EXPLORING GENETIC DIVERSITY UNDERLYING FRUIT DIVERSITY IN TOMATO AND PEPPER

by

NATHAN EDWARD KING TAITANO

(Under the Direction of Esther van der Knaap)

ABSTRACT

Tomato and pepper landraces and cultivars exhibit diverse fruit morphology, contributing to many different uses and market niches. The genetic underpinnings of that diversity are interesting from both applied and basic developmental research perspectives. This dissertation catalogues genetic diversity in a set of pepper and tomato landrace and wild-growing accessions, with a focus on how such genetic diversity underlies fruit shape diversity. Genotyping-by-sequencing, structural variant analysis, QTL-mapping, and coexpression network analysis are all employed to explore that diversity. These analyses result in a genetic description of the population structure of pepper landraces from southern Mexico, a refined understanding of the position, gene action, and phenotype of the pepper fruit-shape QTL *fs3.1*, coexpression networks of several tomato *TRMs* which putatively underlie fruit shape, and a set of structural variants intersecting genes in said coexpression networks. These point to a shared mechanism controlling fruit shape in pepper and tomato, and provide resources for further study of fruit shape in both.

INDEX WORDS: Fruit morphology, TONNEAU1 RECRUITING MOTIF PROTEIN, Population genetics, OVATE FAMILY PROTEIN, Pepper, Tomato, QTL mapping, Comparative genomics.

EXPLORING GENETIC DIVERSITY UNDERLYING FRUIT DIVERSITY IN TOMATO AND PEPPER

by

NATHAN EDWARD KING TAITANO

BS, College of Wooster, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

EXPLORING GENETIC DIVERSITY UNDERLYING FRUIT DIVERSITY IN TOMATO AND PEPPER

by

NATHAN EDWARD KING TAITANO

Major Professor:	Esther van der Knaap
Committee:	Jason Wallace Cecilia McGregor John Burke Chung-Jui Tsai

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
May 2020

DEDICATION

This dissertation is dedicated to the generations of farmers, breeders, and seed-traders, both modern and traditional, who worked for millennia to provide humanity with plentiful, delicious, and nutritious food.

ACKNOWLEDGEMENTS

I first acknowledge my family, especially my wife Mae Manupipatpong, who stayed by my side through episodic long-distance, late dinners after long hours working in the lab and greenhouse, and writing sessions that continued well into the night. Thanks also to my advisor Esther van der Knaap, who provided essential guidance and feedback throughout this process. Thanks also to committee members Jason Wallace, John Burke, CJ Tsai, and Cecilia McGregor for all the training and assistance they provided. I also thank all the members of the van der Knaap lab who have worked with me throughout the past few years: Zhengkun Qiu, Shan Wu, Liang Sun, Yanping Wang, Neda Keyhaninejad, Jammi McClead, Jason Van Houten, Qi Mu, Yi-Hsuan Chu, Eudald Illa-Berenguer, Ashley Snouffer, Katherine Hardigree, Alexis Ramos, Carmen Kraus, Yasin Topcu, Manoj Sapkota, Biyao Zhang, Qiang Li, Lei Zhang, Lara Pereira, Dylan Brumblow, and Shamir Harry. Also, thanks to my fellow graduate students for their camaraderie and friendship - in particular Bill Ricci, John Spiekerman, Vivian Bernau, Lav Yadav, Laxman Adhikari, Maria Ortega, Evan McCoy, Davis and Winnie Gimode. Finally, thank you to the University of Georgia's Glenn and Helen Burton Feeding the Hungry Scholarship, the Center for Applied Plant Sciences at the Ohio State University, and the OARDC SEEDS grant for funding support.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Dissertation Structure and Research Introduction

While the rigorous selection imposed by modern crop breeding programs efficiently improves selected traits in modern cultivars, it decreases crop biodiversity. Less rigorous selection permits greater genetic diversity in crop landraces. Furthermore, unlike in modern breeding that generally targets the same growing environment and market niche, the peculiarities of place and taste govern landrace selection. A landrace must be particularly adapted to its home environment, and often is also adapted to suit a specific use. Landraces growing near the domestication center of a given crop may have been accumulating such adaptations for thousands of years - beginning just after the crop's domestication. These features make landraces useful for two kinds of long-term research: 1) comparing the evolutionary histories of multiple independent populations as they branch off from the original domesticated population, and 2) identifying alleles conferring useful adaptations to harvested organ morphology and helping to narrow down loci to pinpoint candidate genes. Additionally, by drawing on, processing, and synthesizing genomic data in collaboration with others doing similar research in other crop species, particularly the tomato (*Solanum spp.*) model fleshy fruit system, we sought to expand our research into a series of tomato cultivars by exploring structural variants in coexpression networks of a select set of candidate genes thought to participate in a widely conserved fruit shape regulatory network. These aims were addressed in the following research chapters:

Chapter 2 studied the genetic diversity of early chile pepper landraces. The genetic diversity and history of admixture/separation was assessed using genotyping-by-sequencing (GBS) on a population of chile pepper landraces taken from near the center of *C. annuum* domestication in Southern Mexico.

Chapter 3 was aimed at identifying loci conferring shape variation in landrace peppers. QTL-mapping by a syntenic approach was performed to identify, refine and explore shape QTL varying in the landrace peppers. The phenotypes associated with such QTL were further characterized to identify underlying candidate genes potentially responsible.

Chapter 4 followed up on a candidate gene identified in Chapter 3 as one of several interesting genes in a family thought to regulate fruit shape in a wide range of plant species. This final research chapter expanded the dissertation research into tomato to explore that network.

The Mexican chile pepper landraces used in chapters 2 and 3 are long-separated sympatric populations whose adaptive radiation was most likely mediated by human selection for specific use-types. As such, they are well-suited to our research goals. The long history of these use-type populations allows us to meaningfully analyze their evolution as separate, coherent populations over a longer time-window. Their sympatry presents an interesting evolutionary question which begs a human-mediated explanation. Their separate use-types supplies such a potential hypothesis: divergent selection by humans for use-type adaptations.

This research uses chile pepper and fits within the broader effort to identify the genes that explain the diversification in form of solanaceous landrace tubers and fruits under early cultivation – the harvested organs whose morphological diversity enables diverse post-harvest uses.

Origin of Pepper Morphological Diversity

For several reasons, biologists since Darwin (1868) have proposed using the diversification of domesticated plants as a model for understanding adaptive radiation. Cultivated ecosystems are simplified, intimately understood, and often thoroughly documented throughout history in written (Maggioni, Bothmer et al. 2017), oral (Brown, Clement et al. 2013), and archaeological records (Perry, Dickau et al. 2007, Perry and Flannery 2007). Also, insofar as conscious selection causes this diversification, the phenotypic diversity it results in will be human-perceptible.

Bell and chile peppers (*C. annuum*) are ideal for this type of research. The specific organs harvested for use (fruits) exhibit obvious phenotypic differences from one plant to another. Those same fruits contain the plant's germplasm. Unlike in grains or tubers, the germplasm itself is not sought for consumption. Harvesting high-quality fruit for consumption may aid, not compete with, saving high-quality germplasm for subsequent plantings.

The genus *Capsicum* diverged from the *Solanum* potato and tomato genus c. 36 million years ago (Qin, Yu et al. 2014), then spread out from Bolivia across the Andes, where it diversified into over a dozen species (McLeod, Guttman et al. 1982). Five have been domesticated: *C. baccatum, C. pubescens, C. frutescens, C. chinense,* and *C. annuum* (Perry, Dickau et al. 2007). The latter three form the "*C. annuum* complex", a subset with a relatively higher ability to intermate (Pickersgill 1988). *C. annuum* is the most important economic species in the genus, with the most cultivars worldwide (Bosland and Votava 2012).

A combined analysis of archaeological, linguistic, genetic, and ecological evidence pointed to southern Mexico as the most likely hearth of *C. annuum* domestication (Kraft, Brown et al. 2014). The date of domestication, according to archaeological evidence (Smith 1997) is at least 6,000 years ago, and linguistic analysis placed that date as far back as 6,500 years ago (Brown, Clement et al. 2013). To date, limitations in either marker (Hernández-Verdugo, Luna-Reyes et al. 2001, González-Jara, Moreno-Letelier et al. 2011) or population (Hill, Ashrafi et al. 2013, Hulse-Kemp, Ashrafi et al. 2016, Naegele, Mitchell et al. 2016) data resolution have prevented genetic studies from elucidating the processes governing early *C. annuum* diversification under domestication. Despite evidence that pepper landraces have been diversifying for thousands of years, the deep sampling and genomic analysis of long-standing landrace populations needed to dissect the processes of its earliest diversification has not yet been conducted.

Loci Controlling Fruit Shape Diversity in Pepper and Tomato

Seed dispersal in the wild (Bollen *et al.,* 2004) and crop quality under cultivation (Hori *et al.,* 2013), impose strong but distinct pressures for fruit size and shape control. The pressures imposed on cultivated populations by domestication have thus altered allele frequencies of key fruit-shape control genes, relative to wild populations. Among such genes are *SlWUSCHEL* (*SlWUS,* underlying the *LOCULE NUMBER* QTL)*, SlCLAVATA3*, *(SlCLV3,* underlying *FASCIATED*)*, SlSUN* (underlying the *SUN* QTL)*, OVATE* (underlying *OVATE*), and *OVATE FAMILY PROTEIN 20* (*OFP20,* underlying *SUPPRESSOR OF OVATE 1*), which together explain a majority of fruit shape variation in the well-studied tomato fleshy fruit model organism (Rodriguez *et al.,* 2011, Blanca *et al.,* 2015, Wu *et al.,* 2018). For this work, the genes *SlOVATE, SlOFP20,* together with interacting partner *TONNEAU1 RECRUITING MOTIF 5* (*SlTRM5*), found in a yeast 2-hybrid screen, are particularly interesting. These genes act in the developing flower buds long before they open, and explain a significant piece of tomato fruit shape diversity.

Whether the pepper species *C. annuum*'s fruit wide shape diversity (see Bosland and Votava, 2012) is controlled by the same genes as the model tomato is a key question in this work. The diverse shapes of pepper have been classified into nine categories (IPGRI, 1995), but can more quantitatively be described by attributes such as fruit shape index (FSI): the ratio of maximum curved height to width. Major QTL for both types of pepper fruit shape exist on chromosomes 2 (Hill *et al.,* 2017, Chunthawodtiporn *et al.,* 2018), 3 (Han *et al.,* 2016), and 10 (Ben Chaim *et al.,* 2001, Borovsky and Paran, 2011). Early comparisons between fruit shape genes in tomato found QTL distinct for each species, though evidence suggests pepper fruit weight orthologs for tomato weight QTL (Paran *et al.,* 2004, Chakrabarti *et al*., 2013, Paran and van der Knaap, 2007). However, recent developments in pepper genomics - including the assembly of three reference genomes by different research groups – may allow for the reevaluation of fruit shape gene homology between these species.

The first two genome assemblies were of a Mexican landrace Criollo de Morelos 334 (CM334) resistant to *Phytophthora* (Kim *et al.,* 2014), and a cultivar from China named Zunla-1 (Qin *et al.,* 2014). The third assembly was produced by linked-read sequencing, using DNA from the hybrid of CM334 crossed with a UC Davis breeding line (Hulse-Kemp *et al.,* 2018). These assemblies have already been used for mapping, with many previously known fruit shape and weight QTL being positioned on the CM334 reference genome in particular (Kim *et al.,* 2014, Hill *et al.,* 2017). Yet, discrepancies between these reference genomes in difficult-to-assemble regions still complicate the refinement of such QTL to identify causative mutations. A possible solution to this problem may be the presence of many regions of the pepper genome exhibiting synteny with the well-assembled tomato genome (Rinaldi *et al.,* 2016).

TRM-OFP Fruit Shape Regulation

Comparative genomics techniques leveraging synteny are especially promising in an important reproductive trait like fruit shape, where higher-order systems of careful regulation are likely to be conserved across species. OVATE and OVATE FAMILY PROTEINS (OFPs) participate in one such conserved system. OFPs have been found regulating the shapes of fruits and other organs across many plant taxa, such as Arabidopsis (Hackbusch *et al.,* 2005), rice (Schmitz *et al.,* 2015, Ma *et al.,* 2017, Zhao *et al.,* 2018), melon and potato (Wu *et al.,* 2018). However, OFPs were first identified for their role in controlling tomato fruit shape (Liu *et al.,* 2002). Tomato lines where *SlOVATE* has a loss-of-function premature stop mutation and *SlOFP20* has a large promoter deletion have fruits with narrow, pear-like necks at their proximal ends (Liu *et al.,* 2002, Rodriguez *et al.,* 2013, Wu *et al.,* 2018). In flower buds of these mutant plants, a shift in cell division toward more divisions in the proximal-distal direction occurs before seven days post floral meristem initiation (dpi) (Kraus 2019). This change in cell division may be due to protein-protein interactions between OFPs and several TRMs, detected both in yeast 2-hybrid systems and tobacco leaf epidermis (Wu *et al.,* 2018, Keyhaninejad *et al.,* submitted 2020). Also, CRISPR-induced *Sltrm5* null alleles partially rescue the pear-shaped mutant phenotype and cell division patterns (Wu *et al.,* 2018). TRM proteins are part of the TONNEAU1, TONNEAU1 RECRUITING MOTIF, and PROTEIN PHOSPHATASE 2A (TTP) protein complex that controls preprophase band formation and cell division plane direction in plant meristems (Azimzadeh *et al.,* 2008, Spinner *et al.,* 2013, Schaefer *et al.,* 2017). TRM mutants result in organ shape changes in many plants, from rice (Wang Y. *et al.,* 2015, Wang S. *Et al.,* 2015, Zhou *et al.,* 2015), to cucumber (Wu *et al.,* 2018) and Arabidopsis (Wang *et al.,* 2007). Several TRM proteins - all containing a conserved M8 domain - were found to interact with OFPs in tomato, and OFPs were demonstrated capable of relocalizing TRMs from cytoskeleton to cytoplasm (Wu *et al.,* 2018). As will be discussed in Chapter 3, this system also shows promise as a candidate for underlying the fruit shape QTL we found varying between two parents of our pepper landraces.

Research Approach and Objectives

The objective of this work was to explore the genetic diversity that controls fruit morphological diversity – especially fruit shape – and look for parallels between pepper and tomato. Differences in the extent of information available in the pepper and tomato systems necessitated a two-pronged approach. In pepper, the focus was on narrowing QTL down to genes controlling fruit shape, whereas in tomato we worked from known fruit shape genes, expanding into their coexpressed genes, and using structural variant calling to describe new features in coexpressed genes that vary within the tomato population.

Pepper landrace genetic diversity and fruit-shape QTL mapping

Our collection of chile pepper landraces from Oaxaca, Mexico includes several named types with distinct heritable fruit traits. We also have a comparison set of chiles from a global collection of chile pepper cultivars, with seed provided by hobbyists and seed companies. Our hypothesis is that these chile pepper landraces were largely isolated from elite chile pepper cultivar development populations and each other, allowing them to evolve separately under unique selective pressures.

To explore the genetic diversity of landrace and ancestral chile peppers, we assembled a collection that was aimed to cover diverse pepper use types as well as different domestication gradients. Several of the collected Oaxacan landraces were endemic to certain subregions. One of the major cultivated landraces is the Chile de Agua and those populations were exclusively found in the high central valleys of Oaxaca. A less cultivated but still important landrace population is Costeño which is found along the southern coastline of Oaxaca. Tusta accessions were collected from several sites along the North-South transect between the central valleys and southern coast. Taviche were collected only from two sites: San Pablo Coatlán and Ejutla de Crespo. Maax'ic and Dulce were collected only from the Yucatán sites. Some backyard-grown types were eclectic collections of peppers that spanned multiple species. Peppers called Paradito were eclectic, with accessions spanning both the Oaxacan and Yucatán collection regions and both *C. frutescens* and *C. annuum* species.

These landraces formed the basis for our exploration of genetic diversity in pepper along the lines of use-types, and also form the foundation for populations of pepper studied in the third chapter of this dissertation. The pepper fruit shape mapping study in chapter 3 uses synteny with tomato to map a pepper fruit-shape QTL in a poorly assembled segment on pepper chromosome 3 within *fs3.1* (Han *et al.*, 2016). It examines a biparental population originating from a cross between a long-fruited Chile de Arbol and a small, round Chigole. The results point toward potential similarities in the regulation of organ morphology and between these two species in the Solanaceae family.

Exploring OFP-TRM coexpression networks

The research in OFPs and TRMs as controllers of fruit shape across multiple species presents a promising opportunity for further research. The proposed mechanism of these genes also gives us direction as we seek to further illuminate the networks involved in fruit shape regulation. Furthermore, the availability of a higher-quality annotated reference genome, extensive expression data covering key developmental timepoints in floral bud development, and whole-genome sequence data from hundreds of tomato accessions reviewed above, give us a rich set of data available for further analysis.

An important question is: what other proteins aside from OFPs might also serve as interaction partners for TRMs implicated in fruit shape control? Do mutations exist in natural populations that could disrupt these networks and affect shape regulatory mechanisms? To address these questions, we explore coexpression networks for a select set of *TRM* genes implicated in fruit shape control. We further use published whole-genome sequence data to call an underexplored class of mutations - structural variants - in the genes of each TRM’s coexpression network.

CHAPTER 2

GENOME-WIDE GENOTYPING OF A NOVEL MEXICAN CHILE PEPPER COLLECTION ILLUMINATES THE HISTORY OF LANDRACE DIFFERENTIATION AFTER *CAPSICUM ANNUUM* L. DOMESTICATION[1]

---

[1] Taitano, N., Bernau, V., Jardon Barbolla, L., Leckie, B., Mazourek, M., Mercer, K.,, McHale, L., Michel, A., Baumler, D., Kanter, M., van der Knaap, E. 2019. *Evolutionary Applications*, *12*(1), pp.78-92. Reprinted here with permission of the publisher.

Abstract

Studies of genetic diversity among phenotypically distinct crop landraces improve our understanding of fruit evolution and genome structure under domestication. Chile peppers (*Capsicum* spp. L.) are economically valuable and culturally important species, and extensive phenotypic variation among landraces exists in southern Mexico, a center of *C. annuum* diversity. We collected 103 chile pepper seed accessions from 22 named landraces across 27 locations in southern Mexico. We genotyped these accessions with genotyping-by-sequencing (GBS), yielding 32,623 filtered single-nucleotide polymorphisms. Subsequently, we genotyped 32 additional *C. annuum* accessions from a global collection for comparison to the Mexican collection. Within the Mexican collection, genetic assignment analyses showed clear genetic differentiation between landraces and clarified the unique nature of the Tusta landrace. Further clustering analyses indicated that the largest fresh-use Chile de Agua, and dry-use Costeño landraces were part of separate clades, indicating that these two landraces likely represent distinct populations. The global accessions showed considerable admixture and limited clustering, which may be due to the collapse of use-type divisions outside of Central America. The separation of the Mexican landraces in part by fruit morphology related to use, highlights the relevance of this use-type morphological diversity for plant breeders and the utility of fruit development variation for evolutionary biologists.

Introduction

Evolutionary biologists have been interested in domesticated plants to study natural selection for more than a century (Darwin, 1868). Just like any other plant system, crop populations can be subject to the structure-reducing effects of gene flow and the structure-enhancing effects of genetic drift, selection, or assortative mating (Loveless & Hamrick, 1984). However, cultivated populations have unique characteristics as breeding may accelerate local adaptation. In this case, human management can create reproductive isolation, by the removal of phenotypically distinct individuals (rogueing) from homogenous cultivated plant stands, or by purposefully isolating distinct types into separate stands to prevent outcrossing. Landraces grown near a crop center of origin present ideal populations in which to study crop diversification and genetic structure, as well as the presence of long-standing populations derived soon after domestication (Zeven, 1998).

Levels of genetic structure in domesticated populations are largely determined by the diversifying effect of population isolation (e.g. of specialized landraces), balanced against the homogenizing effect of gene flow and the planting of homogenous elite cultivars. Smallholder farmers may also strive to retain traditional varieties on small plots in their farms while participating in germplasm exchange that includes productive modern cultivars (e.g. in the Andes; Brush, 1992). As a result, landraces (e.g. Andean potatoes, *Solanum spp.*) near their center of origin can retain a high level of diversity on individual farms, even as regional diversity diminishes (Zimmerer & Douches, 1991).

Another factor that affects the level of genetic diversity found in a crop population is its level of domestication. The degree of domestication is difficult to measure precisely. For the purposes of this study, which examines chile peppers (*Capsicum spp.* L.) we have grouped the domestication level of different seed accessions into four categories of cultivation. From least domesticated to most domesticated, these are accessions produced in the forest, backyard, *milpa*, and plantation environments. Forest-grown populations may be collected by the community, but they are rarely intentionally planted and exist outside of an intentionally cultivated environment; thus they likely represent the least domesticated types. Backyard populations, which cover let-stand populations and those casually cultivated in backyard gardens, represent a level of human cultivation and domestication that is relatively unrestrictive. Even though there may be a moderate level of selection, these populations are likely not subjected to the rigorous scrutiny imposed on peppers destined for market. In Mesoamerica, a *milpa* is a cultivated maize field that often incorporates intercropping of other species, such as beans, squash, and chile peppers. Some *milpas* include agroforestry components and mirror natural, post-disturbance, forest succession (Nigh & Diemont, 2013). Though actual *milpa* practices may vary among farmers, the higher species diversity and forest proximity inherent to *milpa* environments attracts generalist pollinators that mediate pepper pollination, outcrossing, and fruit set (Raw, 2000; Landaverde-González et al., 2017). We define plantations as agricultural systems where a single crop is planted in rows, usually of a single variety that is saved year to year by the farmer. This is the most restrictive domestication environment.

The chile pepper, especially *Capsicum annuum,* is a widely cultivated species with many phenotypically diverse landraces (Bosland & Votava, 2012), and a well-suited study system for exploring the genetic structure of landraces during their diversification. To date, limitations in data resolution (of markers or populations) have prevented studies from elucidating the genetics of diversification in *C. annuum*. Early studies were limited by genomic resolution since only a dozen or fewer markers were employed for the analyses (Hernández-Verdugo et al., 2001; González-Jara et al., 2011), as it was difficult to generate markers for the large pepper genome (3.48 Gb; Qin et al., 2014). More recent studies sampled many, mainly elite populations with few representatives of each population (Naegele et al., 2016; Hill et al., 2013; Hulse-Kemp et al., 2016). While the latter studies demonstrate the genetic diversity that is relevant to elite germplasm, the degree of genetic diversification among long-standing *C. annuum* landraces is unknown. Thus, despite evidence that pepper has been cultivated for thousands of years (Perry & Flannery, 2007), more comprehensive sampling and genotyping, with improved genomic coverage, may better elucidate the processes of diversification under domestication.

The present study sampled chile pepper populations from the Southern Mexican states of Oaxaca and Yucatan. *C. annuum* is of cultural importance in Mexico, especially in Oaxaca, where it exhibits dramatic genetic and phenotypic diversity. Early chile pepper depictions show that cultivated *C. annuum* fruits were much larger than their wild counterparts and had an array of uses spanning many hundreds of years (Codex Mendoza, 1542). This diversity of uses capitalizes on a diversity of chile pepper “use-types”, *i.e.*: assemblages of potentially related plants bearing fruits with distinct morphological characteristics well-suited for their particular use, which are also found in archaeological remains (Perry & Flannery, 2007). Oaxaca spans a range of climates, owing to its sharp elevation gain inland from the coast, as well as precipitation differences along the coastline (Fick & Hijmans, 2017). Moreover, Oaxaca has been called the most ethnically diverse state in Mexico, home to more than 16 languages (Romero 2000). Thus, high climate and cultural diversity makes Southern Mexico a center of great diversity for chile peppers and a likely center of domestication for chile peppers (Kraft *et al.* 2014; Aguilar-Meléndez *et al.* 2009). In summary, chile peppers from Southern Mexico are an ideal target to analyze long-standing landrace populations for patterns of genetic diversity created since domestication. In order to extend our study of genetic diversity to peppers grown outside its center of diversity in southern Mexico, we included chile peppers collected from around the world (Kantar et al., 2016).

Our objectives were to characterize the genetic diversity among the Oaxacan landraces, and compare them to accessions from around the world. Understanding of the genetic diversity in these chile peppers can lead to populations that may contain potentially useful alleles that were missed when selecting germplasm to include in modern cultivars.

Material and Methods

Study system

The genus *Capsicum* is a member of the agriculturally important Solanaceae family, which also includes potato, tomato, eggplant, and tomatillo. After branching off from the tomato and potato lineage *c.* 36 million years ago (Qin et al., 2014), the *Capsicum* lineage itself diverged into over a dozen species (McLeod et al., 1982). Of these, five species were domesticated in Central and South America: *C. baccatum* L., *C. pubescens* Ruiz & Pav*.*, *C. frutescens* L*.*, *C. chinense* Jacq*.*, and *C. annuum.* The latter three are relatively interfertile with each other and form the “*Capsicum annuum* complex” (Pickersgill, 1988). *C. annuum* makes up the majority of varieties now cultivated worldwide (Bosland & Votava, 2012). All these varieties are descended from *C. annuum* originally domesticated in present-day Mexico (Kraft et al., 2014), with remains in the Tehuacán valley dated to *c.* 6,000 years ago, about 1,000 years after general crop cultivation began in this area (Smith, 1997; Brown et al., 2013).

Plant materials

Pepper accessions were collected in 2013 from two overlapping transects in Oaxaca. These collection sites allowed us to sample the major sources of variation among landraces that are present in the region (Figure 2.1). The first transect encompassed 13 sampling locations and ran north-south along an elevation and temperature gradient, from the central valleys near Oaxaca City, *c.* 1,500 meters above sea level (masl) to the southern tip of the Pacific coastline in Pochutla (< 600 masl). There, it borders the coast and the second transect. The second transect ran east-west along the Pacific coast, which spanned a precipitation gradient and included twelve sites. Both transects spanned ethnic and language groups. The peppers collected from three sites in the Yucatán were from the villages Maní, Acanceh, and Cansahcab. In total, we collected seed from 27 different locations in Mexico. Together, these peppers will be referred to throughout this paper as the "Mexican collection" (Table 2.S1). From the Mexican collection, 103 accessions produced viable seed from which two seedlings were grown where possible, ultimately yielding 190 plants which were genotyped for this study. Those plants were grown in a Columbus, OH greenhouse in 2014 in a completely randomized design.

The pepper collection from around the world was obtained from heirloom seed producers in North America (see Kantar et al., 2016). They originated from multiple geographies and contained *C. annuum* landraces and cultivars. This collection of chile peppers from around the world is henceforth referred to as the "global collection" and only accessions with sufficient read depth were used for this study. After germination and growth indoors for 9-14 weeks, plants from this collection were transplanted into five-gallon containers and grown outdoors in a completely random design with two replications in Madison, WI during the summer of 2013. Young leaves from each plant were harvested and frozen at -20°C for subsequent DNA extraction. Images of fruits were collected for each of the accessions and the major named types.

DNA Extraction

For the global accessions, gDNA from each was extracted by grinding 100 mg of frozen leaf tissue using dry ice, stainless steel beads, and a tissue homogenizer (Bullet Blender), then isolating DNA using the Omega Biotec E.Z.N.A. ® Plant DNA Kit. Extracted DNA was stored at -20ºC and sent to BGI Americas for GBS library construction and sequencing.

For the Mexican collection, DNA was collected from young leaves of adult chile pepper plants. Approximately 50 mg leaf sample from each plant was collected in deep-well tubes in two 96-well plates. The samples were stored on ice until collection was complete, then flash-frozen in liquid nitrogen and lyophilized prior to DNA extraction. After lyophilizing, the samples were ground to a fine powder by adding metal beads and mechanically shaking in a Geno/Grinder 2000® (SPEX, Metuchen, NJ, USA). DNA extraction was performed using QIAGEN’s DNEasy 96 Plant Kit® (Valencia, CA, USA), following the manufacturer’s recommendation. DNAs were eluted into 100 µl TE pH 8.

Genotyping-by-Sequencing library construction

Genotyping-by-sequencing (GBS) libraries were created following the established method (Elshire et al., 2011). To briefly summarize, genomic DNA was digested with the ApeKI methylation-sensitive 5 base-pair (bp) recognition site restriction enzyme. The resulting fragments were ligated to Illumina sequencing adapters, and to adapters with sequence “barcodes” unique to each individual sample, enabling the recovery of source plant identity for each sequenced DNA fragment after multiplexing. For the Mexican collection, GBS libraries were constructed for each genotype and 48 libraries were pooled, size selected to an average size of 350 bp in length and sequenced. Two pools were sequenced on the NextSeq platform from which we obtained 725 million 75 bp single-end sequence reads and an average per-individual coverage of 3.59X. Two additional pools were sequenced on the HiSeq 2500 platform for a total of 397 million 100 bp single-end reads, and an average per-individual coverage of 1.96X. This resulted in an average of 344 million reads obtained per lane (2X expected coverage per reduced genome), with an average PHRED-scaled quality score of 35 for the “Mexican collection”. On the other hand, the “global accessions” sequenced on the Illumina Hiseq 4000 at a 100x2 bp paired-end read length yielded 19.5 million reads total for those individuals retained after quality control, for an average per-individual coverage of 0.219X.

SNP calling

After a quality-control step with FastQC (Andrews et al., 2010) and removal of poor quality reads, the TASSEL GBS Pipeline 5.2.3 (Glaubitz et al., 2014) was used to call single-nucleotide polymorphisms (SNPs) from Illumina sequence data. The *C. annuum* cv. CM334 reference genome was used for read alignment with Bowtie2 (Langmead & Salzberg, 2012), a minor allele count of three reads per minor SNP allele was required to call a SNP (Appendix 2.S1).

This SNP-calling process returned a genotype table, which was then filtered prior to analysis using VCFTools (Danecek et al., 2011) to a list of biallelic SNPs that were excluded on the basis of the proportion of missing data, minor allele frequency, and the proportion of heterozygosity at each locus. Filtering thresholds for these metrics were set at ≤10%, ≥5%, and ≤10%, respectively for the Mexican collection, and ≤20%, ≥1%, and ≤5% for the combined Mexican and global collection dataset. Thresholds were determined by plotting the metric for each SNP on the y-axis, with the SNPs ordered by their value for that metric along the x-axis, and visually identifying an inflection point in the resulting curve, which signified a sudden divergence in value for that metric from the baseline SNPs. Furthermore, to obtain markers for analyzing the combined Mexican and global dataset, SNPs were selected with close to equal coverage between the two datasets. This was done by first removing individuals from the global collection with unusually high (>97%) missing data over all unfiltered SNPs, then selecting those SNPs which had >80% coverage of individuals in the global collection before the final filtering step using the thresholds given above.

A separate SNP filtration step was performed from raw SNPs using the same process to assess the four major landrace subpopulations Tusta, Taviche, Costeño and Chile de Agua. Also, the SNP filtration process was performed separately with and without the incorporation of the global collection, in order to have a set of high-coverage SNPs for higher-resolution genomic analyses of a subset of the accessions.

GBS alignments to the pepper reference genomes

To compare the GBS information with the three referenced genomes, alignments of 150-bp sequences around SNPs to the Zunla and Chiltepin reference genomes were performed. First, we extracted a 150 bp sequence around each SNP in the CM334 reference genome. These sequences were combined in a FASTA file, which was then aligned to the Zunla and Chiltepin reference genomes using Bowtie2 (Langmead & Salzberg, 2012). The SNP genotype was taken from the base call at the Zunla or Chiltepin position aligning to the SNP position in the 150-bp CM334 sequence.

Accession quality control

Three Mexican accessions were removed from the dataset by applying the following criteria. First, individuals could have no more than 30% missing data across all filtered SNPs (returned by VCFTools). Second, accessions could not cluster away from all other plants of the same named type in the initial clustering analysis (below) and be identified as distinct based on fruit phenotype. Individuals #167-1 and #218-1 were removed via the first criterion, and both plants grown from one accession (#122-1 and #122-2) were removed via the second criterion. As #167-1 was the only representative which germinated from accession 167, this quality control left 101 accessions remaining of the original 103. Forest or backyard-grown accessions such as the one guajillo (Table 2.S1) with ambiguous species characteristics that were recorded as *C. annuum* in the field, but grouped together with *C. frutescens*, were reassigned as *C. frutescens* and as such excluded from the in-depth subpopulation structure analysis.

The two-step SNP-filtering (described in SNP-calling above) for the combined Mexican and global collection dataset made a two-step filtration of individuals necessary, to avoid biasing SNPs toward those covering individuals that would be later removed. Thus, prior to the first step of SNP filtering, individuals with unusually high (>97%) missing data among raw SNPs were removed from the global collection. Otherwise, filtration of individuals occurred as described above for the combined dataset.

