campestris* ssp. *chinensis* Makino).

Scientific Reports 9, 1–9. doi:10.1038/s41598-019-44724-z.

Zhu, H., Zhang, M., Sun, S., Yang, S., Li, J., Li, H., et al. (2019b). A Single Nucleotide Deletion in an ABC Transporter Gene Leads to a Dwarf Phenotype in Watermelon. *Frontiers in Plant Science* 10, 1–13. doi:10.3389/fpls.2019.01399.

Zohary, D., Horf, M., and Weiss, E. (2012). Domestication of plants in the Old World, 4th end. Oxford. *Clarendon Press*, 153–155. doi:10.1093/acprof:osobl/9780199549061.003.0001.

Zou, T., Zhou, D., Li, W., Yuan, G., Tao, Y., He, Z., et al. (2020). A silent exonic mutation in a rice integrin-α FG-GAP repeat-containing gene causes male-sterility by affecting mRNA splicing. *International Journal of Molecular Sciences* 21. doi:10.3390/ijms21062018.

Zuckerkandl, E., and Pauling, L. (1965). Molecules as documents of history. *Journal of Theoretical Biology* 8, 357–366.

PLOS ONE

RESEARCH ARTICLE

The Ethylene Biosynthesis Gene *CitACS4* Regulates Monoecy/Andromonoecy in Watermelon (*Citrullus lanatus*)

Susana Manzano, Encarnación Aguado, Cecilia Martínez, Zoraida Megías, Alicia García, Manuel Jamilena*

Departamento de Biología y Geología, Agrifood Campus of International Excellence (ceiA3) and BITAL, Universidad de Almería, La Cañada de San Urbano s/n, 04120, Almería, Spain

* mjamille@ual.es

Citation: Manzano S, Aguado E, Martínez C, Megías Z, García A, Jamilena M (2016) The Ethylene Biosynthesis Gene *CitACS4* Regulates Monoecy/Andromonoecy in Watermelon (*Citrullus lanatus*). PLoS ONE 11(5): e0154362. doi:10.1371/journal.pone.0154362

Editor: Hiroshi Ezura, University of Tsukuba, JAPAN

Received: January 19, 2016

Accepted: April 12, 2016

Published: May 5, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by grant AGL2014-54598-C2-1-R, partly funded by ERDF (European Regional Development Fund) and by the Spanish Ministry of Science and Innovation, and grant P12-AGR-1423, funded by Junta de Andalucía, Spain. Z.M. and E.A. acknowledge FPU and "garantía juvenil" scholarship programmes from MEC, Spain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Monoecious and andromonoecious cultivars of watermelon are characterised by the production of male and female flower or male and hermaphrodite flowers, respectively. The segregation analysis in the offspring of crosses between monoecious and andromonoecious lines has demonstrated that this trait is controlled by a single gene pair, being the monoecious allele *M* semi-dominant to the andromonoecious allele *A*. The two studied F1 hybrids (MA) had a predominantly monoecious phenotype since both produced not only female flowers, but also bisexual flowers with incomplete stamens, and hermaphrodite flowers with pollen. Given that in other cucurbit species andromonoecy is conferred by mutations in the ethylene biosynthesis genes *CmACS7*, *CsACS2* and *CpACS27A* we have cloned and characterised *CitACS4*, the watermelon gene showing the highest similarity with the formers. *CitACS4* encoded for a type ACS type III enzyme that is predominantly expressed in pistillate flowers of watermelon. In the andromonoecious line we have detected a missense mutation in a very conserved residue of CitACS4 (C364W) that cosegregates with the andromonoecious phenotype in two independent F2 populations, concomitantly with a reduction in ethylene production in the floral buds that will develop as hermaphrodite flowers. The gene does not however co-segregates with other sex expression traits regulated by ethylene in this species, including pistillate flowering transition and the number of pistillate flowers per plant. These data indicate that *CitAC4* is likely to be involved in the biosynthesis of the ethylene required for stamen arrest during the development of female flowers. The C364W mutation would reduce the production of ethylene in pistillate floral buds, promoting the conversion of female into hermaphrodite flowers, and therefore of monoecy into andromonoecy.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

The cultivated species of the *Cucurbitaceae* family, including melon, cucumber, watermelon, squash and gourds, are monoecious, developing unisexual male and female flowers on the same individual plant. Evolution has led, however, to a number of sex morphotypes in the species of this family, including andromonoecious (plant produces male and bisexual flowers), gynoecious (only female flowers), androecious (only male flowers) and hermaphrodite (only hermaphrodite flowers) lines. All these sex morphotypes have been detected in melon [1–3]), cucumber [4, 5] and watermelon [6–8]. In squash, the predominant monoecious cultivars coexist with partially andromonoecious ones [9], and some androecious mutants have been also recently described [10, 11], but no gynoecious squash have been identified so far.

Sex determination in this family is mainly controlled by the gaseous hormone ethylene. It has long been known that external treatment with ethylene favours the formation of female flowers in monoecious cultivars of melon, cucumber and squash, while the application of inhibitors of ethylene biosynthesis and response, including aminoethoxyvinylglycine (AVG) or silver thiosulphate (STS), favours the development of male flowers [12–17]. Moreover, in melon and cucumber, the best characterised species of the family, the existence of several of the sexual morphotypes described is controlled by this hormone. Thus, the andromonoecious morphotype in cucumber, melon and zucchini squash, result from mutations in the three orthologous ethylene biosynthesis genes *CmACS7*, *CsACS2* and *CpACS27A*, respectively [9, 18, 19]. These genes are expressed only in pistillate flower primordia and are responsible for the arrest of stamens during the development of unisexual female flowers. The gynoecy of cucumber also depends on an additional ACS gene which is only present in the gynoecious varieties [20–22]. However, in melon gynoecy results from a transposon-mediated mutation in the promoter of the transcription factor *CmWIP1*, a negative regulator of *CmACS7*, responsible of the abortion of carpels and the promotion of stamen development [23]. The genes responsible for androecy in melon and cucumber have been recently characterised. They correspond to *CmACS11* and *CsACS11*, both involved in the biosynthesis of ethylene in the phloem of flowers programmed to become females, and in melon this gene functions as a negative regulator of the male-promoting transcription factor gene *CmWIP1* [24].

Sex determination mechanisms in watermelon have received little attention. Ethylene is also an important regulator of sex in this species, although external treatments with the hormone induce the production of male flowers, [25], while treatments with ethylene inhibitors hasten the appearance of the first female flower and increase the number of female flowers per plant [25–27], which it is contrary to what happens in the other cucurbit species. Recently we have differentiated between two sex related processes: sex expression, i.e. the earliness and production of female flowers per plant, and sex determination, as the mechanism that leads to the proper development and differentiation of unisexual female and male flowers [27]. In contrast to what happens in other cucurbits, ethylene inhibits the transition from male to female flowering and reduces the number of female flowers per plant. Nevertheless, as in other cucurbit species, ethylene is necessary for the arrest of stamen development during the proper development of the female flower, and the reduction of ethylene production or action lead to the transformation of female into bisexual and hermaphrodite flowers [27]. In this paper it is shown that *CitACS4*, an homologous gene to *CmACS7*, *CsACS2* and *CpACS27A* of melon, cucumber and squash, is responsible for the arrest of stamens in female flower development, and that a recessive mutation in this gene reduces the production of ethylene in the floral bud, and leads to the conversion of female into bisexual or hermaphrodite flowers, and therefore monoecy into andromonoecy.

Materials and Methods

Plant material, growing conditions and phenotyping

Three inbred lines of watermelon (*Citrullus. lanatus*) two monoecious lines (P85 and P86) and one andromonoecious line (P87) were characterised in this paper. The F1 and F2 generations from two independent crosses (P85 x P87 and P86 x P87) were used to determine the inheritance of monoecy/andromonoecy in this species. The crosses were performed in spring-summer seasons of 2012 and 2013, and the final phenotyping carried out in plants grown under standard greenhouse conditions in the province of Almería (Spain) in the spring-summer of 2014 and 2015.

To evaluate monoecy in the different inbred lines and populations, the so-called Andromonoecy Index (AI, [9]) were defined for each flower, plant and population. Pistillate flowers were scored from 1 to 3 according to their degree of stamen development. Female flowers with no stamen development were scored as AI = 1, while hermaphrodite flowers with complete stamens and anthers able to produce pollen were scored as AI = 3. A score of 2 was assigned to bisexual flowers not producing pollen with medium-sized stamens and anthers (Fig 1A). Based on the flower scores, the AI of each plant in a population was calculated as the average score for at least five pistillate flowers. The average AI for inbred lines or F1 was then estimated from at least 10 plants with a minimum of 5 pistillate flowers evaluated per plant. Plants and genotypes with an AI = 1–1.2 were considered to be monoecious, while those with AI = 1.2–2.7, partially andromonoecious, and those with AI $\geq$2.7 were phenotyped as andromonoecious.

Sex expression in each plant was assessed by both the number of initial nodes with male flowers before the production of the first pistillate flower in the main shoot (pistillate flowering transition), and the percentage of pistillate flowers per plant in the first 20 nodes of the main shoot. At least 10 plants were phenotyped to assess the sexual expression of each genotype.

Cloning and molecular characterization of *CitACS4*

To identify the watermelon ortholog for *CmACS7*, *CsACS2* and *CpACS27A* [9, 18, 19], we blasted the coding sequences of the known genes on the watermelon genome at Cucurbit Genome Database (http://www.icugi.org). Thereby a watermelon *ACS* gene having the highest homology with the formers was identified. The gene, called *CitACS4*, was cloned and characterised in monoecious and andromonoecious lines. Specific primers *CitACS4gen-F1/R1* and *Fw/Rw* (S1 Table) were designed to amplify a genomic region of 1332 bp, covering the complete sequence of *CitACS4* from P85, P86 and P87 genomes.

For the phylogenetic analysis, alignments were performed using Clustalw at GenomeNet Database Resources (http://www.genome.jp/tools/clustalw/), and the MEGA4 software [28], which allowed the alignment of proteins and the construction of phylogenetic trees using the UPGMA method [29], with 2,000 replicates bootstrap [30]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [31] and are in the units of the number of amino acid substitutions per site. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 519 positions in the final dataset.

Genotyping *M* and *A* alleles of *CitACS4*

We have detected a single nucleotide polymorphism (SNP) between monoecious (P85 and P86) and andromonoecious (P87) lines that produce an amino acid substitution of a cysteine

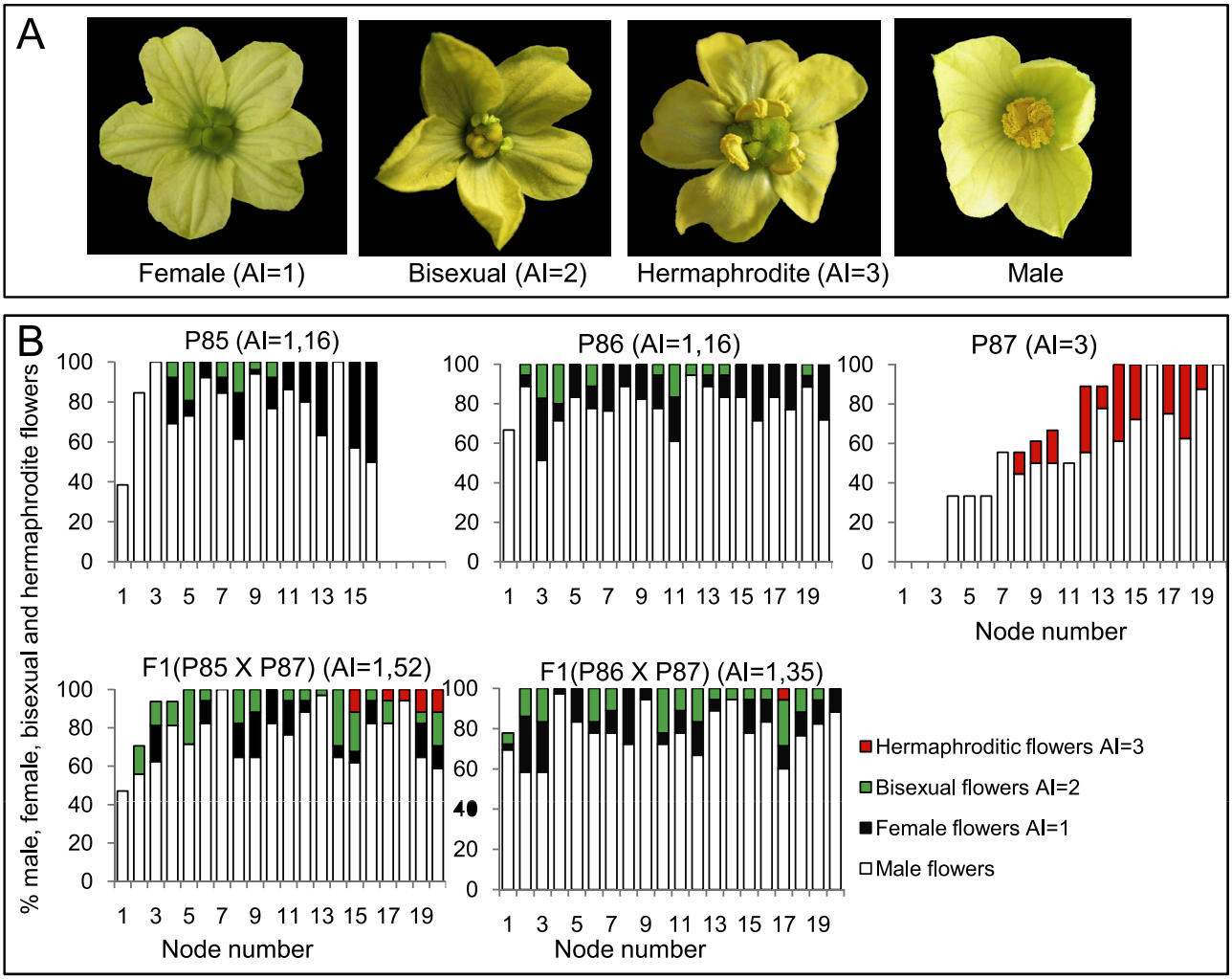


Fig 1. Sexual expression of watermelon lines P86, P86 and P87 and F1 hybrids derived from crosses P85xP87 and P86xP87. (A) Phenotype of watermelon hermaphrodite, bisexual, female and male flowers. (B) Distribution of staminate and pistillate flowers in the 20 first nodes of the main shoot. In each node, white, black, green and red bars represent the percentages of male, female, bisexual and hermaphrodite flowers in the total number of plants analysed ($n \geq 10$ for each genotype). The lack of bar in a node indicates the absence of flower in that node for some of the analysed plants.

doi:10.1371/journal.pone.0154362.g001

to a tryptophan in residue 364 of the CitACS4 protein (C364W). The respective alleles of for *CitACS4* in monoecious and andromonoecious lines were called *M* and *A*, respectively.

To genotype these two alleles in parental lines, and F1 and F2 generations, we used the specific primer pair *CitACS4MF/CitACS4gen-R1* or *CitACS4S-F/CitACS4M-R* (S1 Table), which were designed to specifically amplify the *M* allele, and primer pair *CitACS4A-F/CitACS4gen-R1* or *CitACS4S-F/CitACS4A-R*, that only amplified the allele *A*. DNA was isolated from frozen young leaves using the CTAB method [32]. 15–35 ng of purified DNA was used to amplify by PCR a 253 or 271 bp fragments of *CitACS4* gene. The amplifications were performed using the GeneAmp PCR System 2700 (Applied Biosystems) and PCR reactions consisted of 35 cycles of 30 s at 95°C, 30 s at 60°C and 90 s at 72°C. PCR fragments were resolved in agarose gels at 1.3%.

Ethylene production and quantitative RT-PCR

The production of ethylene and the expression of *CitACS4* gene were studied in flower buds throughout four different stages of floral development (S0 to S3). The different developmental stages were separated on the basis of the corolla length: S0 = 4 ± 1 mm, S1 = 8 ± 2 mm, S2: 12 ± 2 mm, S3: 15 ± 2 mm [33]. Ethylene was determined in three biological replicates per sample, each one containing three female, hermaphrodite or male flowers at the same stage of development. Floral buds were excised from the plant and incubated at room temperature for 6 h in hermetic glass containers in the dark. Ethylene production was determined by analysing 1 ml of gas from the headspace on a Varian 3900 gas chromatograph apparatus, fitted with a flame ionization detector. The instrument was calibrated with standard ethylene gas. At least three technical replicates were made for each biological sample.

Gene expression analysis was performed on three biological replicates per sample. Each replication was the result of an independent extraction of total RNA from 3 different flowers at the same stage of development. RNA extractions were performed according to the protocol of the GeneJET Plant RNA Purification Kit (Thermo). The remaining DNA in RNA samples was eliminated by digestion with RQ1 RNAse free DNAse (Promega). cDNA was then synthesized from 500 ng of total RNA using RevertAid RT Reverse Transcription Kit (Thermo). The expression of genes was evaluated through quantitative RT-PCR by using the Rotorgene thermocycler (Qiagen) and SYBR® Green Master Mix (BioRad). S1 Table shows the different primers used. The q-PCR primers were designed from the 3′ non-coding regions of each gene by using the Primer Express v 2.0 (Applied Biosystem) software. To avoid possible cross-amplification, and before any q-PCR experiment, the size of the PCR products for each pair of primers was tested in agarose gels, and sequenced. Quantitative RT-PCR reactions consisted of 40 cycles of 20 s at 95°C, 15 s at 59°C and 20 s at 60°C.