Population structure

The GBS data from the Mexican collection was used to obtain a population tree. The initial tree was selected using the maximum parsimony method, followed by maximum likelihood optimization based on the general time-reversible model, with 1000 bootstraps as implemented in RAxML (Stamatakis, 2014;

Tavaré, 1986). The *C. chinense* accession 155-1 was used as a midpoint to root the tree. This tree was used to filter individuals and assign individuals to preliminary groups based on their named types, locations, and genetic relatedness as revealed by clustering analysis. We also performed this analysis using less computationally-intensive parameters: creating a neighbor-joining initial tree, followed by maximum-likelihood optimization with the Tamura-Nei mutation model and 100 bootstrap replicates (Figure 2.S1). Finding no substantial differences between the clusters in each analysis, we used the less computationally intensive analysis to explore additional subsets of the data. Population trees were also created separately for each of the four main cultivated Oaxacan subpopulations: Tusta, Taviche, Costeño and Chile de Agua. Additionally, genetic assignment analysis was conducted using the program fastSTRUCTURE (Raj et al., 2014), first, with only the Mexican accessions, then with the combined Mexican and global collection dataset, including the available reference genomes. In both cases, the number of genetic clusters (*K*) was allowed to vary from 2-10. Accessions were assigned to the groups corresponding with their locally-known types except where both clustering analysis and genetic assignment analysis assigned an accession to a group other than the named type, with a threshold of 70% identity in the latter analysis. For the cultivated *C. annuum*, five of the 80 analyzed accessions were reassigned in this way, all of which were locally known as Tusta or Taviche. Mean imputation followed by principal components analysis were performed using the package SNPRelate (Zheng et al., 2012) on the complete population, including global collection and reference genomes.

Genome scans for population origin, selection sweeps, and diversity were, respectively, performed using corrected Wright's F_{ST} (Weir & Cockerham, 1984), Tajima's D (calculated over segregating sites) and the pairwise nucleotide diversity measure π (measured on a per-nucleotide basis calculated using the proportion of the genome included by the GBS reduced-genome methodology), as implemented in VCFTools and performed only on elite Mexican landraces. Pairwise permutation tests were performed by shuffling individuals among population pairs, while keeping population sizes constant in R. The resulting permutated populations were saved into population files for use in VCFTools.

As an internal control and to explore diversity within each accession, two plants were grown for each accession. In all accessions for which genotypes could be recovered from both plants, both plants exhibited the same group membership pattern. Close relatedness was also demonstrated between individuals of the same seedlot, named type, and species (Figure 2.S2).

Results

Distribution and morphology of pepper types

To explore the genetic diversity of landrace and ancestral chile peppers primarily in Oaxaca, Mexico, we assembled a collection to cover diverse pepper use-types, as well as different degrees of domestication. Several of the collected Oaxacan landraces were endemic to specific subregions (Figure 2.1). We collected populations of Chile de Agua (Table 2.S1), a major cultivated landrace, only from the high central valleys of Oaxaca. A less cultivated, but still important, landrace population is Costeño, which we collected along the southern coastline of Oaxaca. Tusta accessions were collected from several sites along the North-South transect between the central valleys and southern coast. Taviche were collected only from San Pablo Coatlán, in the central valleys of Oaxaca. A few accessions were collected from the Yucatán and they were Maax'ik and Dulce. Some backyard-grown named types were eclectic collections of peppers that spanned multiple species. Peppers called Paradito were diverse, with accessions spanning both the Oaxacan and Yucatán collection regions and both *C. frutescens* and *C. annuum* species (Table 2.S1).

The Mexican landrace populations presented a diverse set of fruit phenotypes, ranging from small and round to very narrow, and from elongated to blocky (Figure 2.2). The most intensely cultivated accessions (grown exclusively in *milpas* or on plantations) tended to have larger fruits (Table 2.S1). These included the Chile de Agua (Figure 2.2A), a fresh use-type that was grown most commonly in plantations; Costeño (Figure 2.2B), a dry use-type that was grown in plantations and *milpas*; Taviche (Figure 2.2D), a dry use-type collected only from a *milpa*; Guiña Dahni (Figure 2.2G), a dry use-type collected from a coastal plantation in Oaxaca; and Dulce (Figure 2.2F), a fresh use-type which was grown in several Yucatán *milpas.* The Mexican collection also included chile peppers more commonly grown in backyard or “let-stand” environments, such as the De Arbol, Tusta, Mirasol, Solterito, Mareño, Piquin, Paradito, Chigole, Bolita (Figures 2E, 2C, and 2H-N, respectively). Finally, the Mexican collection included a number of populations growing in forests and uncultivated environments, including some Chigole peppers and those colloquially known as Chile de Monte (any uncultivated peppers growing in mountains or forests). These forest-grown chile peppers were much smaller than the cultivated types, but had seeds which were still similar in size to the cultivated types. Thus, the pericarp around the forest-grown pepper types was little more than a thin coating around the seeds, in contrast to many of the thicker-fleshed cultivated types. In total, the Mexican collection included 19 named types of chile peppers.

Fruit morphologies varied between different named types. All Chile de Agua plants bore fruits with a similar triangular shape with large truncate shoulders tapering into a distal point that was blunt to slightly sunken (IPGRI, 1995). All Costeño plants bore smaller, more elongate, often curved fruits with a pointed distal end (IPGRI, 1995). Accessions bearing the Tusta label (including two from San Pablo Coatlán) were heart-shaped fruits, tending to have high shoulders above a noticeable proximal indentation, while Taviche fruits (including two that fell genetically within the Tusta subpopulation) were more similar to Costeño in size, though they tended toward wider shoulders, making them more triangular than horn-shaped (UPOV, 2006). Both fruit phenotypes and named types from Yucatán accessions were varied. Semi-wild peppers (a pepper where it is unclear if it is a truly wild or a feral) in both *C. annuum* and all *C. frutescens* were much smaller and tended to be rounder than the four main types: Tusta, Taviche, Costeño, and Chile de Agua, as well as the named types with only 1-2 accessions in our Mexican collection: Dulce, Guiña Dahni, and De Arbol (Table 2.S1).

Genetic structure of Mexican chile pepper population

To describe the genetic structure of the Mexican population, we generated a GBS SNP dataset. After filtering, 32,623 SNPs were called among the Mexican accessions, and 3,570 had sufficient coverage for comparisons to the accessions from outside Central America (Figure 2.S3). SNPs called by GBS were distributed mostly in the euchromatic regions, with relatively few being found in the pericentromeric regions as defined by the reference genome (Qin et al., 2014).

Using FastSTRUCTURE (Raj et al., 2014) to assess integrity and admixture in named Mexican landraces, we examined the assignment pattern with the number of sub-populations (*K*) from *K*=3-9 (Figure 2.3). The optimal Δ*K* (Evanno *et al.* 2005) value was predicted to be 7 (Figure 2.3; Figure 2.S4). At *K*=3, there was clear differentiation between *Capsicum frutescens* and Tusta accessions, and the remaining *C. annuum* accessions. At *K*=4, Chile de Agua accessions were a distinct cluster. At *K*=5 the Forest accessions were a distinct cluster. At *K*=5, the single *C. chinense* accession, a Maax'ik accession from the Yucatán, demonstrated admixture between the *C. frutescens* and *C. annuum* accessions. *C. annuum* accessions within the same named type exhibited similarity in the genetic assignment analysis. As demonstrated by the structure plots for *K*=6 through *K*=9, higher levels of *K* created superfluous groups explaining very little variation (Figure 2.3; Table 2.S2) In summary, the analysis identified three main subpopulations among the cultivated accessions, Tusta, Costeno, and Chile de Agua. Taviche accessions did not represent a separate group in the genetic assignment analysis, but instead shared a pattern of admixture between Chile de Agua (~25%) and Costeno (~75%). Based on this, their local names, and their significant bootstrap value in the population tree analysis we show below (Figure 2.4), we found it useful to analyze Taviche separately.

Both plants from Tusta-type accessions (#179 and #185, Table 2.S1) grown in San Pablo Coatlán alongside a Taviche population (indicated by stars Figure 2.3) showed a genetic subpopulation identity pattern that was indistinguishable from peppers in the Taviche subpopulation, and showed no membership in the Tusta subpopulation. All Tusta and Taviche types grown from seed in our greenhouses were phenotypically distinct, each bearing fruits characteristic of their respective parental named types (Figure 2.2). Therefore, despite being morphologically similar to Tusta, these San Pablo Coatlán "Tusta" were genetically more closely related to the Taviche than to Tusta taken from Santa Lucia Miahuatlan (#188), San Baltazar Loxicha (#187), Los Reyes (#105-#108), or Juan Diegal (#93). These accessions were henceforth considered part of the Taviche subpopulation rather than the Tusta subpopulation.

Similarly, two Taviche-named accessions (#181 and #183; Table 2.S1) were closely related to the Tusta (indicated by stars above the corresponding bars in the Tusta subpopulation in Figure 2.3). In all four cases, the two plants derived from each accession (#179, #185, #181, and #183) were paired in clustering (Figure 2.S2), indicating that a DNA mix-up was unlikely and would have required mistakes to have occurred independently in the handling of both plants from each accession. Confirming that the fruits collected from each plant matched the parental type of the accession from which it was grown similarly excluded a seed mix-up (Table 2.S1).

The single de Arból accession appeared genetically in between Costeño and Taviche. Membership in the remaining subpopulations were distributed among two types of accessions: 1) those that were considered "semi-wild" accessions of *C. annuum* collected from the forest understory or backyards known as Chigole or Chile de Monte, and 2) the less restrictively cultivated accessions from Yucatán (belonging to Maax'ik and Xaat'ik) sharing genetic diversity with the semi-wild and Costeño.

To relate the accessions to one another, we reconstructed a population tree of the Mexican chile pepper collection rooted at the midpoint by the single *C. chinense* accession (Figure 2.4). We detected three main groups with high (>95%) bootstrap support that agreed with our previous population assignments from structure. Except for the *C. frutescens*-like Paradito population and the Dulce accession, the Yucatán accessions clustered into basal clades that were sister to the domesticated *C. annuum* (Figure 2.3, 4). Small-fruited accessions grown in the Yucatán such as the Maax'ik and Paradito were closely related to the *C. annuum* accessions grown in backyards in the southern tip of the Oaxacan coast. The accessions within the main fresh and dry use-types in this study, Chile de Agua and Costeño, respectively, formed separate clusters (Figure 2.4).

We further characterized the spatial distribution of genetic diversity within the four main Mexican *C. annuum* landraces (Tusta, Taviche, Chile de Agua, and Costeño). Chile de Agua was only collected in the central valley of Oaxaca. Yet, the various Chile de Agua populations appeared to have retained interpopulation spatial differentiation. Proceeding clockwise from the top of the population tree for Chile de Agua, the most highly domesticated landrace we collected (Figure 2.S5A), the first clade (labeled "i") contained six accessions from two sites on the eastern side of the high central valleys: La Labor and Paraje Coatequillas. These sites were also connected by Federal Highway 175 (Figure 2.S6). The next two clades (ii-iii) were both composed of individuals from southeastern Paraje Coatequillas. Continuing clockwise, the next clade (iv) was composed of two individuals from a northwestern collection site in La Lobera (ID #140, #141, Table 2.S1). Four more accessions from a southern site – Coatecas Altas, near Paraje Coatequillas – formed a fifth clade (v). Accessions from two western sites – southwestern Santa Cruz Nexila (#1, #169, #170, #173, #174) and northwestern La Lobera (#142) – formed a sixth clade (vi), with somewhat weaker bootstrap support and the five Santa Cruz Nexila accessions showing little genetic diversity between them. A seventh clade (vii) was composed of four accessions from the eastern sites La Labor and Coatecas Altas. An eighth clade (viii) was comprised of accessions from western sites Santa Cruz Nexila and La Lobera. Thus, clades vi and viii were both distributed among collection sites in the eastern side of the Oaxacan Central Valleys, connected by highway 131 (Figure 2.S6), whereas clades i and vii were distributed among western collection sites.

Costeño included more backyard accessions than the mostly plantation-grown Chile de Agua, and showed less evidence of subclades (Figure 2.S5B). The exception to this is the first four accessions (ix), which all were taken from the village of Rosedalito near the southern tip of the Oaxacan coast and did constitute a clade. Beyond that, however, there was evidence for admixture, with genetic diversity apportioned more strongly between individual accessions within the same site, and only weak evidence (≤70% bootstrap) of subclades within the Costeño.

Genetic diversity was lowest within Costeño and Chile de Agua landraces compared to the other *C. annuum* landraces (Table 2.1). As defined after our genetic structure analysis, the Tusta and Taviche groups each formed a monophyletic clade (Figure 2.4). One clade comprised Taviche accessions from San Pablo Coatlán and Ejutla de Crespo (Figure 2.S5C), while the other included a mixture of Tusta accessions from various sites (Figure 2.S5D).

Allele frequency differentiation and genetic diversity

The average nucleotide diversity (π) within each use-type group ranged from 0.031% to 0.01% and appeared to decrease with intensity of cultivation (high cultivation for Chile de Agua and low cultivation for Tusta) (Table 2.1). Despite containing the fewest accessions, the Tusta group had the most segregating SNP variation and the highest π. Chile de Agua contained the second-highest number of segregating SNPs. Despite that, Chile de Agua showed the lowest π of all four main groups, indicating a high degree of homogeneity within the accessions. Homogeneity was also evaluated as the percent identity-by-state (% IBS) between same-accession pairs (two plants per seedlot), calculated over all non-missing loci for each accession and averaged over all accessions within each group. The within-accession average IBS percentage was high in each of the four groups, relative to the overall mean IBS of 72% (SD = 22%). Percent IBS differed significantly between groups ($F = 57.56_{3,7}$, $p < 0.001$). Post hoc comparisons using a Bonferroni-corrected LSD test indicated that the Chile de Agua accessions were significantly more homogenous, and the Taviche and Costeño were significantly less homogenous, than Tusta as measured by within-accession average percent IBS (Table 2.1).

Since admixture was apparent between certain subpopulations in our collection, we used F_{ST} to quantify the genetic distance between the admixed populations. Mean F_{ST} averaged over all cultivated landrace populations was 0.821 after correcting for population size and substructure. The highest pairwise F_{ST} was between the Chile de Agua and the Tusta landrace (Table 2.2). Among the *C. annuum* accessions we studied, the allele frequencies of the Tusta landrace were most distinct from those of the other cultivated *C. annuum* landraces (Table 2.2). Pairwise comparisons involving the *C. frutescens* clade, which we defined as monophyletic spanning all individuals identified as *C. frutescens* based on morphological observations, and the population of individuals unambiguously belonging to *C. annuum,* gave consistently high F_{ST} values (0.667 – 0.892), as did comparisons between Tusta and any other subpopulation (Table 2.2). Pairwise F_{ST} demonstrated that all named types were significantly distinct from each other in terms of allele frequencies ($P < 0.001$, permutation test). Pairwise F_{ST} analysis also revealed that Taviche, Costeño, and Chile de Agua were more closely related to each other than to Tusta or to the *C. frutescens* subpopulation (Table 2.2), recapitulating the pattern previously revealed (Figure 2.S2).

Genome scans for selective sweeps

Identifying the Chile de Agua and Costeño as clearly distinct populations allowed us to analyze each of these populations for genomic statistics of diversity. Therefore, we assessed which genomic regions might diverge from neutral evolution for further study of adaptations that are specific to those landraces. We determined that a bin size of 500 kb was reasonable by linkage disequilibrium (LD) decay analysis (Figure 2.S7), and focused on the SNP-rich euchromatic regions. Several regions in the Chile de Agua genome appeared to exhibit clusters of extreme Tajima's D values for nearby bins (Figure 2.S8), including an extreme low cluster at about 225 Mb on the CM334 reference genome chromosome 6 assembly (Kim et al., 2014) (Figure 2.S8). Low Tajima's D values would indicate that minor alleles in a genomic region were rarer than would be expected in a neutrally-evolving population, possibly suggesting the presence of a gene or genes under strong purifying selection at the bottom of chromosome 6. The lack of a corresponding low-diversity region in Costeño suggests this as a potential candidate region for genes that control either Chile de Agua's fresh use-type phenotype or local adaptation.

Genetic comparisons between the Mexican and Global Chile pepper populations

To compare the Mexican chile pepper accessions with a subset of globally grown accessions, we selected 3,570 SNPs with sufficient coverage in both datasets to make comparisons. A principal component analysis (PCA) showed that the Tusta population maintained its distinctive separate clustering pattern and represented a unique portion of the genetic diversity in *C. annuum* (Figure 2.5A). Looking closer at the main *C. annuum* group (excluding Tusta), the relatively greater spread of the 32 global *C. annuum* accessions indicated a greater degree of genetic diversity than amongst our main group of Mexican *C. annuum* accessions (Figure 2.5B). Focusing in on the main group of Mexican accessions and those global accessions clustering closest with them, we observed that the two published reference genomes from accessions Zunla (Qin et al., 2014) and CM334 (Kim et al., 2014) tended to group nearer to the Mexican accessions than those of the global collection (Figure 2.5C). Oaxacan landraces grouped together and encompassed the CM334 reference genome. Some overlap of Taviche with Chile de Agua and Costeño was also observed in the PCA plot (Figure 2.5D), the latter being consistent with the admixture shown in genetic assignment analysis.

Genetic assignment analysis combining both the Mexican and global datasets using fastSTRUCTURE recapitulated the major Mexican landrace identities at an optimum *K* of 7 (Figure 2.6; Figure 2.S9; Table 2.S3). Additionally, it revealed high levels of shared identity between the Costeño, Chile de Agua, and most of the accessions throughout the global collection. A *C. chinense* group was resolved by the addition of the global collection, which included inadvertently some *C. chinense* accessions. Also, partial membership to the same group as the single Mexican *C. chinense* accession was scattered throughout even the *C. annuum* global accessions (Figure 2.6; Table 2.S3), also demonstrating the high level of genetic diversity among the global collection shown in the PCA (Figure 2.5).

Discussion

Genetic structure of the pepper collection population

This study validated the named chile pepper types in Oaxaca as genetically separate landraces that are distinct from each other. We also found strong evidence of separation among the main landraces grown under intensive cultivation for market. Previous studies on genetic structure in chile peppers either focused on few markers in relatively densely-sampled *C. annuum var glabriusculum* populations (Hernández-Verdugo et al., 2001; González-Jara et al., 2011) or many markers but few representatives of many widely-dispersed cultivar populations (Naegele et al., 2016; Hill et al., 2013; Hulse-Kemp et al., 2016). Such studies have found that humans are largely responsible for driving gene flow and therefore population structure in chile pepper populations (González-Jara et al., 2011), and have found evidence that pepper cultivars grown today are descended from a few initial populations (Hulse-Kemp et al., 2016).

By sampling multiple individuals from each of several Mexican landraces, we were able to delve into these genetic patterns differentiating the landraces. In doing so, we determined which landraces are candidates for being major contributors to many cultivars now grown around the world. We found a higher degree of genetic structure in our Mexican collection of chile peppers than had previously been reported for maize landraces grown in the same area (Pressoir & Berthaud, 2004b). This could be due to the relatively limited pollen dispersal in chile peppers (Raw, 2000) leading to greater inbreeding. Genetic assignment patterns were partitioned as expected, grouping together plants from the same species, major landrace types, and our internal biological replicate controls. C. *frutescens* and *C. chinense* were distant from each other and from the main *C. annuum* clade. Both the forest-grown and Yucatán *C. annuum* accessions were located basal to the Mexican landraces in our population tree analysis, supporting the hypothesis that these landraces were derived in common from a broader population including both the forest and Yucatán accessions (Figure 2.4).

Comparing the Mexican collection to the global collection, we found evidence of admixture among the different landraces as components of the highly genetically diverse global collection, consistent with patterns of global admixture from a few initial populations, found previously (Naegele et al., 2016; Hulse-Kemp et al. 2016). Comparing these two datasets to the two independently published reference genomes CM334 and Zunla, we found that the CM334 assembly, created from the Mexican landrace with polygenic resistance to the oomycete pathogen *Phythophthora capsici* (Ogundiwin et al., 2005), clustered more closely with the Mexican collection in principal components analysis (Figure 2.5). Further research, exploring whether some resistance QTLs are shared between CM334 and these or other Mexican landrace, seems promising.

Levels of genetic diversity differed among landraces at various scales

Our study emphasized the four named types of *C. annuum* grown most commonly in Oaxaca: Chile de Agua, Costeño, Tusta, and Taviche. The first two populations were grown almost exclusively in the most restrictively managed plantation environments, whereas the latter two were found in the less-restrictive *milpa* polycultures or in backyards. None of the four were found growing feral in forests. Genetic assignment and clustering analyses revealed that these named types did indeed comprise four major and separate genetic subpopulations of *C. annuum*, and could be considered separate landraces. The major fresh and dry use-types: Chile de Agua and Costeño respectively, were validated as independent, though closely related populations. These two landrace groups show the strictest spatial distribution: no

Chile de Agua in our study was collected from the low coast, and no Costeño was collected from high elevation. Separate Tusta and Taviche populations were also identified, with the exception of four accessions phenotypically resembling one type and genetically resembling the other. As might be expected from a self-pollinating species (Loveless & Hamrick, 1984), genetic diversity was partitioned primarily among, rather than within, these four landrace types.

The Costeño and Chile de Agua landraces had lower overall diversity (π), as would be expected for a restrictively-managed plantation cultivation system, in which off-types would be removed from the population. Within single-accession seedlots, however, we found significantly higher diversity in the Costeño than the Chile de Agua. This result is expected based on the higher degree of structure in the Chile de Agua, relative to the Costeño. Stronger genetic structure, coupled with inbreeding, is expected to depress effective population size and heterozygosity at both the individual plant and population levels (Loveless and Hamrick, 1984). Despite both landraces being found in plantations, and overall genetic diversity being similar for both, the genetic data indicate a larger degree of outcrossing among the Costeño, and more isolation between Chile de Agua subpopulations. A possible explanation for this difference in within-accession diversity lies in the different geographical factors of each landrace's growing region. While the coastline provides a natural trade route for seed sharing among coastal villages, the mountains scattered among the central valleys where Chile de Agua are grown could have historically served as an impediment to trade among mountain villages, and to gene flow among their crop populations.

Though our population was highly structured, we did observe some admixture that can be explained by the large amount of farmers' seed sharing followed by crossing (Gonzalez-Jara et al., 2011). The pattern shown in Chile de Agua reflects a combination of these two forces: the overall landrace population is structured into different subpopulations genetically, but several of those subpopulations are spread among multiple sites rather than being endemic to one location. Not all Chile de Agua subpopulations are present at all Central Valley sites. No subpopulation supported by a bootstrap value of greater than 90% contained two accessions from opposite (e.g. northeast and southwest) corners of the central valley. However, clustering analysis suggested that Chile de Agua populations grown in a single village were combinations of once-separate subpopulations as several well-supported Chile de Agua clades were spread out in the north-south direction along two highways running down each side of the central valleys (Figure 2.S6).

The Tusta population in our dataset was separate from the rest of the *C. annuum* in clustering analysis. The *C. frutescens* accession with Tusta admixture in the genetic assignment analysis may offer a clue to this pattern. As Tusta were found almost exclusively in backyards, they may have a more complicated genealogy than the carefully isolated, row-crop grown Costeño and Chile de Agua. This is consistent with a hybridization event between an ancestor of cultivated *C. annuum* Tusta ancestor and an ancestor to one of the many backyard-grown *C. frutescens*. Such hybridization could explain the relatively large size of Tusta fruits despite its genetic location basal to the *C. annuum* in clustering analysis, including the small-fruited forest-growing *C. annuum*. Also, individuals with ancestry including hybridization between relatively distant lineages often cluster toward the basal parent in clustering analyses (McDade, 1992), as we see in our Tusta subpopulation. However, there was a relatively high level of within-accession homogeneity in the Tusta in this dataset that would not be expected for segregating seeds generated from a highly heterozygous hybrid population. The isolated backyard environments in which Tusta are grown may be responsible for this homogeneity. Several generations of selfing due to such isolation would be sufficient to increase within-accession homogeneity while maintaining the overall genetic pattern of a historic hybrid genealogy, analogous to the production of a recombinant inbred line following a test cross.

The dry-use Taviche exhibits admixture between Costeño and Chile de Agua populations. Taviche also had the lowest within-accession homogeneity of all landraces studied, which could be consistent with relatively recent hybridization. The high levels of heterozygosity resulting from that hybridization event would have been fortified against fixation by their pollinator-friendly *milpa* conditions (Landaverde-González et al., 2017), and geographic proximity to other Taviche stands. In fact, the Taviche in this collection came from just one location in San Pablo Coatlán. Furthermore, despite the fact that Tusta and

Taviche types had clear morphological differences and assorted into two distinct populations in all genetic analyses, two accessions from each type displayed genetic closeness to the other. For example, two accessions displayed the morphological characteristics of Tusta, but were genetically closer to the Taviche. Also, these accessions were collected from the San Pablo Coatlán site from which our Taviche were collected. This could be the result of selection-directed introgression of Tusta morphological traits into the Taviche background. A similar introgression pattern was observed in maize landraces in Oaxaca, where divergent selection caused phenotypic diversification despite overall genetic similarity due to continued gene flow, with genetic effects only discernable as Wahlund effects near the selected loci (Pressoir & Berthaud, 2004b). Perhaps the pollination-permissive *milpas* (Landaverde-González et al., 2017) where Tusta and Taviche are grown together enhance the likelihood for these normally more selfing landraces to outcross and exhibit more maize-like population genetic behavior.

Signals of selection and differentiation across the genome

Using Tajima's D, we identified a region on chromosome 6 in which the Chile de Agua showed evidence of purifying selection (Figure 2.S8). This pattern may be consistent with a QTL allele - such as one conferring a preferable fresh-use phenotype - under selection in the Chile de Agua. Several loci on chromosome 6 are associated with fruit morphology (Han et al., 2016; Hill et al., 2017) including pericarp thickness (Rao et al., 2003), and flowering phenology traits (Yarnes et al., 2013), all of which might be under selective pressure in Chile de Agua. While such QTL could be potentially interesting candidates for follow-up research, a list of candidates from these data would be highly speculative and were thus not included.

Conclusions

In this study, we explored a new collection of chile peppers, which was mostly focused on diverse Mexican landraces that had distinctly different uses. We explored the genetic structure of this collection, identifying that historic use-types formed distinct genetic groups. We found that genetic diversity appeared to be related to the cultivation techniques used for the different landraces. In one landrace cluster (Tusta) there appeared to be a historic hybridization event in an ancestor to one of the many backyard-grown *C. frutescens*, leading to both an interesting genetic and morphological place within the collection. Additionally, we identified signals of selection on chromosomal regions associated with fruit morphology. There was considerable admixture in the global collection as landrace distinction broke down with peppers grown worldwide. This information has provided several hypotheses for future work including exploring differential selection for disease resistance, abiotic stress, and understanding the fine structure of ancient hybridization.

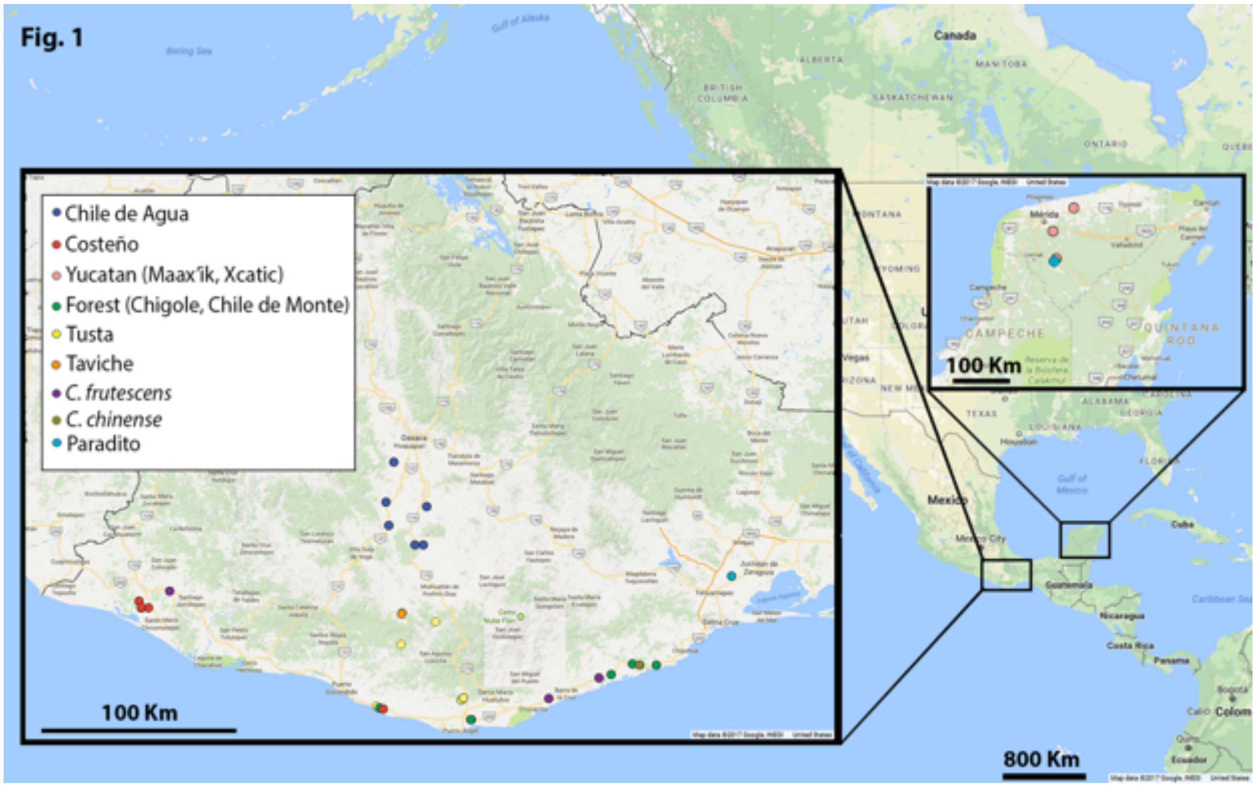

Figure 2.1. Map of chile pepper collection sites in the Mexican states of Oaxaca and Yucatan. The colored circles highlight the most common type grown at that site. Larger image shows Oaxaca and inset map at top right shows Yucatán.

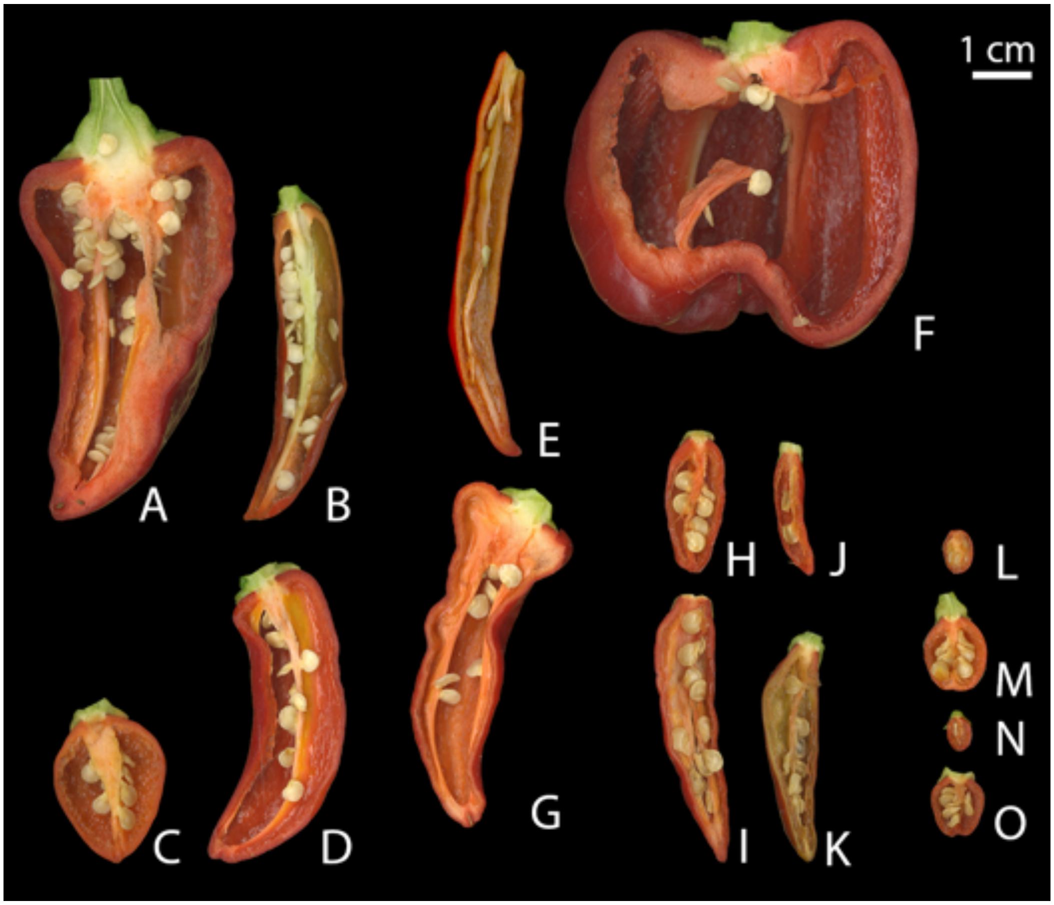

Figure 2.2. Typical pepper fruit types displayed by longitudinal scan of a single fruit. The four major types that predominated this study are on the left. Types are labeled: A) Chile de Agua, B) Costeño, C) Tusta, D) Taviche, E) De Arbol, F) Dulce, G) Guiña Dahni, H) Mirasol, I) Solterito, J) Mareño, K) Piquin, L) Paradito, M) Chigole, N) Bolita

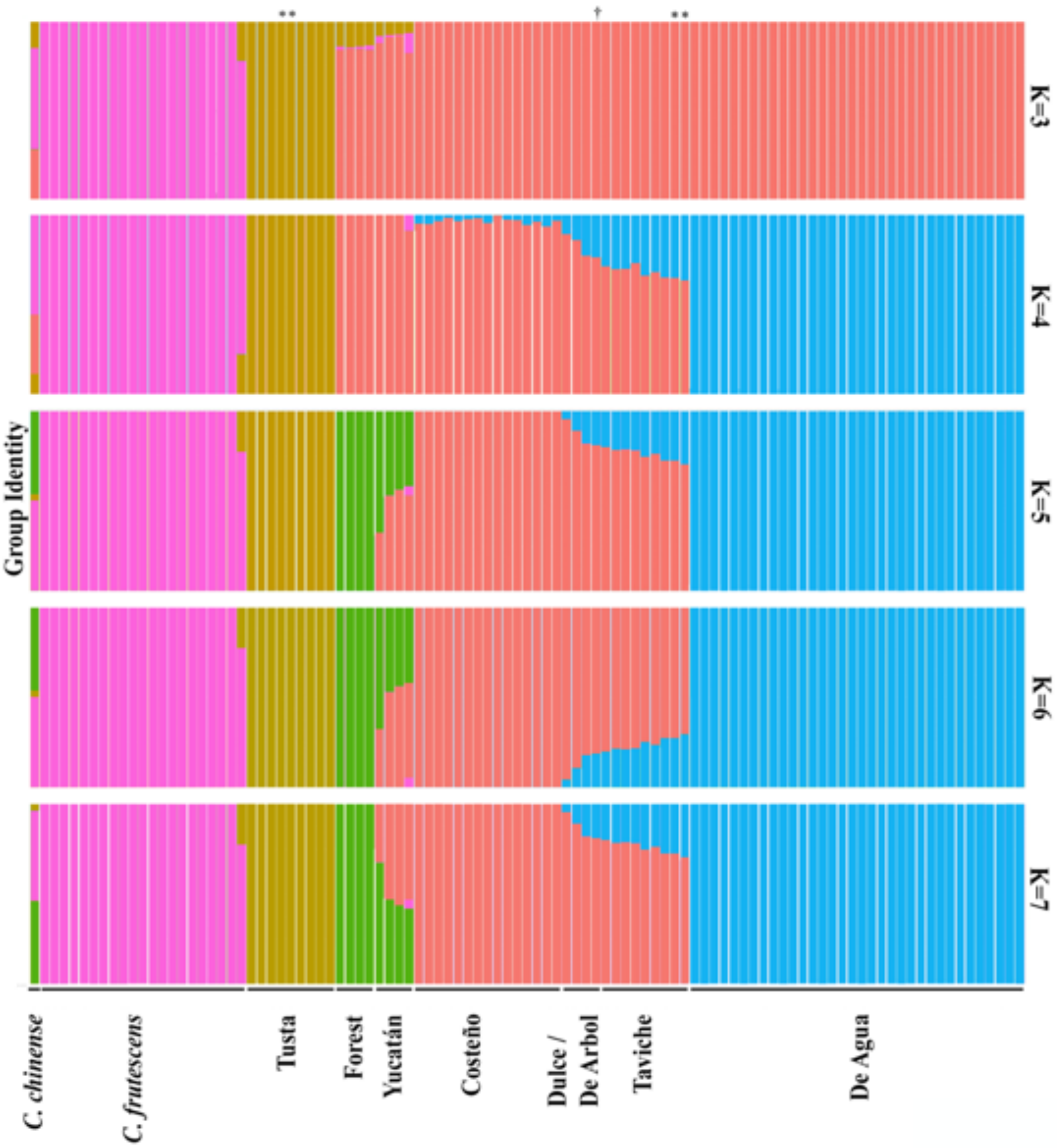

Figure 2.3. Genetic assignment plot. Depicts identity in one of *K* individually colored groups as a stacked barplot, where *K* varies from 3-7. The height of each bar indicates probability of membership for each of the 2013 collection accessions, laid out along the X-axis. Genetic assignment analysis was consistent with the presence of seven subpopulations in our dataset. Clear differences are apparent between named landrace types. After the *C. chinense* accessions, the Costeño and De Agua accessions showed the strongest single-population membership, which is consistent with their being the largest and most restrictively cultivated populations in our collection. Asterisks indicate the Taviche and Tusta accessions bearing high identity with the Tusta and Taviche groups, respectively.