Relative expression of each gene was determined by the comparative Ct (Cycle Threshold) method using *C. pepo* 18S ribosomal RNA and *ACTIN* genes as internal standards. To use this method, it was first demonstrated that the efficiency of amplification for each amplicon was roughly equivalent, regardless of the amount of template cDNA. The absolute value of the slope of ΔCt (Ct of the target gene-Ct of the reference gene) versus serial dilutions of cDNA for a given sample must be less than 0.1. The relative expression of each gene was then calculated relative to a calibrator sample using the formula $2^{-\Delta\Delta Ct}$, where ΔΔCt is the difference between the ΔCt of each sample and the ΔCt of the calibrator sample.

Statistical analysis

Simple and factorial analyses of variance (ANOVA) at $p < 0.05$ were performed by the STATISTIX 8.0 software package, and each two means were compared with the method of Fisher's least significant difference (LSD) or Tukey's multiple comparison test.

Results

Phenotypic and genetic characterisation of monoecious and andromonoecious lines of watermelon

The sexual phenotype of three watermelon-inbred lines (P85, P86 and P87) were studied by phenotyping staminate and pistillate flowers in the first 20 nodes of the main shoot in at least 10 plants per genotype. Given that the development of stamens in pistillate flowers was variable, these flowers were classified and scored according to their stamen development using the Andromonoecy Index (AI, [9]). The female flowers with no stamen development were scored as AI = 1, while hermaphrodite flowers with complete stamens and pollen were scored as

AI = 3. Ovary-bearing flowers with intermediate stamen development and no pollen production were classified as bisexual and scored as AI = 2 (Fig 1A). The AI of each plant, genotype and progeny was then calculated as the average score of a minimum of five pistillate flowers in each plant, and at least 10 plants for each genotype or progeny.

The distribution of staminate and pistillate flowers of the three inbred lines along the 20 first nodes of the plant are shown in Fig 1B. The sexual phenotype of line P87 was very stable for andromonoecy condition (AI = 3). Under our conditions P87 plants only produced staminate and hermaphrodite flowers with complete stamens and pollen (AI = 3). Lines P85 and P86 were monoecious, since the predominantly produced female flowers, but also produced bisexual flowers, which resulted in AI = 1.16 for both P85 and P86. On the basis of these results, plants and genotypes with AI = 1–1.19 were considered to be monoecious, those with AI = 1.2–2.69, partially andromonoecious, and those with AI≥2.7 were considered andromonoecious.

The sexual phenotype of the two F1 hybrids derived from crosses between monoecious and andromonoecious lines (P85xP87 and P86xP87) had an intermediate phenotype between monoecious and andromonoecious, and were therefore classified as partially monoecious (Fig 1B). The two F1 populations had an intermediate AI (1.52 and 1.35), since both produced not only female, but also bisexual and hermaphrodite flowers (Fig 1B), suggesting that the monoecy allele in these two lines of watermelon is a semi-dominant trait in respect of andromonoecy. The segregation of monoecious, andromonoecious and partial andromonoecious plants in the two F2 generations studied demonstrated that the trait is controlled by a single gene pair, being the monoecious allele (M) incompletely dominant over the andromonoecious allele (A). As expected, the segregation of monoecious, partially andromonoecious and andromonoecious plants in the two F2 populations fitted the 1:2:1 ratio, as expected if the homozygous plants MM and AA were monoecious and andromonoecious, respectively, while heterozygous plants MA had an intermediate phenotype between monoecy and andromonoecy, although predominantly monoecious (Table 1).

Cloning and characterisation of *CitACS4*

Since in melon, cucumber and squash the andromonoecious phenotype is caused by mutations in the orthologs *CmACS7*, *CsACS2* and *CpACS27A* [9, 18–19], a homology analysis was performed to identify the watermelon *ACS* gene showing the highest similarity with the former. The nucleotide sequences of these homologous genes were blasted on watermelon genome at

Table 1. Segregation ratio of monoecious, partially andromonoecious and andromonoecious plants in F2 populations derived from two crosses between monoecious and andromonoecious inbred lines.

	No. of plants					
Generation	**Monoecious**	**Partially andromonoecious**	**Andro-monoecious**	**Expected segregation**	**χ^2**	***p-value***
Parental P87	0	0	10	-	-	-
Parental P85	13	0	0	-	-	-
Parental P86	18	0	0	-	-	-
F1 (P85XP87)	0	17	0	-	-	-
F1 (P86XP87)	3	15	0	-	-	-
F2 (P85XP87)	27	41	24	1:2:1	0.34	0.53
F2 (P86XP87)	24	34	13	1:2:1	1.02	0.17

The F2 plants were phenotyped on the basis of their average AI, scored from at least 5 flowers per plant. Monoecious (1<AI<1.2), partially andromonoecious (1.2≤AI<2.7), andromonoecious (2.7≤AI≤3).

doi:10.1371/journal.pone.0154362.t001

Cucurbit Genome Database (http://www.icugi.org), and the highest homology (E-value = 0) was found with Cla011230 gene on chromosome 3, a partial sequence of which was previously reported as *CitACS4* by [34].

The coding sequence of *CitACS4* is 1332 bp, encoding for a protein of 444 amino acids. The gene consists of three exons of 180, 281 and 871 bp, and two introns of 123 and 269 bp, a genomic structure very similar to that found in the orthologs *CmACS7*, *CsACS2* and *CpACS27A* (Fig 2). The CitACS4 protein shares 91–93% similarity with CmACS7, CsACS2 and CpACS27A (Fig 3A). These four enzymes are clustered together with the *Arabidopsis* AtACS7, in the branch corresponding to ACS type III (Fig 3B), lacking the CDPK phosphorylation motif of type I, and the MAPK6 phosphorylation motif of type I and II ACS enzymes [35, 36].

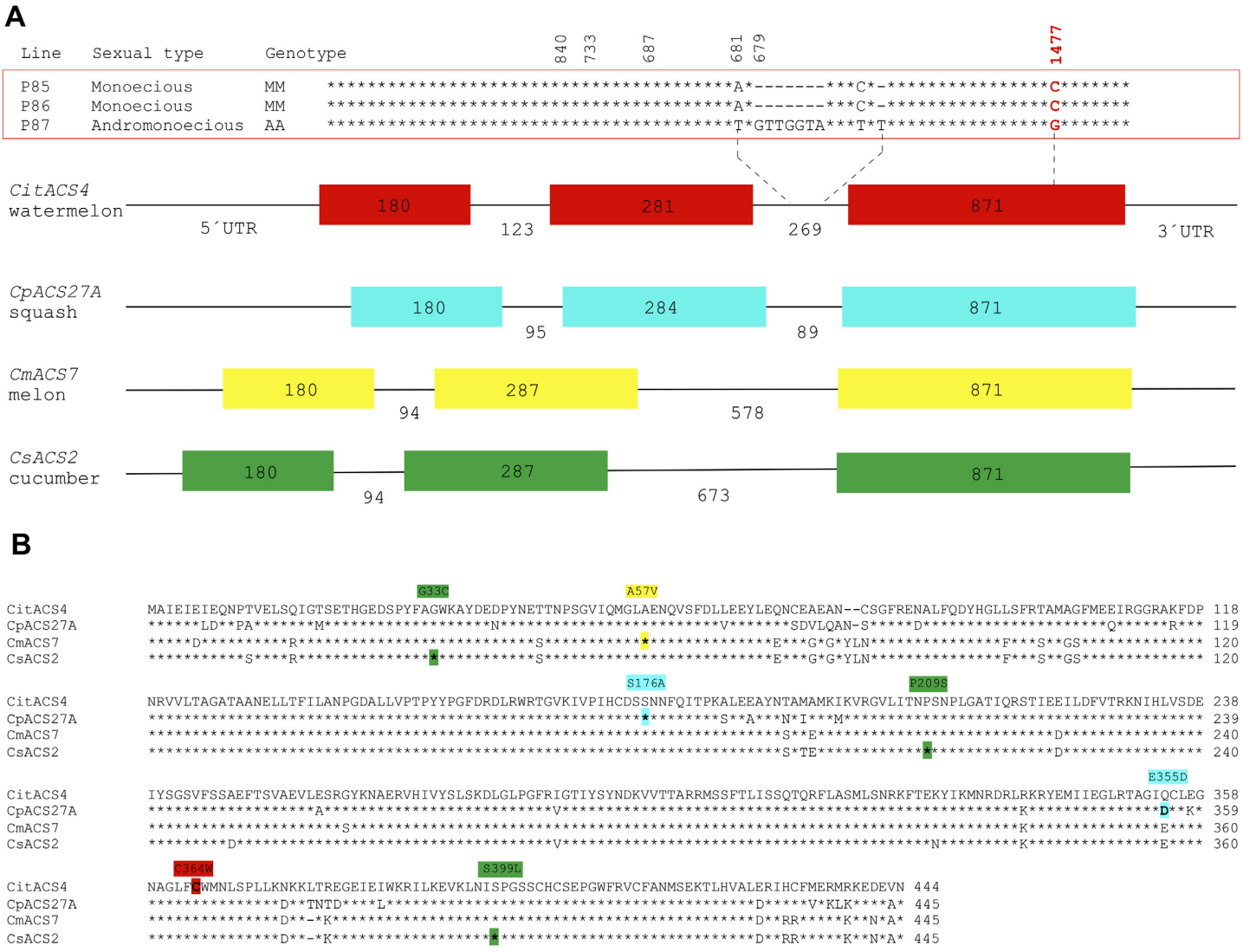


Fig 2. Molecular characterisation of *CitACS4* gene and protein. (A) Gene structure of *CitACS4*, *CpACS27A*, *CmACS7* and *CsACS2* in watermelon, squash, melon and cucumber, respectively. The numbers indicate the size of the three exons (*filled boxes*) and the two introns (*black lines*). The identified polymorphisms between DNA sequences in the monoecious and andromonoecious inbred lines are shown in above *CitACS4*. The missense mutation (C1477G) producing the amino acid substitution C364W in the protein is highlighted in red. (B) Alignment of watermelon CitACS4 with CpACS27A, CmACS7 and CsACS2 in squash, melon and cucumber. The amino acid changes between monoecious and andromonoecious lines in the different species are highlighted in red, blue, yellow and green, respectively.

doi:10.1371/journal.pone.0154362.g002

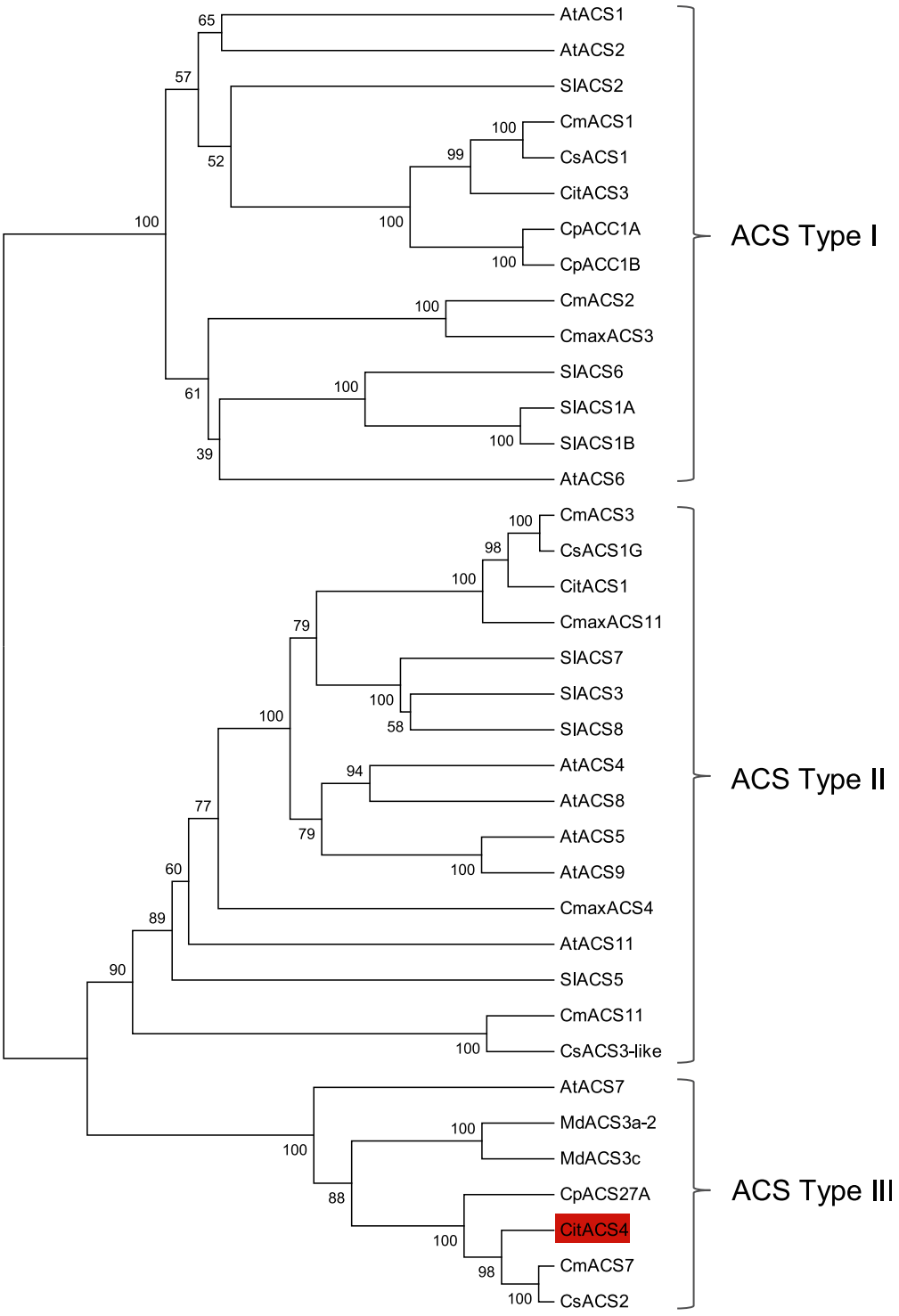


Fig 3. Phylogenetic analysis of CitACS4 protein. Evolutionary tree performed for 37 ACS proteins from different plants: *Arabidopsis thaliana* (AtACS1, AAM91649.1; AtACS2, AAG50097.1; AtACS4, Q43309.1; AtACS5, Q37001.1; AtACS6, Q9SAR0.2; AtACS7, AEE85169.1; AtACS8, Q9T065.1; AtACS9, Q9M2Y8.1; AtACS11, AEE82593.1), *Cucurbita máxima* (CmaxACS3, BAB47124.1; CmaxACS4, BAB47123.1; CmaxACS11, CBAA00839.1), *Cucurbita pepo* (CpACC1A, AAA33111.1; CpACC1B, AAA33112.1; CpACS27A, KF113530), *Cucumis melo* (CmACS1, BAA83618.1; CmACS2, BAB18464.1; CmACS3, ACO83163.1; CmACS7, ACG70849.1; CmACS11, XP_008445556.1), *Cucumis sativus* (CsACS1, BAA93714.1; CsACS1G, ABI33818.1; CsACS2, ACG70849.1; CsACS3-like, XP_004142909.2), *Citrullus*

lanatus (CitACS1, AFI49625.1; CitACS3, ABO76787.1; CitACS4, EF154458.1), *Malus x domestica* (MdACS3a-2, AEP82201.1; MdACS3c, BAE94692.1) and *Solanum lycopersicon* (SlACS1A, AAF97614.1; SlACS1B, AAF97615.1; SlACS2, P18485.2; SlACS3, NP_001234026.1; SlACS5, NP_001234156.1; SlACS6, NP_001234164.1; SlACS7, AAK72432.1; SlACS8, AAK72431.1). The tree was inferred using the UPGMA method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) is shown next to the branches.

doi:10.1371/journal.pone.0154362.g003

Expression of *CitACS4* was determined by quantitative RT-PCR in different plant organs. The gene was found to be specifically expressed in flowers, and predominantly in pistillate flowers (Fig 4). The expression in bisexual flowers was about half of that found in female flowers, and very low expression was detected in the male flowers. No *CitACS4* transcript was detected in the vegetative organs such as leaves or shoots (Fig 4).