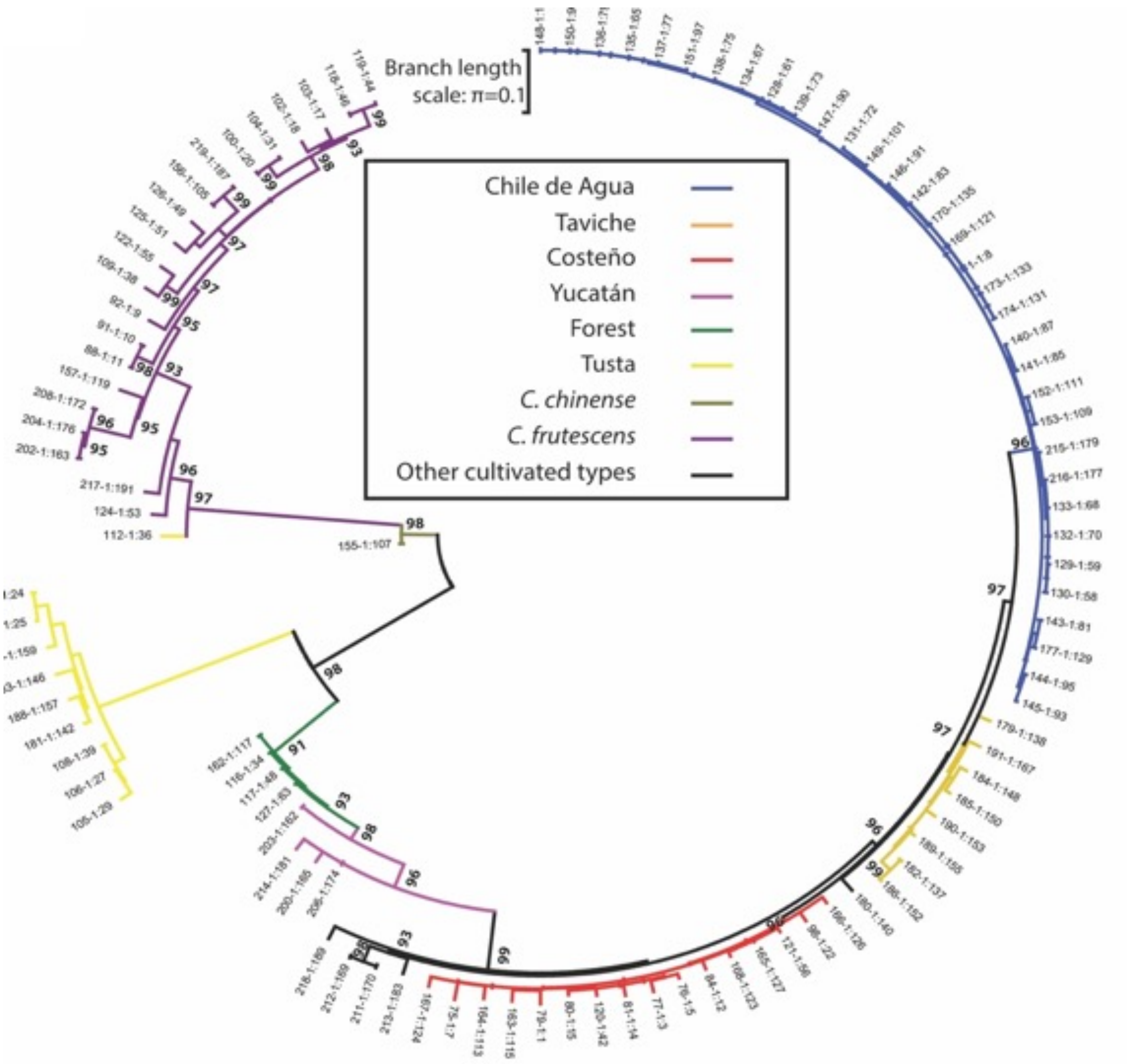

Figure 2.4. Maximum-likelihood population tree with branches color-coded by group. The first clade (shown in purple) included all individuals identified as *C. frutescens*. The second included the Tusta landraces (yellow), and the third contained all other *C. annuum* landraces. This clade also included nearly all accessions from the Yucatán, and many semi-wild accessions collected from forest understories. Both the backyard Yucatán and forest-collected members of the *C. annuum* clade were basal to the other members of this clade. Branch lengths (scale bar shown at top of circle) indicate average pairwise substitutions per site from last common ancestor.

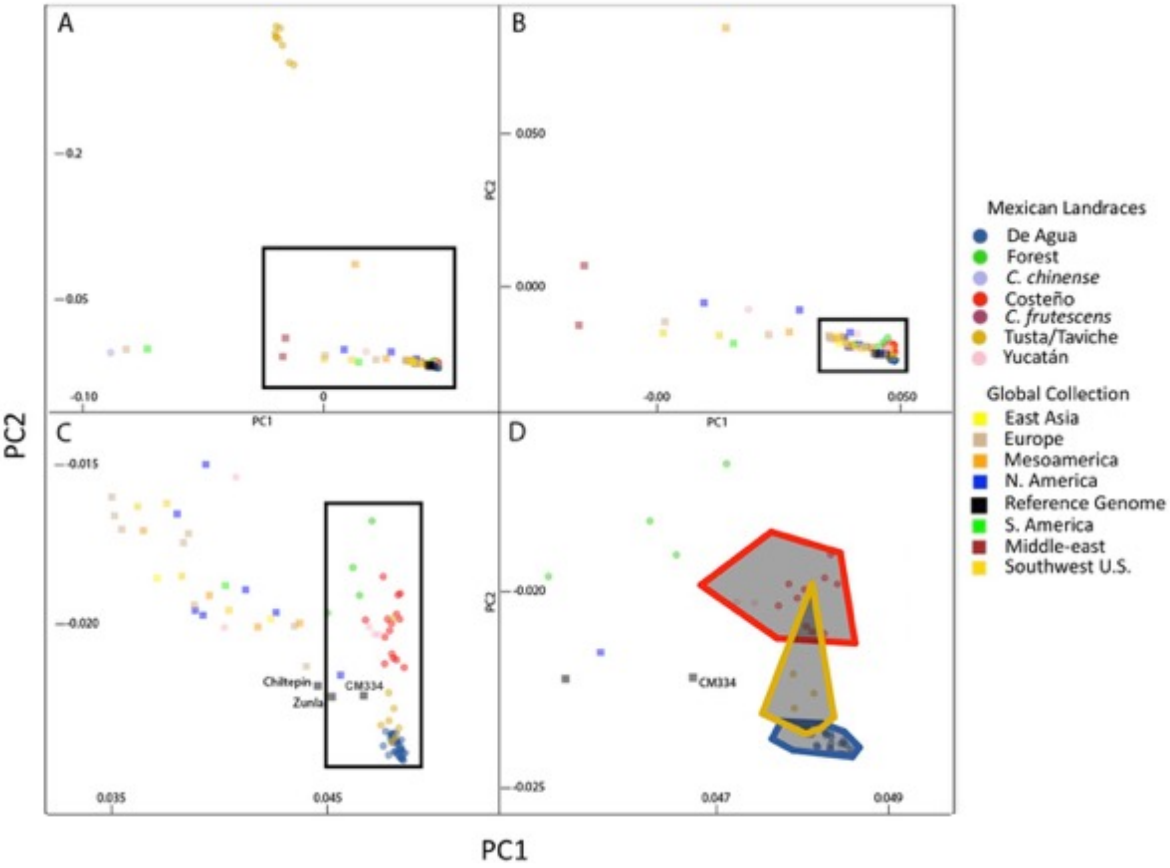

Figure 2.5. Principal components analysis showing genetic structure. Computed over all collections and the three published reference genomes. Gray frames in A, B, and C show the area of extent for the subsequent-lettered panel. Members of our population grown from the 2013 Mexican collection trip are indicated by colored circles, members of our global collection set are indicated by colored squares, and the published reference genomes are indicated by dark grey squares. Oaxacan landraces grouped together, including the Criollo de Morelos 334 (CM334) reference genome. Several Yucatán accessions clustered away, near the global set.

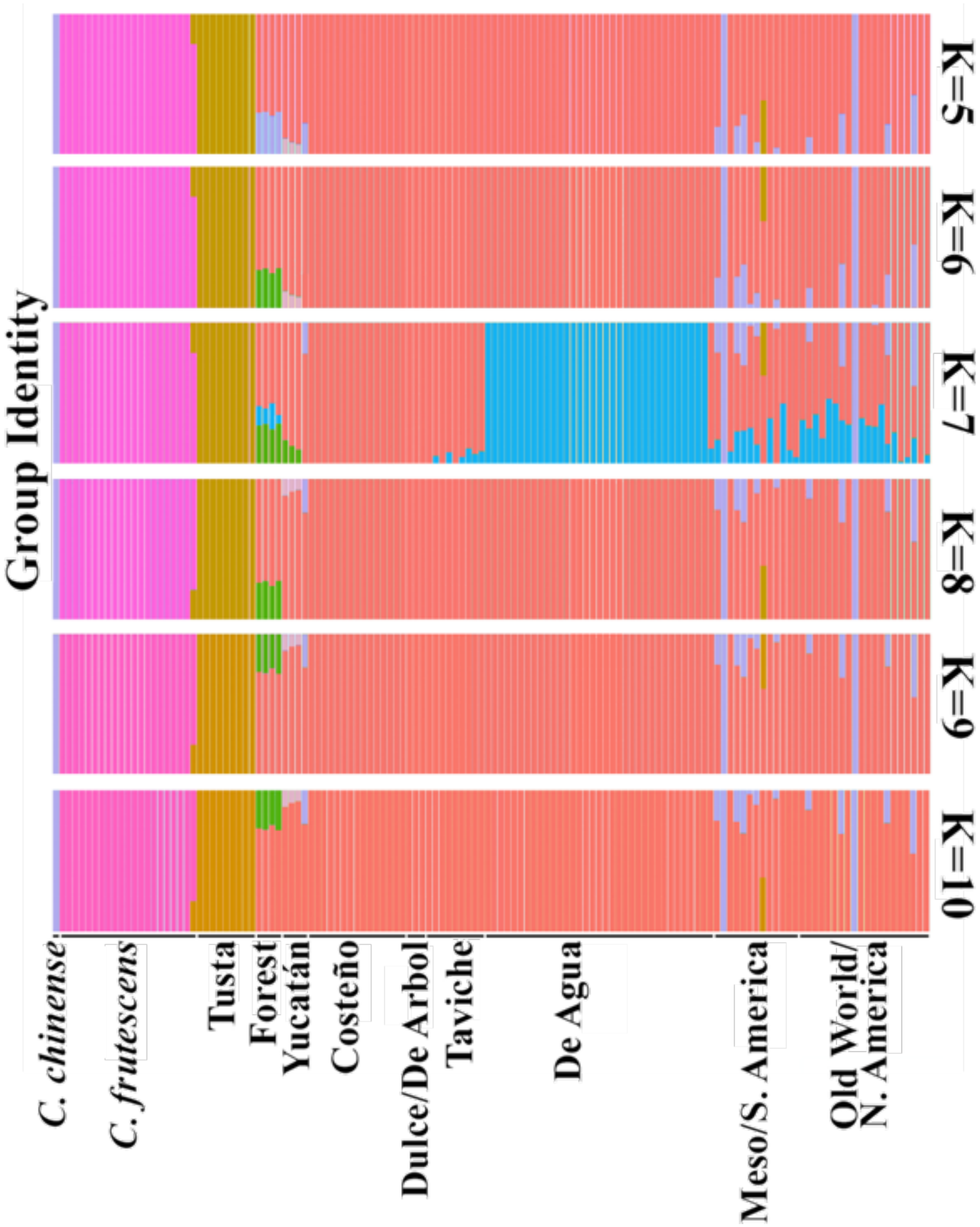

Figure 2.6. Genetic assignment plot including global collection set. Major groups within the 2013 Mexican collections were recapitulated, with the exception of the *C. frutescens*, which had only one group resolved, and the Yucatán accessions, which included a unique group. The genetic structure of the global population was not clearly related to geographic origin, and *C. chinense* identity appeared scattered across the global chile pepper collection.

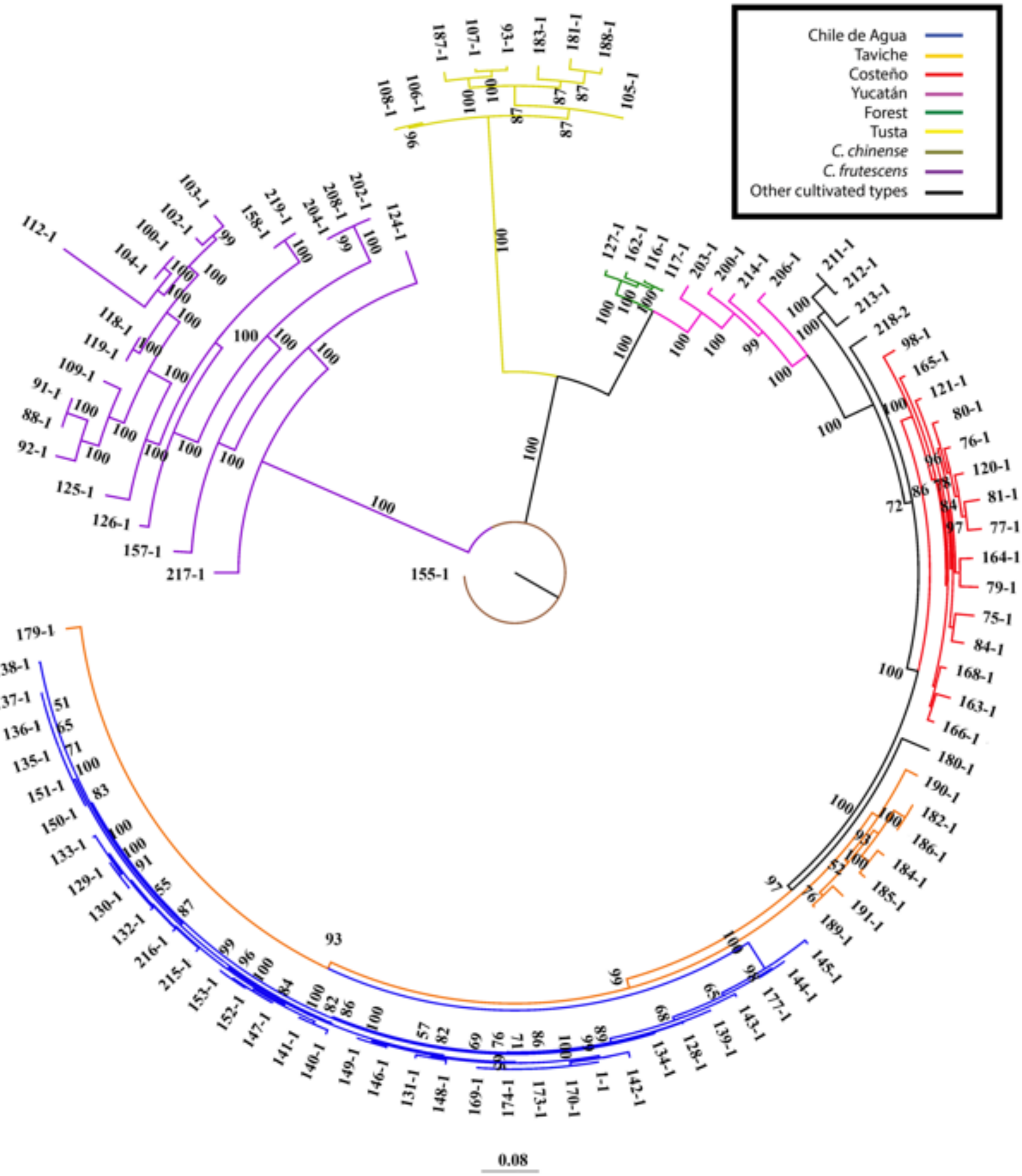

Figure 2.S1. High-bootstrap clustering tree. Clusters created from 1000-bootstrap optimization analysis identified the same groups created with the less computationally intensive analysis (see Figure 2.4).

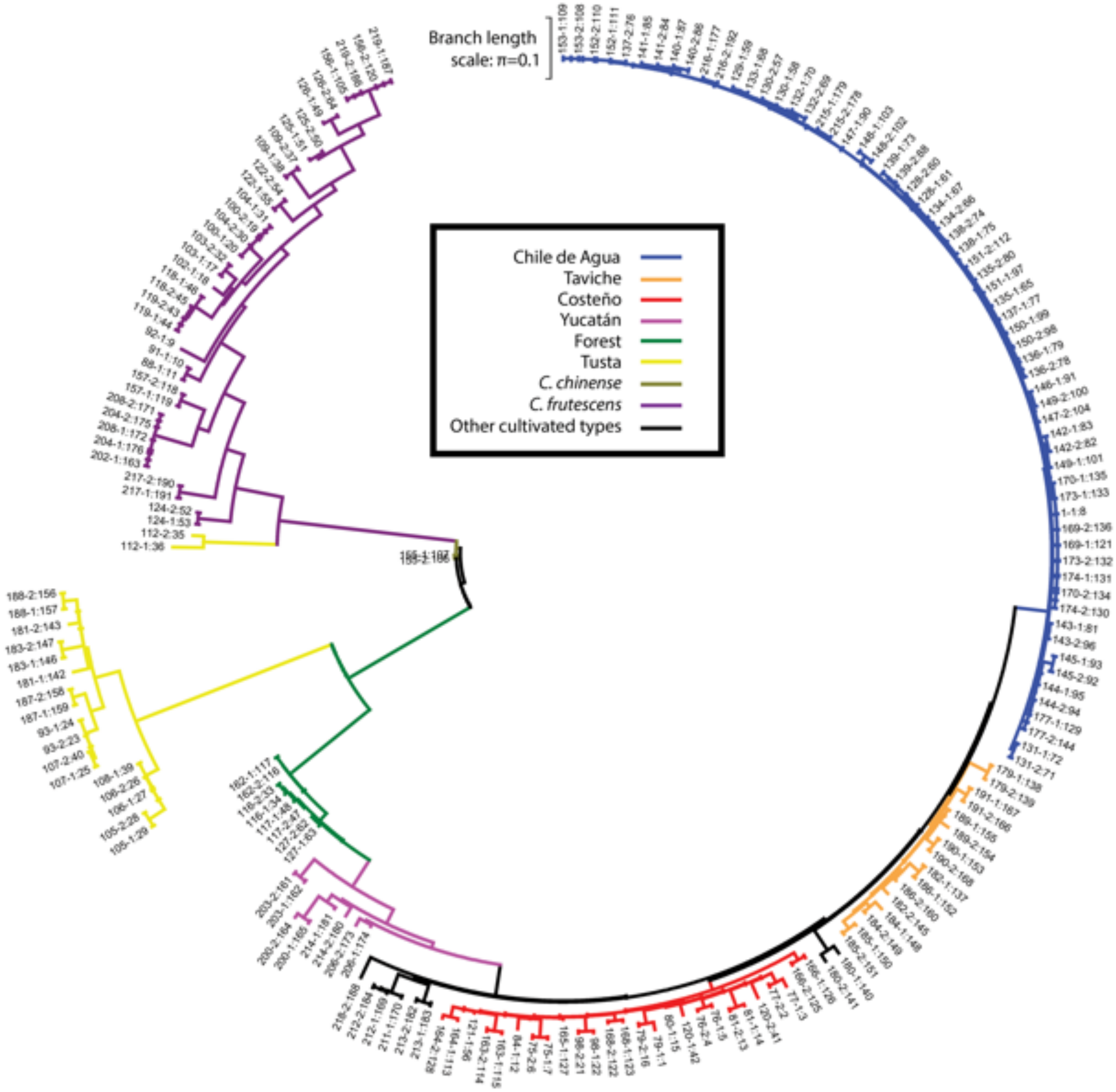

Figure 2.S2. Population tree including both replicate plants from each accession. Approach, color-coding, and branch-length meaning are the same as in Figure 2.4. Branches labeled with the same number preceding the hyphen belong to the same accession. (See Figure 2.4)

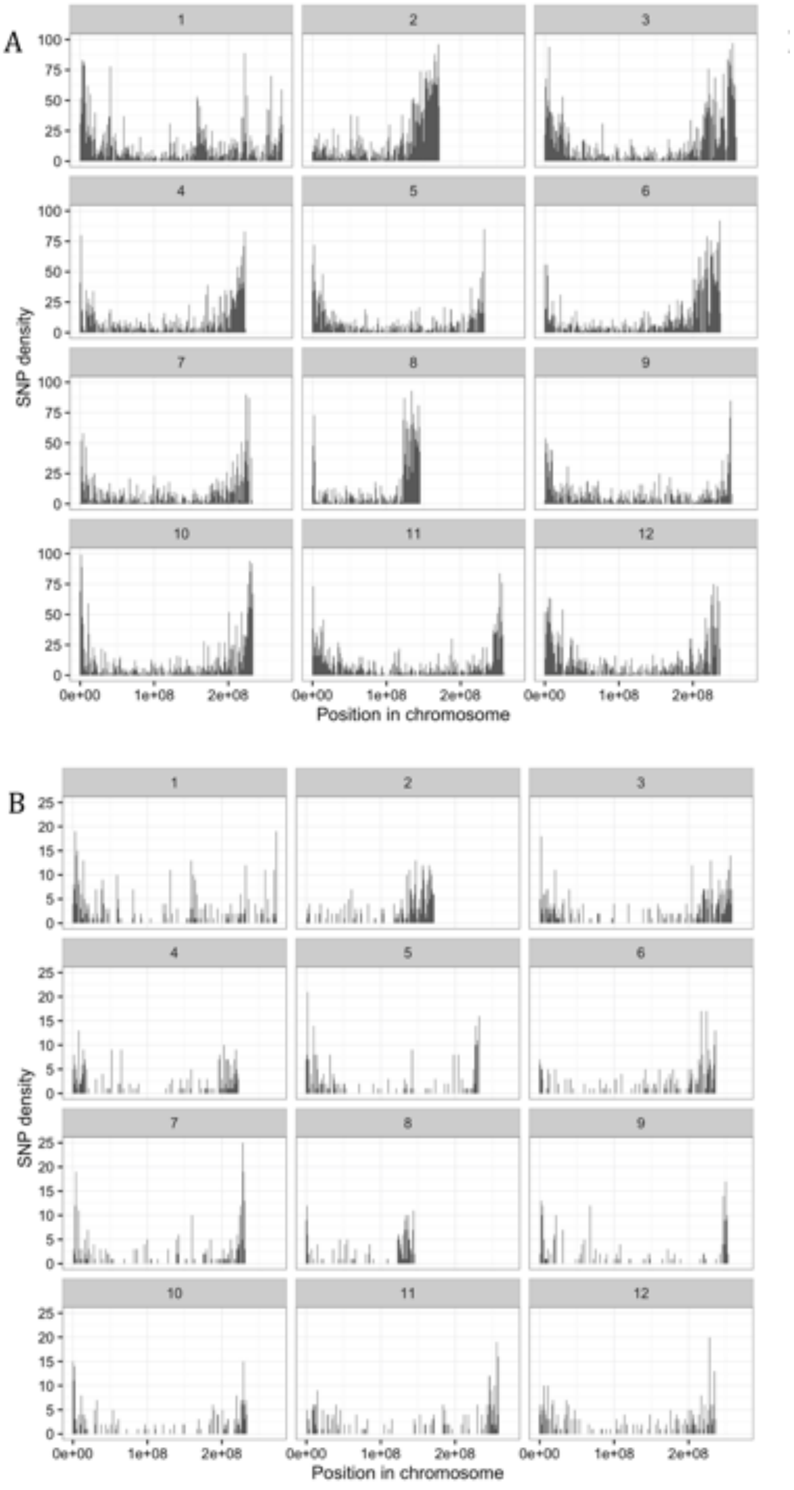

Figure 2.S3. Density maps showing the position of SNPs used over each pepper chromosome. A) SNPs used for analysis of the Mexican collection. B) SNPs with sufficient coverage to analyse the Mexican and global collection. Bar heights indicate number of SNPs called in each megabase window over the entire *C. annuum* genome.

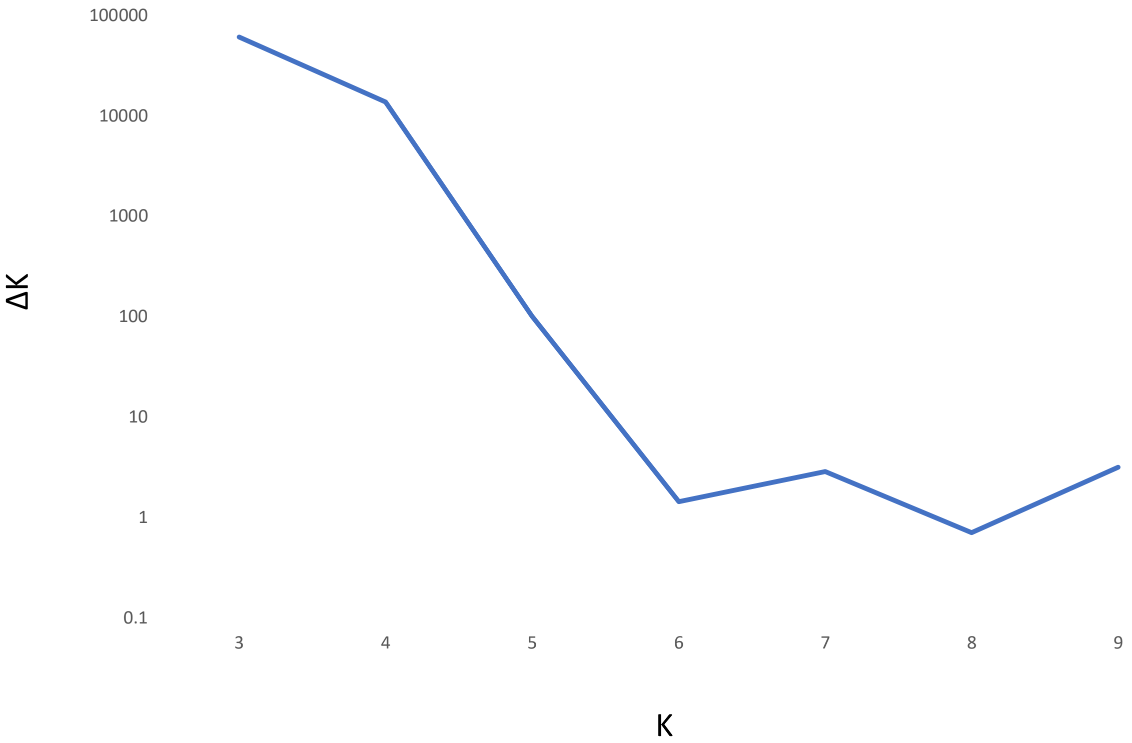

Figure 2.S4. Δ*K* plot for Mexican peppers. Plot of Δ*K* used to determine values of *K* to use in genetic assignment analysis for the Mexican dataset only.

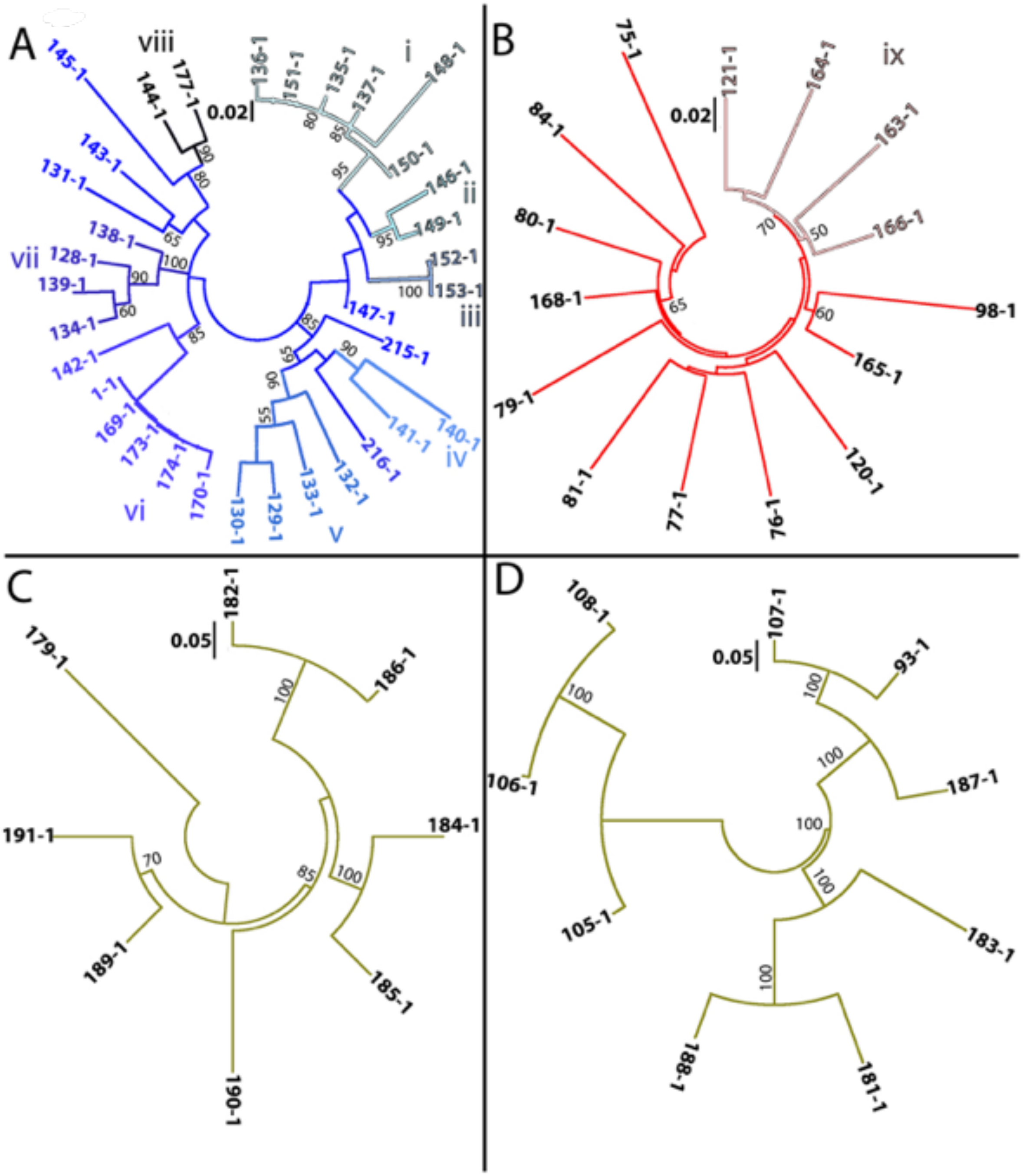

Figure 2.S5. Subtrees of the four major named types. Genetic distance indicated by branch lengths in average substitutions per site, and all bootstrap values exceeding 50% indicated above each node. Types are as follows: A) De Agua, B) Costeño, C) Taviche, except for 179-1 and 185-1 which are Tusta accessions from the same site as the Taviche, and D) Tusta, except for 181-1 and 183-1, which are Taviche.

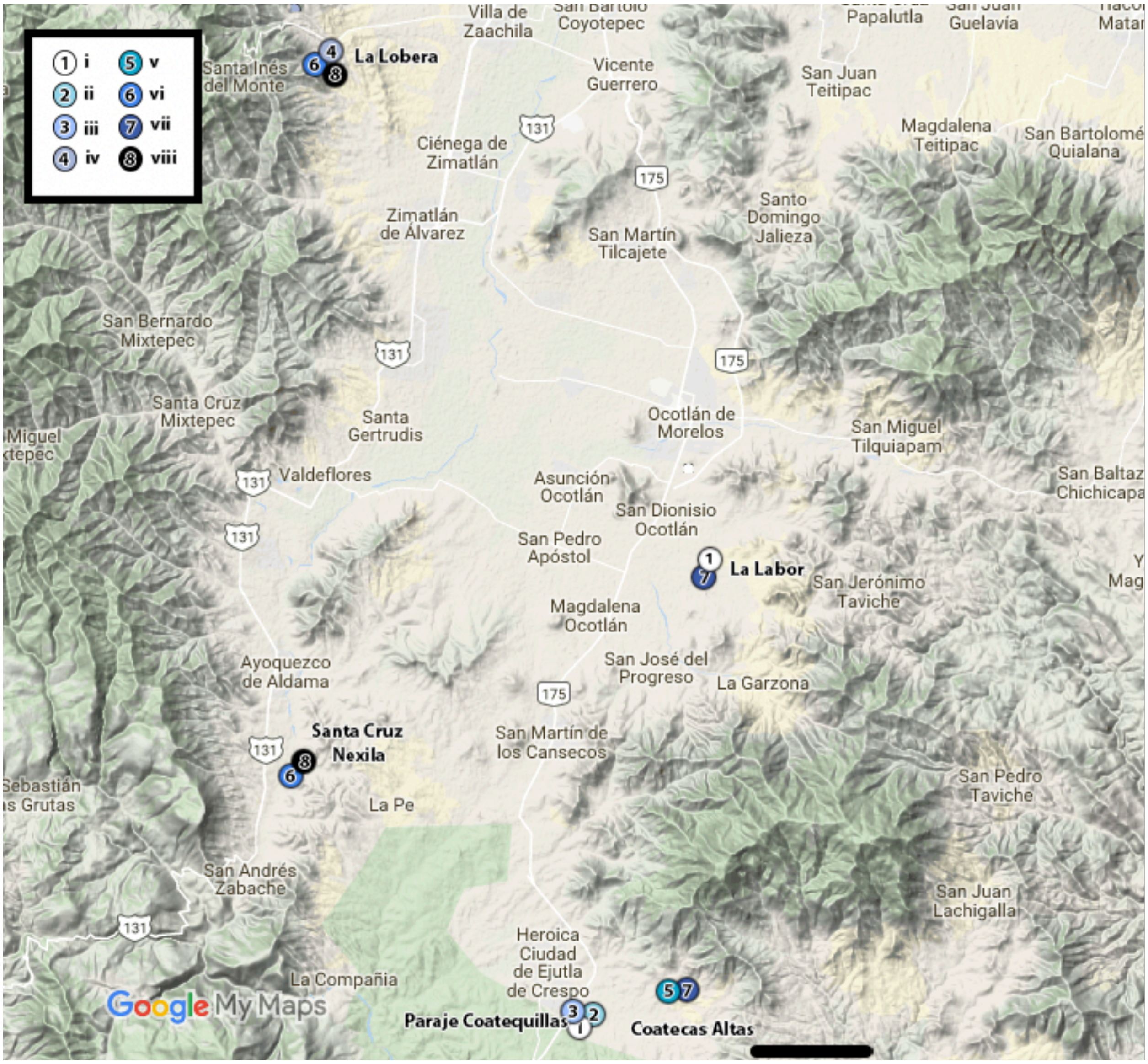

Figure 2.S6. Close-up map of Central Valleys Chile de Agua collection sites. Colored circles indicate the presence of accessions at that site corresponding to clades described in Figure 2.S5, with clade labels in Arabic numerals. Legend indicates how each circle label corresponds to clades in Figure 2.S5.

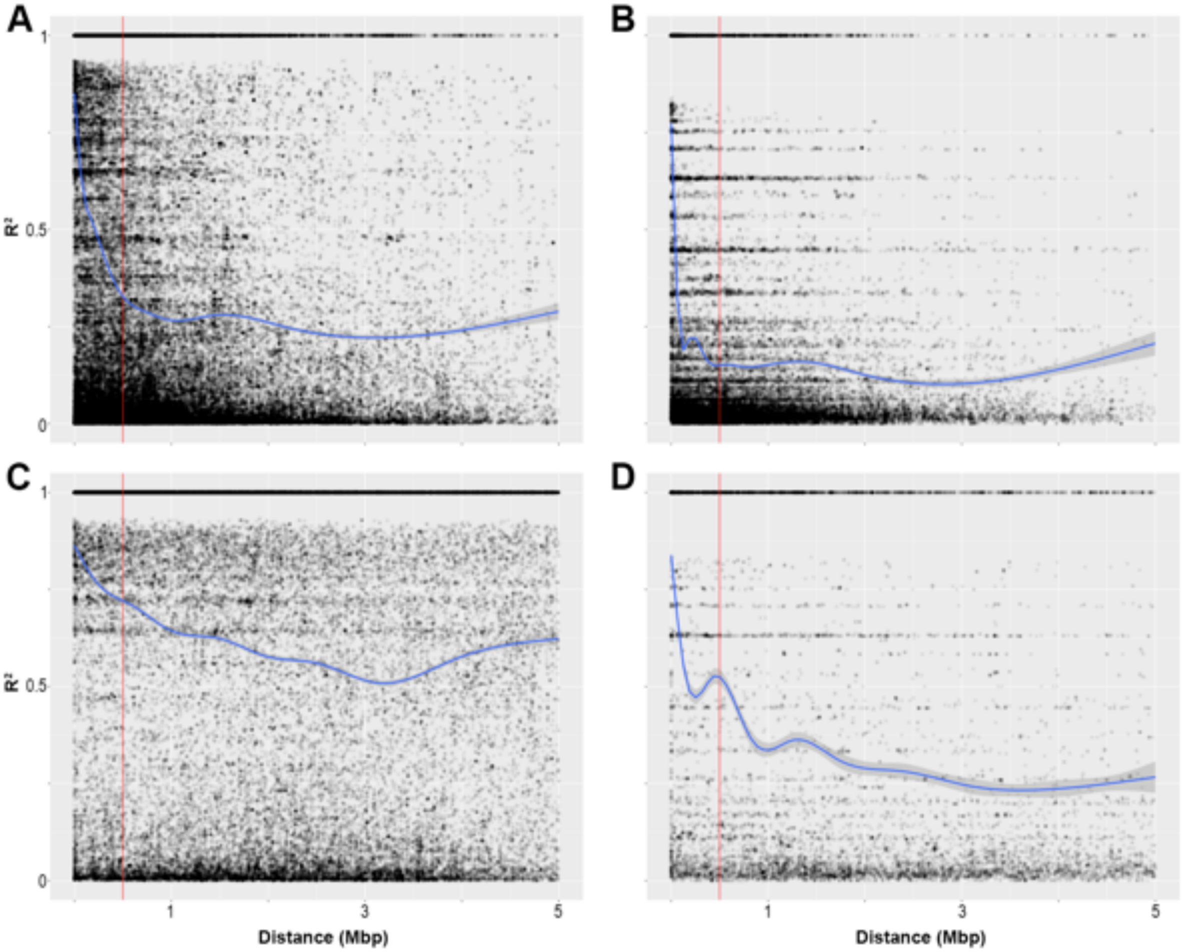

Figure 2.S7. LD-decay plots. Dark circles indicate pairs of SNPs, plotted by distance between each pair and the R^2 value returned by their correlation. Red horizontal line drawn at the 500kb distance between SNPs. Blue line added by a Gaussian smoothing function implemented in R. Top panels A, B display pairs in the euchromatic, while bottom panels C, D display pairs in the pericentromeric, regions. Left panels A, C refer to De Agua subpopulation, and right panels refer to Costeño subpopulation.

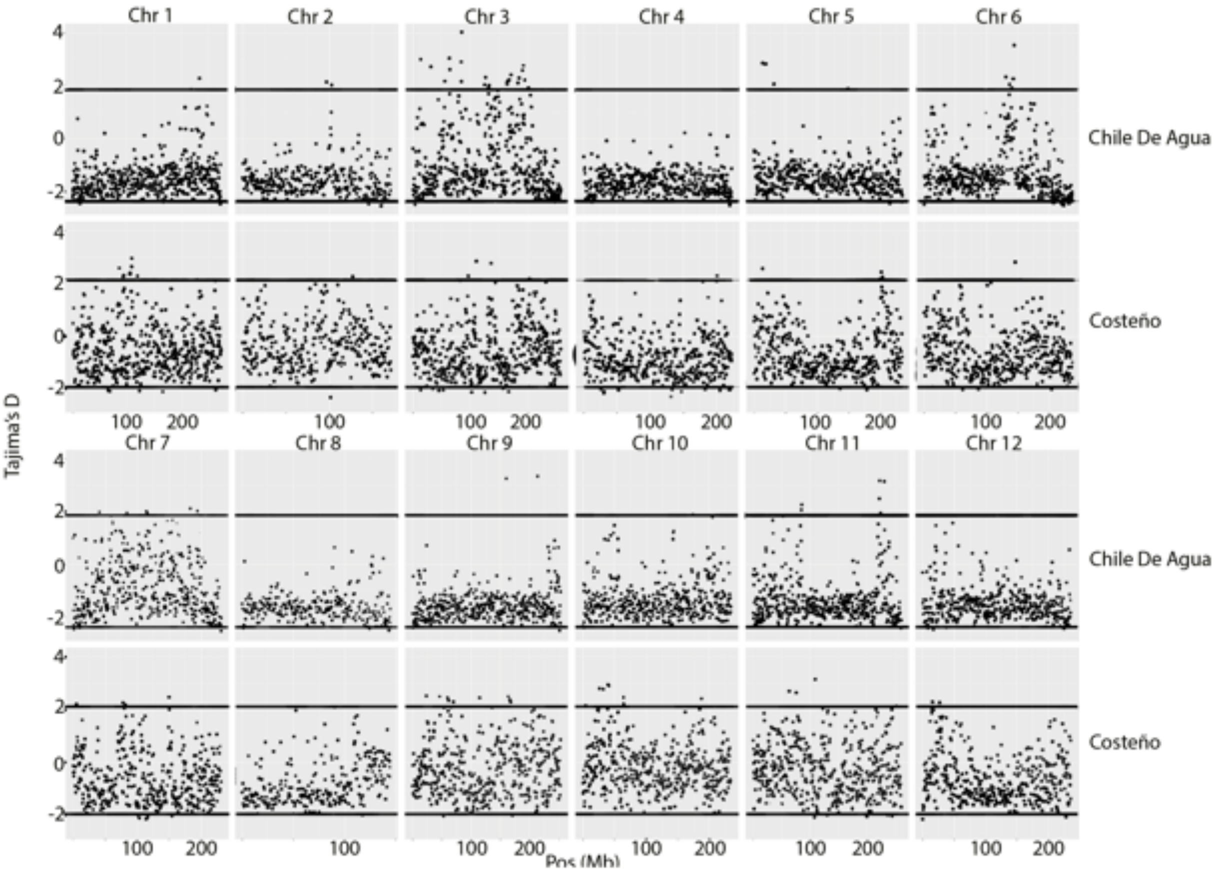

Figure 2.S8. Genome scans for deviation from selectively neutral evolution. Depicts Tajima's D (y-axis), calculated for all 500-kb windows in the genome (start positions on x-axis). Top two rows show chromosomes 1-6, bottom two rows show 7-12. First and third rows from top show Chile de Agua, second and fourth show Costeño values. First and 99th percentile values over all windows indicated by black horizontal lines.

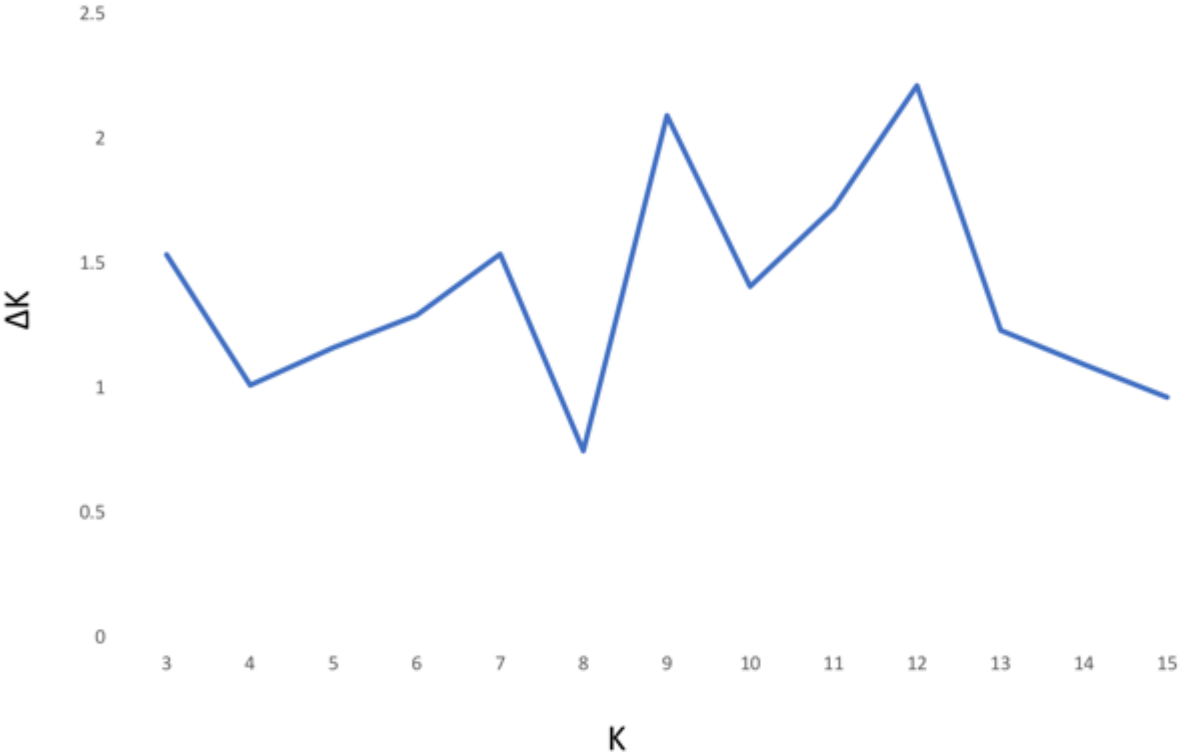

Figure 2.S9. Δ*K* plot for global collection. Plot of Δ*K* used to determine values of *K* to use in genetic assignment analysis including the global chile pepper collection.

Table 2.1: Within-population diversity for main *C. annuum* landraces. Results of re-filtering SNPs within the main C. annuum landraces included in our Mexican collection. Column n: number of accessions within each landrace subpopulation; π: average pairwise differences per nucleotide; % IBS: within-accession identity-by-state (group average using Mexican collection-wide SNPs) with different letters indicating significantly different groups. The last column gives maximum allowed heterozygosity for each SNP during filtering.

Group	*n*	*SNPs*	*π*	*Within-accession IBS (%)*	*Filter Threshold for Heterozygosity (%)*
Taviche	8	4,056	0.00349	96.8% A (SD = 0.8%)	5%
Costeño	15	3,355	0.00222	96.9% A (SD = 1.1%)	2.5%
Tusta	9	9,659	0.00538	98.7% B (SD = 0.9%)	5%
Chile de Agua	34	7,403	0.00144	99.6% C (SD = 0.4%)	2.5%

Table 2.2: Pairwise, corrected FST values for major population pairs. Values above the diagonal are average FST values for all SNPs calculated using Weir and Cockerham's corrected FST (1984). Below the diagonal, levels of significance are indicated. NAs indicate pairwise FST not applicable because populations are nested. All pairs had FST values significantly higher than admixture (*** indicates p < .001, permutation test of 10,000 permutations).

	C. frutescens	*C. annuum*	*Forest*	*Taviche*	*Tusta*	*Costeño*	*De Agua*
C. frutescens	-	0.74231	0.66724	0.75646	0.81053	0.82345	0.89223
C. annuum	***	-	NA	NA	NA	NA	NA
Forest		NA	-	0.19887	0.75809	0.42406	0.59077
Taviche	***	NA		-	0.71391	0.21242	0.46114
Tusta	***	NA		***	-	0.87761	0.95156
Costeño	***	NA		***	***	-	0.63705
De Agua	***	NA		***	***	***	-

Table 2.S1. Collection information for all accessions used in this study. Includes those eventually removed in quality control. Because the Juan Diegal collection location is less than 250m from the nearest Polvorin collection location and they are all located on the same street, the Juan Diegal and Polvorin locations are treated as one collection location.

	State:	OAXACA							
Plant ID	*Location*	*Sample ID*	*Species*	*Region*	*Village*	*Local name*	*Growth environment of accession:*	*Use (fresh /dry)*	*Notes*
219 [-#]	La Tortolita	CanTorto3	C. annuum	COAST	La Tortolita, Santiago Astata				
218	Rancho Llano	CanRan3B	C. annuum	COAST	Rancho Lllano, Tehuantepec	GUIÑA DAHNI	Plantation	Dry	irrigation channels
217	Rancho Llano	CanRan1B	C. annuum	COAST	Rancho Lllano, Tehuantepec	PAYASO	Milpa	Fresh	irrigation channels
216	Abasolo	CanAbasolo2	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	fresh	

215	Abasolo	CanAbasolo1	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	fresh
191	Paraje Coatequillas	CanEjutla3	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh
190	Paraje Coatequillas	CanEjutla1	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh
189	Paraje Coatequillas	CanEjutla2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh
188	Santa Lucía	CanSLucía1	C. annuum	CENTRAL VALLEYS	Santa Lucía Miahuatlán	TUSTA		Fresh-Dry
187	San Baltazar	CanSnBalta1	C. annuum	CENTRAL VALLEYS	San Baltazar Loxicha	TUSTA		Fresh-Dry

186	San Pablo Coatlán	CanSPablo3	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	Dry	local commerce between community-es
185	San Pablo Coatlán	CanSPablo10	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Milpa	Dry	local commerce between community-es
184	San Pablo Coatlán	CanSPablo7	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	Dry	local commerce between community-es
183	San Pablo Coatlán	CanSPablo6	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	Dry	local commerce between community-es
182	San Pablo Coatlán	CanSPablo5	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	Dry	local commerce between community-es
181	San Pablo Coatlán	CanSPablo4	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	Dry	local commerce between community-es

180	San Pablo Coatlán	CanSPablo2	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	DE ÁRBOL	Backyard	Fresh	
179	San Pablo Coatlán	CanSPablo1	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Backyard	Fresh-Dry	
177	Santa Cruz Nexila	CanStaCruz19	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	open field
174	Santa Cruz Nexila	CanStaCruz15	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	open field
173	Santa Cruz Nexila	CanStaCruz14	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	open field
170	Santa Cruz Nexila	CanStaCruz11	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	open field

169	Santa Cruz Nexila	CanStaCruz6	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	open field
168	Rosedal	CanRose10B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	used for salsa
167	Rosedal	CanRose9B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	Failed quality control, removed from analysis
166	Rosedal	CanRose8B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	used for salsa
165	Rosedal	CanRose7B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	used for salsa
164	Rosedal	CanRose6B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	used for salsa
163	Rosedal	CanRose11B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Dry	used for salsa
162	Rosedal	CanRose2B	C. annuum	COAST	Rosedal, close to Tomatal	CHIGOLE	Backyard	Fresh	
157	Copalita	CanCopal1	C. annuum	COAST	Copalita	MIRASOL	Backyard	Fresh	

156	La Tortolita	CanTorto2	C. annuum	COAST	La Tortolita, Santiago Astata	GUAJILLO*	Backyard	Dry	*Referred to as guajillo, but morphology differs from normal Guajillo and genetics shows likely *C. frutescens*
155	La Tortolita	CchTorto1	C. chinense	COAST	La Tortolita, Santiago Astata	HABANERO	Backyard	Fresh	
153	Paraje Coatequillas	CanParaje10	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
152	Paraje Coatequillas	CanParaje9	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
151	Paraje Coatequillas	CanParaje7	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage

150	Paraje Coatequillas	CanParaje6	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
149	Paraje Coatequillas	CanParaje5	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
148	Paraje Coatequillas	CanParaje4	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
147	Paraje Coatequillas	CanParaje2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
146	Paraje Coatequillas	CanParaje1	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
145	La Lobera	CanLobera6	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	

144	La Lobera	CanLobera5	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	
143	La Lobera	CanLobera4	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	
142	La Lobera	CanLobera3	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	
141	La Lobera	CanLobera2	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	
140	La Lobera	CanLobera10	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	Fresh	
139	La Labor	CanLabor6	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage

138	La Labor	CanLabor5	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
137	La Labor	CanLabor4	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
136	La Labor	CanLabor3	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
135	La Labor	CanLabor2	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
134	La Labor	CanLabor1	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	Fresh	Technified, with irrigation, plastic soil coverage
133	Coatecas Altas	CanCoat8	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse

132	Coatecas Altas	CanCoat7	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse
131	Coatecas Altas	CanCoat6	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse
130	Coatecas Altas	CanCoat5	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse
129	Coatecas Altas	CanCoat4	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse
128	Coatecas Altas	CanCoat2	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	Fresh	Highly technified organic greenhouse
127	Saachilac	CanSaa5	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	Fresh	Deciduous tropical forest
126	Saachilac	CanSaa3	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	Fresh	Deciduous tropical forest

125	Saachilac	CanSaa2	C. annuum	COAST	Saachilac, Santiago Astata	SOLTERITO	Backyard	Dry	
124	Saachilac	CanSaa1	C. annuum	COAST	Saachilac, Santiago Astata	MAREÑO	Backyard	Fresh	
122	Rosedal	CanRose3	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO	Plantation	Dry	Failed quality control, removed from analysis
121	Rosedal	CanRose2	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO FLAQUITO	Plantation	Dry	used for salsa
120	Rosedal	CanRose1	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO	Plantation	Dry	used for salsa
119	Río Seco	CanRS2	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	Fresh	
118	Río Seco	CanRS1	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	Fresh	
117	Polvorín	CanPol10	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Forest	Fresh	with voucher exemplar

116	Polvorín	CanPol9	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Milpa	Fresh	
109	Paso de Los Indios	CfrPind1	C. frutescens	COAST	Paso de los Indios, Pochutla	MIRASOL	Backyard	Fresh	
108	Los Reyes	CanReyes4	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Milpa	Fresh-Dry	sample of the seed of his milpa
107	Los Reyes	CanReyes3	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	Fresh-Dry	
106	Los Reyes	CanReyes2	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	Fresh-Dry	
105	Los Reyes	CanReyes1	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	Fresh-Dry	
104	Los Reyes	CfrReyes4	C. frutescens	COAST	Los Reyes, Pochutla		Backyard		
103	Los Reyes	CfrReyes3	C. frutescens	COAST	Los Reyes, Pochutla		Backyard	Fresh	
102	Los Reyes	CfrReyes2	C. frutescens	COAST	Los Reyes, Pochutla		Backyard	Fresh	
100	Lagunilla	CfrLaguna1	C. frutescens	COAST	Lagunilla, Pochutla	CHILE BOLITA	Backyard	Fresh	used in salsa
98	Lagartero	CanLag4	C. annuum	COAST	Lagartero, Pinotepa Nacional	COSTEÑO ROJO	Milpa	Dry	very complex polyculture

93	JuanDiegal	CanJuan1	C. annuum	COAST	Juan Diegal, Pochutla	TUSTA	Milpa	Fresh-Dry
92	Huaxpaltepec	CfrHuax6	C. frutescens?	COAST	Huaxpaltepec	PIQUIN	Backyard	Fresh
91	Huaxpaltepec	CfrHuax5	C. frutescens?	COAST	Huaxpaltepec	PIQUIN	Backyard	Fresh
88	Huaxpaltepec	CanHuax2	C. frutescens?	COAST	Huaxpaltepec	PIQUIN	Backyard	Fresh
84	GpeVictoria	CanGpe1	C. anuum	COAST	Guadalupe Victoria, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Dry
81	Añil	CanAnil7	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Dry
80	Añil	CanAnil6	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Dry
79	Añil	CanAnil5	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Dry
77	Añil	CanAnil3	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Dry

76	Añil	CanAnil2	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Dry	
75	Añil	CanAnil10	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Dry	
1	Santa Cruz Nexila	CanStaCruz20	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	Fresh	
112	Paso de Los Indios	CanPind1	C. annuum	COAST	Paso de los Indios, Pochutla	TUSTA	Backyard	Fresh-Dry	
112	Paso de Los Indios	CanPind2	C.annuum	COAST	Paso de los Indios, Pochutla	TUSTA	Backyard	Fresh-Dry	
	YUCATAN								
214	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	Fresh	Synonymous with CanDzib1
213	Dzibtzantun	CanDzib14	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	Fresh	
212	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	Fresh	
211	Maní	CanMani14	C. annuum	YUCATAN	Maní	DULCE	Backyard	Fresh	traspatio casco urbano
208	Maní	CanMani11	C. annuum	YUCATAN	Maní	PARADITO	Backyard		traspatio casco

									urbano
206	Maní	CanMani7	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Fresh	Semilla propia de muchas generaciones
204	Maní	CanMani3	C. annuum	YUCATAN	Maní	PARADITO	Backyard		traspatio casco urbano
203	Maní	CanMani2	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Fresh	traspatio casco urbano
202	Maní	CanMani1	C. annuum	YUCATAN	Maní	PARADITO	Backyard		traspatio casco urbano
200	Acanceh	CanAcan20	C. annuum	YUCATAN	Acanceh	MAAX'IK		Fresh	

Table 2.S2. **Identity scores output by FastStructure (varQ).** For Mexican chile pepper collection, at each level of K. Higher K values (>5) create groups with very low identity across all accessions.

	K = 2	
Plant ID	Group 1	Group 2
155	35213	22022
208	59952	0.7268
202	59874	0.6781
204	59860	0.6779
124	58960	0.9285
217	58428	0.9036
157	59680	0.7056
119	59922	0.6474
103	59926	0.6569
219	59926	0.6637
92	59910	0.6868
118	59764	0.6969
88	59710	0.6935
100	59656	0.683
102	59570	0.6614
109	59428	0.6767
104	58752	0.6693
156	59954	0.6712
91	59526	0.7375
125	59888	0.7012
126	59946	0.6564
112	55789	4129.6
93	34863	25056
107	35082	24843
187	34734	25115
183	34987	24798
181	34511	25122
188	34732	24867
106	33431	26026
105	34095	24826
108	32592	26149

116	5523.7	53797
117	5303.6	53645
127	5309.3	53314
162	5574.3	52415
203	4669.8	54211
200	2857	56516
214	2627.4	56802
206	8604.4	51231
168	0.6346	60010
165	0.6299	59992
121	0.6423	59916
164	0.6273	59904
80	0.6299	59906
163	0.6293	59866
81	0.6288	59820
84	0.6306	59806
75	0.6386	59726
77	0.6272	59698
120	0.64	59950
79	0.6314	59468
98	0.6381	59376
166	0.6296	58630
76	0.6328	58626
218	0.6957	59918
213	0.6682	59728
211	0.6471	59874
212	0.6501	59894
180	0.6477	59616
182	0.6162	59860
189	0.6219	59962
186	0.6253	59680
184	0.6121	59850
190	0.6325	58318
191	0.6287	59838
185	0.6169	59852
179	0.6273	59896
138	0.5844	60036
150	0.5841	60030
148	0.5849	60030
151	0.5841	60026

134	0.5864	60028
133	0.5857	60024
169	0.5856	60020
174	0.5856	60016
1	0.5865	60012
215	0.5861	60014
149	0.5887	60012
143	0.5915	60006
130	0.5927	60004
147	0.5863	60002
142	0.5877	60002
152	0.5871	59988
128	0.6051	59980
177	0.5865	59978
136	0.5952	59950
137	0.5872	59948
135	0.5964	59926
132	0.6063	59952
139	0.5902	59908
216	0.5864	59888
131	0.5916	59848
146	0.5903	59838
129	0.6005	59720
141	0.6014	59644
145	0.5891	59614
144	0.5917	59592
140	0.5952	59384
173	0.5868	59302
153	0.5883	58556
170	0.5904	55116

$K = 3$

Plant ID	Group 1	Group 2	Group 3
155	15896	32482	8856.9
208	0.3772	59952	0.3734
202	0.3712	59874	0.3707
204	0.3713	59860	0.3708
124	0.4125	58960	0.4069
217	0.4118	58428	0.4063

157	0.3767	59680	0.3755
119	0.371	59922	0.3716
103	0.3694	59926	0.3695
219	0.3695	59926	0.3694
92	0.3721	59910	0.3707
118	0.373	59764	0.3703
88	0.3721	59710	0.3701
100	0.3714	59656	0.37
102	0.3694	59570	0.3696
109	0.3703	59428	0.3694
104	0.3697	58752	0.3694
156	0.3702	59954	0.3696
91	0.3842	59526	0.3938
125	0.3755	59888	0.3758
126	0.3693	59946	0.3694
112	0.3951	46312	13607
93	0.3782	0.3741	59918
107	0.3746	0.3817	59924
187	0.379	0.3735	59848
183	0.379	0.3701	59784
181	0.382	0.3694	59632
188	0.3796	0.3698	59598
106	0.3806	0.3723	59456
105	0.3756	0.3785	58920
108	0.3852	0.3721	58740
116	49462	1151.5	8707.8
117	49346	601.07	9002
127	49112	803.99	8707.3
162	48197	1573.2	8219.1
203	51239	2457.1	5184.5
200	53729	657.51	4986.4
214	54144	535.39	4749.4
206	48714	6843	4278.2
168	60010	0.3713	0.3854
165	59992	0.3709	0.3839
121	59916	0.3722	0.3878
164	59904	0.3703	0.3833
80	59906	0.3707	0.3836
163	59866	0.3707	0.3843
81	59820	0.3704	0.384

84	59806	0.3709	0.3833
75	59726	0.3718	0.3859
77	59698	0.3702	0.3837
120	59950	0.3719	0.3863
79	59468	0.3707	0.385
98	59376	0.3718	0.385
166	58630	0.3707	0.3846
76	58626	0.371	0.3854
218	59918	0.3814	0.4024
213	59728	0.3769	0.4001
211	59874	0.3733	0.3905
212	59894	0.3737	0.3919
180	59616	0.3739	0.3852
182	59860	0.3694	0.3772
189	59962	0.3698	0.3789
186	59680	0.3701	0.3797
184	59850	0.3695	0.3773
190	58318	0.371	0.3831
191	59838	0.3705	0.3812
185	59852	0.3694	0.378
179	59896	0.3704	0.3811
138	60036	0.4008	0.3703
150	60030	0.4016	0.3702
148	60030	0.3993	0.3705
151	60026	0.4017	0.3702
134	60028	0.3948	0.3705
133	60024	0.3969	0.3705
169	60020	0.3972	0.3705
174	60016	0.3971	0.3705
1	60012	0.395	0.3706
215	60014	0.3956	0.3705
149	60012	0.3895	0.3708
143	60006	0.3848	0.3708
130	60004	0.383	0.3707
147	60002	0.3959	0.3706
142	60002	0.3922	0.3707
152	59988	0.3934	0.3707
128	59980	0.3719	0.372
177	59978	0.3952	0.3705
136	59950	0.3804	0.3709

137	59948	0.3943	0.3706
135	59926	0.3797	0.3706
132	59952	0.371	0.3723
139	59908	0.3878	0.3708
216	59888	0.3952	0.3704
131	59848	0.3852	0.371
146	59838	0.3877	0.3706
129	59720	0.375	0.3714
141	59644	0.3738	0.3713
145	59614	0.3885	0.3708
144	59592	0.3847	0.371
140	59384	0.3796	0.372
173	59302	0.3944	0.3705
153	58556	0.3909	0.371
170	55116	0.3871	0.3712