We have also compared the expression of *CitACS4* during the development of pistillate flowers in the monoecious (P85 and P86) and andromonoecious (P87) lines of watermelon (Fig 5). The maximum expression was found in the female flowers of the monoecious lines P85 and especially in the P86 at very early stages of development (stage S0, floral buds of about 4 mm). Subsequently gene expression decreased until cessation at stage S3 (floral buds of about 15 mm). In the hermaphrodite flowers of the andromonoecious line P87, *CitACS4* showed the same expression profile, although with a lower level at the earliest stage of development (Fig 5B). No expression was detected in pistillate flowers at anthesis or post-anthesis stages of

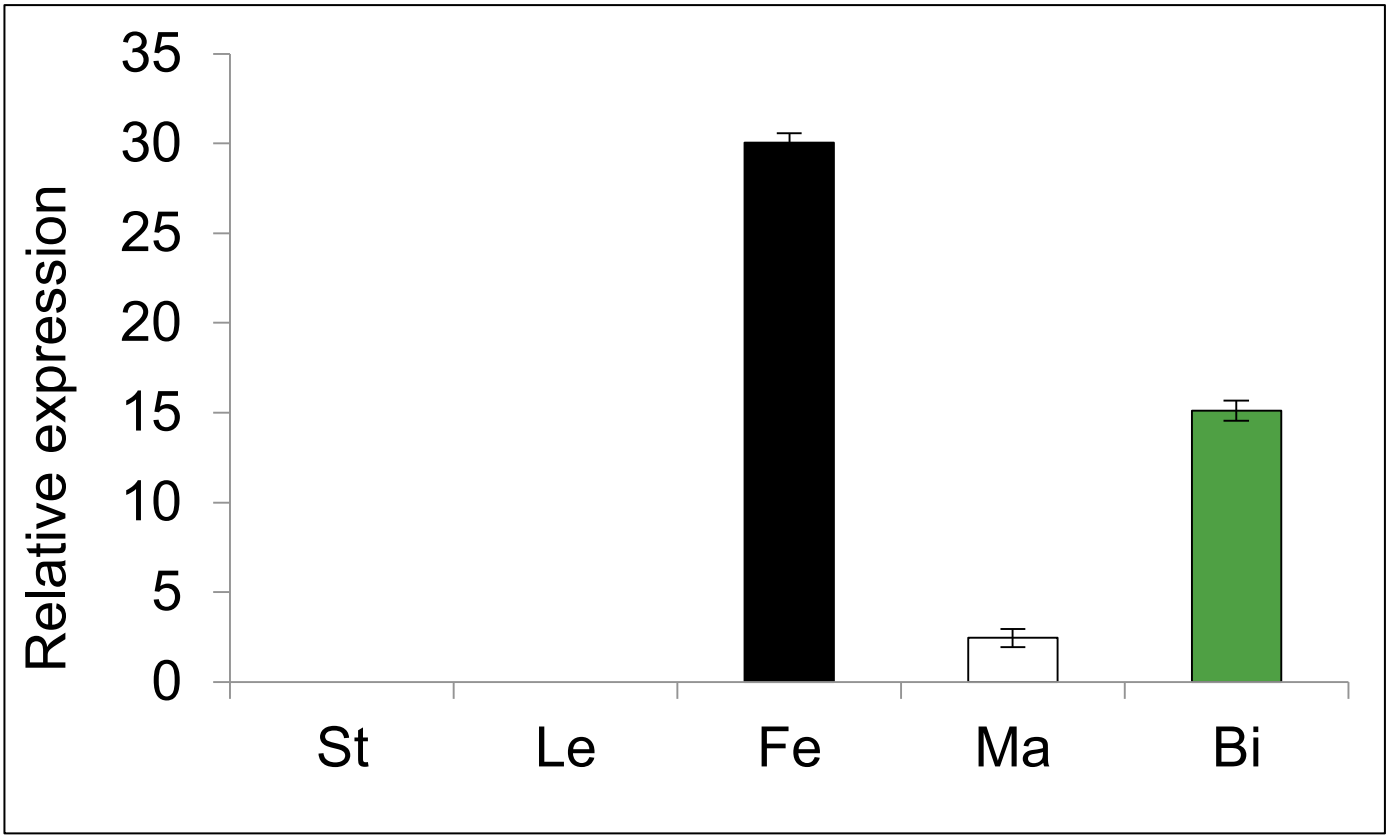


Fig 4. Relative expression of *CitACS4* in different tissues of watermelon cv. Premium. The values are the average and standard deviation of three biological replicates. St stem, Le leaves, Fe female flowers, Ma male flowers, Bi bisexual flowers. The utilized flowers were at early stages of development (S0).

doi:10.1371/journal.pone.0154362.g004

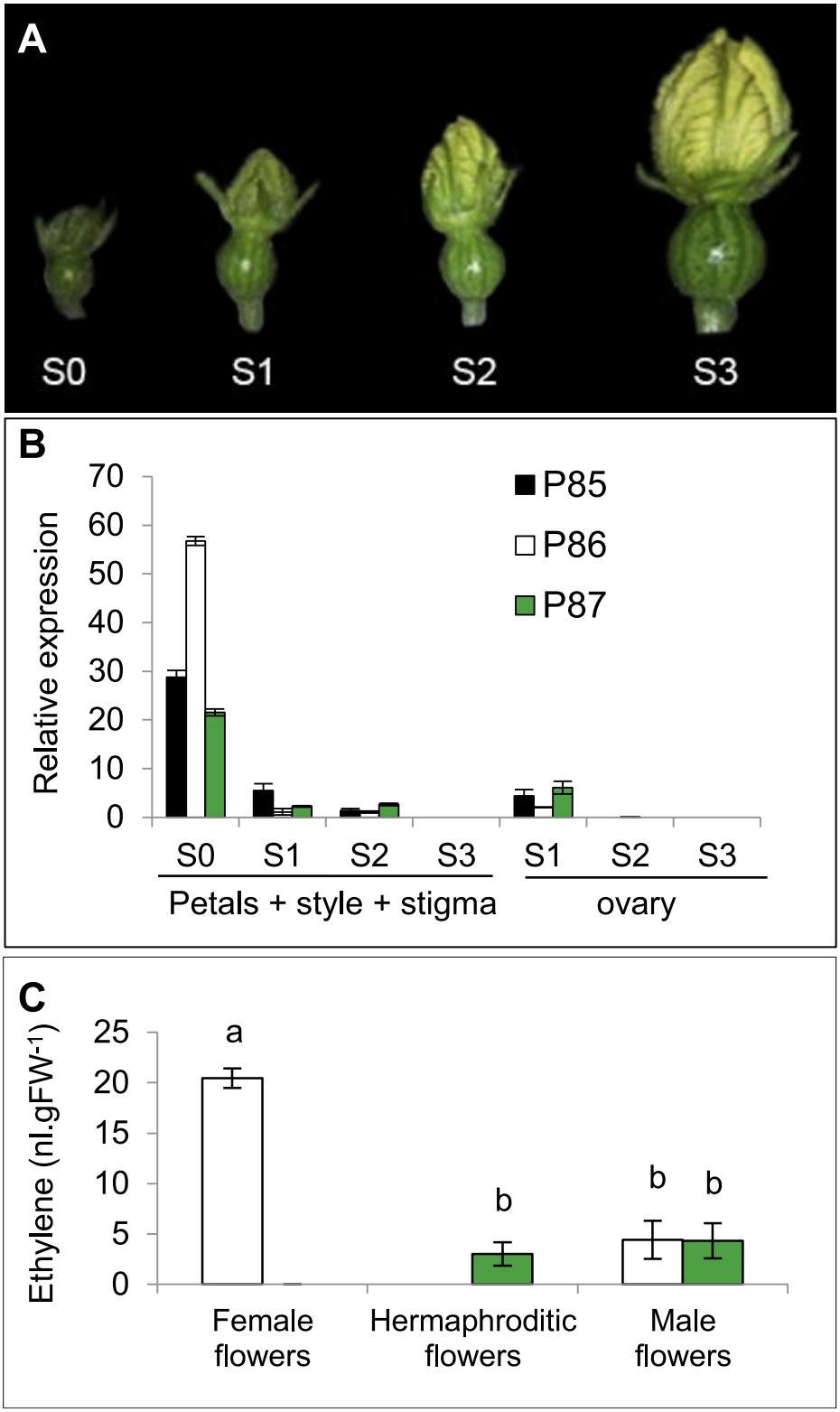


Fig 5. Expression of *CitACS4* and ethylene production during the development of pistillate flowers in monoecious and andromonoecious lines of watermelon. (A) Stages of development studied. (B) Relative expression of the gene in female flowers of monoecious (P85 and P86) and in the hermaphrodite flowers of andromonoecious (P87) lines. At S0, the expression corresponds to complete flowers, but in the other stages

(S1 to S3), the expression in the ovary was separated to that in the rest of the floral organs (petals, style and stigma, and stamens). (C) Ethylene production in female, hermaphrodite and male flowers of monoecious and andromonoecious lines was measured at earlier developmental stages (S0-S1). Each value is the average from at least three biological replicates. Error bars indicate standard deviation.

doi:10.1371/journal.pone.0154362.g005

development (data not shown). In S1-S3 floral buds, where it was possible to separate the ovary from petals, style and stigma, it was found that the accumulation of *CitACS4* transcripts in the ovary was lower than that found in the other floral organs, including petals, style and stigma (Fig 5B). Ethylene production in flowers correlated to *CitACS4* expression. In comparison with female flowers of monoecious line P86, the hermaphrodite flowers of andromonoecious line P87 showed reduced ethylene production to a level that is similar to that produced by male flowers (Fig 5C).

Co-segregation analysis of *CitACS4* with monoecious/ andromonoecious phenotypes

Polymorphisms between the *CitACS4* gene in monoecious (P85 and P86) and andromonecious (P87) lines have been searched for, and the possible co-segregation of the alleles with the monoecious and andromonoecious phenotypes in segregating populations derived from crosses P85xP87 and P86xP87 have been analysed. In comparison with the monoecious lines, the andromonoecious line one displayed not only two SNPs and two insertions of 8 nucleotides in the second intron of the gene, but also a SNP in the third exon (C1477G) that produced an amino acid substitution of cysteine (C) by tryptophan (W) at the residue 364 of the protein (Fig 2A). The residue C^{364} in the monoecious lines was conserved not only in the orthologs CmAC7, CsACS2 and CpACS27A (Fig 3), but also in other ACS enzymes from different plant species (data not shown), indicating that it is likely an essential residue for the enzymatic activity.

To study the possible involvement of *CitACS4* in the control of andromonoecy in watermelon, the C364W mutation in 163 plants from the F2 populations derived from two crosses P85xP87 and P86xP87 were genotyped (Table 2). All F2 plants homozygous for the mutated allele (genotype AA) were andromonoecious (average AI = 2.87±0.24 and 1.76±0.15 in each F2 population), while those homozygous for the WT allele (MM) were monoecious (average AI = 1.11±0.13 and 1.13±0.17 for each F2 population). The heterozygous plants (MA) showed a partially andromonoecious phenotype (average AI = 1.67±0.45 and 1.51±0.49 for each population), although some plants had a monoecious phenotype (Table 2). These data demonstrated that the andromonoecious phenotype in watermelon co-segregated with the mutated allele *A* of *CitACS4*, and therefore that the mutation C364W is likely the responsible for the andromonoecious phenotype in watermelon.

A linkage analysis was also performed for two other sex expression traits that are also regulated by ethylene [27]: the number of nodes before the production of the first pistillate flower (pistillate flowering transition) and the number of pistillate flowers per plant (Table 3). The andromonoecious parental line P87 had a later pistillate flowering transition (average node = 12.55) in comparison with the monoecious lines P85 (average node = 4.77) and P86 (average node 2.05) (Table 3). The two F1 generations had an early flowering phenotype (Table 3) but, in the F2 generations, the plants with the andromonoecious allele (genotype AA) did not flower later than those with the M allele (genotype MM). In fact, no significant differences were detected among F2 plants for three genotypes MM, MA and AA (Table 3). For the number of pistillate flowers per plant, no significant differences were detected between andromonoecious (P87) and monoecious (P85 and P86) parental lines, nor between genotypes MM,

Table 2. Segregation of the *M* and *A* alleles of *CitACS4* with sex monoecy/andromonoecy phenotype in the two F2 populations derived from crosses monoecious x andromonoeicous.

			No. of plants		
Generation	***CitACS4*genotype**	**Andromonoecious index (mean±sd)**	**Monoecious**	**Partially Andro-monoecious**	**Andro-monoecious**
P87	*AA*	3±0 a	0	0	9
F1(P85xP87)	*MA*	1.52±0.19 b	0	17	0
F1(P86xP87)	*MA*	1.35±0.21 b	3	15	0
P85	*MM*	1.16±0.12 c	13	0	0
P86	*MM*	1.16±0.12 c	18	0	0
F2(P85xP87)	*AA*	2.87±0.24 a	0	0	24
	MA	1.67±0.45 b	5	41	0
	MM	1.11±0.13 c	22	0	0
F2(P86xP87)	*AA*	2.76±0.15 a	0	0	13
	MA	1.51±0.49 b	7	34	0
	MM	1.13±0.17 c	17	0	0

[a-c]. Different letters indicate significant differences between genotypes.

Note that the A allele of *CitACS4* co-segregates with andromonoecy phenotype in the 169 F2 plants analysed.

doi:10.1371/journal.pone.0154362.t002

MA and AA in the F2 generation (Table 3). These data indicate that pistillate flowering transition and the percentage of female flowers, although controlled by ethylene, are not regulated the *CitACS4* gene.

Discussion

Studies on the inheritance of watermelon sex morphotypes have indicated that monoecy is dominant to andromonoecy and controlled by a single gene with two alleles [6, 8, 37]. The results from two crosses between monoecious and andromonoecious lines indicate that the F1 offspring has a predominantly monoecious phenotype. Nevertheless the higher production of

Table 3. Evaluation of sex expression (transition to pistillate flowering and % pistillate flowers per plant) in F1 and F2 populations derived from crosses monoecious x andromonoecious.

Generation	***CitACS4* genotype**	**Pistillate flowering transition**	**Percentage pistillate flowers**
P87	*AA*	12.55±4.12 a	16.66±7.9 ab
P85	*MM*	4.77±1.92 b	13.84±5.46 b
P86	*MM*	2.05±0.72 cd	19.44±5.66 ab
F1(P85XP87)	*MA*	3.64±2.23 bc	22.05±5.15 a
F1(P86XP87)	*MA*	1.56±1.19 d	19.16±4.28 ab
F2(P85XP87)	*MM*	5.18±3.16 b	16.09±4.25 b
	MA	4.04±2.63 b	16.85±5.8 ab
	AA	4.26±2.54 b	20±7.07 ab
F2(P86XP87)	*MM*	4.65±2.54 b	15±4.3 b
	MA	3.83±2.61 b	16.78±5.27 ab
	AA	4.2±3.09 b	18.33±4.49 ab

[a-d]. For each trait, different letters indicate significant differences between genotypes.

No significant differences was detected among *MM*, *MA* and *AA* genotypes for the two traits in the two F2 generations analysed, indicating that the gene *CitACS4* des not cosegregate with these two traits.

doi:10.1371/journal.pone.0154362.t003

bisexual and hermaphrodite flowers in the F1 suggests that the monoecy of these two lines is not actually dominant but semi-dominant to andromonoecy. The contrasting data may reflect the existence of different monoecious or andromonoecious alleles in watermelon. Differences in the average AI between F1 offspring of the two crosses performed (monoecious x andromonoecious), should be caused by two distinct monoecious alleles in the parental lines P85 and P86, as the andromonoecious parental lines were the same in both cases. The two F1 generations produced female, bisexual and male flowers, but the F1 derived from the cross P85xP87 had a higher number of bisexual and hermaphrodite flowers and a higher AI value (AI = 1.52) than the F1 derived from the cross P86xP87 (AI = 1.35). This suggests that the monoecious allele derived from P85 is less dominant to andromonoecy than that derived from P86. Therefore the existence of completely dominant alleles for monoecy in other genotypes of watermelon is not excluded.

Different monoecious alleles may explain differences in the expression of *CitACS4* such as has been observed for P85 and P86 at the earliest stage of development. The higher expression of *CitACS4* in P86 and the higher production of ethylene in the pistillate floral bud can result in a higher monoecy stability and a higher dominance of the monoecious over the andromonoecious allele in the F1 generation. It is known that ethylene regulates sex determination in watermelon, not only controlling the number of floral buds that will be developed as male or pistillate flowers, but also the differentiation and development of individual floral buds as male or female flowers [27]. The arrest of stamens during the development of female flowers requires ethylene, since external treatments with ethylene inhibitors induce the transformation of female into bisexual flowers with variable stamen size and even into hermaphrodite flowers with viable pollen [27]. In this paper it is found that the ethylene required to arrest stamen development in pistillate flowers is likely to be produced by the action of *CitACS4*, a major ethylene biosynthesis gene, already proposed as a candidate for the control of monoecy/andromonoecy in watermelon [34, 38]. *CitACS4*, as other orthologs in melon, cucumber and squash [9, 18, 19], is mainly expressed in pistillate flowers. Moreover, the mutation C364W is a very conserved residue of CitACS4 that co-segregates with the andromonoecious phenotype in two independent F2 populations, concomitantly with a reduction in ethylene production in the floral buds that will develop as hermaphrodite flowers in andromonoecious plants of the F2 segregating populations. These data indicate therefore that the abortion of stamen during female flowers development in watermelon requires the production of ethylene mediated by *CitACS4*.

The genomic structure, nucleotide and protein sequence, and the expression profile of *CitACS4* also support that it is the orthologous gene to *CmACS7*, *CsACS2* and *CpACS27A*. Similar to the other three genes, *CitACS4* is composed of 3 exons and 2 introns of similar size, suggesting that the different genes have evolved from the same ancestral sequence. Moreover, the phylogenetic analysis carried out with different ACS enzymes on a variety of plant species has demonstrated that CitACS4 is a type-III ACC synthase with a short C-terminal tail, showing none of the identifiable phosphorylation sites in type-I and type-II ACS enzymes [39]. The expression pattern of these orthologous genes has also been conserved through evolution. In melon, cucumber and squash the gene is specifically transcribed in the pistillate flowers, with a higher expression in female than in hermaphrodite flowers [9, 18, 19]. In watermelon the expression of *CitACS4* is also higher in female than in hermaphrodite flowers, but a low level of transcripts were also detected in male flowers, indicating that the function of the gene is dosage-dependent. The differential expression of *CitACS4* gene in the two analysed monoecious lines, and the phenotype of F1 hybrids, also indicate that the level of *CitACS4* gene expression is essential to control the abortion of stamen development and monoecy stability through plant development.

Apart from andromonoecy, no co-segregation between *CitACS4* gene and other sex expression traits regulated by ethylene in this species, including pistillate flowering transition and the number of pistillate flowers per plant have been detected. These two sex expression traits should be regulated by other ethylene genes, which supporting previous data indicating that sexual expression of watermelon is an independent mechanism from sex determination of individual floral buds [27]. In fact, an increase of ethylene in the apical shoot does not induce the production of pistillate flowers, as occurs in melon, cucumber and squash, but on the contrary it reduces the number of pistillate flowers in the shoot [27]. This paper confirms therefore that there is a conserved molecular mechanism that makes use of the hormone ethylene for promoting the transformation of hermaphrodite to female flowers at the origin of monoecy in cucurbit species. The mechanisms that regulate the formation of male and female flower along main and lateral shoots, although still dependent on ethylene production and sensitivity, has diverged in watermelon [27] from what occurs in other cucurbit cultivated species such as *Cucumis* [40] and *Cucurbita* [41].