K = 4

Plant ID	Group 1	Group 2	Group 3	Group 4
155	6652.1	18774	31809	0.2507
208	0.2503	0.2501	59952	0.2501
202	0.2507	0.2504	59874	0.2506
204	0.2507	0.2503	59860	0.2506
124	0.2503	0.2506	58960	0.2504
217	0.2503	0.2505	58428	0.2504
157	0.2502	0.2501	59680	0.2502
119	0.2503	0.2518	59922	0.2503
103	0.2512	0.251	59926	0.2518
219	0.2515	0.2508	59926	0.2514
92	0.2507	0.2503	59910	0.2505
118	0.2508	0.2502	59764	0.2503
88	0.2509	0.2503	59710	0.2505
100	0.2509	0.2504	59656	0.2506
102	0.2519	0.251	59570	0.2517
109	0.2515	0.2505	59428	0.2509
104	0.2515	0.2507	58752	0.2513
156	0.2512	0.2505	59954	0.251
91	0.2501	0.2501	59526	0.2501
125	0.2502	0.2501	59888	0.2502
126	0.2516	0.2509	59946	0.2518

112	13548	0.2502	46370	0.2501
93	59918	0.2501	0.2503	0.2501
107	59924	0.2502	0.2504	0.2502
187	59848	0.2501	0.2503	0.2501
183	59784	0.2501	0.252	0.2501
181	59631	0.2501	0.2512	1.4809
188	59598	0.2501	0.2518	0.2501
106	59456	0.2501	0.2503	0.2501
105	58920	0.2501	0.2504	0.2502
108	58740	0.2501	0.2503	0.2501
116	0.2505	59320	0.2503	0.2505
117	0.2505	58948	0.2502	0.2505
127	0.2507	58622	0.2502	0.2506
162	0.2506	57988	0.2504	0.2505
203	0.2504	58880	0.2503	0.2505
200	0.2503	59372	0.2501	0.2506
214	0.2503	59428	0.2501	0.2506
206	1499	53000	5335.5	0.2508
168	0.2502	57101	0.2503	2909
165	0.2502	56922	0.2503	3070.9
121	0.2502	57828	0.2503	2088.1
164	0.2502	58913	0.2503	991.4
80	0.2502	57810	0.2503	2096.2
163	0.2502	58391	0.2503	1475.4
81	0.2502	58654	0.2503	1166.8
84	0.2502	57142	0.2503	2664.3
75	0.2502	59572	0.2503	154.73
77	0.2502	58091	0.2503	1607.4
120	0.2502	58225	0.2503	1725.9
79	0.2502	56114	0.2503	3354.9
98	0.2502	57207	0.2503	2169.1
166	0.2502	54914	0.2503	3716.5
76	0.2502	56779	0.2503	1848
218	0.2501	53433	0.2512	6485.5
213	0.2508	51440	0.2517	8288.7
211	0.2501	46180	0.2518	13695
212	0.2501	45839	0.2516	14056
180	0.2502	43644	0.2517	15973
182	0.2503	42692	0.2503	17169
189	0.2502	41738	0.2503	18225

186	0.2502	41741	0.2637	17939
184	0.2502	39617	0.2503	20234
190	0.2501	39732	0.2522	18587
191	0.2502	38967	0.2522	20871
185	0.2502	38805	0.2503	21048
179	0.2502	38119	0.2521	21777
138	0.251	0.2504	0.2508	60036
150	0.2512	0.2503	0.2509	60030
148	0.2508	0.2504	0.2507	60030
151	0.2512	0.2503	0.2509	60026
134	0.2509	0.2503	0.2507	60028
133	0.2508	0.2504	0.2507	60024
169	0.251	0.2503	0.2508	60020
174	0.251	0.2503	0.2508	60016
1	0.2509	0.2503	0.2507	60012
215	0.2509	0.2504	0.2507	60014
149	0.2508	0.2504	0.2506	60012
143	0.2507	0.2504	0.2505	60006
130	0.2507	0.2504	0.2505	60004
147	0.2508	0.2504	0.2507	60002
142	0.2507	0.2504	0.2506	60002
152	0.2508	0.2504	0.2507	59988
128	0.2506	0.2504	0.2504	59980
177	0.2507	0.2504	0.2506	59978
136	0.2509	0.2503	0.2506	59950
137	0.251	0.2503	0.2508	59948
135	0.251	0.2503	0.2505	59926
132	0.2503	0.2505	0.2503	59952
139	0.2508	0.2504	0.2506	59908
216	0.2506	0.2505	0.2505	59888
131	0.2506	0.2504	0.2505	59848
146	0.2508	0.2504	0.2506	59838
129	0.2505	0.2504	0.2504	59720
141	0.2504	0.2505	0.2503	59644
145	0.2504	0.2505	0.2504	59614
144	0.2505	0.2505	0.2504	59592
140	0.2503	0.2505	0.2503	59384
173	0.2509	0.2503	0.2507	59302
153	0.2508	0.2504	0.2506	58556
170	0.2507	0.2504	0.2506	55116

K = 5

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5
155	0.2052	28890	2008.5	26336	0.2052
208	0.2038	59952	0.2035	0.204	0.2037
202	0.2033	59874	0.2032	0.2037	0.2032
204	0.2033	59860	0.2032	0.2037	0.2032
124	0.2052	58960	0.205	0.2053	0.2052
217	0.2052	58428	0.2049	0.2053	0.2052
157	0.2037	59680	0.2036	0.2041	0.2037
119	0.2029	59922	0.2028	0.2032	0.2028
103	0.203	59926	0.2031	0.2034	0.2029
219	0.2031	59926	0.203	0.2035	0.203
92	0.2034	59910	0.2032	0.2037	0.2033
118	0.2034	59764	0.2032	0.2037	0.2034
88	0.2034	59710	0.2031	0.2037	0.2033
100	0.2033	59656	0.2031	0.2036	0.2032
102	0.203	59570	0.2029	0.2034	0.203
109	0.2032	59428	0.203	0.2035	0.2031
104	0.2031	58752	0.203	0.2034	0.2031
156	0.2032	59954	0.2031	0.2036	0.2031
91	0.2042	59526	0.2046	0.2044	0.2041
125	0.2037	59888	0.2036	0.204	0.2036
126	0.203	59946	0.203	0.2034	0.203
112	0.2047	46377	13541	0.2052	0.2046
93	0.2038	0.2028	59918	0.204	0.2038
107	0.2036	0.2026	59924	0.2038	0.2035
187	0.2039	0.2028	59848	0.204	0.2038
183	0.2039	0.2029	59784	0.2041	0.2038
181	0.2041	0.2031	59632	0.2042	0.204
188	0.2039	0.203	59598	0.2041	0.2039
106	0.204	0.2028	59456	0.2041	0.2039
105	0.2037	0.2027	58920	0.2039	0.2036
108	0.2042	0.2028	58740	0.2043	0.2042
116	0.2052	0.2032	0.2039	59320	0.2052
117	0.2052	0.2031	0.2039	58948	0.2052
127	0.2053	0.2032	0.2039	58622	0.2053
162	0.2052	0.2032	0.2039	57988	0.2052
203	19160	0.2037	0.2042	39721	0.2057

200	31497	0.2037	0.2044	27875	0.2023
214	33527	0.2036	0.2044	25902	0.2052
206	31721	3137.7	0.2053	24976	0.2054
168	60010	0.2029	0.2037	0.2049	0.2054
165	59992	0.2029	0.2037	0.2049	0.2021
121	59916	0.2029	0.2037	0.2049	0.2054
164	59904	0.2028	0.2036	0.2049	0.2053
80	59906	0.2029	0.2037	0.2049	0.2062
163	59866	0.2028	0.2036	0.2049	0.2053
81	59820	0.2028	0.2036	0.2048	0.2053
84	59806	0.2029	0.2036	0.2049	0.2054
75	59726	0.2029	0.2037	0.2049	0.2053
77	59698	0.2028	0.2036	0.2038	0.2053
120	59950	0.2029	0.2037	0.2049	0.2053
79	59468	0.2028	0.2037	0.2049	0.2054
98	59376	0.2029	0.2037	0.2049	0.2054
166	58630	0.2028	0.2037	0.2049	0.2047
76	58626	0.2028	0.2037	0.2049	0.2054
218	57307	0.2033	0.2041	0.2052	2611.7
213	53343	0.2033	0.2042	0.2053	6385.8
211	49139	0.2031	0.204	0.2052	10735
212	48737	0.2032	0.204	0.2052	11157
180	46681	0.2033	0.2039	0.2053	12936
182	48014	0.2029	0.2035	0.2049	11847
189	47111	0.2029	0.2036	0.2049	12852
186	47058	0.2029	0.2036	0.2049	12623
184	44802	0.2028	0.2036	0.2049	15049
190	44552	0.203	0.2038	0.205	13766
191	43524	0.203	0.2037	0.205	16314
185	43487	0.2029	0.2036	0.205	16366
179	42221	0.2031	0.2038	0.2052	17675
138	0.2052	0.2024	0.2031	0.2046	60036
150	0.205	0.2023	0.2031	0.2046	60030
148	0.2052	0.2024	0.2032	0.2047	60030
151	0.205	0.2023	0.2031	0.2046	60026
134	0.2051	0.2024	0.2031	0.2046	60028
133	0.2052	0.2024	0.2032	0.2047	60024
169	0.2051	0.2024	0.2031	0.2046	60020
174	0.2051	0.2024	0.2031	0.2046	60016
1	0.2051	0.2024	0.2031	0.2046	60012

215	0.2052	0.2024	0.2032	0.2046	60014
149	0.2052	0.2025	0.2032	0.2046	60012
143	0.2052	0.2025	0.2032	0.2047	60006
130	0.2052	0.2026	0.2032	0.2047	60004
147	0.2052	0.2024	0.2032	0.2046	60002
142	0.2052	0.2025	0.2032	0.2047	60002
152	0.2051	0.2024	0.2032	0.2047	59988
128	0.2052	0.2026	0.2033	0.2047	59980
177	0.2052	0.2025	0.2032	0.2047	59978
136	0.2055	0.2025	0.2031	0.2046	59950
137	0.2262	0.2024	0.2031	0.2046	59948
135	0.2026	0.2025	0.2031	0.2046	59926
132	0.2052	0.2028	0.2034	0.2048	59952
139	0.2052	0.2025	0.2032	0.2046	59908
216	0.2052	0.2025	0.2033	0.2048	59888
131	0.2052	0.2026	0.2033	0.2047	59848
146	0.2052	0.2025	0.2032	0.2047	59838
129	0.2052	0.2026	0.2033	0.2047	59720
141	0.2053	0.2027	0.2033	0.2049	59644
145	0.2053	0.2026	0.2034	0.2049	59614
144	0.2053	0.2026	0.2033	0.2048	59592
140	0.2053	0.2027	0.2034	0.2049	59384
173	0.2051	0.2024	0.2031	0.2046	59302
153	0.2051	0.2025	0.2032	0.2047	58556
170	0.2052	0.2025	0.2032	0.2047	55116

$K = 6$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
155	28890	2008.4	0.1684	0.1678	0.1685	26336
208	59952	0.1678	0.1679	0.1676	0.1679	0.168
202	59874	0.1678	0.1678	0.1676	0.1678	0.1679
204	59860	0.1678	0.1678	0.1676	0.1678	0.1679
124	58960	0.1684	0.1684	0.1678	0.1685	0.1685
217	58428	0.1684	0.1684	0.1678	0.1684	0.1685
157	59680	0.1679	0.1679	0.1676	0.1679	0.168
119	59922	0.1676	0.1676	0.1675	0.1676	0.1678
103	59926	0.1677	0.1677	0.1675	0.1677	0.1678
219	59926	0.1677	0.1677	0.1675	0.1677	0.1678
92	59910	0.1678	0.1678	0.1676	0.1678	0.1679

118	59764	0.1677	0.1678	0.1676	0.1678	0.1679
88	59710	0.1677	0.1678	0.1675	0.1678	0.1679
100	59656	0.1677	0.1678	0.1675	0.1678	0.1679
102	59570	0.1677	0.1677	0.1675	0.1677	0.1678
109	59428	0.1677	0.1677	0.1675	0.1677	0.1679
104	58752	0.1677	0.1677	0.1675	0.1677	0.1678
156	59954	0.1677	0.1677	0.1675	0.1677	0.1679
91	59526	0.1682	0.168	0.1677	0.1681	0.1681
125	59888	0.1679	0.1679	0.1676	0.1679	0.168
126	59946	0.1677	0.1677	0.1675	0.1677	0.1678
112	46377	13542	0.1682	0.1679	0.1682	0.1683
93	0.1676	59918	0.1679	0.1675	0.168	0.168
107	0.1676	59924	0.1678	0.1674	0.1679	0.1679
187	0.1676	59848	0.1679	0.1675	0.168	0.168
183	0.1677	59784	0.1679	0.1675	0.168	0.168
181	0.1677	59632	0.168	0.1675	0.168	0.1681
188	0.1677	59598	0.168	0.1675	0.168	0.168
106	0.1676	59456	0.168	0.1675	0.168	0.168
105	0.1676	58920	0.1679	0.1674	0.1679	0.168
108	0.1676	58740	0.1681	0.1675	0.1681	0.1681
116	0.1677	0.168	0.1685	0.1674	0.1685	59320
117	0.1677	0.168	0.1685	0.1674	0.1685	58948
127	0.1677	0.168	0.1685	0.1675	0.1685	58622
162	0.1678	0.168	0.1685	0.1675	0.1685	57988
203	0.1679	0.1681	0.1685	0.1676	19159	39721
200	0.1679	0.1681	0.1686	0.1676	31497	27875
214	0.1679	0.1681	0.1685	0.1676	33527	25901
206	3137.6	0.1685	0.1686	0.1679	31721	24976
168	0.1676	0.1679	0.1685	0.1675	60010	0.1684
165	0.1676	0.1679	0.1686	0.1675	59992	0.1684
121	0.1677	0.1679	0.1685	0.1675	59916	0.1683
164	0.1676	0.1679	0.1685	0.1675	59904	0.1683
80	0.1676	0.1679	0.1685	0.1675	59906	0.1684
163	0.1676	0.1679	0.1685	0.1675	59866	0.1683
81	0.1676	0.1679	0.1685	0.1675	59820	0.1683
84	0.1676	0.1679	0.1686	0.1675	59806	0.1684
75	0.1676	0.1679	0.1685	0.1675	59726	0.1683
77	0.1676	0.1679	0.1685	0.1675	59698	0.1683
120	0.1676	0.1679	0.1685	0.1675	59950	0.1683
79	0.1676	0.1679	0.1681	0.1675	59468	0.1683

98	0.1676	0.1679	0.1686	0.1675	59376	0.1684
166	0.1676	0.1679	0.1668	0.1675	58630	0.1683
76	0.1676	0.1679	0.1685	0.1675	58626	0.1683
218	0.1678	0.168	2604.4	0.1676	57314	0.1684
213	0.1678	0.1681	6383.9	0.1676	53344	0.1685
211	0.1677	0.168	10734	0.1675	49141	0.1684
212	0.1677	0.168	11156	0.1675	48739	0.1684
180	0.1678	0.168	12934	0.1676	46682	0.1685
182	0.1676	0.1679	11845	0.1675	48015	0.1684
189	0.1677	0.1679	12850	0.1675	47112	0.1684
186	0.1677	0.1679	12621	0.1675	47059	0.1684
184	0.1676	0.1679	15047	0.1675	44803	0.1683
190	0.1677	0.1679	13765	0.1675	44554	0.1684
191	0.1677	0.1679	16313	0.1675	43525	0.1684
185	0.1677	0.1679	16365	0.1675	43488	0.1684
179	0.1677	0.1679	17674	0.1675	42222	0.1684
138	0.1675	0.1677	60036	0.1674	0.1684	0.1682
150	0.1675	0.1677	60030	0.1674	0.1684	0.1682
148	0.1675	0.1677	60030	0.1674	0.1684	0.1682
151	0.1675	0.1677	60026	0.1674	0.1684	0.1682
134	0.1675	0.1677	60028	0.1674	0.1684	0.1682
133	0.1675	0.1677	60024	0.1674	0.1684	0.1682
169	0.1675	0.1677	60020	0.1674	0.1684	0.1682
174	0.1675	0.1677	60016	0.1674	0.1684	0.1682
1	0.1675	0.1677	60012	0.1674	0.1684	0.1682
215	0.1675	0.1677	60014	0.1674	0.1684	0.1682
149	0.1675	0.1677	60012	0.1674	0.1684	0.1682
143	0.1675	0.1678	60006	0.1674	0.1685	0.1682
130	0.1675	0.1678	60004	0.1674	0.1685	0.1682
147	0.1675	0.1677	60002	0.1674	0.1684	0.1682
142	0.1675	0.1678	60002	0.1674	0.1685	0.1682
152	0.1675	0.1677	59988	0.1674	0.1684	0.1682
128	0.1676	0.1678	59980	0.1674	0.1684	0.1682
177	0.1675	0.1677	59978	0.1674	0.1685	0.1682
136	0.1675	0.1677	59950	0.1674	0.1684	0.1682
137	0.1675	0.1677	59948	0.1674	0.1684	0.1682
135	0.1675	0.1677	59926	0.1674	0.1684	0.1682
132	0.1676	0.1678	59952	0.1674	0.1685	0.1683
139	0.1675	0.1677	59908	0.1674	0.1684	0.1682
216	0.1675	0.1678	59888	0.1674	0.1685	0.1683

131	0.1675	0.1678	59848	0.1674	0.1685	0.1682
146	0.1675	0.1677	59838	0.1674	0.1684	0.1682
129	0.1676	0.1678	59720	0.1674	0.1685	0.1682
141	0.1676	0.1678	59644	0.1674	0.1685	0.1683
145	0.1676	0.1678	59614	0.1674	0.1685	0.1683
144	0.1676	0.1678	59592	0.1674	0.1685	0.1683
140	0.1676	0.1678	59384	0.1674	0.1685	0.1683
173	0.1675	0.1677	59302	0.1674	0.1684	0.1682
153	0.1675	0.1677	58556	0.1674	0.1684	0.1682
170	0.1675	0.1677	55116	0.1674	0.1684	0.1682

$K = 7$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
155	0.1433	26338	28890	0.1433	2006.3	0.1435	0.1435
208	0.1432	0.1433	59952	0.1432	0.1433	0.1433	0.1433
202	0.1432	0.1433	59874	0.1432	0.1432	0.1432	0.1432
204	0.1432	0.1433	59860	0.1432	0.1432	0.1432	0.1432
124	0.1433	0.1435	58960	0.1433	0.1434	0.1435	0.1435
217	0.1433	0.1435	58428	0.1433	0.1434	0.1435	0.1435
157	0.1432	0.1433	59680	0.1432	0.1433	0.1433	0.1433
119	0.1431	0.1432	59922	0.1431	0.1432	0.1432	0.1432
103	0.1432	0.1433	59926	0.1432	0.1432	0.1432	0.1432
219	0.1432	0.1433	59926	0.1432	0.1432	0.1432	0.1432
92	0.1432	0.1433	59910	0.1432	0.1432	0.1433	0.1432
118	0.1432	0.1433	59764	0.1432	0.1432	0.1433	0.1433
88	0.1432	0.1433	59710	0.1432	0.1432	0.1433	0.1432
100	0.1432	0.1433	59656	0.1432	0.1432	0.1432	0.1432
102	0.1432	0.1433	59570	0.1432	0.1432	0.1432	0.1432
109	0.1432	0.1433	59428	0.1432	0.1432	0.1432	0.1432
104	0.1432	0.1433	58752	0.1432	0.1432	0.1432	0.1432
156	0.1432	0.1433	59954	0.1432	0.1432	0.1432	0.1432
91	0.1432	0.1434	59526	0.1432	0.1434	0.1433	0.1433
125	0.1432	0.1433	59888	0.1432	0.1433	0.1433	0.1433
126	0.1432	0.1433	59946	0.1432	0.1432	0.1432	0.1432
112	0.1433	0.1434	46377	0.1433	13541	0.1434	0.1434
93	0.1431	0.1433	0.1432	0.1431	59918	0.1433	0.1433
107	0.1431	0.1433	0.1432	0.1431	59924	0.1433	0.1433
187	0.1431	0.1433	0.1432	0.1431	59848	0.1433	0.1433
183	0.1431	0.1433	0.1432	0.1431	59784	0.1433	0.1433

181	0.1432	0.1433	0.1432	0.1432	59632	0.1433	0.1433
188	0.1431	0.1433	0.1432	0.1431	59598	0.1433	0.1433
106	0.1431	0.1433	0.1432	0.1431	59456	0.1433	0.1433
105	0.1431	0.1433	0.1432	0.1431	58920	0.1433	0.1433
108	0.1432	0.1434	0.1432	0.1432	58740	0.1433	0.1433
116	0.1431	59320	0.1432	0.1431	0.1433	0.1435	0.1435
117	0.1431	58948	0.1432	0.1431	0.1433	0.1435	0.1435
127	0.1431	58622	0.1432	0.1431	0.1433	0.1435	0.1435
162	0.1431	57988	0.1432	0.1431	0.1433	0.1435	0.1435
203	0.1432	39721	0.1433	0.1432	0.1433	19159	0.1435
200	0.1432	27875	0.1433	0.1432	0.1434	31497	0.1435
214	0.1432	25901	0.1433	0.1432	0.1434	33527	0.1435
206	0.1433	24977	3137.2	0.1433	0.1435	31721	0.1435
168	0.1431	0.1434	0.1432	0.1431	0.1433	60010	0.1435
165	0.1431	0.1434	0.1432	0.1431	0.1433	59992	0.1435
121	0.1431	0.1434	0.1432	0.1431	0.1433	59916	0.1435
164	0.1431	0.1434	0.1432	0.1431	0.1433	59904	0.1435
80	0.1431	0.1434	0.1432	0.1431	0.1433	59906	0.1435
163	0.1431	0.1434	0.1432	0.1431	0.1433	59866	0.1435
81	0.1431	0.1434	0.1432	0.1431	0.1433	59820	0.1435
84	0.1431	0.1434	0.1432	0.1431	0.1433	59806	0.1435
75	0.1431	0.1434	0.1432	0.1431	0.1433	59726	0.1435
77	0.1431	0.1434	0.1432	0.1431	0.1433	59698	0.1435
120	0.1431	0.1434	0.1432	0.1431	0.1433	59950	0.1435
79	0.1431	0.1434	0.1432	0.1431	0.1433	59468	0.1435
98	0.1431	0.1434	0.1432	0.1431	0.1433	59376	0.1435
166	0.1431	0.1434	0.1432	0.1431	0.1433	58630	0.1435
76	0.1431	0.1434	0.1432	0.1431	0.1433	58626	0.1435
218	0.1432	0.1435	0.1432	0.1432	0.1433	57309	2609.2
213	0.1432	0.1435	0.1432	0.1432	0.1433	53344	6384
211	0.1432	0.1435	0.1432	0.1432	0.1433	49140	10734
212	0.1432	0.1435	0.1432	0.1432	0.1433	48738	11156
180	0.1432	0.1435	0.1432	0.1432	0.1433	46682	12935
182	0.1431	0.1434	0.1432	0.1431	0.1433	48015	11845
189	0.1431	0.1434	0.1432	0.1431	0.1433	47112	12850
186	0.1432	0.1434	0.1432	0.1432	0.1433	47059	12621
184	0.1431	0.1434	0.1432	0.1431	0.1433	44803	15047
190	0.1432	0.1435	0.1432	0.1432	0.1433	44553	13765
191	0.1432	0.1435	0.1432	0.1432	0.1433	43525	16313
185	0.1432	0.1435	0.1432	0.1432	0.1433	43488	16365

179	0.1432	0.1435	0.1432	0.1432	0.1433	42222	17674
138	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60036
150	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60030
148	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60030
151	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60026
134	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60028
133	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60024
169	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60020
174	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60016
1	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60012
215	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60014
149	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	60012
143	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	60006
130	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	60004
147	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	60002
142	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	60002
152	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59988
128	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59980
177	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59978
136	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59950
137	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	59948
135	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59926
132	0.1431	0.1434	0.1432	0.1431	0.1433	0.1435	59952
139	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59908
216	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59888
131	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59848
146	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59838
129	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59720
141	0.1431	0.1434	0.1432	0.1431	0.1433	0.1435	59644
145	0.1431	0.1434	0.1432	0.1431	0.1433	0.1435	59614
144	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	59592
140	0.1431	0.1434	0.1432	0.1431	0.1433	0.1435	59384
173	0.1431	0.1434	0.1431	0.1431	0.1432	0.1435	59302
153	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	58556
170	0.1431	0.1434	0.1432	0.1431	0.1432	0.1435	55116

$K = 8$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
155	28890	2008.4	26336	0.1252	0.1251	0.1252	0.1251	0.1251

208	59952	0.1251	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
202	59874	0.1251	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
204	59860	0.1251	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
124	58960	0.1252	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
217	58428	0.1252	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
157	59680	0.1252	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
119	59922	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
103	59926	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
219	59926	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
92	59910	0.1251	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
118	59764	0.1251	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
88	59710	0.1251	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
100	59656	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
102	59570	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
109	59428	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
104	58752	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
156	59954	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
91	59526	0.1252	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
125	59888	0.1252	0.1252	0.1252	0.1251	0.1251	0.1251	0.1251
126	59946	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251	0.1251
112	46377	13541	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
93	0.1251	59918	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
107	0.1251	59924	0.1252	0.1251	0.1251	0.1251	0.1251	0.1251
187	0.1251	59848	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
183	0.1251	59784	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
181	0.1251	59632	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
188	0.1251	59598	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
106	0.1251	59456	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
105	0.1251	58920	0.1252	0.1252	0.1251	0.1251	0.1251	0.1251
108	0.1251	58740	0.1252	0.1252	0.1251	0.1252	0.1251	0.1251
116	0.1251	0.1252	59320	0.1252	0.1251	0.1252	0.1251	0.1251
117	0.1251	0.1252	58948	0.1252	0.1251	0.1252	0.1251	0.1251
127	0.1251	0.1252	58622	0.1252	0.1251	0.1252	0.1251	0.1251
162	0.1251	0.1252	57988	0.1252	0.1251	0.1252	0.1251	0.1251
203	0.1252	0.1252	39720	19160	0.1251	0.1252	0.1251	0.1251
200	0.1252	0.1252	27874	31498	0.1251	0.1258	0.1251	0.1251
214	0.1252	0.1252	25901	33528	0.1251	0.1252	0.1251	0.1251
206	3137.7	0.1252	24974	31723	0.1252	0.1252	0.1252	0.1252
168	0.1251	0.1252	0.1252	60010	0.1251	0.1252	0.1251	0.1251
165	0.1251	0.1252	0.1252	59992	0.1251	0.1252	0.1251	0.1251

121	0.1251	0.1252	0.1252	59916	0.1251	0.1252	0.1251	0.1251
164	0.1251	0.1251	0.1252	59904	0.1251	0.1252	0.1251	0.1251
80	0.1251	0.1252	0.1252	59906	0.1251	0.1252	0.1251	0.1251
163	0.1251	0.1252	0.1252	59866	0.1251	0.1252	0.1251	0.1251
81	0.1251	0.1252	0.1252	59820	0.1251	0.1252	0.1251	0.1251
84	0.1251	0.1252	0.1252	59806	0.1251	0.1252	0.1251	0.1251
75	0.1251	0.1252	0.1252	59726	0.1251	0.1252	0.1251	0.1251
77	0.1251	0.1252	0.1252	59698	0.1251	0.1252	0.1251	0.1251
120	0.1251	0.1252	0.1252	59950	0.1251	0.1252	0.1251	0.1251
79	0.1251	0.1252	0.1252	59468	0.1251	0.1252	0.1251	0.1251
98	0.1251	0.1252	0.1252	59376	0.1251	0.1252	0.1251	0.1251
166	0.1251	0.1252	0.1252	58630	0.1251	0.1252	0.1251	0.1251
76	0.1251	0.1252	0.1252	58626	0.1251	0.1252	0.1251	0.1251
218	0.1251	0.1252	0.1252	57347	0.1251	2571.6	0.1251	0.1251
213	0.1251	0.1252	0.1252	53352	0.1251	6376.6	0.1251	0.1251
211	0.1251	0.1252	0.1252	49146	0.1251	10728	0.1251	0.1251
212	0.1251	0.1252	0.1252	48744	0.1251	11150	0.1251	0.1251
180	0.1251	0.1252	0.1252	46686	0.1251	12930	0.1251	0.1251
182	0.1251	0.1251	0.1252	48022	0.1251	11839	0.1251	0.1251
189	0.1251	0.1251	0.1252	47118	0.1251	12844	0.1251	0.1251
186	0.1251	0.1252	0.1252	47065	0.1251	12615	0.1251	0.1251
184	0.1251	0.1251	0.1252	44808	0.1251	15042	0.1251	0.1251
190	0.1251	0.1252	0.1252	44559	0.1251	13760	0.1251	0.1251
191	0.1251	0.1252	0.1252	43530	0.1251	16308	0.1251	0.1251
185	0.1251	0.1252	0.1252	43493	0.1251	16360	0.1251	0.1251
179	0.1251	0.1252	0.1252	42227	0.1251	17669	0.1251	0.1251
138	0.1251	0.1251	0.1252	0.1252	0.1251	60036	0.1251	0.1251
150	0.1251	0.1251	0.1252	0.1252	0.1251	60030	0.1251	0.1251
148	0.1251	0.1251	0.1252	0.1252	0.1251	60030	0.1251	0.1251
151	0.1251	0.1251	0.1252	0.1252	0.1251	60026	0.1251	0.1251
134	0.1251	0.1251	0.1252	0.1252	0.1251	60028	0.1251	0.1251
133	0.1251	0.1251	0.1252	0.1252	0.1251	60024	0.1251	0.1251
169	0.1251	0.1251	0.1252	0.1252	0.1251	60020	0.1251	0.1251
174	0.1251	0.1251	0.1252	0.1252	0.1251	60016	0.1251	0.1251
1	0.1251	0.1251	0.1252	0.1252	0.1251	60012	0.1251	0.1251
215	0.1251	0.1251	0.1252	0.1252	0.1251	60014	0.1251	0.1251
149	0.1251	0.1251	0.1252	0.1252	0.1251	60012	0.1251	0.1251
143	0.1251	0.1251	0.1252	0.1252	0.1251	60006	0.1251	0.1251
130	0.1251	0.1251	0.1252	0.1252	0.1251	60004	0.1251	0.1251
147	0.1251	0.1251	0.1252	0.1252	0.1251	60002	0.1251	0.1251

142	0.1251	0.1251	0.1252	0.1252	0.1251	60002	0.1251	0.1251
152	0.1251	0.1251	0.1252	0.1252	0.1251	59988	0.1251	0.1251
128	0.1251	0.1251	0.1252	0.1252	0.1251	59980	0.1251	0.1251
177	0.1251	0.1251	0.1252	0.1252	0.1251	59978	0.1251	0.1251
136	0.1251	0.1251	0.1252	0.1252	0.1251	59950	0.1251	0.1251
137	0.1251	0.1251	0.1252	0.1252	0.1251	59948	0.1251	0.1251
135	0.1251	0.1251	0.1252	0.1252	0.1251	59926	0.1251	0.1251
132	0.1251	0.1251	0.1252	0.1252	0.1251	59952	0.1251	0.1251
139	0.1251	0.1251	0.1252	0.1252	0.1251	59908	0.1251	0.1251
216	0.1251	0.1251	0.1252	0.1252	0.1251	59888	0.1251	0.1251
131	0.1251	0.1251	0.1252	0.1252	0.1251	59848	0.1251	0.1251
146	0.1251	0.1251	0.1252	0.1252	0.1251	59838	0.1251	0.1251
129	0.1251	0.1251	0.1252	0.1252	0.1251	59720	0.1251	0.1251
141	0.1251	0.1251	0.1252	0.1252	0.1251	59644	0.1251	0.1251
145	0.1251	0.1251	0.1252	0.1252	0.1251	59614	0.1251	0.1251
144	0.1251	0.1251	0.1252	0.1252	0.1251	59592	0.1251	0.1251
140	0.1251	0.1251	0.1252	0.1252	0.1251	59384	0.1251	0.1251
173	0.1251	0.1251	0.1252	0.1252	0.1251	59302	0.1251	0.1251
153	0.1251	0.1251	0.1252	0.1252	0.1251	58556	0.1251	0.1251
170	0.1251	0.1251	0.1252	0.1252	0.1251	55116	0.1251	0.1251