Supporting Information

S1 Table. Primers used in quantitative real time RT-PCR reactions and to amplify a full sequence of *CitACS4* gene.
(PDF)

Acknowledgments

This work was supported by grant AGL2014-54598-C2-1-R, partly funded by ERDF (European Regional Development Fund) and by the Spanish Ministry of Science and Innovation, and grant P12-AGR-1423, funded by Junta de Andalucía, Spain. Z.M. and E.A. acknowledge FPU and "garantía juvenil" scholarship programmes from MEC, Spain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceived and designed the experiments: MJ SM. Performed the experiments: EA SM AG ZM. Analyzed the data: SM EA CM. Contributed reagents/materials/analysis tools: MJ SM EA AG. Wrote the paper: MJ SM.

References

1. Poole CF, Grimball PC. Inheritance of new sex forms in *Cucumis melo* L. J Heredity. 1938; 30: 21–25.
2. Kenigsbuch D, Cohen Y. Inheritance of gynoecious sex type in muskmelon. Cucurbit Genet Coop Rep. 1987; 10: 47–48.
3. Kenigsbuch D, Cohen Y. The inheritance of gynoecy in muskmelon. Genome. 1990; 33:317–320.
4. Kubicki B. Investigations on sex determination in cucumber (*Cucumis sativus* L.). III. Variability of sex expression in the monoecious and gynoecious lines. Genet Pol. 1969; 10: 5–22.
5. Kubicki B. Investigations on sex determination in cucumber (*Cucumis sativus* L.). V. Genes controlling intensity of femaleness. Genet Pol. 1969; 10: 69–86.
6. Rosa JT. The inheritance of flower types in *Cucumis* and *Citrullus*. Hilgardia. 1928; 3: 233–250.
7. Jiang X, Lin D. Discovery of watermelon gynoecious gene *gy*. Ac Hort Sinica. 2007; 34: 141–142.
8. Ji G, Zhang J, Gong G, Shi J, Zhang H, Ren Y, et al. Inheritance of sex forms in watermelon (*Citrullus lanatus*). Sci Hort. 2015; 193: 367–373.
9. Martínez C, Manzano S, Megías Z, Boualem A, Garrido D, Bendahmane A, et al. Molecular and functional characterization of *CpACS27A* gene reveals its involvement in monoecy instability and other

associated traits in squash (*Cucurbita pepo* L.). Planta. 2014; 239: 1201–1215. doi: 10.1007/s00425-014-2043-0 PMID: 24595516

10. Manzano S, Martínez C, Domínguez V, Avalos E, Garrido D, Gómez P, et al. A major gene conferring reduced ethylene sensitivity and maleness in *Cucurbita pepo*. J Plant Growth Regul. 2010; 29:73–80.
11. García A, Manzano S, Martínez C, Megías Z, Gázquez JC, Jamilena M. Isolation and characterisation of ethylene insensitive mutants from a collection of Cucurbita pepo L. V ISHS International Symposium on Cucurbits; 2015. pp. 22–26.
12. Byers RE, Baker LR, Sell HM, Herner RC, Dilley DR. Ethylene: a natural regulator of sex expression of *Cucumis melo* L. Proc Natl Acad Sci USA. 1972; 69:717–720. PMID: 16591971
13. Den Nijs APM, Visser DL. Induction of male flowering in gynoecious cucumbers (*Cucumis sativus* L.) by silver ions. Euphytica. 1980; 29:273–280
14. Manzano S, Martínez C, Megías Z, Gómez P, Garrido D, Jamilena M. The role of ethylene and brassinosteroids in the control of sex expression and flower development in *Cucurbita pepo*. Plant Growth Regul. 2011; 65: 213–221.
15. Owens KW, Peterson CE, Tolla GE. Production of hermaphrodite flowers on gynoecious muskmelon by silver nitrate and aminoethyoxyvinylglycine. HortScience. 1980; 15:654–655.
16. Rudich J. Biochemical aspects of hormonal regulation of sex expression in Cucurbits. In: Biology and utilization of the Cucurbitaceae, editors. Bates DM, Robinson RW. Ithaca-NY: Cornell University Press; 1990. pp. 269–280.
17. Rudich J, Halevy AH, Kedar N. Increase in femaleness of three cucurbits by treatment with ethrel, an ethylene releasing compound. Planta. 1969; 86:69–76. doi: 10.1007/BF00385305 PMID: 24515743
18. Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, et al. A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science. 2008; 321: 836–838. doi: 10.1126/science.1159023 PMID: 18687965
19. Boualem A, Troadec C, Kovalski I, Sari M, Treves R, Bendahmane A. A conserved ethylene biosynthesis enzyme leads to andromonoecy in two *Cucumis* species. PLoS One. 2009; 4: e6144. doi: 10.1371/journal.pone.0006144 PMID: 19578542
20. Knopf RR, Trebitsh T. The female-specific *Cs-ACS1G* gene of cucumber. A case of gene duplication and recombination between the non-sex-specific *1-aminocyclopropane-1-carboxylate synthase* gene and a branched-chain amino acid transaminase gene. Plant Cell Physiol. 2006; 47:1217–1228. PMID: 16887844
21. Mibus H, Tatlioglu T. Molecular characterization and isolation of the *F/f* gene for femaleness in cucumber (*Cucumis sativus* L.). Theor Appl Genet. 2004; 109:1669–1676. PMID: 15490106
22. Trebitsh T, Staub JE, O'Neill SD. Identification of a *1-Aminocyclopropane- 1-carboxylic acid synthase* gene linked to the female (*F*) locus that enhances female sex expression in cucumber. Plant Physiol. 1997; 113:987–995. PMID: 9085580
23. Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, et al. A transposon-induced epigenetic change leads to sex determination in melon. Nature. 2009; 461: 1135–1138. doi: 10.1038/nature08498 PMID: 19847267
24. Boualem A, Troadec C, Camps C, Lemhemdi A, Morin H, Sari MA, et al. A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. Science. 2015; 350: 688–691. doi: 10.1126/science.aac8370 PMID: 26542573
25. Rudich J, Zamski E. *Citrullus lanatus*. In: Halevy A, editor. Handbook of Flowering, vol. 2. CRC Press, Boca Raton, FL; 1985. pp. 272–274.
26. Sugiyama K, Kanno T, Morishita M. Evaluation method of female flower bearing ability in watermelon using silver thiosulfate (STS). J Jpn Soc Hortic Sci. 1998; 67: 185–189.
27. Manzano S, Martínez C, García JM, Megías Z, Jamilena M. Involvement of ethylene in sex expression and female flower development in watermelon (*Citrullus lanatus*). Plant Physiol Bioch. 2014; 85: 96–104.
28. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 2007; 24:1596–1599. PMID: 17488738
29. Sneath PHA, Sokal RR. Numerical Taxonomy. Freeman, San Francisco; 1973.
30. Felsenstein J. Phylogenies and the comparative method. Am Nat. 1985; 125:1–15
31. Zuckerkandl E, Pauling L. Molecules as documents of evolutionary history. J Theory Biol. 1965; 8:357–366.
32. Levi A, Thomas C. An improved procedure for isolation of high quality DNA from watermelon and melon leaves. Cucurbit Genet Coop. Rep. 1999; 22:41–42.

33. Manzano S, Martínez C, Gómez P, Garrido D, Jamilena M. Cloning and characterisation of two *CTR1-like* genes in *Cucurbita pepo*: regulation of their expression during male and female flower development. Sex Plant Reprod. 2010; 23:301–313. doi: 10.1007/s00497-010-0140-1 PMID: 20390430

34. Salman-Minkov A, Levi A, Wolf S, Trebitsh T. *ACC synthase* genes are polymorphic in watermelon (*Citrullus* spp.) and differentially expressed in flowers and in response to auxin and gibberellin. Plant Cell Physiol. 2008; 49:740–750. doi: 10.1093/pcp/pcn045 PMID: 18367517

35. Yoshida H, Nagata M, Saito K, Wang KLC, Ecker JR. *Arabidopsis ETO1* specifically interacts with and negatively regulates type 2 1-Aminocyclopropane-1-carboxylate synthases. BMC Plant Biol. 2005; 5:14. PMID: 16091151

36. Zhang T, Qiao Q, Zhong Y. Detecting adaptive evolution and functional divergence in 1-Aminocyclopropane-1-carboxylate synthase (*ACS*) gene family. Comput Biol Chem. 2012; 38:10–16. doi: 10.1016/j.compbiolchem.2012.04.001 PMID: 22543105

37. Poole C.F, Grimball PC. Interaction of sex, shape, and weight genes in watermelon. J. Agric. Res. 1944; 71:533–552.

38. Prothro J, Abdel-Haleem H, Bachlava E, White V, Knapp S, McGregor C. Quantitative trait loci associated with sex expression in an inter-subspecific watermelon population. J Am Soc Hortic Sci. 2013; 138:125–130.

39. Argueso CT, Hansen M, Kieber JJ. Regulation of ethylene biosynthesis. J Plant Growth Regul. 2007; 26:92–105.

40. Yamasaki S, Fujii N, Matsuura S, Mizusawa H, Takahashi H. The *M* locus and ethylene-controlled sex determination in andromonoecious cucumber plants. Plant Cell Physiol. 2001; 42:608–619. PMID: 11427680

41. Manzano S, Martínez C, Megías Z, Garrido D, Jamilena M. Involvement of ethylene biosynthesis and signalling in the transition from male to female flowering in the monoecious *Cucurbita pepo*. J Plant Growth Regul. 2013; 32:789–798.

Plant Reproduction
https://doi.org/10.1007/s00497-018-0346-1

ORIGINAL ARTICLE

The sex-determining gene *CitACS4* is a pleiotropic regulator of flower and fruit development in watermelon (*Citrullus lanatus*)

Encarnación Aguado[1] · Alicia García[1] · Susana Manzano[1] · Juan Luis Valenzuela[1] · Julián Cuevas[2] · Virginia Pinillos[2] · Manuel Jamilena[1]

Received: 12 April 2018 / Accepted: 13 August 2018

Abstract

In the species of the *Cucurbitaceae* family, the occurrence of separate male and female flowers in the same plant (monoecy) is controlled by an ethylene biosynthesis *ACS* gene, which specifically suppresses the development of stamen in the female flower. In watermelon, a mutation of loss of function in *CitACS4* promotes the conversion of female into hermaphrodite flowers, and of monoecious into andromonoecious plants. We have studied whether the ethylene produced by CitACS4 enzyme could also be involved in other ethylene-regulated traits, including pistillate flowering transition and the number of female flowers per plant, the development of floral organs other than stamens, as well as fruit and seed set, and fruit development. A linkage analysis approach was performed in three independent F2 populations segregating for the two alleles of the gene (*M*, monoecious; *m*, andromonoecious), and the different traits under study. The *CitACS4m* allele not only cosegregated with andromonoecy, but also with earlier pistillate transition, an increased number of pistillate flowers per plant, and a slower growth and maturation of petals and carpels, which delayed anthesis time in hermaphrodite flowers. The *m* allele was also found to be linked to a reduced fruit set, which was not caused by a deficiency in pollination or fertilization. The gene also affected the longitudinal and transverse growth rates of the ovary and fruit, which means that fruits from andromonoecious plants (*mm*) were rounder than those from monoecious (*MM*) ones. Taken together, these data indicate that the locus defined by the ethylene biosynthesis and sex-determining gene *CitACS4* acts as a pleiotropic regulator of the complete development of the pistillate flower and the earlier development of the fruit.

Keywords Watermelon · Monoecious · Andromonoecious · *CitACS4* · Fruit set · Fruit shape

Introduction

Watermelon (*Citrullus lanatus*) is a major horticultural crop worldwide, with a production of over 111 million tons in 2014 (FAOSTAT 2017). Production-related traits, including pollination efficiency and fruit set, are quite dependent on the sexual expression of the cultivar. The flowering pattern of watermelon *Citrullus* spp. is either monoecious (male and female flowers in the same plant), andromonoecious (male and hermaphrodite flowers in the same plant) or trimonoecious (female, hermaphrodite and male flowers in the same plant) (Rudich and Zamski 1985; Ji et al. 2015). Andromonoecy and trimonoecy are undesirable traits in cucurbits, since hermaphrodite flowers need to be emasculated when acting as female parents in the production of hybrid seed (Prothro et al. 2013), and also because the trait is usually associated with a reduction in fruit set and fruit quality (Monforte et al. 2005; Abdelmohsin and Pitrat 2008; Martínez et al. 2014).

Sex expression and flower development in watermelon are known to be regulated by several environmental factors and phytohormones such as ethylene and gibberellins. External treatments with ethylene and GA_3 inhibit the transition from

Communicated by Dolf Weijers.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00497-018-0346-1) contains supplementary material, which is available to authorized users.

✉ Manuel Jamilena
mjamille@ual.es

[1] Department of Biology and Geology, Research Centres CIAIMBITAL and CeiA3, University of Almeria, 04120 Almería, Spain

[2] Department of Agronomy, Research Centres CIAIMBITAL and CeiA3, University of Almeria, 04120 Almería, Spain

Published online: 20 August 2018

male to female flowering and reduce the production of pistillate flowers, while treatments with the ethylene inhibitors AVG promote female flowering transition and increase the number of pistillate flowers per plant (Manzano et al. 2014; Zhang et al. 2017). High temperatures, and the concomitant reduction in ethylene production, are also responsible for the conversion of monoecious into partially andromonoecious plants. Treatments with silver sulphate, an inhibitor of ethylene action, also produce a total or partial transformation of female into hermaphrodite flowers (Zhang et al. 2017), indicating that ethylene, as occurs in other cucurbit species, is responsible for the arrest of stamen growth during female flower development (Manzano et al. 2014, 2016). Studies on the inheritance of watermelon sex morphotypes have indicated that monoecy is dominant to andromonoecy and controlled by a single gene with two alleles (Rosa 1928; Poole and Grimball 1945; Rudich and Zamski 1985; Salman-Minkov et al. 2008). It has recently been demonstrated that monoecy is actually controlled by a single semi-dominant gene called *CitACS4* (Boualem et al. 2016; Manzano et al. 2016; Ji et al. 2016). The gene encodes for a flower-specific ACS enzyme involved in the biosynthesis of the ethylene required for stamen arrest in the female flowers. A single missense mutation in the coding region of this gene produces an amino acid substitution of cysteine to tryptophan in residue 364 of the *CitACS4* protein (C364 W), reducing the production of ethylene in pistillate floral buds and promoting a complete conversion of female into hermaphrodite flowers, and therefore of monoecy into andromonoecy (Boualem et al. 2016; Manzano et al. 2016; Ji et al. 2016). The andromonoecious trait in other cucurbit species, including cucumber, melon and zucchini squash, also results from mutations in the orthologous ethylene biosynthesis genes *CmACS7, CsACS2,* and *CpACS27A*, respectively (Boualem et al. 2008, 2009; Martínez et al. 2014).

Besides sex determination, ethylene regulates several developmental processes associated with flower and fruit development. After pollination, the induction of ethylene production in the ovaries and petals appears to be responsible for coordinating ovary growth and petal senescence (Larsen et al. 1993; Balbi and Lomax 2003; Wang et al. 2005; Stepnova et al. 2008). Recent studies have shown an interconnection between early ovule abortion and the size of the silique in *Arabidopsis* ethylene mutants (Carbonell-Bejerano et al. 2011). In squash, Martínez et al. (2013) found that a reduction in ethylene production or signalling in the flower induces fruit set and early fruit development. Similarly, pollination and gibberellin treatments downregulate ethylene biosynthesis and signalling genes in tomato immediately after fruit set (Pandolfini et al. 2007; Stepnova et al. 2008). Fruit set in watermelon is unstable at low temperatures and under cloudy or rainy weather, as the activity of flower-visiting insects is sluggish and the dehiscence of anthers is hindered (Tsukahara 1988). Whether this fruit set is dependent on ethylene is unknown, but there are some data suggesting that fruit set improves in monoecious cultivars (those producing more ethylene in the female flower) in comparison with andromonoecious ones (those producing less ethylene in the female flower) (Wechter et al. 2008; Manzano et al. 2014).

Fruit shape is also related to sex expression in the species of *Cucurbitaceae*, which also suggests the potential involvement of ethylene in this developmental process. In cucumber and melon, the fruits developed from hermaphrodite flowers on andromonoecious plants are rounder than those derived from female flowers (Loy 2006; Abdelmohsin and Pitrat 2008; Sakata et al. 2013; Díaz et al. 2014). In watermelon, Rosa (1928) also reported that andromonoecious plants produced fruit that was rounder, and Poole and Grimball (1945) detected a genetic linkage between round fruits and andromonoecy, and between oval-shaped fruits and monoecy.

In the present study, we used watermelon populations that segregate for two alleles of the *CitACS4* gene, and therefore for monoecy and andromonoecy, to study whether *CitACS4*, and consequently the production of ethylene in the female flower, not only controls sex determination, but is also responsible for the regulation of the following traits: number of male and female flowers per plant, floral organ maturation, fruit and seed set, growth rate and shape of the watermelon ovary and fruit.