$K = 9$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9
155	0.1112	0.1112	0.1112	28890	0.1112	0.1112	0.1112	2003.7	26341
208	0.1112	0.1112	0.1112	59952	0.1112	0.1112	0.1112	0.1112	0.1112
202	0.1112	0.1112	0.1112	59874	0.1112	0.1112	0.1112	0.1112	0.1112
204	0.1112	0.1112	0.1112	59860	0.1112	0.1112	0.1112	0.1112	0.1112
124	0.1112	0.1112	0.1112	58960	0.1112	0.1112	0.1112	0.1112	0.1112
217	0.1112	0.1112	0.1112	58428	0.1112	0.1112	0.1112	0.1112	0.1112
157	0.1112	0.1112	0.1112	59680	0.1112	0.1112	0.1112	0.1112	0.1112
119	0.1111	0.1111	0.1112	59922	0.1111	0.1112	0.1111	0.1112	0.1112
103	0.1112	0.1112	0.1112	59926	0.1112	0.1112	0.1112	0.1112	0.1112
219	0.1112	0.1112	0.1112	59926	0.1112	0.1112	0.1112	0.1112	0.1112
92	0.1112	0.1112	0.1112	59910	0.1112	0.1112	0.1112	0.1112	0.1112
118	0.1112	0.1112	0.1112	59764	0.1112	0.1112	0.1112	0.1112	0.1112
88	0.1112	0.1112	0.1112	59710	0.1112	0.1112	0.1112	0.1112	0.1112
100	0.1112	0.1112	0.1112	59656	0.1112	0.1112	0.1112	0.1112	0.1112
102	0.1111	0.1111	0.1112	59570	0.1111	0.1112	0.1111	0.1112	0.1112
109	0.1112	0.1112	0.1112	59428	0.1112	0.1112	0.1112	0.1112	0.1112

104	0.1112	0.1112	0.1112	58752	0.1112	0.1112	0.1112	0.1112	0.1112
156	0.1112	0.1112	0.1112	59954	0.1112	0.1112	0.1112	0.1112	0.1112
91	0.1112	0.1112	0.1112	59526	0.1112	0.1112	0.1112	0.1112	0.1112
125	0.1112	0.1112	0.1112	59888	0.1112	0.1112	0.1112	0.1112	0.1112
126	0.1112	0.1112	0.1112	59946	0.1112	0.1112	0.1112	0.1112	0.1112
112	0.1112	0.1112	0.1112	46377	0.1112	0.1112	0.1112	13541	0.1112
93	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59918	0.1112
107	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59924	0.1112
187	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59848	0.1112
183	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59784	0.1112
181	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59632	0.1112
188	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59598	0.1112
106	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	59456	0.1112
105	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	58920	0.1112
108	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	58740	0.1112
116	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	0.1112	59320
117	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	0.1112	58948
127	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	0.1112	58622
162	0.1111	0.1111	0.1112	0.1112	0.1111	0.1112	0.1111	0.1112	57988
203	0.1112	0.1112	0.1112	0.1112	0.1112	19160	0.1112	0.1112	39720
200	0.1112	0.1112	0.1112	0.1112	0.1112	31498	0.1112	0.1112	27874
214	0.1112	0.1112	0.1112	0.1112	0.1112	33528	0.1112	0.1112	25900
206	0.1112	0.1112	0.1112	3136.2	0.1112	31720	0.1112	0.1112	24978
168	0.1111	0.1111	0.1112	0.1112	0.1111	60010	0.1111	0.1112	0.1112
165	0.1111	0.1111	0.1112	0.1112	0.1111	59992	0.1111	0.1112	0.1112
121	0.1111	0.1111	0.1112	0.1112	0.1111	59916	0.1111	0.1112	0.1112
164	0.1111	0.1111	0.1112	0.1112	0.1111	59904	0.1111	0.1112	0.1112
80	0.1111	0.1111	0.1112	0.1112	0.1111	59906	0.1111	0.1112	0.1112
163	0.1111	0.1111	0.1112	0.1112	0.1111	59866	0.1111	0.1112	0.1112
81	0.1111	0.1111	0.1112	0.1112	0.1111	59820	0.1111	0.1112	0.1112
84	0.1111	0.1111	0.1112	0.1112	0.1111	59806	0.1111	0.1112	0.1112
75	0.1111	0.1111	0.1112	0.1112	0.1111	59726	0.1111	0.1112	0.1112
77	0.1111	0.1111	0.1112	0.1112	0.1111	59698	0.1111	0.1112	0.1112
120	0.1111	0.1111	0.1112	0.1112	0.1111	59950	0.1111	0.1112	0.1112
79	0.1111	0.1111	0.1112	0.1112	0.1111	59468	0.1111	0.1112	0.1112
98	0.1111	0.1111	0.1112	0.1112	0.1111	59376	0.1111	0.1112	0.1112
166	0.1111	0.1111	0.1112	0.1112	0.1111	58630	0.1111	0.1112	0.1112
76	0.1111	0.1111	0.1112	0.1112	0.1111	58626	0.1111	0.1112	0.1112
218	0.1112	0.1112	2563	0.1112	0.1112	57355	0.1112	0.1112	0.1112
213	0.1112	0.1112	6375.3	0.1112	0.1112	53353	0.1112	0.1112	0.1112

211	0.1112	0.1112	10727	0.1112	0.1112	49147	0.1112	0.1112	0.1112
212	0.1112	0.1112	11149	0.1112	0.1112	48745	0.1112	0.1112	0.1112
180	0.1112	0.1112	12929	0.1112	0.1112	46687	0.1112	0.1112	0.1112
182	0.1111	0.1111	11837	0.1112	0.1111	48023	0.1111	0.1112	0.1112
189	0.1111	0.1111	12843	0.1112	0.1111	47119	0.1111	0.1112	0.1112
186	0.1111	0.1111	12614	0.1112	0.1111	47066	0.1111	0.1112	0.1112
184	0.1111	0.1111	15041	0.1112	0.1111	44809	0.1111	0.1112	0.1112
190	0.1112	0.1112	13759	0.1112	0.1112	44560	0.1112	0.1112	0.1112
191	0.1112	0.1112	16307	0.1112	0.1112	43531	0.1112	0.1112	0.1112
185	0.1111	0.1111	16359	0.1112	0.1111	43494	0.1111	0.1112	0.1112
179	0.1112	0.1112	17668	0.1112	0.1112	42228	0.1112	0.1112	0.1112
138	0.1111	0.1111	60036	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
150	0.1111	0.1111	60030	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
148	0.1111	0.1111	60030	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
151	0.1111	0.1111	60026	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
134	0.1111	0.1111	60028	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
133	0.1111	0.1111	60024	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
169	0.1111	0.1111	60020	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
174	0.1111	0.1111	60016	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
1	0.1111	0.1111	60012	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
215	0.1111	0.1111	60014	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
149	0.1111	0.1111	60012	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
143	0.1111	0.1111	60006	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
130	0.1111	0.1111	60004	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
147	0.1111	0.1111	60002	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
142	0.1111	0.1111	60002	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
152	0.1111	0.1111	59988	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
128	0.1111	0.1111	59980	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
177	0.1111	0.1111	59978	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
136	0.1111	0.1111	59950	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
137	0.1111	0.1111	59948	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
135	0.1111	0.1111	59926	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
132	0.1111	0.1111	59952	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
139	0.1111	0.1111	59908	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
216	0.1111	0.1111	59888	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
131	0.1111	0.1111	59848	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
146	0.1111	0.1111	59838	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
129	0.1111	0.1111	59720	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
141	0.1111	0.1111	59644	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
145	0.1111	0.1111	59614	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112

144	0.1111	0.1111	59592	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
140	0.1111	0.1111	59384	0.1112	0.1111	0.1112	0.1111	0.1112	0.1112
173	0.1111	0.1111	59302	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
153	0.1111	0.1111	58556	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112
170	0.1111	0.1111	55116	0.1111	0.1111	0.1112	0.1111	0.1112	0.1112

$K = 10$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
155	0.1	0.1	26340	0.1	0.1	2004.6	0.1	28890	0.1	0.1
208	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59952	0.1	0.1
202	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59874	0.1	0.1
204	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59860	0.1	0.1
124	0.1	0.1	0.1	0.1	0.1	0.1	0.1	58960	0.1	0.1
217	0.1	0.1	0.1	0.1	0.1	0.1	0.1	58428	0.1	0.1
157	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59680	0.1	0.1
119	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59922	0.1	0.1
103	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59926	0.1	0.1
219	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59926	0.1	0.1
92	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59910	0.1	0.1
118	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59764	0.1	0.1
88	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59710	0.1	0.1
100	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59656	0.1	0.1
102	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59570	0.1	0.1
109	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59428	0.1	0.1
104	0.1	0.1	0.1	0.1	0.1	0.1	0.1	58752	0.1	0.1
156	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59954	0.1	0.1
91	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59526	0.1	0.1
125	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59888	0.1	0.1
126	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59946	0.1	0.1
112	0.1	0.1	0.1	0.1	0.1	13541	0.1	46377	0.1	0.1
93	0.1	0.1	0.1	0.1	0.1	59918	0.1	0.1	0.1	0.1
107	0.1	0.1	0.1	0.1	0.1	59924	0.1	0.1	0.1	0.1
187	0.1	0.1	0.1	0.1	0.1	59848	0.1	0.1	0.1	0.1
183	0.1	0.1	0.1	0.1	0.1	59784	0.1	0.1	0.1	0.1
181	0.1	0.1	0.1	0.1	0.1	59632	0.1	0.1	0.1	0.1
188	0.1	0.1	0.1	0.1	0.1	59598	0.1	0.1	0.1	0.1
106	0.1	0.1	0.1	0.1	0.1	59456	0.1	0.1	0.1	0.1
105	0.1	0.1	0.1	0.1	0.1	58920	0.1	0.1	0.1	0.1
108	0.1	0.1	0.1	0.1	0.1	58740	0.1	0.1	0.1	0.1

116	0.1	0.1	59320	0.1	0.1	0.1	0.1	0.1	0.1	0.1
117	0.1	0.1	58948	0.1	0.1	0.1	0.1	0.1	0.1	0.1
127	0.1	0.1	58622	0.1	0.1	0.1	0.1	0.1	0.1	0.1
162	0.1	0.1	57988	0.1	0.1	0.1	0.1	0.1	0.1	0.1
203	0.1	0.1	39721	0.1	0.1	0.1	0.1	0.1	0.1	19159
200	0.1	0.1	27875	0.1	0.1	0.1	0.1	0.1	0.1	31497
214	0.1	0.1	25901	0.1	0.1	0.1	0.1	0.1	0.1	33527
206	0.1	0.1	24977	0.1	0.1	0.1	0.1	3136.8	0.1	31720
168	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	60010
165	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59992
121	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59916
164	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59904
80	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59906
163	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59866
81	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59820
84	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59806
75	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59726
77	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59698
120	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59950
79	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59468
98	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	59376
166	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	58630
76	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	58626
218	0.1	0.1	0.1	0.1	2604.1	0.1	0.1	0.1	0.1	57314
213	0.1	0.1	0.1	0.1	6382	0.1	0.1	0.1	0.1	53346
211	0.1	0.1	0.1	0.1	10732	0.1	0.1	0.1	0.1	49142
212	0.1	0.1	0.1	0.1	11154	0.1	0.1	0.1	0.1	48740
180	0.1	0.1	0.1	0.1	12933	0.1	0.1	0.1	0.1	46683
182	0.1	0.1	0.1	0.1	11844	0.1	0.1	0.1	0.1	48017
189	0.1	0.1	0.1	0.1	12849	0.1	0.1	0.1	0.1	47113
186	0.1	0.1	0.1	0.1	12620	0.1	0.1	0.1	0.1	47061
184	0.1	0.1	0.1	0.1	15046	0.1	0.1	0.1	0.1	44804
190	0.1	0.1	0.1	0.1	13764	0.1	0.1	0.1	0.1	44555
191	0.1	0.1	0.1	0.1	16312	0.1	0.1	0.1	0.1	43526
185	0.1	0.1	0.1	0.1	16364	0.1	0.1	0.1	0.1	43489
179	0.1	0.1	0.1	0.1	17673	0.1	0.1	0.1	0.1	42223
138	0.1	0.1	0.1	0.1	60036	0.1	0.1	0.1	0.1	0.1
150	0.1	0.1	0.1	0.1	60030	0.1	0.1	0.1	0.1	0.1
148	0.1	0.1	0.1	0.1	60030	0.1	0.1	0.1	0.1	0.1
151	0.1	0.1	0.1	0.1	60026	0.1	0.1	0.1	0.1	0.1

134	0.1	0.1	0.1	0.1	60028	0.1	0.1	0.1	0.1	0.1
133	0.1	0.1	0.1	0.1	60024	0.1	0.1	0.1	0.1	0.1
169	0.1	0.1	0.1	0.1	60020	0.1	0.1	0.1	0.1	0.1
174	0.1	0.1	0.1	0.1	60016	0.1	0.1	0.1	0.1	0.1
1	0.1	0.1	0.1	0.1	60012	0.1	0.1	0.1	0.1	0.1
215	0.1	0.1	0.1	0.1	60014	0.1	0.1	0.1	0.1	0.1
149	0.1	0.1	0.1	0.1	60012	0.1	0.1	0.1	0.1	0.1
143	0.1	0.1	0.1	0.1	60006	0.1	0.1	0.1	0.1	0.1
130	0.1	0.1	0.1	0.1	60004	0.1	0.1	0.1	0.1	0.1
147	0.1	0.1	0.1	0.1	60002	0.1	0.1	0.1	0.1	0.1
142	0.1	0.1	0.1	0.1	60002	0.1	0.1	0.1	0.1	0.1
152	0.1	0.1	0.1	0.1	59988	0.1	0.1	0.1	0.1	0.1
128	0.1	0.1	0.1	0.1	59980	0.1	0.1	0.1	0.1	0.1
177	0.1	0.1	0.1	0.1	59978	0.1	0.1	0.1	0.1	0.1
136	0.1	0.1	0.1	0.1	59950	0.1	0.1	0.1	0.1	0.1
137	0.1	0.1	0.1	0.1	59948	0.1	0.1	0.1	0.1	0.1
135	0.1	0.1	0.1	0.1	59926	0.1	0.1	0.1	0.1	0.1
132	0.1	0.1	0.1	0.1	59952	0.1	0.1	0.1	0.1	0.1
139	0.1	0.1	0.1	0.1	59908	0.1	0.1	0.1	0.1	0.1
216	0.1	0.1	0.1	0.1	59888	0.1	0.1	0.1	0.1	0.1
131	0.1	0.1	0.1	0.1	59848	0.1	0.1	0.1	0.1	0.1
146	0.1	0.1	0.1	0.1	59838	0.1	0.1	0.1	0.1	0.1
129	0.1	0.1	0.1	0.1	59720	0.1	0.1	0.1	0.1	0.1
141	0.1	0.1	0.1	0.1	59644	0.1	0.1	0.1	0.1	0.1
145	0.1	0.1	0.1	0.1	59614	0.1	0.1	0.1	0.1	0.1
144	0.1	0.1	0.1	0.1	59592	0.1	0.1	0.1	0.1	0.1
140	0.1	0.1	0.1	0.1	59384	0.1	0.1	0.1	0.1	0.1
173	0.1	0.1	0.1	0.1	59302	0.1	0.1	0.1	0.1	0.1
153	0.1	0.1	0.1	0.1	58556	0.1	0.1	0.1	0.1	0.1
170	0.1	0.1	0.1	0.1	55116	0.1	0.1	0.1	0.1	0.1

Table 2.S3. **Entire collection identity scores.** Output by FastStructure (varQ) at each level of K. Higher K values (>7) create groups with very low identity across all accessions.

	$K = 2$	
Plant ID	Group 1	Group 2
155	4998.6	1920.4
208	7068	1.0224
202	7036.1	0.9349
204	7028.1	0.9374
124	6997.8	1.1949
217	6891.8	1.1748
157	7008	0.9691
119	7080.1	0.888
103	7056.1	0.895
219	7036.1	0.8991
92	7040.1	0.9409
118	7088.1	0.9417
88	7046.1	0.9068
100	7052	0.9503
102	7050.1	0.88
109	7034.1	0.8753
104	6986.1	0.8986
156	7038.1	0.8962
91	7046	0.9528
125	7065.9	1.0831
126	7060.1	0.9031

112	6445.4	607.62
93	3808.9	3238.1
107	3759.5	3257.5
187	3769.1	3233.9
183	3707.1	3257.9
181	3754.7	3256.3
188	3710.2	3254.8
106	3501.4	3533.6
105	3666.6	3348.4
108	3442	3591
116	26.387	6972.6
117	65.939	6907.1
127	28.967	6958
162	71.807	6889.2
203	210.58	6764.4
200	178.09	6860.9
214	96.17	6940.8
206	1110.4	5966.6
168	0.7471	7082.3
165	0.7409	7086.3
121	0.7525	7080.2
164	0.7429	7042.3
80	0.7471	7094.3
163	0.7406	7080.3
81	0.7422	7058.3
84	0.747	7028.

	3	3
75	0.739	7028.3
77	0.7423	7054.3
120	0.7356	7078.3
79	0.7509	7078.2
98	0.7443	7054.3
166	0.7485	7002.3
76	0.7518	7070.2
213	0.7915	7068.2
211	0.7617	7060.2
212	0.7715	7042.2
180	0.7379	7070.3
182	0.7389	7088.3
189	0.7388	7092.3
186	0.7402	7074.3
184	0.7373	7082.3
190	0.7502	7090.2
191	0.7368	7088.3
185	0.7373	7080.3
179	0.7445	7084.3
138	0.7305	7094.3
150	0.7304	7092.3
148	0.7317	7094.3
151	0.7303	7092.3
134	0.7338	7094.3
133	0.731	7094.

	3	3
169	0.7329	7092.3
174	0.7327	7090.3
1	0.7328	7094.3
215	0.7312	7086.3
149	0.7304	7094.3
143	0.7335	7082.3
130	0.7318	7094.3
147	0.7311	7094.3
142	0.7317	7094.3
152	0.7307	7080.3
128	0.7574	7094.2
177	0.7318	7092.3
136	0.7305	7094.3
137	0.7302	7094.3
135	0.7304	7092.3
132	0.7316	7094.3
139	0.7506	7086.2
216	0.7359	7086.3
131	0.7321	7088.3
146	0.7302	7090.3
129	0.732	7086.3
141	0.7342	7090.3
145	0.7314	7088.3
144	0.7409	7092.3
140	0.755	7084.

	7	2
173	0.7328	7094.3
153	0.7311	7086.3
170	0.7314	7054.3
Calico	0.8331	6036.2
Columbian Rainbow	1125	5762
Pimento de Neyde	4262.5	1964.5
Sandia	0.8117	6206.2
sunrise eclipse	1189.9	5551.1
California Mild	1434.7	4388.3
Chili De Arbol	101.23	6551.8
Filius Blue	379.04	5444
Mayan Cobanero	1413.8	4711.2
Mulato	0.7402	6218.3
Pequin	241.47	6717.5
Sweet Chocolate	0.7584	6626.2
CM334	0.759	6980.2
chiltepin	0.8027	6726.2
Apple Pimento	1.2106	5739.8
Bulgarian Carrot 2	724.32	5942.7
Bulgarian Carrot	1.0897	6039.9
Bulls heart 2	1.186	6851.8

Buran	0.8399	6046.2
Dulcetta Orange	0.9906	6096
Feher Ozon Paprika	1514.8	4548.2
Marseilles Sweet Yellow	0.7451	6584.3
Piment Vegetarian	4569.7	1889.3
Succette de Provence	0.8684	5634.1
Banana Sweet	0.8193	6144.2
Big Red	57.834	6533.2
China Giant Sweet	0.7624	6022.2
Chinese Ching Choo	1113.1	4767.9
Chocolate Beauty	0.9918	6484
Peter Pepper Red	0.7645	6388.2
Riot	0.7828	6794.2
Mams Biber	2350.1	4294.9
Shishito	0.822	6286.2
Zunla	0.795	6776.2

	$K = 3$		
Plant ID	Group 1	Group 2	Group 3
155	4942.2	0.417	1976.4
208	7068.2	0.3911	0.4036
202	7036.2	0.3883	0.3983

204	7028.2	0.3884	0.3984
124	6998.2	0.4041	0.4178
217	6892.2	0.4017	0.4166
157	7008.2	0.3897	0.4007
119	7080.2	0.3847	0.3944
103	7056.2	0.387	0.3953
219	7036.2	0.3861	0.3947
92	7040.2	0.3862	0.3976
118	7088.2	0.3863	0.3979
88	7046.2	0.3852	0.3955
100	7052.2	0.3882	0.3985
102	7050.2	0.3835	0.3932
109	7034.2	0.384	0.3931
104	6986.2	0.3864	0.3951
156	7038.2	0.3867	0.3949
91	7046.2	0.3996	0.4013
125	7066.2	0.3933	0.4065
126	7060.2	0.3846	0.3947
112	5583.9	1468.7	0.4166
93	0.3868	7046.2	0.4053
107	0.3882	7016.2	0.4071
187	0.3886	7002.2	0.4076
183	0.389	6964.2	0.4095
181	0.3906	7010.2	0.4093
188	0.3878	6964.2	0.4088

106	0.3908	7034.2	0.4147
105	0.3882	7014.2	0.4078
108	0.3897	7032.2	0.4159
116	0.3944	2488.4	4510.2
117	0.3955	2535.9	4436.7
127	0.396	2319.8	4666.8
162	0.3957	2530.8	4429.8
203	0.4098	1685	5289.6
200	0.4078	1296.8	5741.8
214	0.4026	1080.7	5955.9
206	956.36	637.26	5483.4
168	0.3852	0.3929	7082.2
165	0.384	0.394	7086.2
121	0.3858	0.3939	7080.2
164	0.3845	0.3926	7042.2
80	0.3848	0.394	7094.2
163	0.384	0.3924	7080.2
81	0.3837	0.3956	7058.2
84	0.3858	0.3952	7028.2
75	0.3832	0.3959	7028.2
77	0.384	0.3928	7054.2
120	0.3835	0.3911	7078.2
79	0.3864	0.3938	7078.2
98	0.3861	0.3932	7054.2
166	0.3859	0.3945	7002.2

76	0.3857	0.399	7070.2
213	0.3917	0.4022	7068.2
211	0.3873	0.3962	7060.2
212	0.3884	0.3982	7042.2
180	0.3839	0.3912	7070.2
182	0.3841	0.3902	7088.2
189	0.3839	0.39	7092.2
186	0.3843	0.3901	7074.2
184	0.3837	0.3895	7082.2
190	0.3843	0.3987	7090.2
191	0.3835	0.3904	7088.2
185	0.3837	0.3893	7080.2
179	0.3848	0.3927	7084.2
138	0.3825	0.388	7094.2
150	0.3825	0.3878	7092.2
148	0.3827	0.3883	7094.2
151	0.3825	0.3878	7092.2
134	0.3833	0.3886	7094.2
133	0.3827	0.3881	7094.2
169	0.3829	0.3884	7092.2
174	0.3829	0.3883	7090.2
1	0.3829	0.3883	7094.2
215	0.3827	0.388	7086.2
149	0.3825	0.3879	7094.2
143	0.3833	0.3886	7082.2

130	0.3827	0.3882	7094.2
147	0.3826	0.388	7094.2
142	0.3827	0.3882	7094.2
152	0.3826	0.3879	7080.2
128	0.3862	0.3895	7094.2
177	0.3828	0.3883	7092.2
136	0.3825	0.3879	7094.2
137	0.3825	0.3877	7094.2
135	0.3825	0.3878	7092.2
132	0.3828	0.3883	7094.2
139	0.3853	0.389	7086.2
216	0.3838	0.3885	7086.2
131	0.3828	0.3882	7088.2
146	0.3825	0.388	7090.2
129	0.3828	0.3882	7086.2
141	0.3831	0.3886	7090.2
145	0.3827	0.3884	7088.2
144	0.3841	0.3882	7092.2
140	0.386	0.3884	7084.2
173	0.3829	0.3883	7094.2
153	0.3826	0.3879	7086.2
170	0.3826	0.3883	7054.2
Calico	0.3941	0.3935	6036.2
Columbian Rainbow	1223.7	0.409	5662.9

Pimento de Neyde	4226.4	0.4176	2000.2
Sandia	0.3931	0.3976	6206.2
sunrise eclipse	1285.9	0.409	5454.6
California Mild	1521.1	0.4053	4301.5
Chili De Arbol	109.97	0.4071	6542.6
Filius Blue	438.63	0.4124	5384
Mayan Cobanero	0.4052	2218.8	3905.8
Mulato	0.3857	0.3921	6218.2
Pequin	294.04	0.4193	6664.5
Sweet Chocolate	0.3861	0.3912	6626.2
CM334	0.387	0.3936	6980.2
chiltepin	0.3922	0.3949	6726.2
Apple Pimento	0.4195	0.4098	5740.2
Bulgarian Carrot 2	797.87	0.4149	5868.7
Bulgarian Carrot	0.4125	0.3987	6040.2
Bulls heart 2	0.4144	0.4004	6852.2
Buran	0.3953	0.3946	6046.2
Dulcetta Orange	0.406	0.4038	6096.2
Feher Ozon Paprika	1604	0.4129	4458.6
Marseilles Sweet Yellow	0.385	0.3907	6584.2

Piment Vegetarian	4513.6	0.4189	1945
Succette de Provence	0.3986	0.3992	5634.2
Banana Sweet	0.394	0.3927	6144.2
Big Red	83.548	0.4074	6507
China Giant Sweet	0.3859	0.3959	6022.2
Chinese Ching Choo	1069	168.93	4643.1
Chocolate Beauty	0.4064	0.3981	6484.2
Peter Pepper Red	0.3862	0.3973	6388.2
Riot	0.3909	0.3958	6794.2
Mams Biber	2453.7	0.4112	4190.9
Shishito	0.3959	0.3998	6286.2
Zunla	0.3913	0.3947	6776.2

K = 4

Plant ID	Group 1	Group 2	Group 3	Group 4
155	3937.3	2836.1	145.38	0.2511
208	0.2649	7068.2	0.2664	0.2647
202	0.2642	7036.2	0.2655	0.264
204	0.2642	7028.2	0.2655	0.264
124	0.2668	6998.2	0.2684	0.2669
217	0.2665	6892.2	0.2681	0.2664
157	0.2643	7008.2	0.2658	0.2646
119	0.2634	7080.2	0.265	0.2634
103	0.263	7056.	0.265	0.263

	9	2	1	6
219	0.2638	7036.2	0.2649	0.2633
92	0.2637	7040.2	0.2654	0.2639
118	0.2637	7088.2	0.2655	0.2639
88	0.2635	7046.2	0.2651	0.2636
100	0.264	7052.2	0.2656	0.2643
102	0.2631	7050.2	0.2647	0.2632
109	0.2633	7034.2	0.2647	0.2632
104	0.2637	6986.2	0.2651	0.2637
156	0.2639	7038.2	0.2649	0.2634
91	0.2665	7046.2	0.266	0.2646
125	0.2652	7066.2	0.2669	0.2654
126	0.2634	7060.2	0.2649	0.2634
112	1459.9	5592.6	0.2683	0.2668
93	7046.2	0.2641	0.2665	0.2649
107	7016.2	0.2643	0.2666	0.2649
187	7002.2	0.2644	0.2667	0.265
183	6964.2	0.2646	0.2669	0.2655
181	7010.2	0.2647	0.2671	0.2653
188	6964.2	0.2643	0.267	0.2653
106	7034.2	0.2649	0.268	0.266
105	7014.2	0.2643	0.2668	0.2652
108	7032.2	0.2646	0.268	0.266
116	0.2656	0.2641	0.2701	6998.2
117	0.2656	0.2641	0.27	6972.2
127	0.265	0.263	0.270	6986.

	4	9	8	2
162	0.2655	0.2639	0.27	6960.2
203	0.2665	0.2657	4260.7	2713.8
200	0.2664	0.2652	4854.7	2183.8
214	0.2671	0.2654	5132.1	1904.4
206	1071.6	333.42	4842.5	829.45
168	0.266	0.2642	7082.2	0.2685
165	0.2663	0.2639	7086.2	0.2675
121	0.266	0.2643	7080.2	0.2679
164	0.266	0.264	7042.2	0.2678
80	0.2663	0.2641	7094.2	0.2683
163	0.266	0.2639	7080.2	0.2678
81	0.2665	0.2638	7058.2	0.2674
84	0.2666	0.2645	7028.2	0.2682
75	0.2664	0.2637	7028.2	0.2679
77	0.266	0.2639	7054.2	0.2674
120	0.2657	0.2638	7078.2	0.2677
79	0.2662	0.2645	7078.2	0.2686
98	0.2663	0.2645	7054.2	0.2683
166	0.2665	0.2644	7002.2	0.2678
76	0.2671	0.2643	7070.2	0.2689
213	0.2677	0.2657	7068.2	0.2511
211	0.2666	0.2647	7060.2	0.2536
212	0.2668	0.2649	7042.2	0.2527
180	0.2657	0.2639	7070.2	0.2674
182	0.265	0.263	7088.	0.267

	5	9	2	6
189	0.2654	0.2638	7092.2	0.2674
186	0.2656	0.264	7074.2	0.2676
184	0.2653	0.2638	7082.2	0.2676
190	0.2668	0.2639	7090.2	0.2677
191	0.2655	0.2638	7088.2	0.2676
185	0.2653	0.2638	7080.2	0.2676
179	0.2659	0.264	7084.2	0.2674
138	0.2651	0.2635	7094.2	0.2667
150	0.265	0.2635	7092.2	0.2667
148	0.2651	0.2636	7094.2	0.2668
151	0.265	0.2635	7092.2	0.2667
134	0.2652	0.2637	7094.2	0.2669
133	0.2651	0.2635	7094.2	0.2668
169	0.2651	0.2636	7092.2	0.2668
174	0.2651	0.2636	7090.2	0.2668
1	0.2651	0.2636	7094.2	0.2668
215	0.2651	0.2635	7086.2	0.2667
149	0.265	0.2635	7094.2	0.2667
143	0.2652	0.2637	7082.2	0.2669
130	0.2651	0.2636	7094.2	0.2668
147	0.2651	0.2635	7094.2	0.2667
142	0.2651	0.2635	7094.2	0.2667
152	0.265	0.2635	7080.2	0.2667
128	0.2654	0.2642	7094.2	0.2667
177	0.265	0.263	7092.	0.266

	1	6	2	8
136	0.265	0.2635	7094.2	0.2667
137	0.265	0.2635	7094.2	0.2667
135	0.265	0.2635	7092.2	0.2667
132	0.2651	0.2636	7094.2	0.2668
139	0.2653	0.2642	7086.2	0.267
216	0.2651	0.264	7086.2	0.2676
131	0.2651	0.2636	7088.2	0.2667
146	0.2651	0.2635	7090.2	0.2667
129	0.2651	0.2636	7086.2	0.2668
141	0.2652	0.2636	7090.2	0.2669
145	0.2651	0.2635	7088.2	0.2667
144	0.2651	0.2638	7092.2	0.2668
140	0.2651	0.2643	7084.2	0.2669
173	0.2651	0.2636	7094.2	0.2668
153	0.265	0.2635	7086.2	0.2667
170	0.2651	0.2635	7054.2	0.2671
Calico	0.266	0.2661	6036.2	0.2741
Columbian Rainbow	948.78	575.97	5362	0.2512
Pimento de Neyde	3133.4	2592.4	500.9	0.2512
Sandia	0.2671	0.2657	6206.2	0.2682
sunrise eclipse	1015.8	580.46	5144.5	0.2512
California Mild	1085.8	744.59	3992.3	0.2512
Chili De Arbol	0.2698	87.251	6565.2	0.2649

Filius Blue	394.1	188.32	5240.3	0.2512
Mayan Cobanero	2246.8	0.2678	3877.6	0.2689
Mulato	0.2661	0.2644	6218.2	0.268
Pequin	298.76	116.46	6543.5	0.2512
Sweet Chocolate	0.2659	0.2643	6626.2	0.2674
CM334	0.2664	0.2646	6980.2	0.2676
chiltepin	0.2666	0.2656	6726.2	0.2533
Apple Pimento	0.2693	0.2704	5740.2	0.2695
Bulgarian Carrot 2	566.69	405.27	5694.8	0.2512
Bulgarian Carrot	0.2675	0.269	6040.2	0.269
Bulls heart 2	0.2678	0.2694	6852.2	0.2685
Buran	0.2667	0.2661	6046.2	0.2678
Dulcetta Orange	0.267	0.2676	6096.2	0.2675
Feher Ozon Paprika	1266.8	691.31	4104.5	0.4591
Marseilles Sweet Yellow	0.2657	0.264	6584.2	0.2672
Piment Vegetarian	3306.5	2878.6	273.49	0.4213
Succette de Provence	0.2673	0.2666	5634.2	0.2681
Banana Sweet	0.2669	0.2658	6144.2	0.2684
Big Red	0.2714	0.2743	6590.2	0.2757
China Giant Sweet	0.2665	0.2643	6022.2	0.2675

Chinese Ching Choo	909.18	561.99	4409.6	0.2512
Chocolate Beauty	0.2671	0.2679	6484.2	0.2685
Peter Pepper Red	0.2667	0.2643	6388.2	0.2682
Riot	0.2667	0.2656	6794.2	0.2687
Mams Biber	1847.4	1153.3	3644	0.2512
Shishito	0.2673	0.266	6286.2	0.2684
Zunla	0.2666	0.2654	6776.2	0.2524

$K = 5$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5
155	0.2051	0.2046	0.2027	6918.2	0.2051
208	7068.2	0.2046	0.2027	0.2044	0.205
202	7036.2	0.2044	0.2027	0.2043	0.2048
204	7028.2	0.2045	0.2027	0.2043	0.2048
124	6998.2	0.2055	0.2039	0.2054	0.2058
217	6892.2	0.2053	0.2034	0.2054	0.2057
157	7008.2	0.2045	0.2029	0.2045	0.205
119	7080.2	0.2042	0.2025	0.2042	0.2046
103	7056.2	0.2044	0.2025	0.2042	0.2047
219	7036.2	0.2043	0.2026	0.2042	0.2046
92	7040.2	0.2043	0.2025	0.2043	0.2048
118	7088.2	0.2043	0.2026	0.2042	0.2048
88	7046.2	0.2043	0.2025	0.2042	0.2047
100	7052.2	0.2044	0.2026	0.2042	0.2048
102	7050.	0.204	0.202	0.204	0.204

	2	2	5	2	6
109	7034.2	0.2042	0.2025	0.2042	0.2046
104	6986.2	0.2043	0.2025	0.2042	0.2046
156	7038.2	0.2043	0.2026	0.2042	0.2046
91	7046.2	0.205	0.2028	0.2045	0.205
125	7066.2	0.2047	0.2028	0.2044	0.2052
126	7060.2	0.2042	0.2026	0.2042	0.2046
112	5522.1	1530.3	0.2032	0.2048	0.2055
93	0.2042	7046.2	0.2025	0.2039	0.205
107	0.2043	7016.2	0.2025	0.204	0.2051
187	0.2043	7002.2	0.2025	0.204	0.2051
183	0.2044	6964.2	0.2026	0.2041	0.2052
181	0.2044	7010.2	0.2026	0.2041	0.2052
188	0.2043	6964.2	0.2026	0.204	0.2052
106	0.2045	7034.2	0.2026	0.2041	0.2054
105	0.2043	7014.2	0.2025	0.204	0.2051
108	0.2044	7032.2	0.2026	0.2041	0.2055
116	0.2046	0.2052	0.2041	2100.1	4898.3
117	0.2047	0.2052	0.2041	2139	4833.4
127	0.2047	0.2052	0.2044	1952	5034.4
162	0.2046	0.2052	0.2043	2152	4808.4
203	0.2053	0.2055	828.66	0.2051	6145.7
200	0.2054	0.2057	644.28	0.2053	6394.1
214	0.2053	0.2056	559.58	0.2052	6476.8
206	0.2054	0.2058	0.2054	1591.8	5484.6
168	0.204	0.204	0.202	0.204	7082.

	3	7	9	1	2
165	0.2042	0.2048	0.2028	0.2041	7086.2
121	0.2043	0.2047	0.2028	0.2041	7080.2
164	0.2042	0.2047	0.2028	0.204	7042.2
80	0.2043	0.2048	0.2029	0.2041	7094.2
163	0.2042	0.2047	0.2026	0.204	7080.2
81	0.2042	0.2049	0.2029	0.204	7058.2
84	0.2043	0.2048	0.2029	0.2041	7028.2
75	0.2042	0.2048	0.2029	0.204	7028.2
77	0.2042	0.2047	0.2027	0.204	7054.2
120	0.2042	0.2046	0.2027	0.204	7078.2
79	0.2043	0.2048	0.203	0.2043	7078.2
98	0.2043	0.2047	0.2028	0.2041	7054.2
166	0.2043	0.2048	0.2028	0.2041	7002.2
76	0.2043	0.205	0.2031	0.2043	7070.2
213	0.2046	0.2051	0.2035	0.2044	7068.2
211	0.2044	0.2049	0.2031	0.2042	7060.2
212	0.2045	0.2049	0.2033	0.2043	7042.2
180	0.2042	0.2047	0.2031	0.2041	7070.2
182	0.2042	0.2046	0.2026	0.204	7088.2
189	0.2042	0.2046	0.2026	0.204	7092.2
186	0.2042	0.2046	0.2026	0.204	7074.2
184	0.2042	0.2045	0.2027	0.204	7082.2
190	0.2042	0.205	0.2029	0.2041	7090.2
191	0.2042	0.2046	0.2028	0.2041	7088.2
185	0.204	0.204	0.202	0.204	7080.

	2	5	7		2
179	0.2043	0.2047	0.2028	0.2041	7084.2
138	0.2041	0.2045	0.2023	0.2039	7094.2
150	0.2041	0.2045	0.2023	0.2039	7092.2
148	0.2041	0.2045	0.2023	0.2039	7094.2
151	0.2041	0.2045	0.2023	0.2039	7092.2
134	0.2042	0.2045	0.2023	0.2039	7094.2
133	0.2041	0.2045	0.2023	0.2039	7094.2
169	0.2042	0.2045	0.2023	0.2039	7092.2
174	0.2042	0.2045	0.2023	0.2039	7090.2
1	0.2042	0.2045	0.2023	0.2039	7094.2
215	0.2041	0.2045	0.2023	0.2039	7086.2
149	0.2041	0.2045	0.2024	0.2039	7094.2
143	0.2042	0.2045	0.2024	0.2039	7082.2
130	0.2041	0.2045	0.2024	0.2039	7094.2
147	0.2041	0.2045	0.2023	0.2039	7094.2
142	0.2041	0.2045	0.2023	0.2039	7094.2
152	0.2041	0.2045	0.2023	0.2039	7080.2
128	0.2043	0.2046	0.2024	0.204	7094.2
177	0.2041	0.2045	0.2023	0.2039	7092.2
136	0.2041	0.2045	0.2023	0.2039	7094.2
137	0.2041	0.2045	0.2023	0.2039	7094.2
135	0.2041	0.2045	0.2023	0.2039	7092.2
132	0.2041	0.2045	0.2023	0.2039	7094.2
139	0.2043	0.2045	0.2024	0.204	7086.2
216	0.204	0.204	0.202	0.204	7086.

	2	5	5		2
131	0.2041	0.2045	0.2024	0.204	7088.2
146	0.2041	0.2045	0.2024	0.2039	7090.2
129	0.2041	0.2045	0.2023	0.2039	7086.2
141	0.2042	0.2045	0.2024	0.2039	7090.2
145	0.2041	0.2045	0.2025	0.2039	7088.2
144	0.2042	0.2045	0.2025	0.2039	7092.2
140	0.2043	0.2045	0.2024	0.204	7084.2
173	0.2042	0.2045	0.2023	0.2039	7094.2
153	0.2041	0.2045	0.2023	0.2039	7086.2
170	0.2041	0.2045	0.2023	0.2039	7054.2
Calico	0.2047	0.2047	0.203	0.2046	6036.2
Columbian Rainbow	0.2053	0.2051	0.2033	1398.8	5487.6
Pimento de Neyde	0.2054	0.2049	0.2028	6226.2	0.2055
Sandia	0.2047	0.205	0.2029	0.2045	6206.2
sunrise eclipse	0.2054	0.2052	0.2032	1412.8	5327.6
California Mild	0.2055	0.2051	0.2032	1679	4143.4
Chili De Arbol	0.2071	0.2058	0.2042	0.2089	6652.2
Filius Blue	0.2053	0.2052	0.2033	569.29	5253.1
Mayan Cobanero	0.2051	2368.5	0.2033	0.2048	3755.9
Mulato	0.2043	0.2047	0.2028	0.2042	6218.2
Pequin	0.2055	0.2057	0.2038	380.42	6578
Sweet Chocolate	0.2043	0.2046	0.2026	0.2043	6626.2

CM334	0.2044	0.2047	0.2028	0.2042	6980.2
chiltepin	0.2046	0.2048	0.203	0.2044	6726.2
Apple Pimento	0.2057	0.2054	0.2032	0.2056	5740.2
Bulgarian Carrot 2	0.2058	0.2054	0.2033	869.43	5797
Bulgarian Carrot	0.2055	0.205	0.2032	0.2059	6040.2
Bulls heart 2	0.2055	0.2051	0.2033	0.206	6852.2
Buran	0.2048	0.2048	0.2026	0.2046	6046.2
Dulcetta Orange	0.2052	0.2052	0.2028	0.205	6096.2
Feher Ozon Paprika	0.2054	0.2052	0.2031	1762.7	4299.6
Marseilles Sweet Yellow	0.2043	0.2046	0.2025	0.2041	6584.2
Piment Vegetarian	0.2054	0.2049	0.2028	6458.2	0.2054
Succette de Provence	0.2049	0.205	0.2029	0.2051	5634.2
Banana Sweet	0.2047	0.2048	0.2029	0.2051	6144.2
Big Red	0.2068	0.2058	0.2039	0.2079	6590.2
China Giant Sweet	0.2043	0.2049	0.2027	0.2042	6022.2
Chinese Ching Choo	0.2055	0.2059	0.2034	1293.4	4586.9
Chocolate Beauty	0.2052	0.205	0.2032	0.2054	6484.2
Peter Pepper Red	0.2043	0.2049	0.2028	0.2043	6388.2
Riot	0.2046	0.2048	0.203	0.2046	6794.2
Mams Biber	0.2055	0.2052	0.2031	2812.1	3832.3

Shishito	0.2048	0.205	0.2029	0.2047	6286.2
Zunla	0.2046	0.2048	0.203	0.2044	6776.2

K = 6

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
155	6918.2	0.1675	0.1677	0.1683	0.1682	0.1681
208	0.1682	0.1676	0.1678	0.1683	7068.2	0.1682
202	0.1682	0.1676	0.1678	0.1683	7036.2	0.1681
204	0.1682	0.1676	0.1678	0.1683	7028.2	0.1681
124	0.1685	0.168	0.1681	0.1686	6998.2	0.1685
217	0.1685	0.1678	0.168	0.1685	6892.2	0.1684
157	0.1682	0.1676	0.1678	0.1683	7008.2	0.1682
119	0.1681	0.1675	0.1677	0.1682	7080.2	0.1681
103	0.1681	0.1675	0.1677	0.1682	7056.2	0.1681
219	0.1681	0.1676	0.1677	0.1682	7036.2	0.1681
92	0.1681	0.1675	0.1677	0.1682	7040.2	0.1681
118	0.1681	0.1675	0.1677	0.1682	7088.2	0.1681
88	0.1681	0.1675	0.1677	0.1682	7046.2	0.1681
100	0.1681	0.1675	0.1677	0.1683	7052.2	0.1681
102	0.1681	0.1675	0.1677	0.1682	7050.2	0.1681
109	0.1681	0.1675	0.1677	0.1682	7034.2	0.1681
104	0.1681	0.1675	0.1677	0.1682	6986.2	0.1681
156	0.1681	0.1676	0.1677	0.1682	7038.2	0.1681
91	0.1682	0.1676	0.1678	0.1683	7046.2	0.1683
125	0.1682	0.1676	0.1678	0.1684	7066.2	0.1682
126	0.1681	0.1676	0.1677	0.1682	7060.2	0.1681

112	0.1683	0.1677	0.1679	0.1685	5522.1	1530.2
93	0.168	0.1675	0.1677	0.1683	0.1681	7046.2
107	0.168	0.1675	0.1677	0.1683	0.1681	7016.2
187	0.168	0.1675	0.1677	0.1684	0.1681	7002.2
183	0.1681	0.1675	0.1678	0.1684	0.1681	6964.2
181	0.1681	0.1676	0.1678	0.1684	0.1681	7010.2
188	0.168	0.1675	0.1678	0.1684	0.1681	6964.2
106	0.1681	0.1675	0.1678	0.1684	0.1681	7034.2
105	0.168	0.1675	0.1678	0.1684	0.1681	7014.2
108	0.1681	0.1675	0.1678	0.1685	0.1681	7032.2
116	0.1682	0.1678	1885.1	5113.2	0.1682	0.1683
117	0.1682	0.1678	1934.6	5037.8	0.1682	0.1683
127	0.1682	0.1678	1712.8	5273.5	0.1682	0.1684
162	0.1682	0.1678	1968.7	4991.6	0.1682	0.1683
203	0.1684	828.99	0.1682	6145.3	0.1684	0.1685
200	0.1684	645.25	0.1682	6393.1	0.1684	0.1685
214	0.1684	560.61	0.1682	6475.7	0.1684	0.1685
206	1704.3	0.1683	0.1682	5372	0.1684	0.1685
168	0.1681	0.1677	0.1679	7082.2	0.1681	0.1682
165	0.1681	0.1676	0.1679	7086.2	0.1681	0.1682
121	0.1681	0.1676	0.1679	7080.2	0.1681	0.1682
164	0.168	0.1676	0.1679	7042.2	0.1681	0.1682
80	0.1681	0.1677	0.1679	7094.2	0.1681	0.1682
163	0.168	0.1676	0.1678	7080.2	0.1681	0.1682
81	0.168	0.1676	0.1679	7058.2	0.1681	0.1683

84	0.1681	0.1676	0.1679	7028.2	0.1681	0.1683
75	0.168	0.1676	0.1679	7028.2	0.1681	0.1683
77	0.168	0.1676	0.1678	7054.2	0.1681	0.1682
120	0.168	0.1676	0.1678	7078.2	0.1681	0.1682
79	0.1681	0.1677	0.1679	7078.2	0.1681	0.1682
98	0.1681	0.1676	0.1679	7054.2	0.1681	0.1682
166	0.1681	0.1676	0.1679	7002.2	0.1681	0.1683
76	0.1681	0.1677	0.168	7070.2	0.1681	0.1683
213	0.1682	0.1679	0.168	7068.2	0.1682	0.1683
211	0.1681	0.1677	0.1679	7060.2	0.1681	0.1683
212	0.1681	0.1678	0.1679	7042.2	0.1681	0.1683
180	0.1681	0.1677	0.1679	7070.2	0.1681	0.1682
182	0.168	0.1676	0.1678	7088.2	0.1681	0.1682
189	0.168	0.1676	0.1678	7092.2	0.1681	0.1682
186	0.168	0.1676	0.1678	7074.2	0.1681	0.1682
184	0.168	0.1676	0.1679	7082.2	0.1681	0.1682
190	0.1681	0.1677	0.1679	7090.2	0.1681	0.1683
191	0.168	0.1676	0.1679	7088.2	0.1681	0.1682
185	0.168	0.1676	0.1679	7080.2	0.1681	0.1682
179	0.1681	0.1676	0.1679	7084.2	0.1681	0.1682
138	0.168	0.1675	0.1678	7094.2	0.168	0.1682
150	0.168	0.1675	0.1678	7092.2	0.168	0.1681
148	0.168	0.1675	0.1678	7094.2	0.168	0.1682
151	0.168	0.1675	0.1678	7092.2	0.168	0.1681
134	0.168	0.1675	0.1678	7094.2	0.1681	0.1682

133	0.168	0.1675	0.1678	7094.2	0.168	0.1682
169	0.168	0.1675	0.1678	7092.2	0.1681	0.1682
174	0.168	0.1675	0.1678	7090.2	0.1681	0.1682
1	0.168	0.1675	0.1678	7094.2	0.1681	0.1682
215	0.168	0.1675	0.1678	7086.2	0.168	0.1682
149	0.168	0.1675	0.1678	7094.2	0.168	0.1682
143	0.168	0.1675	0.1678	7082.2	0.1681	0.1682
130	0.168	0.1675	0.1678	7094.2	0.168	0.1682
147	0.168	0.1675	0.1678	7094.2	0.168	0.1682
142	0.168	0.1675	0.1678	7094.2	0.168	0.1682
152	0.168	0.1675	0.1678	7080.2	0.168	0.1682
128	0.168	0.1675	0.1678	7094.2	0.1681	0.1682
177	0.168	0.1675	0.1678	7092.2	0.168	0.1682
136	0.168	0.1675	0.1678	7094.2	0.168	0.1682
137	0.168	0.1675	0.1678	7094.2	0.168	0.1681
135	0.168	0.1675	0.1678	7092.2	0.168	0.1681
132	0.168	0.1675	0.1678	7094.2	0.168	0.1682
139	0.168	0.1675	0.1678	7086.2	0.1681	0.1682
216	0.168	0.1675	0.1678	7086.2	0.1681	0.1682
131	0.168	0.1675	0.1678	7088.2	0.168	0.1682
146	0.168	0.1675	0.1678	7090.2	0.168	0.1682
129	0.168	0.1675	0.1678	7086.2	0.168	0.1682
141	0.168	0.1675	0.1678	7090.2	0.1681	0.1682
145	0.168	0.1675	0.1678	7088.2	0.168	0.1682
144	0.168	0.1675	0.1678	7092.2	0.1681	0.1682

140	0.168	0.1675	0.1678	7084.2	0.1681	0.1682
173	0.168	0.1675	0.1678	7094.2	0.1681	0.1682
153	0.168	0.1675	0.1678	7086.2	0.168	0.1682
170	0.168	0.1675	0.1678	7054.2	0.168	0.1682
Calico	0.1682	0.1677	0.1679	6036.2	0.1682	0.1682
Columbian Rainbow	1492.1	0.1678	0.168	5394.2	0.1684	0.1683
Pimento de Neyde	6226.2	0.1676	0.1678	0.1684	0.1683	0.1682
Sandia	0.1682	0.1677	0.1679	6206.2	0.1682	0.1683
sunrise eclipse	1502.2	0.1677	0.1679	5238.1	0.1684	0.1684
California Mild	1787.1	0.1677	0.1679	4035.2	0.1684	0.1683
Chili De Arbol	203.74	0.1677	0.168	6448.6	0.1684	0.1684
Filius Blue	615.89	0.1678	0.168	5206.4	0.1684	0.1684
Mayan Cobanero	0.1683	0.1678	0.168	3755.7	0.1684	2368.6
Mulato	0.1681	0.1676	0.1679	6218.2	0.1681	0.1682
Pequin	422.97	0.1679	0.168	6535.4	0.1684	0.1685
Sweet Chocolate	0.1681	0.1676	0.1679	6626.2	0.1681	0.1682
CM334	0.1681	0.1676	0.1679	6980.2	0.1681	0.1682
chiltepin	0.1681	0.1677	0.168	6726.2	0.1682	0.1683
Apple Pimento	0.1686	0.1678	0.168	5740.2	0.1685	0.1684
Bulgarian Carrot 2	939.74	0.1678	0.168	5726.6	0.1685	0.1684
Bulgarian Carrot	0.1686	0.1678	0.168	6040.2	0.1685	0.1683
Bulls heart 2	0.1687	0.1678	0.168	6852.2	0.1685	0.1683

Buran	0.1682	0.1676	0.1678	6046.2	0.1682	0.1683
Dulcetta Orange	0.1684	0.1676	0.1679	6096.2	0.1684	0.1684
Feher Ozon Paprika	1877.3	0.1677	0.1679	4185	0.1684	0.1684
Marseilles Sweet Yellow	0.1681	0.1675	0.1678	6584.2	0.1681	0.1682
Piment Vegetarian	6458.2	0.1676	0.1678	0.1683	0.1684	0.1682
Succette de Provence	0.1684	0.1677	0.1679	5634.2	0.1683	0.1683
Banana Sweet	0.1684	0.1676	0.1679	6144.2	0.1682	0.1682
Big Red	159.96	0.1677	0.1679	6430.4	0.1684	0.1684
China Giant Sweet	0.1681	0.1676	0.1679	6022.2	0.1681	0.1683
Chinese Ching Choo	1378.2	0.1678	0.168	4502.1	0.1684	0.1685
Chocolate Beauty	0.1685	0.1678	0.168	6484.2	0.1684	0.1683
Peter Pepper Red	0.1681	0.1676	0.1679	6388.2	0.1681	0.1683
Riot	0.1682	0.1677	0.168	6794.2	0.1682	0.1683
Mams Biber	2968	0.1677	0.1679	3676.3	0.1684	0.1683
Shishito	0.1683	0.1677	0.1679	6286.2	0.1683	0.1683
Zunla	0.1681	0.1677	0.168	6776.2	0.1682	0.1682

$K = 7$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
155	0.1432	0.1434	0.1434	0.1434	6918.1	0.1431	0.1434
208	0.1433	0.1434	7068.1	0.1434	0.1434	0.1432	0.1434
202	0.1432	0.1434	7036.1	0.1434	0.1434	0.1432	0.1434

204	0.1432	0.1434	7028.1	0.1434	0.1434	0.1432	0.1434
124	0.1434	0.1435	6998.1	0.1435	0.1435	0.1433	0.1435
217	0.1433	0.1435	6892.1	0.1435	0.1435	0.1432	0.1435
157	0.1433	0.1434	7008.1	0.1434	0.1434	0.1432	0.1434
119	0.1432	0.1434	7080.1	0.1434	0.1434	0.1431	0.1433
103	0.1432	0.1434	7056.1	0.1434	0.1434	0.1431	0.1434
219	0.1432	0.1434	7036.1	0.1434	0.1434	0.1431	0.1434
92	0.1432	0.1434	7040.1	0.1434	0.1434	0.1431	0.1434
118	0.1432	0.1434	7088.1	0.1434	0.1434	0.1431	0.1434
88	0.1432	0.1434	7046.1	0.1434	0.1434	0.1431	0.1433
100	0.1432	0.1434	7052.1	0.1434	0.1434	0.1431	0.1434
102	0.1432	0.1434	7050.1	0.1434	0.1434	0.1431	0.1433
109	0.1432	0.1434	7034.1	0.1434	0.1434	0.1431	0.1433
104	0.1432	0.1434	6986.1	0.1434	0.1434	0.1431	0.1434
156	0.1432	0.1434	7038.1	0.1434	0.1434	0.1431	0.1434
91	0.1433	0.1434	7046.1	0.1434	0.1434	0.1432	0.1434
125	0.1433	0.1434	7066.1	0.1434	0.1434	0.1432	0.1434
126	0.1432	0.1434	7060.1	0.1434	0.1434	0.1431	0.1433
112	0.1433	0.1435	5521.7	0.1435	0.1434	0.1432	1530.6
93	0.1432	0.1434	0.1433	0.1434	0.1433	0.1431	7046.1
107	0.1432	0.1434	0.1433	0.1434	0.1433	0.1431	7016.1
187	0.1432	0.1434	0.1434	0.1434	0.1433	0.1431	7002.1
183	0.1433	0.1434	0.1434	0.1434	0.1433	0.1431	6964.1
181	0.1433	0.1434	0.1434	0.1434	0.1433	0.1431	7010.1
188	0.1433	0.1434	0.1434	0.1434	0.1433	0.1431	6964.1

106	0.1433	0.1435	0.1434	0.1435	0.1433	0.1431	7034.1
105	0.1432	0.1434	0.1434	0.1434	0.1433	0.1431	7014.1
108	0.1433	0.1435	0.1434	0.1435	0.1433	0.1431	7032.1
116	1926	955.95	0.1434	4116.4	0.1434	0.1432	0.1434
117	1976.4	766.7	0.1434	4229.3	0.1434	0.1432	0.1434
127	1743.3	1270.4	0.1434	3972.7	0.1434	0.1432	0.1434
162	1995.3	416.9	0.1434	4548.2	0.1434	0.1432	0.1434
203	1197.5	0.1435	0.1434	5776.7	0.1434	0.1432	0.1434
200	899.12	0.1435	0.1434	6139.2	0.1434	0.1433	0.1435
214	757.7	0.1435	0.1434	6278.6	0.1434	0.1433	0.1435
206	0.1435	0.1435	0.1435	5509.6	1566.7	0.1433	0.1435
168	0.1433	0.1431	0.1433	7082.1	0.1433	0.1431	0.1434
165	0.1433	0.1435	0.1433	7086.1	0.1433	0.1431	0.1434
121	0.1433	0.1435	0.1433	7080.1	0.1433	0.1431	0.1434
164	0.1433	0.1435	0.1433	7042.1	0.1433	0.1431	0.1434
80	0.1433	0.1435	0.1433	7094.1	0.1433	0.1431	0.1434
163	0.1433	0.1435	0.1433	7080.1	0.1433	0.1431	0.1434
81	0.1433	0.1435	0.1433	7058.1	0.1433	0.1431	0.1434
84	0.1433	0.1435	0.1433	7028.1	0.1433	0.1431	0.1434
75	0.1433	0.1435	0.1433	7028.1	0.1433	0.1431	0.1434
77	0.1433	0.1435	0.1433	7054.1	0.1433	0.1431	0.1434
120	0.1433	0.1435	0.1433	7078.1	0.1433	0.1431	0.1434
79	0.1433	0.1435	0.1433	7078.1	0.1433	0.1432	0.1434
98	0.1433	0.1435	0.1433	7054.1	0.1433	0.1431	0.1434
166	0.1433	0.1435	0.1433	7002.1	0.1433	0.1431	0.1434

76	0.1433	0.1435	0.1433	7070.1	0.1433	0.1432	0.1434
213	0.1433	0.1435	0.1434	7068.1	0.1434	0.1432	0.1434
211	0.1433	0.1435	0.1434	7060.1	0.1433	0.1431	0.1434
212	0.1433	0.1435	0.1434	7042.1	0.1433	0.1432	0.1434
180	0.1433	585.81	0.1433	6484.5	0.1433	0.1432	0.1434
182	0.1433	0.1437	0.1433	7088.1	0.1433	0.1431	0.1434
189	0.1433	414.57	0.1433	6677.7	0.1433	0.1431	0.1434
186	0.1433	0.1435	0.1433	7074.1	0.1433	0.1431	0.1434
184	0.1433	0.1435	0.1433	7082.1	0.1433	0.1431	0.1434
190	0.1433	364.62	0.1433	6725.7	0.1433	0.1431	0.1434
191	0.1433	804.46	0.1433	6283.8	0.1433	0.1431	0.1434
185	0.1433	515.17	0.1433	6565.1	0.1433	0.1431	0.1434
179	0.1433	660	0.1433	6424.3	0.1433	0.1431	0.1434
138	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
150	0.1432	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
148	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
151	0.1432	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
134	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
133	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
169	0.1432	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
174	0.1432	7090.1	0.1433	0.1435	0.1433	0.1431	0.1434
1	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
215	0.1433	7086.1	0.1433	0.1435	0.1433	0.1431	0.1434
149	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
143	0.1433	7082.1	0.1433	0.1435	0.1433	0.1431	0.1434

130	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
147	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
142	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
152	0.1433	7080.1	0.1433	0.1435	0.1433	0.1431	0.1434
128	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
177	0.1433	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
136	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
137	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
135	0.1432	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
132	0.1433	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
139	0.1433	7086.1	0.1433	0.1435	0.1433	0.1431	0.1434
216	0.1433	7086.1	0.1433	0.1435	0.1433	0.1431	0.1434
131	0.1433	7088.1	0.1433	0.1435	0.1433	0.1431	0.1434
146	0.1433	7090.1	0.1433	0.1435	0.1433	0.1431	0.1434
129	0.1433	7086.1	0.1433	0.1435	0.1433	0.1431	0.1434
141	0.1433	7090.1	0.1433	0.1435	0.1433	0.1431	0.1434
145	0.1433	7088.1	0.1433	0.1435	0.1433	0.1431	0.1434
144	0.1433	7092.1	0.1433	0.1435	0.1433	0.1431	0.1434
140	0.1433	7084.1	0.1434	0.1435	0.1433	0.1431	0.1434
173	0.1432	7094.1	0.1433	0.1435	0.1433	0.1431	0.1434
153	0.1432	7086.1	0.1433	0.1435	0.1433	0.1431	0.1434
170	0.1433	7054.1	0.1433	0.1435	0.1433	0.1431	0.1434
Calico	0.1433	660.55	0.1434	5375.7	0.1434	0.1431	0.1434
Columbian Rainbow	0.1433	1194.5	0.1434	4234.2	1457.7	0.1432	0.1434

Pimento de Neyde	0.1433	0.1434	0.1434	0.1434	6226.1	0.1431	0.1434
Sandia	0.1433	548.04	0.1434	5658.2	0.1434	0.1431	0.1434
sunrise eclipse	0.1433	1555.7	0.1435	3696.7	1488	0.1432	0.1434
California Mild	0.1433	1366	0.1435	2676.9	1779.6	0.1432	0.1434
Chili De Arbol	0.1433	1692.9	0.1434	4794.2	165.39	0.1432	0.1434
Filius Blue	0.1433	807.92	0.1434	4451.1	563.45	0.1432	0.1434
Mayan Cobanero	0.1433	0.1435	0.1434	3825.2	0.1434	0.1432	2299.1
Mulato	0.1433	2022.7	0.1434	4195.6	0.1433	0.1431	0.1434
Pequin	0.1433	0.1435	0.1435	6627.9	330.35	0.1432	0.1435
Sweet Chocolate	0.1433	2831.6	0.1434	3794.7	0.1434	0.1431	0.1434
CM334	0.1433	683.42	0.1434	6296.9	0.1433	0.1431	0.1434
chiltepin	0.1433	343.96	0.1434	6382.3	0.1433	0.1432	0.1434
Apple Pimento	0.1433	1787.4	0.1435	3952.9	0.1435	0.1432	0.1434
Bulgarian Carrot 2	0.1433	1682	0.1435	4079.8	904.62	0.1432	0.1435
Bulgarian Carrot	0.1433	2110	0.1434	3930.3	0.1435	0.1432	0.1434
Bulls heart 2	0.1433	1247.6	0.1434	5604.7	0.1435	0.1432	0.1434
Buran	0.1433	2777	0.1434	3269.3	0.1434	0.1431	0.1434
Dulcetta Orange	0.1433	2597.6	0.1434	3498.7	0.1434	0.1432	0.1434
Feher Ozon Paprika	0.1433	1867.6	0.1435	2310.2	1884.6	0.1432	0.1434
Marseilles Sweet Yellow	0.1433	1839.9	0.1433	4744.4	0.1433	0.1431	0.1434

Piment Vegetarian	0.1432	0.1434	0.1434	0.1434	6458.1	0.1431	0.1434
Succette de Provence	0.1433	1825.8	0.1434	3808.5	0.1434	0.1432	0.1434
Banana Sweet	0.1433	1662.9	0.1434	4481.4	0.1434	0.1432	0.1434
Big Red	0.1433	1742.2	0.1435	4735.5	112.68	0.1432	0.1434
China Giant Sweet	0.1433	2533.2	0.1434	3489.1	0.1434	0.1431	0.1434
Chinese Ching Choo	0.1433	842.59	0.1435	3692.3	1345.6	0.1432	0.1435
Chocolate Beauty	0.1433	1454.2	0.1434	5030.1	0.1434	0.1432	0.1434
Peter Pepper Red	0.1433	137.28	0.1433	6251	0.1433	0.1431	0.1434
Riot	0.1433	328.7	0.1434	6465.6	0.1434	0.1432	0.1434
Mams Biber	0.1433	1234.2	0.1435	2432.9	2977.3	0.1432	0.1434
Shishito	0.1433	0.1435	0.1434	6286.1	0.1434	0.1431	0.1434
Zunla	0.1433	447.09	0.1434	6329.2	0.1433	0.1432	0.1434

K = 8

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
155	0.1252	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	6918.1
208	0.1252	0.1251	0.1251	0.1252	0.1251	7068.1	0.1251	0.1252
202	0.1252	0.1251	0.1251	0.1252	0.1251	7036.1	0.1251	0.1252
204	0.1252	0.1251	0.1251	0.1252	0.1251	7028.1	0.1251	0.1252
124	0.1252	0.1252	0.1252	0.1252	0.1252	6998.1	0.1252	0.1252
217	0.1252	0.1251	0.1252	0.1252	0.1251	6892.1	0.1251	0.1252
157	0.1252	0.1251	0.1251	0.1252	0.1251	7008.1	0.1251	0.1252
119	0.1252	0.1251	0.1251	0.1252	0.1251	7080.1	0.1251	0.1252
103	0.125	0.125	0.125	0.125	0.125	7056.	0.1251	0.125

	2	1	1	2	1	1		2
219	0.1252	0.1251	0.1251	0.1252	0.1251	7036.1	0.1251	0.1252
92	0.1252	0.1251	0.1251	0.1252	0.1251	7040.1	0.1251	0.1252
118	0.1252	0.1251	0.1251	0.1252	0.1251	7088.1	0.1251	0.1252
88	0.1252	0.1251	0.1251	0.1252	0.1251	7046.1	0.1251	0.1252
100	0.1252	0.1251	0.1251	0.1252	0.1251	7052.1	0.1251	0.1252
102	0.1252	0.1251	0.1251	0.1252	0.1251	7050.1	0.1251	0.1252
109	0.1252	0.1251	0.1251	0.1252	0.1251	7034.1	0.1251	0.1252
104	0.1252	0.1251	0.1251	0.1252	0.1251	6986.1	0.1251	0.1252
156	0.1252	0.1251	0.1251	0.1252	0.1251	7038.1	0.1251	0.1252
91	0.1252	0.1251	0.1251	0.1252	0.1251	7046.1	0.1251	0.1252
125	0.1252	0.1251	0.1251	0.1252	0.1251	7066.1	0.1251	0.1252
126	0.1252	0.1251	0.1251	0.1252	0.1251	7060.1	0.1251	0.1252
112	1530.1	0.1251	0.1252	0.1252	0.1251	5522.2	0.1251	0.1252
93	7046.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
107	7016.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
187	7002.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
183	6964.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
181	7010.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
188	6964.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
106	7034.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
105	7014.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
108	7032.1	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	0.1252
116	0.1252	0.1251	1885	5113.2	0.1251	0.1252	0.1251	0.1252
117	0.1252	0.1251	1934.5	5037.8	0.1251	0.1252	0.1251	0.1252
127	0.125	0.125	1712.	5273.	0.125	0.125	0.1251	0.125