Materials and methods

Plant material and growing conditions

Four inbred lines of watermelon (*C. lanatus*), three monoecious lines (P84, P85 and P86) and one andromonoecious line (P87), as well as the F2 generations derived from crosses between monoecious and andromonoecious lines (P84XP87, P85XP87 and P86XP87), were characterized. The number of phenotyped plants in the parent lines and in the plants genotyped as *MM*, *Mm* and *mm* of the F2 generations is shown in Table S1. Sex determination and sex expression in the crosses P85XP87 and P86XP87 were previously studied by Manzano et al. (2016). In this paper, we analysed the sex expression and sex determination in a new cross between the monoecy unstable line P84 and the andromonoecious line P87, and studied the floral and fruit traits detailed below in the three crosses.

Seeds of the different lines were simultaneously germinated in seed trays in both spring/summer and autumn/winter seasons, and seedlings transplanted in a greenhouse at the experimental station of the University of Almería (Spain), and grown under the same standard crop management of

the region. Phenotypic evaluations were performed in the spring–summer and autumn–winter seasons of 2014, 2015 and 2016.

Genotyping for *CitACS4* alleles

The F2 seedlings from the three independent crosses (P84XP87, P85XP87 and P86XP87) were genotyped for *CitACS4* alleles before being transplanted to the greenhouse. The specific primer pairs CitACS4MF/CitACS4gen-R1 or CitACS4S-F/CitACS4M-R, designed to specifically amplify the *M* allele, and CitACS4A-F/CitACS4gen-R1 or CitACS4S-F/CitACS4A-R, which only amplified the *m* allele, were used for genotyping. These primers pair resulted in a 253- or 271-bp PCR fragment of the *CitACS4* gene, respectively. Plant DNA was extracted from frozen young leaves using the CTAB method (Manzano et al. 2016), and the PCRs were performed in the GeneAmp PCR System 2700 (Applied Biosystems). PCRs consisted of 35 cycles of 30 s at 95 °C, 30 s at 60 °C and 90 s at 72 °C. PCR fragments were resolved in agarose gels at 1% and plants classified in *MM*, *Mm* and *mm*. At least 15 seedlings of each *CitACS4* genotype were transplanted to the greenhouse for phenotyping (Table S1).

Phenotyping for monoecy stability and sex expression

To assess the level of monoecy in the different inbred lines and populations, the so-called Andromonoecy Index (AI; Martínez et al. 2014) was defined for each flower, plant and population. Pistillate flowers were scored from 1 to 3 according to their degree of stamen development. Female flowers with no stamen were scored as AI = 1, while hermaphrodite flowers with complete stamens and anthers able to produce pollen were scored as AI = 3. A score of 2 was assigned to bisexual flowers with medium-sized stamens and anthers (Fig. 1). Based on the flower scores, the AI of each plant was assessed as the average AI of at least five pistillate flowers. The average AI in inbred lines was estimated from at least 15 plants. Plants with an AI = 1–1.2 were considered monoecious, those with AI = 1.2–2.7 were considered partially andromonoecious, and those with AI ≥ 2.7 were phenotyped as andromonoecious (Manzano et al. 2016).

Sex expression in each plant was assessed by both the number of initial nodes with male flowers before the production of the first pistillate flower in the main shoot (female flowering transition), and the percentage of pistillate flowers per plant in the first 20 nodes of the main shoot. At least 15 plants were phenotyped to assess the AI and the sexual expression of both parental lines, and *MM*, *Mm* and *mm* plants of three F2 populations (Table S1).

Phenotyping for floral and fruit traits

To assess floral organ development, the growth rates of ovaries and petals of each *CitACS4* genotype were determined by measuring the length of these floral organs every 2 days until anthesis, starting with flower buds of about 2 mm in length.

The evaluation of fruit set and early fruit development was conducted in 15 pistillate flowers for each genotype (*MM*, *Mm* or *mm*) in each of the three analysed F2 populations. Plants were hand-pollinated with fresh pollen of the same plant for a total of 12 consecutive days, when the environmental conditions were similar, and always at the same time of the day (9:00–10:00 in spring and 10:00–11:00 in fall). Pollination was done on the day of anthesis for both pistillate and male flowers. To prevent flower damage and abortion, the hermaphrodite and bisexual flowers in *Mm* and

Fig. 1 Phenotypes of watermelon hermaphrodite, bisexual and female flowers

mm plants were not emasculated before pollination. After hand pollination, the length and diameter of at least 15 ovaries/fruitlets were measured from anthesis to 14 days post-anthesis (DPA). The ratio between the number of fruits that continued growing and the number of fruits whose growth aborted over this period of time was used to calculate the percentage of fruit set. When the number of abortions was very high, many more flowers were pollinated to reach a minimum of 10 fruits for seed set analyses (Table S2).

The ovary/fruitlet shape (FS) throughout development was assessed by calculating the ratio of fruit length (FL) over maximum fruit diameter (FD) at anthesis, at 14 DPA and in mature fruit (Díaz et al. 2014).

In at least 10 fully mature fruits for each *CitACS4* genotype, harvested 60 days after pollination, the number of viable and non-viable seeds in 1/4 of each fruit was assessed, after which the number of seeds per kilogram of fresh fruit was calculated. The viability of seeds was determined using the floatation test. When newly extracted seeds were placed in a container with water, the submerged and floating seeds were classified as viable and non-viable, respectively. We verified that the floating seeds contained no embryo and did not germinate, while the submerged seeds contained embryos and most of them germinated under our conditions.

Evaluation of pollination and fertilization

Pollen–pistil interaction was analysed in female and hermaphrodite flowers of P84, P86 and P87 lines, determining the best fertility period for setting fruits in each *CitACS4* genotype, but also the possible failure associated with reduced fruit set in P87. Pistillate or hermaphrodite flower was hand-pollinated with its own fresh pollen at anthesis and at -1, -2, $+1$ and $+2$ days from anthesis, always at the same time of day. Hermaphrodite flowers from P87 line were previously emasculated to avoid self-pollination before scheduled date. Flowers were fixed in FAE (formalin/glacial acetic acid/ethanol 70%, in a ratio 1:2:17 v/v) 24 h after pollination. Fixed flowers were processed as explained in Cuevas et al. (1994) and stained using 0.1% aniline blue in phosphate buffer for observations under fluorescence microscopy (Martin 1959) in a Nikon Labophot epifluorescence microscope. Pollen adhesion, germination, pollen tube growth, fertilization levels were determined in each flower and the results averaged for each pollination date. Pollen adhesion was estimated by counting the number of pollen grains in three different areas of the stigma (3.8 mm^2 each). Pollen germination was expressed as the ratio between pollen grains adhered and those germinated forming a pollen tube and penetrating the stigma. Pollen tubes in the style were observed, and an approximate number range was indicated: 0–5 (very few pollen tubes), 5–25 (scarce number of pollen tubes) or > 25 (high number of pollen tubes). In order to estimate fertilization, ovules were extracted from a section of the ovary of flowers pollinated at anthesis. Fertilization rates were calculated as the percentage of fertilized ovules. An ovule was considered fertilized if a pollen tube was present at the micropyle (Fig. 5). The presence of pollen tubes in the ovary was also observed.

Linkage and statistical analysis

As the *CitACS4* gene is involved in ethylene production, we have studied whether this gene might also regulate other ethylene-regulated processes and traits in flowers and fruits. The expression of each trait was compared in *MM* and *mm* parental lines, as well as in the *MM*, *Mm* and *mm* plants of the three segregating F2 generations. When differences between parental lines are maintained between *MM* and *mm* genotypes in the F2 generation, we conclude that *CitACS4* is cosegregating with the trait and therefore that the gene is likely involved in its regulation.

For statistical comparison, simple and factorial analyses of variance (ANOVA) at $p < 0.05$ were performed by the Statistix 8.0 software package, and each two means were compared by Fisher's least significant difference (LSD) test. The Tukey's multiple comparison test was mainly used when the number of samples per comparing group was the same (only in the comparisons of pollination and fertilization events).

Results

Involvement of the *CitACS4* gene in sex determination and sex expression

Four types of flowers can be found in watermelon: female flowers, which develop carpels but no stamens, male flowers, developing stamens but no carpels, and hermaphrodite and bisexual flowers which are flowers producing both complete carpels and stamens, or complete carpels and partially developed stamens, respectively (Fig. 1). To assess the sex phenotype of watermelon plants, we defined the andromonoecious index (AI). AI ranges from 1 to 3 and assesses the degree of development of stamens in each pistillate flower and therefore the level of monoecy–andromonoecy per plant and population. Plants and lines with AI = 1–1.19 produced predominantly female flowers and were considered monoecious, those with AI = 1.2–2.69 produced female, bisexual and hermaphrodite flowers and were considered partially andromonoecious, and those with AI ≥ 2.7 were considered andromonoecious because they produced predominantly hermaphrodite flowers.

Table 1 summarizes AI and other sex-related traits in the four parental lines, as well as in F2 generations derived

Table 1 Comparison of andromonoecious index (AI), pistillate flowering transition and percentage of pistillate flowers per plant in monoecious and andromonoecious plants from parental, F1 and F2 generations

Generation	*CitACS4* genotype	AI	Sex phenotype	Pistillate flowering transition	Percentage pistillate flowers/plant
P86	*MM*	1.16 cd	Mono	4.70ab	20.0 ab
P87	*mm*	3a	Andro	5.58a	25.42a
F1	*Mm*	1.35bc	PA	1.56c	19.16b
F2	*MM*	1.11d	Mono	4.65ab	15.00c
	Mm	1.51b	PA	3.83b	16.78bc
	mm	2.76a	Andro	4.20ab	18.33bc
P85	*MM*	1.16c	Mono	4.77ab	13.84c
P87	*mm*	3a	Andro	5.58a	25.42a
F1	*Mm*	1.52b	PA	3.64b	22.05ab
F2	*MM*	1.11c	Mono	5.18ab	16.09c
	Mm	1.67b	PA	4.04ab	16.85c
	mm	2.87a	Andro	4.26ab	20.00b
P84	*MM*	1.26c	Mono	6.78ab	22.78bc
P87	*mm*	3a	Andro	5.58b	25.42ab
F1	*Mm*	1.94b	PA	5.92b	22.92bc
F2	*MM*	1.22c	Mono	8.22a	21.22c
	Mm	1.8b	PA	7.46a	22.54bc
	mm	2.81a	Andro	6.52b	27.58a

The traits were assessed in monoecious (*MM*) and andromonoecious (*mm*) parental lines, and in *MM*, *Mm* and *mm* F2 plants derived from monoecious x andromonoecious crosses. Mono, monoecious; PA, partially andromonoecious; Andro, andromonoecious. Statistical analysis was performed using LSD test ($p \leq 0.05$), and the different letters indicate significant differences between genotypes of the same cross

from crosses between monoecious (P86, P85 and P84) and andromonoecious (P87) lines. The P87 line has a very stable andromonoecy, producing only male and hermaphrodite flowers with complete stamens and pollen (AI = 3). The monoecious P85 and P86 lines produced predominantly female flowers, although they also produced some bisexual flowers (AI = 1.16). The sexual phenotype of the two F1 hybrids P85XP87 and P86XP87 had an intermediate phenotype (AI = 1.52 and AI = 1.35) and were therefore classified as partially andromonoecious. The monoecious line P84 showed a higher andromonoecious index (AI = 1.26), suggesting that its monoecy is less stable than that of P85 and P86 lines. The AI of the F1 from P84XP87 (AI = 1.94) was also intermediate, although more biased to andromonoecy (Table 1). As previously demonstrated for P85 and P86 (Manzano et al. 2016), the monoecy of P84 was also controlled by a single semi-dominant gene (Table S3). Among the 137 phenotyped F2 plants (P84XP87), 31 were monoecious, 81 were partially andromonoecious and 25 were andromonoecious, which fits the segregation ratio 1:2:1 ($\chi^2 = 5.087$, $p = 0.078$), expected for a single semi-dominant gene controlling the trait (Table S3).

Segregation data from P84XP87 confirmed that *CitACS4* regulates monoecy/andromonoecy in watermelon. Indeed, the *M* and *m* alleles of the gene cosegregated with either monoecious or andromonoecious phenotypes in all analysed F2 plants. Homozygous *MM* and *mm* plants were monoecious and andromonoecious, respectively, while heterozygous *Mm* plants had a partially andromonoecious phenotype (Table 1).

A linkage analysis was also performed for two other sex expression traits that are known to be regulated by ethylene: the number of nodes before the production of the first pistillate flower (pistillate flowering transition) and the number of pistillate flowers per plant (Table 1). For pistillate flowering transition, no difference was detected among the four parental lines. For the number of pistillate flowers per plant, however, monoecious lines P85 and P86 showed less number of female flowers than the andromonoecious P87, and in F2 populations derived from crosses between these lines, *MM* plants also produced fewer female flowers than *mm* plants (Table 1).

Involvement of the *CitACS4* in floral organ maturation

The anthesis time was measured as the number of days it takes a floral bud of 2 mm in length to reach anthesis in a minimum of 10 flowers for each *CitACS4* genotype. Female and male flowers in monoecious (*MM*) plants differed in the time to reach complete maturation at anthesis, but no statistical difference was found between hermaphrodite and male flowers in andromonoecious plants

Table 2 Comparison of anthesis time (days) in pistillate and male flowers of monoecious (*MM*) and andromonoecious (*mm*) plants from parental and F2 generations

Generation	*CitACS4* genotype	Anthesis time (days)	
		Female flowers	Male flowers
P86	*MM*	6.0d	7.8c*
P87	*mm*	9.3ab	9.8b
F2 (P86XP87)	*MM*	8.4c	11.0a*
	Mm	8.6bc	9.8b*
	mm	9.8a	11.0a
P85	*MM*	6.1c	7.8a*
P87	*mm*	9.3b	9.8b
F2 (P85XP87)	*MM*	7.2c	10.0b*
	Mm	8.9b	10.0b*
	mm	11.7a	11.2b
P84	*MM*	6.0c	8.5d*
P87	*mm*	9.3a	9.8bc
F2 (P84XP87)	*MM*	8.4b	10.4a*
	Mm	9.0ab	9.4c
	mm	9.0ab	10.0ab

The trait was assessed in monoecious (*MM*) and andromonoecious (*mm*) parental lines, and in F2 plants (*MM*, *Mm* and *mm*) derived from monoecious x andromonoecious crosses. Statistical analysis was performed using the LSD test ($p \leq 0.05$). Different letters specify significant differences between genotypes within the same cross; * indicates significant differences between male and pistillate flowers of the same genotype and generation

(Table 2), suggesting that the presence of stamens delayed the aperture of both male and hermaphrodite flowers.

Given that male and female flowers differ in their production of ethylene (Manzano et al. 2016), we studied whether the ethylene derived from *CitACS4* expression could also regulate anthesis time in pistillate and male flowers. The anthesis time of pistillate flowers in the andromonoecious line P87 was delayed (average = 9.3 days) in comparison with that in the monoecious lines P86 (average = 6.0 days), P85 (average = 6.01 days) and P84 (average = 6.0 days) (Table 2). In the F2 generation of P86XP87 and P85XP87, the hermaphrodite flowers of andromonoecious *mm* plants also delayed anthesis in comparison with female flowers of *MM* plants (Table 2).

Differences in anthesis times were also found in male flowers of the monoecious (P84, P85 and P86) and andromonoecious (P87) lines, but those differences were not maintained among the *MM*, *Mm* and *mm* genotypes in the F2 generations (Table 2), suggesting that *CitACS4* and ethylene could control the maturation time of the pistillate flower, but not that of the male flower.

Involvement of the *CitACS4* gene in ovary and fruit development

A linkage analysis was performed between the *CitACS4* gene and floral organ size throughout development, including ovary and fruit. At earlier stages of pistillate flower development, ovary growth rate in *MM*, *Mm* and *mm* flowers was very similar (Fig. 2; Table S4). Significant differences in the petal and ovary size were detected, however, after 6 days, when the ovary and the corolla of *MM* flowers in both parental lines and F2 plants were larger than those of *mm* flowers (Fig. 2; Table S4). This higher growth rate in *MM* flowers was maintained up to anthesis, but given that *MM* flowers reached anthesis earlier than *mm* flowers, the size of the ovary at anthesis was smaller in the female flowers of monoecious *MM* plants than in the hermaphrodite flowers of *mm* plants (Fig. 2, Table S4). The ovary growth rate of heterozygous *Mm* flowers was intermediate with respect to that of the two homozygous genotypes in the three crosses (Fig. 2). These data demonstrated that the larger ovary size of *mm* flowers at anthesis is not due to a higher growth rate of the organ throughout development, but rather because the full maturation of floral organs and anthesis is delayed in hermaphrodite flowers.

Immediately after anthesis, pollinated fruits of monoecious *MM* lines (P86, P85 and P84) also grew at a higher rate than those of the andromonoecious *mm* line, but these differences were not detected between *MM* and *mm* fruits in the F2 populations of the three crosses (Fig. 3, Table S4). At 14 DPA, *MM* fruits were larger than *mm* fruits, but in the F2 generations significant differences between *MM* and *mm* fruits were only detected in the P84XP87 cross (Fig. 3, Table S4).

We also found a close linkage between the gene and fruit shape, in that *mm* fruits were rounder than *MM* ones. The fruit shape, estimated as the ratio between fruit length and width (FS), did not change between the two experimental seasons (Table 3). The P87 line produced round-shaped fruits at anthesis, at 14 DPA and at the mature stage, while the monoecious lines P85 and P86 displayed oval-shaped fruits at anthesis (Fig. 3; Table 3) which became rounded 14 DPA and at maturation (Table 3). In the F2 generations derived from P85XP87 and P86XP87, *MM* fruits at anthesis were also more elongated than *mm* fruits, but no significant difference was detected between *MM* and *mm* fruits at 14 DPA or at the mature stage. Fruits of the monoecious line P84, on the other hand, displayed a more elongated shape at anthesis, becoming oval shape at 14 DPA and at maturation, in both experimental seasons (Table 3), and in the F2 from cross P84XP87, *MM* plants produced a more elongated fruit than *mm* plants (Table 3), suggesting that the elongated fruit shape of P84 is linked to *CitACS4*.

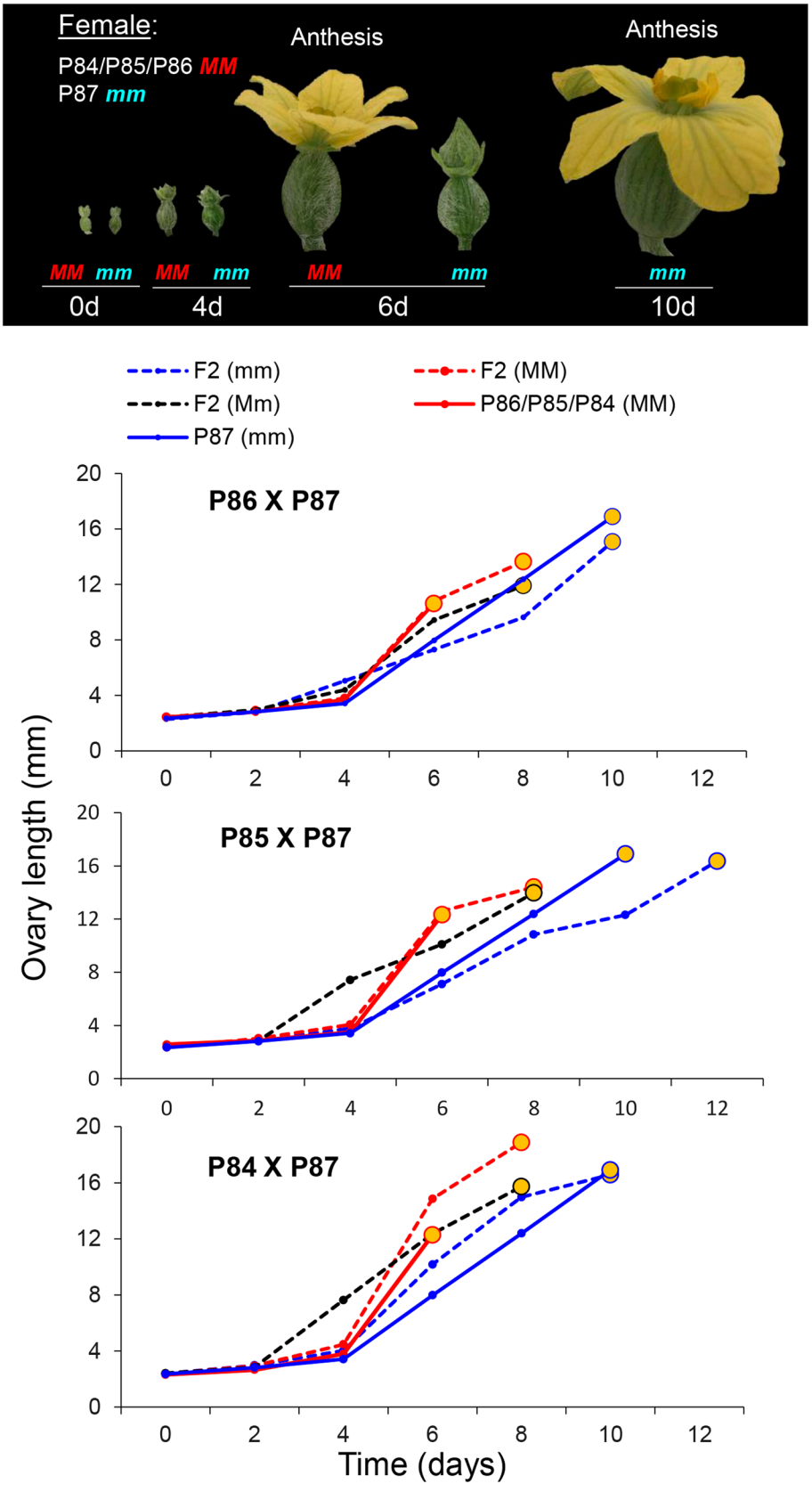


Fig. 2 Ovary growth rate in monoecious *MM* (P84, P85 and P86) and andromonoecious *mm* (P87) lines, and in *MM*, *Mm* and *mm* plants of three F2 generations derived from crosses between monoecious and andromonoecious lines. Flowers were labelled when they were 2 mm in length, and ovaries were measured every 2 days until anthesis. Average of 10–15 flowers and fruits for each *Cit-ACS4* genotype. Yellow circles show the average anthesis day. Significant differences between genotypes on each sampled day are shown in Table S4

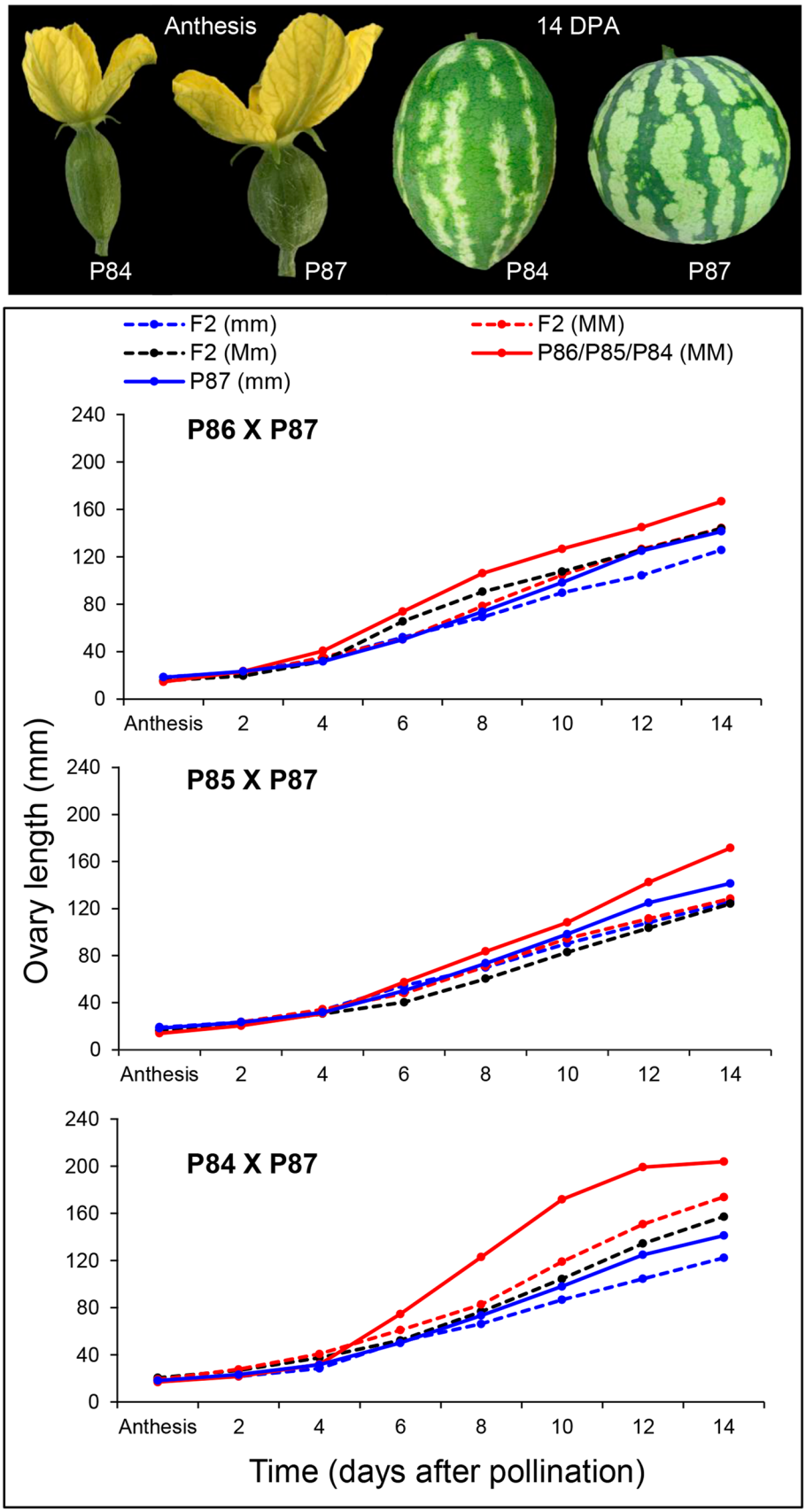


Fig. 3 Fruit growth rate in monoecious *MM* (P84, P85 and P86) and andromonoecious *mm* (P87) lines, and in *MM*, *Mm* and *mm* plants of three F2 generations derived from crosses between monoecious and andromonoecious lines. Fruit size was recorded every 2 days from anthesis up to 14 DPA. Average of 10–15 flowers and fruits for each *MM*, *Mm* and *mm* genotypes. Significant differences between genotypes on each sampled day are shown in Table S4

Table 3 Fruit shape index (FS) in *MM*, *Mm* and *mm* plants of parental lines and F2 generations

Generation	*CitACS4* genotype	FS (length/width ratio)					
		Spring/summer			Autumn/winter		
		Anthesis	14dpa	Mature	Anthesis	14dpa	Mature
P86	*MM*	1.46b	1.17a	1.15abc	1.33a	1.16a	1.06a
P87	*mm*	1.26c	1.12a	1.07c	1.22b	1.06b	0.99b
F2 (P86XP87)	*MM*	1.63a	1.15a	1.21a	1.28a	1.16a	1.09a
	Mm	1.41b	1.18a	1.17ab	1.30a	1.13ab	1.15a
	mm	1.26c	1.09a	1.11abc	1.23b	1.13ab	1.05ab
P85	*MM*	1.39a	1.15a	1.13ab	1.30a	1.14b	1.03bc
P87	*mm*	1.26b	1.12a	1.07b	1.22b	1.06c	0.99c
F2 (P85XP87)	*MM*	1.42a	1.18a	1.14ab	1.31a	1.20ab	1.13a
	Mm	1.42a	1.14a	1.19a	1.26ab	1.22a	1.14a
	mm	1.22b	1.10a	1.09ab	1.24ab	1.15ab	1.07ab
P84	*MM*	1.48bc	1.21ab	ND	1.59b	1.35b	1.32b
P87	*mm*	1.26d	1.12c		1.22d	1.06c	0.99c
F2 (P84XP87)	*MM*	1.82a	1.34a		1.82a	1.58a	1.69a
	Mm	1.50b	1.17bc		1.45c	1.45ab	1.31b
	mm	1.31 cd	0.94c		1.17d	1.05c	0.96c

The trait was assessed in parental monoecious (*MM*) and andromonoecious (*mm*) lines, and in F2 plants (*MM*, *Mm* and *mm*) derived from monoecious x andromonoecious crosses. FS was calculated as the ratio of fruit length to width. Data are the average of a minimum of 10 fruits per genotype. Statistical analysis was performed using the LSD method ($p \leq 0.05$), and the different letters indicate significant differences between *CitACS4* genotypes. ND, non-determined

Involvement of the *CitACS4* in fruit and seed set

Fruit and seed set were determined by hand-pollinating a minimum of 15 flowers for each *CitACS4* genotype and then assessing the number of setting fruits and viable seeds in at least 10 mature fruits. Since emasculation could decrease fruit and seed set, none of the bisexual or hermaphrodite flowers were emasculated before hand pollination.

Fruit set varied between monoecious and andromonoecious lines under the two studied conditions (spring/summer and autumn/winter), with P85 and P86 showing significantly higher fruit set than P84 and P87. We have determined whether these differences cosegregated with *CitACS4* alleles in the F2 generations. As expected, no difference was detected between *MM* and *mm* plants in the F2 generation of the cross P84XP87 (Fig. 4). In the P85XP87 and P86XP87 crosses, however, the higher fruit set of the monoecious parental lines was also observed in the monoecious *MM* plants of the segregating F2 generations (Fig. 4). Heterozygous *Mm* plants showed an intermediate percentage of fruit set (Fig. 4).

Table 4 compares the production of viable seeds in *MM*, *Mm* and *mm* fruits in parental lines and F2 generations of plants growing under autumn/winter conditions. The higher number of seeds in the monoecious *MM* lines (P84, P85 and P86) cosegregated with *MM* plants only in the F2 population of the cross P86XP87 (Table 4). In the crosses P84XP87 and P85XP87, however, both *MM* and *mm* F2 fruits produced a very low number of seeds, displaying no significant difference in seed set (Table 4).

Pollination and fertilization in female and hermaphrodite flowers

To investigate the possible factors accounting for the differences in fruit and seed set between female and hermaphrodite flowers, we compared pollen–stigma interaction, pollen tube germination and growth, and ovule fertilization in pistillate flowers of monoecious (P86 and P84) and andromonoecious (P87) lines at anthesis, and at −2, −1, +1 and +2 days post-anthesis (DPA). The results are shown in Table 5 and Fig. 5. In female flowers of P84 and P86, pollen adhesion and germination occur similarly between flowers at different phenological stages of the flowers, reaching a maximum around anthesis. In hermaphrodite flowers of P87, however, pollen adhesion was clearly reduced at −1 and −2 DPA, and pollen germination was almost nil in floral buds at −2 DPA (Table 5). The dynamic of pollen tube growth followed the same trend in female and hermaphrodite flowers of P86, P84 and P87, with a maximum number of pollen tubes in styles of flowers that were pollinated at anthesis (Table 5; Fig. 5). Nor were any differences found in pollen tube penetration and fertilization between female and hermaphrodite flowers at any of the floral stages at which

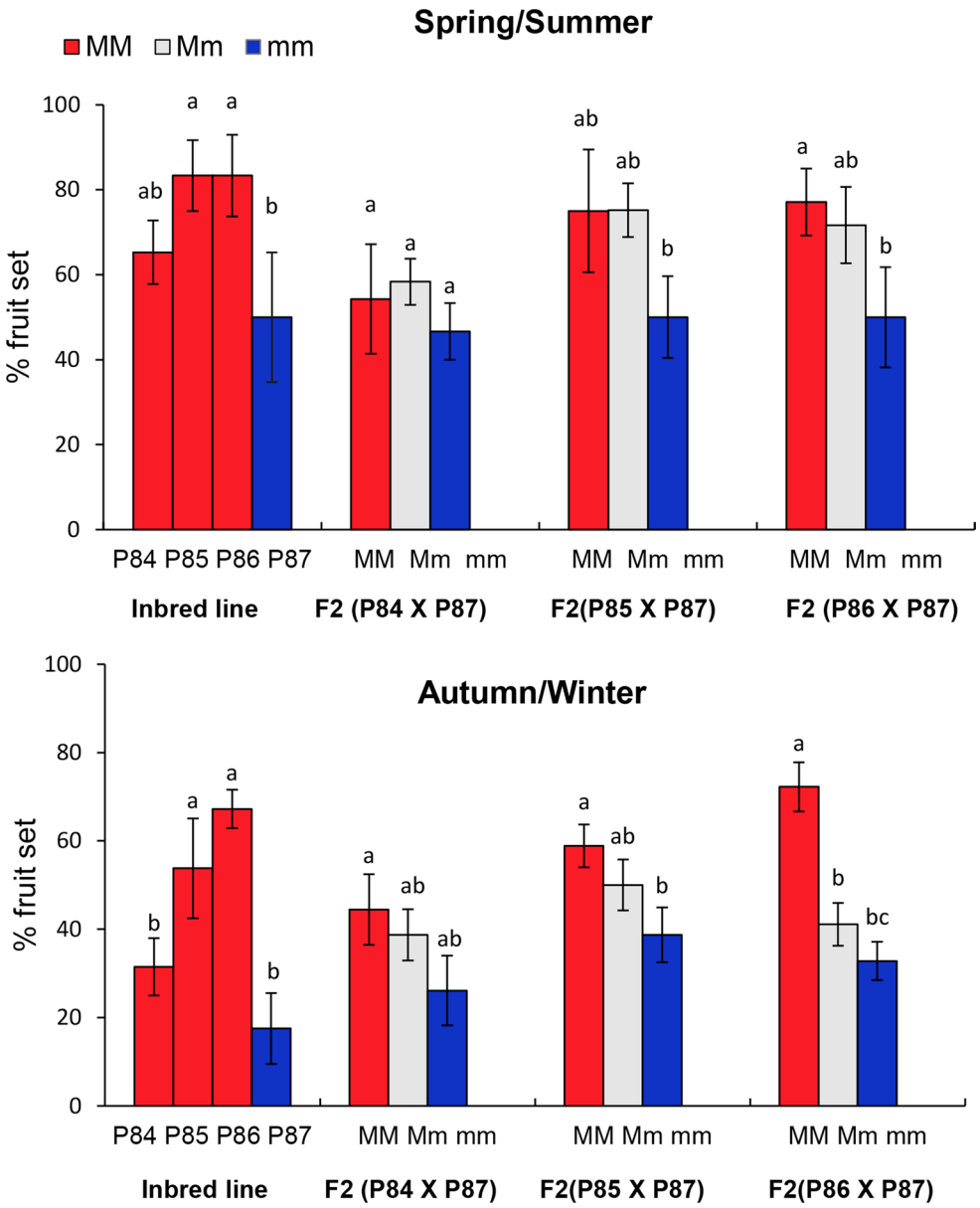


Fig. 4 Percentage of fruit set in monoecious (P84, P85 and P86) and andromonoecious (P87) lines, and in *MM*, *Mm* and *mm* plants of three F2 generations derived from crosses P84XP87, P85XP87 and P86XP87. Bars represent SE of at least 15 fruits. Different letters indicate significant differences between genotypes ($p \leq 0.05$)

they were pollinated (Table 5; Fig. 5). In the ovary of flowers pollinated at anthesis and + 1 DPA, pollen tubes were frequently observed close to the ovules. When flowers were pollinated at anthesis, fertilization rates were similar in the three *CitACS4* genotypes, although slightly higher in P87 flowers (Table 5).

Discussion

In the monoecious species of the *Cucurbitaceae* family, sex determination, i.e. the conversion of a putative hermaphrodite floral meristem into a female or a male flower, is known to be regulated by ethylene (Manzano et al. 2014). The key regulator is an ACS enzyme encoded by the orthologous genes *CmACS7*, *CsACS2*, *CpACS2/7* and *CitACS4* in melon, cucumber, zucchini and watermelon, respectively. These genes are specifically expressed in female floral buds at early stages of development, and their function results in the arrest of stamen development during the formation of the female flower (Boualem et al. 2008, 2009; Li et al. 2009; Martínez et al. 2014; Manzano et al. 2016; Ji et al. 2016). Although sex determination seems to be the main function of these genes, they could also control other developmental processes regulated by ethylene. In this paper, we have studied whether the watermelon *CitACS4* gene could also be involved in sex expression, floral organ development, including petals, ovaries and fruits, as well as in fruit and seed set. Results demonstrated that the ethylene produced in earlier pistillate floral buds is enough

Table 4 Number of viable seeds per kg of fruit pulp in *MM*, *Mm* and *mm* plants of parental lines and F2 generations growing under autumn/winter conditions

Generation	*CitACS4* genotype	Number of seeds
P84	*MM*	53.34a
P87	*Mm*	0b
F2 (P84XP87)	*MM*	0.69b
	Mm	0.59b
	mm	0b
P85	*MM*	50.04a
P87	*Mm*	0b
F2 (P85XP87)	*MM*	12.48b
	Mm	1.35b
	mm	4.12b
P86	*MM*	70.32a
P87	*Mm*	0 c
F2 (P86XP87)	*MM*	50.73ab
	Mm	27.6bc
	mm	10.61c

SE of at least 10 fruits per line and generation. Statistical analysis was performed using the LSD method ($p \leq 0.05$), and the different letters indicate statistical differences between genotypes within the same cross

to control the entire development of pistillate flowers and fruits.

In *CitACS4* loss of function mutants (*mm*), female flowers are converted into hermaphrodite flowers, and monoecious into andromonoecious plants (Manzano et al. 2016; Boualem et al. 2016; Ji et al. 2016). Crosses indicated that the monoecious allele *M* of P84, P85 and P86 is semi-dominant to the andromonoecious allele *m* of P87. The stability of sex determination, however, varied among monoecious lines (*MM*), showing P85 and P86 a more stable monoecy than P84. Consequently, the P84 line not only produced male and female flowers, but also several bisexual flowers with stamens at different developmental stages.

Regulation of sex expression

Sex expression, i.e. the transition from male to female phases of development (pistillate flowering transition), and the female-to-male flower ratio are regulated by ethylene in watermelon and other cucurbit species, although in a completely opposite way. In *Cucumis* and *Cucurbita* species, ethylene promotes femaleness (Rudich 1990; Perl-Treves 1999; Manzano et al. 2011, 2013), while in watermelon it stimulates maleness, delaying female flowering transition and reducing the number of pistillate flowers per plant (Sugimaya et al. 1998; Manzano et al. 2014; Zhang et al. 2017). Our data indicate that the reduced ethylene in *mm* pistillate floral buds could affect the number of pistillate flower per plant but not the pistillate flowering transition. In the previous data, we found that sex expression traits were not affected by *CitACS4* gene. The differences between the previous and current results are likely due to the influence of grafts. In the previous experiments, watermelon plants were grafted on *Cucurbita* rootstocks, which surely altered their sexual expression (Manzano et al. 2014). In the present

Table 5 Pollen–pistil interaction 24 h after pollination at different dates in P84, P86 and P87 lines

Line	Pollination date (days from anthesis)	Pollen adhesion (grains/mm^2)	Pollen germination (%)	Pollen tube number in style	Pollen tubes in ovary and fertilization[2] (%)
P84	− 2	10.8b[1]	37.7b	5–25	Not observed
	− 1	18.8ab	57.8ab	> 25	Some pollen tubes
	0	22.5a	69.0a	> 25	20 ± 1
	+ 1	11.4b	69.1a	> 25	Many pollen tubes
	+ 2	20.6ab	67.8a	> 25	Many pollen tubes
P86	− 2	10.6b	40b	5–25/> 25	Very few pollen tubes
	− 1	15.3b	42.8ab	> 25	Few pollen tubes
	0	11.0b	69.5a	> 25	27 ± 1
	+ 1	29.8a	37.8ab	> 25	Many tubes near ovules
	+ 2	17.1b	26.7b	< 25	Some pollen tubes
P87	− 2	7.1ab	0.76b	5–25/> 25	Not observed
	− 1	5.8b	60.3a	> 25	Few pollen tubes
	0	25.4a	60.2a	> 25	33 ± 8
	+ 1	23.2ab	62.1a	> 25	P.t. near ovules
	+ 2	16.9ab	49.8a	< 25	P.t. near ovules

[1]Mean values ($n = 10$) followed by different letters in each line and column indicate means significant differences at $p < 0.05$ by Tukey test

[2]Fertilization expressed as mean value ± standard error (%)

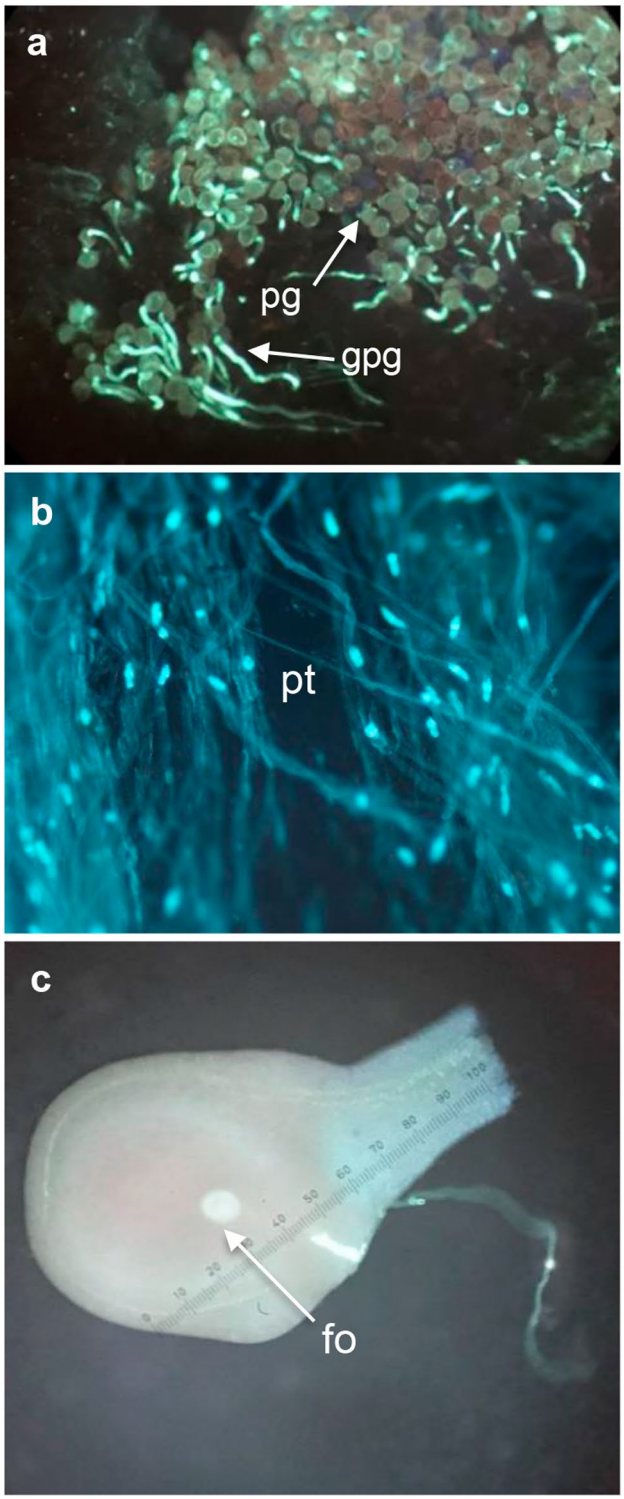


Fig. 5 Pollination and fertilization in andromonoecious line P87. **a** Pollen adhesion and germination in the stigma. **b** Pollen tubes growing on the style. **c** Fertilized ovule. pg, pollen grain; gpg, germinated pollen grain; pt, pollen tubes; fo, fertilized ovule

study, however, none of the plants were grafted, and monoecious and andromonoecious plants showed more noticeable differences in the number of pistillate and male flowers per plant. Even so, P86 and P87 parental lines showed no difference in sex expression, which suggests the influence of the genetic background or the existence of other genetic factors regulating this trait. Therefore, although sex determination and sex expression are different mechanisms (Manzano et al. 2014), our present data demonstrate that the ethylene required for sex determination in the pistillate flowers, converting the putative hermaphrodite floral buds into female flowers, can influence the number of female flower per plant. Given that sex expression mechanisms should be controlled from the apical meristem, it is likely that the ethylene produced in pistillate floral buds (encoded by *CitACS4*) can influence the production and/or the action of ethylene in the apical shoot, as proposed by Manzano et al. (2013) in zucchini squash.

Coordination of floral organ development

Our data also demonstrate that *CitACS4* is a coordinator of floral organ development, acting as a repressor of stamen development, but also as a promoter of ovary and corolla growth, and consequently of the maturation and aperture of the pistillate flower. The growth rates of both petals and ovaries were significantly higher in female flowers of *MM* plants than in hermaphrodite flowers of *mm* plants. Consequently, flower maturation and anthesis time are delayed in hermaphrodite flowers in comparison with female ones, which leads to larger *mm* ovaries than *MM* ones at anthesis, despite their reduced growth rates. The differences in flower maturation between *mm* and *MM* plants did not affect male flowers, which showed a significantly longer anthesis time than female flowers in *MM* but not in *mm* plants (Table 2). These data strongly suggest that the masculinization of the pistillate flower in the andromonoecious *mm* plants decreases the growth rate of the flower and delays its maturation and aperture for about 2 days, in a similar way as occurs in male flowers. Similar results have been found in zucchini, where the delayed anthesis of bisexual and hermaphrodite flowers resulted in ovaries much larger than those of female flowers (Martínez et al. 2013). Since female flowering occurs at later stages of the plant development, it is likely that the acceleration of the anthesis in female flowers was a coevolutionary mechanism that ensured pollination during the evolution of monoecy in the *Cucurbitaceae* family.

The role of ethylene as a promoter of carpel development, but also as a repressor of stamen development has been found not only in the unisexual flowers of cucurbit species (Boualem et al. 2008, 2009; Li et al. 2009; Martínez et al. 2014; Manzano et al. 2016), but also in the hermaphrodite flowers of *Arabidopsis* and other species. Overexpression of

the ethylene biosynthesis cucumber gene *CsACO2* represses stamen development in *Arabidopsis* (Duan et al. 2008), while downregulation of the ethylene receptor gene *ETR1* reduces the ethylene signalling repressor CTR1, resulting in the production of female flowers in *Arabidopsis* (Wang et al. 2010). Transgenic tobacco silencing an *ACO* gene has shown female sterility due to an arrest of megasporogenesis (De Martinis et al. 1999). Here we demonstrate that ethylene is also a positive regulator of petal development and maturation and therefore of corolla aperture and anthesis time.

Regulation of fruit set and development

Data indicated that *CitACS4* affects developmental events occurring before or at anthesis, including fruit shape and fruit setting, but not those occurring after anthesis and pollination such as fruit growth rate and final size. In the P84XP87 cross, the monoecious *M* allele was linked to elongated fruits, while the andromonoecious *m* allele was rather linked to round-shaped fruits. Since the final shape of a fruit depends on ovary shape at anthesis (Perin et al. 2002), it is not difficult to realize that a gene like *CitACS4*, which is specifically expressed in floral buds at earlier stages of development, can regulate the final shape of the fruit. The association of hermaphrodite flowers with round-shaped fruits was first reported by Rosa (1928) and later by Kubicki (1962) and Wall (1967) in melon. Among QTLs controlling fruit shape in melon, that in LGII seems to be a pleiotropic effect of the sex-determining gene *CmACS7* (Perin et al. 2002; Díaz et al. 2014). The monoecious (*M*) and andromonoecious (*m*) alleles of the cucumber gene *CsACS2* also cosegregate with elongated and round fruits, respectively, although a novel allele of the gene (*m1*), encoding for a truncated protein, is responsible for elongated fruit shape and andromonoecy (Tan et al. 2015). In zucchini, hermaphrodite flowers of monoecious unstable cultivars produce larger ovaries and fruits, but fruit shape is not altered (Martínez et al. 2014).

The cosegregation between *m* allele and a reduced fruit set in P85XP87 and P86XP87 crosses could be the result of *CitACS4*, but the existence of other linked genes cannot be ruled out. Moreover, the fact that the monoecious line P84 does not differ in fruit set with respect to the andromonoecious one P87 also suggests that the trait is influenced by other unlinked loci. The role of ethylene in fruit setting has not been studied in depth. A downregulation of ethylene biosynthesis and signalling genes has been observed immediately after anthesis in pollinated, GA3-treated and parthenocarpic fruits of tomato (Vriezen et al. 2008) and zucchini (Martínez et al. 2013). Ethylene produced in the ovules appears to be responsible for both the ovule lifespan and the fate of the ovary/fruit in tomato (Olimpieri et al. 2007) and *Arabidopsis* (Carbonell-Bejerano et al. 2010, 2011), controlling fruit set in response to GA in *Arabidopsis*-unfertilized ovaries. In zucchini, the inhibition of ethylene biosynthesis or response is sufficient to induce the set and early development of the fruit in the absence of pollination, demonstrating a direct involvement of ethylene in fruit set (Martínez et al. 2013). This post-anthesis ethylene, which could be involved in ovule senescence and fruit abortion, does not appear to be the same as the one responsible for *mm* fruit abortion in watermelon. Although this interesting finding requires more research, it seems that fruit set requires higher ethylene production in the immature flower buds, probably for a coordinated development and maturation of floral organs at anthesis, but lower ethylene production in ovules and fruits immediately after anthesis, because at this later stage ethylene could trigger ovule senescence and consequently fruit abortion.

Under unfavourable environmental conditions of autumn/winter, the correct set of seeds in fruits was found to be linked to the M/m locus only in one of the analysed F2 populations (P86XP87), where monoecious *MM* fruits had higher seed yield than andromonoecious *mm* ones. This trait is very influenced by environmental conditions, especially temperature. The lack of linkage between seed set and M/m locus in the other two crosses could indicate the existence of other major loci regulating this trait in watermelon. Therefore, the role of the M/m locus in the regulation of watermelon seed set will require further experimental work. Comparison of pollen adhesion, pollen tube growth and ovule fertilization in monoecious and andromonoecious lines shed also no light on seed abortive mechanisms in fruits derived from hermaphroditic *mm* flowers. The pollination window at which pollen adheres to the stigma is slightly delayed in *mm* flowers, but pollen adhesion and pollen tubes observed at anthesis and + 1 and + 2 DPA ensured fertilization in both *MM* and *mm* flowers. Therefore, it is likely that the loss of seeds in the P87 line is not due to a lack of pollination or fertilization events, but rather to a premature abortion of fertilized ovules. Ethylene plays a significant role in ovule development and female gametophyte fertility (Tsai et al. 2008; Clark et al. 2010). In *Arabidopsis*, the onset of ovule senescence and the time window for the pistil to respond to GA treatments is modulated by ethylene (Carbonell-Bejerano et al. 2011). Silencing of ethylene biosynthesis genes in transgenic tobacco plants results in a reversible inhibition of ovule development (De Martinis and Mariani 1999). Moreover, ethylene biosynthesis and signalling genes have been found to be hypomethylated in the female sterile rice mutant *fsv*, in which the ethylene genes were upregulated and then downregulated during ovule development (Yang et al. 2016; Liu et al. 2017).

Taken together, the results presented in this paper indicate that, in addition to arresting stamen growth and development and determining the sex identity of the female flower, the ethylene biosynthesis gene *CitACS4* is capable of regulating

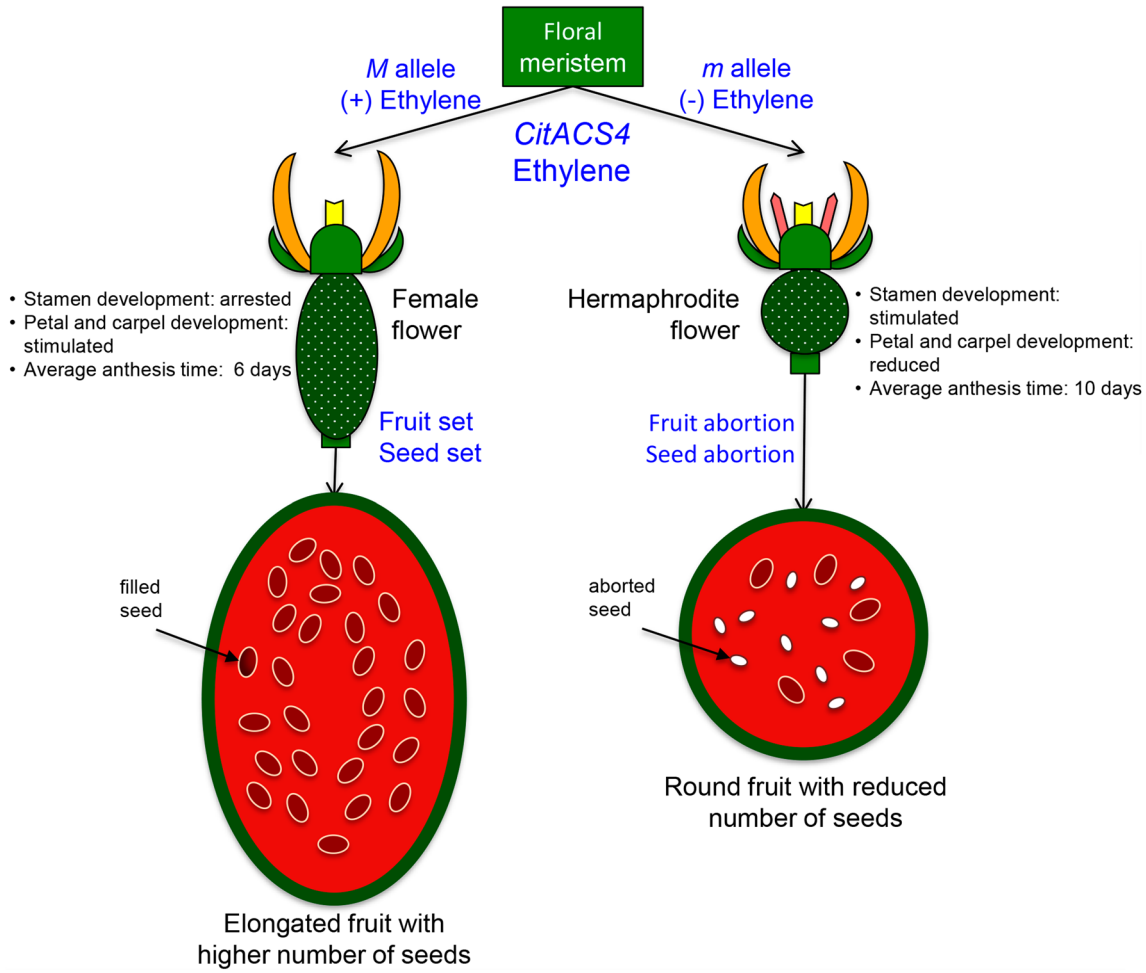


Fig. 6 Involvement of the ethylene biosynthesis gene *CitACS4* in watermelon flower and fruit development. In the monoecious *MM* plants, the production of ethylene in the floral meristem arrests the development of stamens, but stimulates the growth of petals and carpels, which results in a female flower with an elongated ovary. In the *mm* plants, the lack of ethylene production prevents stamen arrest and reduces the growth rate of petals and carpels, which results in a hermaphrodite flower with a round-shaped ovary. After pollination, the reduced ethylene production in the *mm* flowers could be also responsible of fruit and seed abortion observed in the andromonoecious mm line

several developmental processes that occur in the pistillate flower and in the early development of the fruit (Fig. 6). The decrease in the production of ethylene associated with the loss of function *m* allele prevents stamen arrest, but inhibits the development of the petals and carpels, making the flower reach anthesis about 4 days later. The result at anthesis is a hermaphrodite flower with a rounder and larger ovary than that of the female flower. The lack of ethylene during the development of the hermaphrodite flower could also explain the reduced fruit set found in the andromonoecious *mm* plants (Fig. 6).

Author contribution statement MJ and EA conceived and designed the experiments. EA, AG, SM, VP and JC performed the experiments. EA and MJ analysed the data. MJ, EA, AG, JV and SM contributed reagents/materials/analysis tools. MJ and EA wrote the paper. All authors read and approved the manuscript.

Acknowledgements AG and EA acknowledge FPI and "Garantía juvenil" scholarship programmes from MEC, Spain. This work has been supported by Grants AGL2014-54598-C2-1-R and AGL2017-82885-C2-1-R, partly funded by the ERDF (European Regional Development Fund) and by the Spanish Ministry of Science and Innovation, and Grant P12-AGR-1423, funded by Junta de Andalucía, Spain.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abdelmohsin M, Pitrat M (2008) Pleiotropic effect of sex expression on fruit shape in melon. In: Pitrat M (ed) Cucurbitaceae 2008, Proceedings of the IXth EUCARPIA meeting on genetics and breeding of Cucurbitaceae. INRA, Avignon, pp 551–556

Balbi V, Lomax T (2003) Regulation of early tomato fruit development by the diageotropica gene. Plant Physiol 131:186–197. https://doi.org/10.1104/pp.010132.auxin

Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science 321:836–838. https://doi.org/10.1126/science.1159023

Boualem A, Troadec C, Kovalski I, Sari MA, Perl- Treves R, Bendahmane A (2009) A conserved ethylene biosynthesis enzyme leads to andromonoecy in two *Cucumis* species. PLoS ONE 4:e6144. https://doi.org/10.1371/journal.pone.0006144

Boualem A, Lemhemdi A, Sari MA, Pignoly S, Troadec C, Choucha FA, Solmaz I, Sari N, Dogimont C, Bendahmane A (2016) The andromonoecious sex determination gene predates the separation of *Cucumis* and *Citrullus* genera. PLoS ONE 11:1–13. https://doi.org/10.1371/journal.pone.0155444

Carbonell-Bejerano P, Urbez C, Carbonell J, Granell A, Perez-Amador MA (2010) A fertilization-independent developmental program triggers partial fruit development and senescence processes in pistils of *Arabidopsis*. Plant Physiol 154:163–172. https://doi.org/10.1104/pp.110.160044

Carbonell-Bejerano P, Urbez C, Granell A, Carbonell J, Perez-Amador MA (2011) Ethylene is involved in pistil fate by modulating the onset of ovule senescence and the GA-mediated fruit set in *Arabidopsis*. BMC Plant Biol 11:84. https://doi.org/10.1186/1471-2229-11-84

Clark G, Torres J, Finlayson S, Guan X, Handley C, Lee J, Kays JE, Chen ZJ, Roux SJ (2010) Apyrase (nucleoside triphosphate-diphosphohydrolase) and extracellular nucleotides regulate cotton fiber elongation in cultured ovules. Plant Physiol 152:1073–1083. https://doi.org/10.1104/pp.109.147637

Cuevas J, Rallo L, Rapoport HF (1994) Procedure to study ovule senescence in olive. Acta Hortic 356:252–256

De Martinis D, Mariani C (1999) Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants. Plant Cell 11:1061–1072. https://doi.org/10.1105/tpc.11.6.1061

De Martinis D, Haenen I, Pezzotti M, Benvenuto E, Mariani C (1999) Ethylene and flower development in tobacco plants. In: Kanellis AK, Chang C, Klee H, Bleecker AB, Pech JC, Grierson D (eds) Biology and biotechnology of the plant hormone ethylene II. Springer, Dordrecht, pp 157–164. https://doi.org/10.1007/978-94-011-4453-7_26

Díaz A, Zarouri B, Fergany M, Eduardo I, Álvarez JM, Picó B, Monforte AJ (2014) Mapping and introgression of QTL involved in fruit shape transgressive segregation into 'Piel de Sapo' melon (*Cucumis melo* L.). PLoS ONE 9(8):e104188. https://doi.org/10.1371/journal.pone.0104188

Duan QH, Wang DH, Xu ZH, Bai SN (2008) Stamen development in *Arabidopsis* is arrested by organ-specific overexpression of a cucumber ethylene synthesis gene CsACO2. Planta 228:537–543. https://doi.org/10.1007/s00425-008-0756-7

FAOSTAT (2017) http://www.fao.org/faostat/es/#data/QC. Accessed 1 Dec 2017

Ji G, Zhang J, Gong G, Shi J, Zhang H, Ren Y, Guo S, Gao J, Shen H, Xu Y (2015) Inheritance of sex forms in watermelon (*Citrullus lanatus*). Sci Hortic 193 (2015):367–373. https://doi.org/10.1016/j.scienta.2015.07.039

Ji G, Zhang J, Zhang H, Sun H, Gong G, Shi J, Tian S, Guo S, Ren Y, Shen H, Gao J, Xu Y (2016) Mutation in the gene encoding 1-aminocyclopropane-1-carboxylate synthase 4 (CitACS4) led to andromonoecy in watermelon. J Integr Plant Biol 58(9):762–765. https://doi.org/10.1111/jipb.12466

Kubicki B (1962) Inheritance of some characters in muskmelons (*Cucumis melo*). Genet Pol 3:265–274

Larsen PB, Woltering EJ, Woodson WR (1993) Ethylene and inter-organ signaling in flowers following pollination. In: Schultz J, Raskin I (eds) Plant signals in interactions with others organisms. American Society of Plant Physiologists, Rockville, pp 112–122

Li Z, Huang S, Liu S, Pan J, Zhang Z, Tao Q, Shi Q, Jia Z, Zhang W, Chen H, Si L, Zhu L, Cai R (2009) Molecular isolation of the *M* gene suggests that a conserved-residue conversation induces the formation of bisexual flowers in cucumber plants. Genetics 182:1381–1385. https://doi.org/10.1534/genetics.109.104737

Liu H, Wu Y, Cao A, Mao B, Zhao B, Wang J (2017) Genome-wide analysis of DNA methylation during ovule development of female-sterile rice fsv1. G3 (Bethesda, Md) 7:3621–3635. https://doi.org/10.1534/g3.117.300243

Loy JB (2006) Fruit size in melon in monoecious and andromonoecious isolines. Cucurbits Genet Coop Rep 28–29:12–13

Manzano S, Martínez C, Megías Z, Gómez P, Garrido D, Jamilena M (2011) The role of ethylene and brassinosteroids in the control of sex expression and flower development in *Cucurbita pepo*. Plant Growth Regul 65:213–221. https://doi.org/10.1007/s10725-011-9589-7

Manzano S, Martínez C, Megías Z, Garrido D, Jamilena M (2013) Involvement of ethylene biosynthesis and signalling in the transition from male to female flowering in the monoecious *Cucurbita pepo*. J Plant Growth Regul 32:789–798. https://doi.org/10.1007/s00344-013-9344-6

Manzano S, Martínez C, García JM, Megías Z, Jamilena M (2014) Involvement of ethylene in sex expression and female flower development in watermelon (*Citrullus lanatus*). Plant Physiol Biochem 85:96–104. https://doi.org/10.1016/j.plaphy.2014.11.004

Manzano S, Aguado E, Martínez C, Megías Z, García A, Jamilena M (2016) The ethylene biosynthesis gene CitACS4 regulates monoecy/andromonoecy in watermelon (*Citrullus lanatus*). PLoS ONE 11(5):e0154362. https://doi.org/10.1371/journal.pone.0154362

Martin FW (1959) Staining and observing pollen tubes in the style by means of fluorescence. Stain Technol 34:125–128. https://doi.org/10.3109/10520295909114663

Martínez C, Manzano S, Megías Z, Garrido D, Picó B, Jamilena M (2013) Involvement of ethylene biosynthesis and signalling in fruit set and early fruit development in zucchini squash (*Cucurbita pepo* L.). BMC Plant Biol 13:139. https://doi.org/10.1186/1471-2229-13-139

Martínez C, Manzano S, Megías Z, Barrera A, Boualem A, Garrido D, Bendahmane A, Jamilena M (2014) Molecular and functional characterization of CpACS27A gene reveals its involvement in monoecy instability and other associated traits in squash (*Cucurbita pepo* L.). Planta 239:1201–1215. https://doi.org/10.1007/s00425-014-2043-0

Monforte AJ, Eduardo I, Abad S, Arús P (2005) Inheritance mode of fruit traits in melon: heterosis for fruit shape and its correlation with genetic distance. Euphytica 144:31–38. https://doi.org/10.1007/s10681-005-0201-y

Olimpieri I, Siligato F, Caccia R, Mariotti L, Ceccarelli N, Soressi GP, Mazzucato A (2007) Tomato fruit set driven by pollination or by the parthenocarpic fruit allele are mediated by transcriptionally regulated gibberellin biosynthesis. Planta 226:877–888. https://doi.org/10.1007/s00425-007-0533-z

Pandolfini T, Molesini B, Spena A (2007) Molecular dissection of the role of auxin in fruit initiation. Trends Plant Sci 12:327–329. https://doi.org/10.1016/j.tplants.2007.06.011

Perin C, Gomez-Jimenez MC, Hagen L, Dogimont C, Pech JC, Latché A, Pitrat M, Lelièvre JM (2002) Molecular and genetic characterization of a non-climacteric phenotype in melon reveals two loci conferring altered ethylene response in fruit. Plant Physiol 129:300–309. https://doi.org/10.1104/pp.010613

Perl-Treves R (1999) Male to female conversion along the cucumber shoot: approaches to studying sex genes and floral development in *Cucumis sativus*. In: Ainsworth CC (ed) Sex determination in plants. BIOS Scientific, Oxford, pp 193–221

Poole CF, Grimball PC (1945) Interaction of sex, shape and weight genes in watermelon. J Agric Res 71:533–552

Prothro J, Abdel-Haleem H, Bachlava E, White V, Knapp S, McGregor C (2013) Quantitative trait loci associated with sex expression in an inter-subspecific watermelon population. J Am Soc Hortic Sci 138(2):125–130

Rosa J (1928) The inheritance of flower types in *Cucumis* and *Citrullus*. Hilgardia 3(9):233–250. https://doi.org/10.3733/hilg.v03n09p233

Rudich J (1990) Biochemical aspects of hormonal regulation of sex expression in Cucurbits. In: Bates DM, Robinson RW, Jeffrey C (eds) Biology and utilization of the cucurbitaceae. Cornell University Press, Ithaca, pp 269–280

Rudich J, Zamski E (1985) *Citrullus lanatus*. In: Halevy A (ed) Handbook of flowering. CRC Press, Boca Raton, pp 272–274

Sakata Y, Fukino N, Ohara T, Sugiyama M, Yoshioka Y (2013) Effect of monoecious trait on fruit shape and total soluble solid contents of melon. Hort Res (Japan) 12:15–22. https://doi.org/10.2503/hrj.12.15

Salman-Minkov A, Levi A, Wolf S, Trebitsh T (2008) ACC synthase genes are polymorphic in watermelon (*Citrullus* spp.) are differentially expressed in flowers and response to auxin and gibberellin. Plant Cell Physiol 49:740–750. https://doi.org/10.1093/pcp/pcn045

Stepnova AN, Robertson-Hoy J, Yun J, Benavente LM, Xie DY, Doležal K, Schlereth A, Jürgens G, Alonso JM (2008) TAA1-Mediated auxin biosynthesis is essential for hormone crosstalk and plant development. Cell 133:177–191. https://doi.org/10.1016/j.cell.2008.01.047

Sugimaya K, Kanno T, Morishita M (1998) Evaluation method of female flower bearing ability in watermelon using silver thiosulfate (STS). J Jpn Soc Hortic Sci 67(2):185–189. https://doi.org/10.1248/cpb.37.3229

Tan J, Tao Q, Niu H, Zhang Z, Li D, Gong Z, Weng Y, Li Z (2015) *A* novel allele of monoecious (m) locus is responsible for elongated fruit shape and perfect flowers in cucumber (*Cucumis sativus* L.). Theor Appl Genet 128:2483. https://doi.org/10.1007/s00122-015-2603-0

Tsai WC, Hsiao YY, Pan ZJ et al (2008) The role of ethylene in orchid ovule development. Plant Sci 175:98–105. https://doi.org/10.1016/j.plantsci.2008.02.011

Tsukahara A (1988) Semi-forcing culture in watermelon. The system of agricultural technology, melon and watermelon. Soc Cult Agr Vill Press, Tokyo, pp 194–196 (**In Japanese**)

Vriezen WH, Feron R, Maretto F, Keijman J, Mariani C (2008) Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytol 177:60–76. https://doi.org/10.1111/j.1469-8137.2007.02254.x

Wall JR (1967) Correlated inheritance of sex expression and fruit in *Cucumis*. Euphytica 16:199–208. https://doi.org/10.1007/BF00043455

Wang X, Goshe MB, Soderblom EJ, Phinney BS, Kuchar JA, Li J, Asami T, Yoshida S, Huber SC, Clouse SD (2005) Identification and functional analysis of in vivo phosphorylation sites of the *Arabidopsis* BRASSINOSTEROID-INSENSITIVE1 receptor Kinase. Plant Cell 17:1685–1703. https://doi.org/10.1105/tpc.105.031393.tion

Wang DH, Li F, Duan QH, Han T, Xu ZH, Bai SN (2010) Ethylene perceptions is involved in female cucumber flower development. Plant J 61:862–872. https://doi.org/10.1111/j.1365-313X.2009.04114.x

Wechter WP, Levi A, Harris KR, Davis AR, Fei Z, Katzir N, Giovannoni JJ, Salman-Minkov A, Hernandez A, Thimmapuram J, Tadmor Y, Portnoy V, Trebitsh T (2008) Gene expression in developing watermelon fruit. BMC Genom 9:1–13. https://doi.org/10.1186/1471-2164-9-275

Yang L, Wu Y, Yu M, Mao B, Zhao B, Wang J (2016) Genome-wide transcriptome analysis of female-sterile rice ovule shed light on its abortive mechanism. Planta 244(5):1011–1028. https://doi.org/10.1007/s00425-016-2563-x

Zhang J, Shi J, Ji G, Zhang H, Gong G, Guo S, Ren Y, Fan J, Tian S, Xu Y (2017) Modulation of sex expression in four forms of watermelon by gibberellin, ethephone and silver nitrate. Hortic Plant J 3(3):91–100. https://doi.org/10.1016/j.hpj.2017.07.010