	2	1	6	6	1	2		2
162	0.1252	0.1251	1968.5	4991.7	0.1251	0.1252	0.1251	0.1252
203	0.1252	0.1252	0.1252	6145.7	0.1252	0.1252	828.51	0.1252
200	0.1252	0.1252	0.1252	6393.6	0.1252	0.1252	644.66	0.1252
214	0.1252	0.1252	0.1252	6476.3	0.1252	0.1252	559.97	0.1252
206	0.1252	0.1252	0.1252	5380.3	0.1252	0.1252	0.1252	1696
168	0.1252	0.1251	0.1252	7082.1	0.1251	0.1252	0.1251	0.1252
165	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
121	0.1252	0.1251	0.1252	7080.1	0.1251	0.1252	0.1251	0.1252
164	0.1252	0.1251	0.1251	7042.1	0.1251	0.1252	0.1251	0.1252
80	0.1252	0.1251	0.1252	7094.1	0.1251	0.1252	0.1251	0.1252
163	0.1252	0.1251	0.1251	7080.1	0.1251	0.1252	0.1251	0.1252
81	0.1252	0.1251	0.1252	7058.1	0.1251	0.1252	0.1251	0.1252
84	0.1252	0.1251	0.1252	7028.1	0.1251	0.1252	0.1251	0.1252
75	0.1252	0.1251	0.1252	7028.1	0.1251	0.1252	0.1251	0.1252
77	0.1252	0.1251	0.1251	7054.1	0.1251	0.1252	0.1251	0.1252
120	0.1252	0.1251	0.1251	7078.1	0.1251	0.1252	0.1251	0.1252
79	0.1252	0.1251	0.1252	7078.1	0.1251	0.1252	0.1251	0.1252
98	0.1252	0.1251	0.1252	7054.1	0.1251	0.1252	0.1251	0.1252
166	0.1252	0.1251	0.1251	7002.1	0.1251	0.1252	0.1251	0.1252
76	0.1252	0.1251	0.1252	7070.1	0.1251	0.1252	0.1251	0.1252
213	0.1252	0.1251	0.1252	7068.1	0.1251	0.1252	0.1251	0.1252
211	0.1252	0.1251	0.1252	7060.1	0.1251	0.1252	0.1251	0.1252
212	0.1252	0.1251	0.1252	7042.1	0.1251	0.1252	0.1251	0.1252
180	0.1252	0.1251	0.1252	7070.1	0.1251	0.1252	0.1251	0.1252
182	0.125	0.125	0.125	7088.	0.125	0.125	0.1251	0.125

	2	1	1	1	1	2		2
189	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
186	0.1252	0.1251	0.1251	7074.1	0.1251	0.1252	0.1251	0.1252
184	0.1252	0.1251	0.1251	7082.1	0.1251	0.1252	0.1251	0.1252
190	0.1252	0.1251	0.1252	7090.1	0.1251	0.1252	0.1251	0.1252
191	0.1252	0.1251	0.1252	7088.1	0.1251	0.1252	0.1251	0.1252
185	0.1252	0.1251	0.1251	7080.1	0.1251	0.1252	0.1251	0.1252
179	0.1252	0.1251	0.1251	7084.1	0.1251	0.1252	0.1251	0.1252
138	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
150	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
148	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
151	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
134	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
133	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
169	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
174	0.1252	0.1251	0.1251	7090.1	0.1251	0.1252	0.1251	0.1252
1	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
215	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
149	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
143	0.1252	0.1251	0.1251	7082.1	0.1251	0.1252	0.1251	0.1252
130	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
147	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
142	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
152	0.1252	0.1251	0.1251	7080.1	0.1251	0.1252	0.1251	0.1252
128	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
177	0.125	0.125	0.125	7092.	0.125	0.125	0.1251	0.125

	2	1	1	1	1	2		2
136	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
137	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
135	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
132	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
139	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
216	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
131	0.1252	0.1251	0.1251	7088.1	0.1251	0.1252	0.1251	0.1252
146	0.1252	0.1251	0.1251	7090.1	0.1251	0.1252	0.1251	0.1252
129	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
141	0.1252	0.1251	0.1251	7090.1	0.1251	0.1252	0.1251	0.1252
145	0.1252	0.1251	0.1251	7088.1	0.1251	0.1252	0.1251	0.1252
144	0.1252	0.1251	0.1251	7092.1	0.1251	0.1252	0.1251	0.1252
140	0.1252	0.1251	0.1251	7084.1	0.1251	0.1252	0.1251	0.1252
173	0.1252	0.1251	0.1251	7094.1	0.1251	0.1252	0.1251	0.1252
153	0.1252	0.1251	0.1251	7086.1	0.1251	0.1252	0.1251	0.1252
170	0.1252	0.1251	0.1251	7054.1	0.1251	0.1252	0.1251	0.1252
Calico	0.1252	0.1251	0.1252	6036.1	0.1251	0.1252	0.1251	0.1252
Columbian Rainbow	0.1252	0.1251	0.1252	5402.7	0.1251	0.1252	0.1251	1483.5
Pimento de Neyde	0.1252	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	6226.1
Sandia	0.1252	0.1251	0.1252	6206.1	0.1251	0.1252	0.1251	0.1252
sunrise eclipse	0.1252	0.1251	0.1252	5248.5	0.1251	0.1252	0.1251	1491.7
California Mild	0.1252	0.1251	0.1252	4045.3	0.1251	0.1252	0.1251	1776.9
Chili De Arbol	0.1252	0.1252	0.1252	6652.1	0.1252	0.1253	0.1252	0.1253

Filius Blue	0.1252	0.1251	0.1252	5212.8	0.1251	0.1252	0.1251	609.41
Mayan Cobanero	2367.7	0.1251	0.1252	3756.6	0.1251	0.1252	0.1251	0.1252
Mulato	0.1252	0.1251	0.1252	6218.1	0.1251	0.1252	0.1251	0.1252
Pequin	0.1252	0.1251	0.1252	6539.5	0.1251	0.1252	0.1252	418.77
Sweet Chocolate	0.1252	0.1251	0.1251	6626.1	0.1251	0.1252	0.1251	0.1252
CM334	0.1252	0.1251	0.1252	6980.1	0.1251	0.1252	0.1251	0.1252
chiltepin	0.1252	0.1251	0.1252	6726.1	0.1251	0.1252	0.1251	0.1252
Apple Pimento	0.1252	0.1251	0.1252	5740.1	0.1251	0.1252	0.1251	0.1252
Bulgarian Carrot 2	0.1252	0.1251	0.1252	5735.3	0.1251	0.1252	0.1251	930.91
Bulgarian Carrot	0.1252	0.1251	0.1252	6040.1	0.1251	0.1252	0.1251	0.1252
Bulls heart 2	0.1252	0.1251	0.1252	6852.1	0.1251	0.1252	0.1251	0.1252
Buran	0.1252	0.1251	0.1251	6046.1	0.1251	0.1252	0.1251	0.1252
Dulcetta Orange	0.1252	0.1251	0.1251	6096.1	0.1251	0.1252	0.1251	0.1252
Feher Ozon Paprika	0.1252	0.1251	0.1252	4194.2	0.1251	0.1252	0.1251	1868.1
Marseilles Sweet Yellow	0.1252	0.1251	0.1251	6584.1	0.1251	0.1252	0.1251	0.1252
Piment Vegetarian	0.1252	0.1251	0.1251	0.1252	0.1251	0.1252	0.1251	6458.1
Succette de Provence	0.1252	0.1251	0.1252	5634.1	0.1251	0.1252	0.1251	0.1252
Banana Sweet	0.1252	0.1251	0.1252	6144.1	0.1251	0.1252	0.1251	0.1252
Big Red	0.1252	0.1252	0.1252	6590.1	0.1252	0.1253	0.1252	0.1253
China Giant Sweet	0.1252	0.1251	0.1251	6022.1	0.1251	0.1252	0.1251	0.1252

Chinese Ching Choo	0.1252	0.1251	0.1252	4511.8	0.1251	0.1252	0.1251	1368.4
Chocolate Beauty	0.1252	0.1251	0.1252	6484.1	0.1251	0.1252	0.1251	0.1252
Peter Pepper Red	0.1252	0.1251	0.1251	6388.1	0.1251	0.1252	0.1251	0.1252
Riot	0.1252	0.1251	0.1252	6794.1	0.1251	0.1252	0.1251	0.1252
Mams Biber	0.1252	0.1251	0.1252	3687.5	0.1251	0.1252	0.1251	2956.7
Shishito	0.1252	0.1251	0.1252	6286.1	0.1251	0.1252	0.1251	0.1252
Zunla	0.1252	0.1251	0.1252	6776.1	0.1251	0.1252	0.1251	0.1252

$K = 9$

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9
155	0.1112	0.1112	6918.1	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
208	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	7068.1
202	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7036.1
204	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7028.1
124	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	6998.1
217	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	6892.1
157	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	7008.1
119	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	7080.1
103	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7056.1
219	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7036.1
92	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7040.1
118	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7088.1
88	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7046.1
100	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7052.1
102	0.111	0.111	0.111	0.111	0.111	0.111	0.1111	0.111	7050.

	2	2	2	1	1	1		2	1
109	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7034.1
104	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	6986.1
156	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7038.1
91	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	7046.1
125	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	7066.1
126	0.1112	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	7060.1
112	0.1112	1530	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	5522.2
93	0.1112	7046.1	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
107	0.1112	7016.1	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
187	0.1112	7002.1	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
183	0.1112	6964.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
181	0.1112	7010.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
188	0.1112	6964.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
106	0.1112	7034.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
105	0.1112	7014.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
108	0.1112	7032.1	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
116	5114	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	1884.3	0.1112
117	5038.5	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	1933.7	0.1112
127	5274.2	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	1712	0.1112
162	4992.4	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	1967.8	0.1112
203	6145.8	0.1112	0.1112	0.1112	0.1112	0.1112	828.41	0.1112	0.1112
200	6393.5	0.1112	0.1112	0.1112	0.1112	0.1112	644.76	0.1112	0.1112
214	6476.4	0.1112	0.1112	0.1112	0.1112	0.1112	559.86	0.1112	0.1112
206	5376.1	0.1112	1700.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
168	7082.	0.111	0.111	0.111	0.111	0.111	0.1112	0.111	0.111

	1	2	2	2	2	2		2	2
165	7086.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
121	7080.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
164	7042.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
80	7094.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
163	7080.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
81	7058.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
84	7028.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
75	7028.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
77	7054.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
120	7078.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
79	7078.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
98	7054.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
166	7002.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
76	7070.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
213	7068.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
211	7060.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
212	7042.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
180	7070.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
182	7088.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
189	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
186	7074.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
184	7082.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
190	7090.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
191	7088.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
185	7080.	0.111	0.111	0.111	0.111	0.111	0.1112	0.111	0.111

	1	2	2	1	1	1		2	2
179	7084.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
138	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
150	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
148	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
151	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
134	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
133	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
169	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
174	7090.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
1	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
215	7086.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
149	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
143	7082.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
130	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
147	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
142	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
152	7080.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
128	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
177	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
136	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
137	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
135	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
132	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
139	7086.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
216	7086.	0.111	0.111	0.111	0.111	0.111	0.1111	0.111	0.111

	1	2	2	1	1	1		2	2
131	7088.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
146	7090.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
129	7086.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
141	7090.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
145	7088.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
144	7092.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
140	7084.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
173	7094.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
153	7086.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
170	7054.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1111	0.1112	0.1112
Calico	6036.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Columbian Rainbow	5398.6	0.1112	1487.6	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Pimento de Neyde	0.1112	0.1112	6226.1	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
Sandia	6206.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
sunrise eclipse	5243	0.1112	1497.2	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
California Mild	4039.3	0.1112	1782.9	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Chili De Arbol	6450.8	0.1112	201.4	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Filius Blue	5210.3	0.1112	611.92	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Mayan Cobanero	3756	2368.2	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Mulato	6218.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Pequin	6537.9	0.1112	420.28	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Sweet Chocolate	6626.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112

CM334	6980.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
chiltepin	6726.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Apple Pimento	5740.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Bulgarian Carrot 2	5730.6	0.1112	935.64	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Bulgarian Carrot	6040.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Bulls heart 2	6852.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Buran	6046.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
Dulcetta Orange	6096.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Feher Ozon Paprika	4189.5	0.1112	1872.7	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Marseilles Sweet Yellow	6584.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
Piment Vegetarian	0.1112	0.1112	6458.1	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
Succette de Provence	5634.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Banana Sweet	6144.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Big Red	6590.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
China Giant Sweet	6022.1	0.1112	0.1112	0.1111	0.1111	0.1111	0.1112	0.1112	0.1112
Chinese Ching Choo	4506.5	0.1112	1373.8	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Chocolate Beauty	6484.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Peter Pepper Red	6388.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Riot	6794.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Mams Biber	3681.3	0.1112	2962.9	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112

Shishito	6286.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112
Zunla	6776.1	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112	0.1112

K = 10

Plant ID	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
155	0.1	0.1	0.1	0.1	0.1	0.1	6918.1	0.1	0.1	0.1
208	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7068.1	0.1	0.1
202	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7036.1	0.1	0.1
204	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7028.1	0.1	0.1
124	0.1	0.1	0.1	0.1	0.1	0.1	0.1	6998.1	0.1	0.1
217	0.1	0.1	0.1	0.1	0.1	0.1	0.1	6892.1	0.1	0.1
157	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7008.1	0.1	0.1
119	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7080.1	0.1	0.1
103	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7056.1	0.1	0.1
219	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7036.1	0.1	0.1
92	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7040.1	0.1	0.1
118	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7088.1	0.1	0.1
88	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7046.1	0.1	0.1
100	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7052.1	0.1	0.1
102	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7050.1	0.1	0.1
109	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7034.1	0.1	0.1
104	0.1	0.1	0.1	0.1	0.1	0.1	0.1	6986.1	0.1	0.1
156	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7038.1	0.1	0.1
91	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7046.1	0.1	0.1
125	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7066.1	0.1	0.1
126	0.1	0.1	0.1	0.1	0.1	0.1	0.1	7060.1	0.1	0.1
112	1530	0.1	0.1	0.1	0.1	0.1	0.1	5522.	0.1	0.1

								2		
93	7046.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
107	7016.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
187	7002.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
183	6964.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
181	7010.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
188	6964.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
106	7034.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
105	7014.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
108	7032.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
116	0.1	5114	0.1	0.1	0.1	1884.2	0.1	0.1	0.1	0.1
117	0.1	5038.6	0.1	0.1	0.1	1933.6	0.1	0.1	0.1	0.1
127	0.1	5274.2	0.1	0.1	0.1	1712	0.1	0.1	0.1	0.1
162	0.1	4992.5	0.1	0.1	0.1	1967.7	0.1	0.1	0.1	0.1
203	0.1	6145.8	0.1	828.38	0.1	0.1	0.1	0.1	0.1	0.1
200	0.1	6393.5	0.1	644.74	0.1	0.1	0.1	0.1	0.1	0.1
214	0.1	6476.4	0.1	559.83	0.1	0.1	0.1	0.1	0.1	0.1
206	0.1	5376.1	0.1	0.1	0.1	0.1	1700.1	0.1	0.1	0.1
168	0.1	7082.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
165	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
121	0.1	7080.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
164	0.1	7042.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
80	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
163	0.1	7080.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
81	0.1	7058.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
84	0.1	7028.	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

		1								
75	0.1	7028.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
77	0.1	7054.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
120	0.1	7078.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
79	0.1	7078.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
98	0.1	7054.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
166	0.1	7002.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
76	0.1	7070.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
213	0.1	7068.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
211	0.1	7060.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
212	0.1	7042.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
180	0.1	7070.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
182	0.1	7088.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
189	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
186	0.1	7074.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
184	0.1	7082.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
190	0.1	7090.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
191	0.1	7088.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
185	0.1	7080.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
179	0.1	7084.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
138	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
150	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
148	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
151	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
134	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
133	0.1	7094.	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

		1								
169	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
174	0.1	7090.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
1	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
215	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
149	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
143	0.1	7082.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
130	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
147	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
142	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
152	0.1	7080.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
128	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
177	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
136	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
137	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
135	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
132	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
139	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
216	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
131	0.1	7088.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
146	0.1	7090.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
129	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
141	0.1	7090.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
145	0.1	7088.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
144	0.1	7092.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
140	0.1	7084.	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

		1								
173	0.1	7094.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
153	0.1	7086.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
170	0.1	7054.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Calico	0.1	6036.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Columbian Rainbow	0.1	5398.7	0.1	0.1	0.1	0.1	1487.5	0.1	0.1	0.1
Pimento de Neyde	0.1	0.1	0.1	0.1	0.1	0.1	6226.1	0.1	0.1	0.1
Sandia	0.1	6206.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
sunrise eclipse	0.1	5243	0.1	0.1	0.1	0.1	1497.2	0.1	0.1	0.1
California Mild	0.1	4039.3	0.1	0.1	0.1	0.1	1782.9	0.1	0.1	0.1
Chili De Arbol	0.1	6450.9	0.1	0.1	0.1	0.1	201.35	0.1	0.1	0.1
Filius Blue	0.1	5210.3	0.1	0.1	0.1	0.1	611.88	0.1	0.1	0.1
Mayan Cobanero	2368.2	3756	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Mulato	0.1	6218.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Pequin	0.1	6538	0.1	0.1	0.1	0.1	420.23	0.1	0.1	0.1
Sweet Chocolate	0.1	6626.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
CM334	0.1	6980.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
chiltepin	0.1	6726.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Apple Pimento	0.1	5740.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Bulgarian Carrot 2	0.1	5730.6	0.1	0.1	0.1	0.1	935.6	0.1	0.1	0.1
Bulgarian Carrot	0.1	6040.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Bulls heart 2	0.1	6852.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Buran	0.1	6046.	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

		1								
Dulcetta Orange	0.1	6096.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Feher Ozon Paprika	0.1	4189.5	0.1	0.1	0.1	0.1	1872.7	0.1	0.1	0.1
Marseilles Sweet Yellow	0.1	6584.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Piment Vegetarian	0.1	0.1	0.1	0.1	0.1	0.1	6458.1	0.1	0.1	0.1
Succette de Provence	0.1	5634.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Banana Sweet	0.1	6144.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Big Red	0.1	6590.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
China Giant Sweet	0.1	6022.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Chinese Ching Choo	0.1	4506.5	0.1	0.1	0.1	0.1	1373.7	0.1	0.1	0.1
Chocolate Beauty	0.1	6484.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Peter Pepper Red	0.1	6388.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Riot	0.1	6794.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Mams Biber	0.1	3681.3	0.1	0.1	0.1	0.1	2962.9	0.1	0.1	0.1
Shishito	0.1	6286.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Zunla	0.1	6776.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

CHAPTER 3

SYNTENIC MAPPING IDENTIFIED A FRUIT SHAPE QTL AND A *TRM25* CANDIDATE GENE ON CAPSICUM ANNUUM CHROMOSOME 3[1]

---

[1] Taitano, N., Qiu, Z., van der Knaap, E. 2020. *To be submitted to Theoretical and Applied Genetics*

Abstract

The shape of harvested organs in crop species is important for ease of harvest, processing, and shipping, and can limit supply in vegetables and other horticultural crops. In recent years, mounting evidence has shown that orthologous genes can regulate organ shape similarly across plant species. The present study sought to extend these findings in the bell/chile pepper *Capsicum annuum*. Chile pepper is a Solanaceous crop with regions of strong synteny with the tomato (*Solanum lycopersicum*), another member of the family and a model for fleshy fruit development. Following a candidate gene approach, marker-trait association tests were performed with organ shape genes known from tomato in a population of F_2 pepper plants segregating for fruit shape. Markers on candidate genes *CaOVATE*, and *CaWUSCHEL* on chromosome 2 were not associated with fruit shape in the population. A marker for the candidate pepper fruit shape QTL on chromosome 10, *fs10.1*, was also not associated with shape. Instead, a marker on *CaELONGATA3* (*ft3.1*) was significantly associated with pepper fruit shape in the segregating population. Due to the lack of high-quality contiguous sequences in the pepper reference genomes, we compared the pepper chromosome 3 region to the syntenic position in tomato. This finding enabled us to narrow another fruit shape QTL, *fs3.1,* from most of chromosome 3 to a 27 Mb region containing 364 genes. These genes included an *OVATE*-interacting *TONNEAU1 RECRUITING MOTIF* (*TRM*) ortholog, *CaTRM25*. A marker on this gene showed the strongest association with *fs3.1* and therefore *CaTRM25* was considered a likely fruit shape candidate gene in the segregating population. The tight linkage of *CaTRM25* to *fs3.1* was further supported by results from progeny testing. Semi-quantitative PCR results demonstrated that *CaTRM25* was differentially expressed, which might underlie the molecular basis of altered fruit shape. One of the SNPs in the coding region of *CaTRM25* resulted in a conservative amino acid change and showed association with fruit shape in some landrace populations but not others, suggesting that the associated mutation was not causative. Histological analysis of the ovary at anthesis revealed a difference in the number of pericarp cells in the proximal-distal direction which could underlie the histological basis of the fruit shape difference. These findings combined showed that pepper *TRM25*

was a likely gene to control fruit shape in this species, in support of previous findings that the TRM-OVATE family protein network is central to organ shape regulation in several crop species.

Introduction

The control of fruit size and shape is important for crop quality under cultivation (Hori *et al.,* 2013) and seed dispersal in the wild (Bollen *et al.,* 2004), but often in opposing directions. As a result, allele frequencies of key genes controlling fruit shape have changed over the course of domestication. In tomato (*Solanum lycopersicum* var *lycopersicum*), this consisted of a diversification of shapes based on enrichment for a few mutant alleles that are rare in the semi-domesticated and wild germplasm. The loci *LOCULE NUMBER* (*lc* underlying *SlWUSCHEL*), *OVATE, SUPPRESSOR OF OVATE1* (*sov1*, underlying *SlOFP20*), *SUN,* and *FASCIATED* (*fas*, underlying *SlCLAVATA3*) control most of the natural variation in modern tomato (Rodriguez *et al.,* 2011, Blanca *et al.,* 2015, Wu *et al*, 2018). Specifically, the genes *OVATE, SlOFP20* and *TONNEAU1 RECRUITING MOTIF5* (*SlTRM5*) - the latter encodes a protein that was identified in a Yeast 2-Hybrid screen to interact with OVATE - act well before anthesis, and together account for a significant part of the morphological diversity in tomato (Wu *et al.,* 2018).

The species *Capsicum annuum,* which comprises most bell and chile pepper varieties, contains remarkable diversity in fruit shape (Bosland and Votava, 2012). Pepper fruit shape can be divided into nine visual categories (IPGRI, 1995) or quantified by attributes of shape such as fruit shape index (FSI): the ratio of maximum curved height to width. To date, the major pepper fruit shape QTLs have been mapped on chromosomes 2, 3 and 10 (Ben Chaim *et al.,* 2001, Borovsky and Paran, 2011, Han *et al*., 2016, Hill *et al.,* 2017, Chunthawodtiporn *et al*., 2018). Early comparisons between the now well-established tomato fruit morphology genes and those in pepper identified separate QTL acting on shape in the two species, though there is evidence that orthologous fruit weight QTL were shared between them (Paran *et al.,* 2004, Chakrabarti *et al*., 2013, Paran and van der Knaap, 2007). However, this view may need to be reevaluated in light of several recently published pepper reference genome assemblies that allows for the discovery of additional homologs to known fruit shape genes.

Three independent genome assemblies of *C. annuum* are publicly available and they are the Illumina-based assemblies of the *Phytophthora*-resistant Mexican landrace Criollo de Morelos 334 (CM334, Kim *et al.,* 2014), the elite Chinese cultivar Zunla-1 (Qin *et al.,* 2014), and the linked-read sequence-based assembly UCD10X, sequenced from a hybrid between CM334 and a UC Davis breeding line. Many fruit shape and weight QTL have been positioned on the CM334 pepper reference genome (Kim *et al.,* 2014, Hill *et al.,* 2017). However, further narrowing of QTLs and the inference of specific candidate genes has been hampered due to discrepancies between the assemblies in regions that are poorly assembled. Fortunately, select regions of the pepper genome exhibit sufficient synteny with the well-assembled and annotated tomato genome (Rinaldi *et al.,* 2016).

Evidence of the importance of diversity in morphology implies that genetic mechanisms regulating fruit shape may be conserved. Particularly noteworthy are OVATE and other OVATE family proteins (OFPs). OFPs regulate fruit shape and that of other organs in species spanning a wide range of plants including Arabidopsis (Hackbusch *et al.,* 2005, Wang *et al.,* 2007, Wang *et al.,* 2011, Wang *et al.,* 2019), rice (Schmitz *et al.,* 2015, Ma *et al.,* 2017, Zhao *et al.,* 2018), melon, potato, and tomato (Liu *et al.,* 2002, Wu *et al.,* 2018). Taking tomato as a well-studied example, natural variants carrying both a large promoter deletion in *SlOFP20* at the *suppressor of ovate1* (*sov1*) locus and a premature-stop loss-of-function mutation in *OVATE* develop fruits with narrowed proximal ends leading to a longer, pear-shaped fruit without changing fruit weight (Liu *et al.,* 2002, Rodriguez *et al.,* 2013, Wu *et al.,* 2018). This shape-specific phenotype is driven by a change in cell division orientation occurring before the seventh day after floral meristem initiation which is more than 10 days before anthesis (Kraus 2019). OVATE and other members of the family interact with a subset of the TRMs in a Yeast 2-Hybrid study and in tobacco leaf epidermal cells (Wu *et al.*, 2018, Keyhaninejad *et al.,* submitted). Importantly, the CRISPR-induced *Sltrm5* null leads to the partial rescue to a round fruit in the *ovate/sov1* double NIL background implying that *OFPs* and *TRMs* interact genetically as well (Wu *et al.,* 2018). In the triple mutant, the cell division orientation is partially restored to wild type division patterns which would explain the partial restoration of fruit shape. The TRMs are thought to affect organ shape by directing subcellular localization of the

TONNEAU1 protein subunit to the pre-prophase band (PPB) to regulate cell division patterns (Spinner *et al.*, 2013, Schaefer *et al.,* 2017). In tomato, TRMs interact with OFPs through the conserved M8 domain that is found in certain TRM family members. We further hypothesized that the relocalization of TRMs by OVATE and SlOFP20 to different subcellular compartment regulates cell-division patterning and ultimately fruit shape (Wu *et al.,* 2018). TRMs have also been found to affect organ shape across a wide range of species including Arabidopsis (Lee *et al.,* 2006, Wang *et al.,* 2007, Drevensek *et al.,* 2012), rice (Wang Y. *et al.,* 2015, Wang S. *et al.,* 2015, Zhou *et al.,* 2015), cucumber (Wu *et al.,* 2018).

This study sought to use synteny with tomato to map a pepper fruit-shape QTL in a poorly assembled segment on pepper chromosome 3 overlapping with *fs3.1* (Ben Chaim *et al.*, 2001). The results allowed us to elucidate potential similarities in the regulation of organ morphology and between these two species in the Solanaceae family.

Materials and Methods

Plant Materials

The parents of the F_2 population are the long-fruited Mexican Chile de Arbol (14P9-1) and round Mexican Chigole (14P15-1) accessions (Taitano *et al.,* 2019). Seeds from a single F_1 plant were germinated and grown into a population of 178 F_2 plants, together with a control population consisting of 10 plants from each parental seed lot in a greenhouse in Athens, GA in 2016 in a completely randomized design. The parents were part of a previous study that had identified the genetic relationship and classification among the landraces from the Mexican states of Oaxaca and Yucatan. A description of the accessions used in this study follows.

The Chigole accession is a backyard-grown pepper from near the Oaxacan coast in Rosedal. It is part of the “Forest” group, which also included a set of forest-growing peppers, known locally as Chile de Monte. The Chile de Arbol is another Oaxacan accession, which genetically grouped together with a subpopulation containing Tusta and Taviche landrace accessions. Chile de Arbol, Tusta and Taviche peppers are all either grown in backyards or traditional *milpa* polycultures in the Oaxacan Central Valleys and surrounding mountainside villages.

In addition to the two accessions used as parents for the biparental mapping population, accessions from four additional subpopulations were used for sequencing and genotyping. The first of these is the Chile de Agua landrace. Chile de Agua is the largest-fruited and most intensively cultivated of the Oaxacan landraces used in this study and exhibited the strongest evidence of selection in a previous landrace study (Taitano *et al.,* 2019). The second subpopulation is the Costeño landrace: this population of long-fruited, drying peppers grown on the Oaxacan coast exhibited the second-strongest bottleneck and was also commonly grown under intensive cultivation systems. A third population consists of a small collection of backyard and milpa-grown accessions from the Yucatan peninsula. Finally, a group of accessions from another cultivated *Capsicum* species, *C. frutescens,* which is mostly grown in backyards and sometimes found in wild forest environments, was also used for genotyping (Taitano *et al.* 2019).

One pair of accessions from each of the four main *C. annuum* groups was selected for sequencing in the *CaTRM25* locus. The first pair are 14Chg162-1 and 14CdA180-1, two plants grown from the same original seedlot as the Chigole and Chile de Arbol parent plants, respectively. The second pair were two Chile de Agua: 14CAg132-1 from a plantation in Coatecas Altas, and 14CAg144-1 from a milpa in Lobera. The third pair consisted of two Costeño: CRo80-1 from the coastal town of Añil, and CRo168-1 from Rosedalito, both plantation-grown. The last pair were one Tusta (14Tus179-1) and one Taviche (14Tus186-1), both of which belong to the same Tusta/Taviche subpopulation (TusTav1) as our Chile de Arbol parent. 14Tus179-1 and 14Tus186-1 originate from San Pablo Coatlán and Santa Cruz Nexila: neighboring towns in the Central Valleys, where they were grown in a plantation and milpa setting, respectively.

For progeny testing fruit shape in the F_3 families, each F_2 plant was selfed by hand-pollination, with outcross contamination prevented by staking and pruning to keep plants separated, and collecting seed from fruits located close to the central stem. F_3 seedlings from seed thus obtained were genotyped for markers flanking the *fs3.1* QTL, selected from selfed F_3 progeny of recombinant F_2 plants.

For fruit developmental analysis and the RNA expression studies, one F_3 family, 19P1, was grown as above and assessed for fruit shape during development starting at anthesis until mature fruit.

DNA and RNA extractions

For DNA, approximately 100 mg of young-leaf tissue from seedlings with 4-6 true leaves were collected on ice into deep-well tubes in 96-well plates. These were then flash-frozen in liquid nitrogen and lyophilized prior to DNA extraction. After lyophilizing, the samples were ground to a fine powder by adding metal beads and mechanically shaking in a Geno/Grinder 2000® (SPEX, Metuchen, NJ, USA). DNA extraction was performed using QIAGEN's DNEasy 96 Plant Kit® (Valencia, CA, USA), following the manufacturer's recommendation. DNAs were eluted into 100 µl TE pH 8. F_3 progeny DNA was extracted from lyophilized young leaf tissue in a phenol extraction protocol (Prince *et al.,* 1997) adapted for use with approximately 50 mg leaf tissue in 2 mL microcentrifuge tubes.

For RNA, multiple young flower buds of approximately 2 mm were collected from 4 to 5 plants of each homozygous genotype, flash-frozen in liquid nitrogen and stored at -80ºC. Total RNA from the buds from each plant was extracted using Trizol reagent following the protocol provided by the manufacturer (Invitrogen, USA) and treated with DNase I (NEB, USA).

Plant Phenotyping

For mature fruit shape and weight evaluations in the F_2, eight fruits were collected from each plant, individually weighed, bisected longitudinally, and scanned to collect fruit morphology data. Heritability was calculated by the difference method for inbred populations, with equal weighting given to F_1 and parental inbred lines (Marsh *et al.,* 1985 as cited in Nyquist 1991). Curved fruit shape index, the length of a curve bisecting the fruit vertically divided by maximum fruit width measured perpendicular to this curve, was calculated for each fruit. In F_3 progeny test phenotyping, mature fruits were collected, weighed, scanned, and measured as for the F_2 population. For each group of F_3 homozygotes, Welch's unequal variances t-test was performed to assess whether the two homozygous genotypes differed for the trait within each family.

To assess ovary shape at anthesis, five flowers from each of 4 to 7 plants per genotype in F_3 population 19P1 were bisected longitudinally and scanned at 600 dpi. The ovaries in these scans were evaluated for shape index (ovary length : width ratio), distal shape index (ratio of distance from the distal end of the locules to the distal tip : ovary width at distal end of locules), distal end macro (angle of tangents at 10% along perimeter from distal tip), distal end blockiness (ratio of width mid-height to width 10% from distal tip), and triangle (ratio of widths 10% from proximal : distal ends) (Brewer *et al.,* 2006), quantified in ImageJ (Rasband, 2012). To monitor fruit development, 2 to 3 anthesis stage flowers from each plant were tagged, hand-pollinated, and their corollas were removed. Images were taken of the developing fruits three times weekly until elongation ceased in all fruits for three consecutive measurements. Flowers with poor or delayed fruit set were excluded from subsequent analysis. Constant perspective was carefully controlled through the use of a rigid index card with both a scale bar and perspective square. Maximum widths and curved lengths following an apical-distal line bisecting the fruit were then measured in ImageJ (Rasband, 2012).

Synteny-informed marker development and QTL mapping

KASP (Cuppen, 2007) or dCAPS (Neff *et al.,* 1998) markers were designed based on several known fruit shape genes as well as on GBS SNPs between the parents (Taitano et al., 2019) located within previously published QTL regions for fruit shape on pepper chromosomes 3 and 10 (Table 3.S1). Interval mapping in R/QTL was performed using each of these markers as an initial search for known QTL varying in the population. Using the CM334 Pepper1.55 reference genome as a physical map, a series of flanking markers were developed in both directions outside *CaELONGATA3* on chromosome 3 using SNPs that differed between the two parents.

To resolve discrepancies between the reference genomes for the top of chromosome 3, the genomes were assessed for synteny with tomato by protein BLAST. First, orthologous tomato sequences were obtained by using 1) BLASTn of the primer sequence and 2) BLASTp of the nearest protein-coding gene to each marker within an ungapped pepper CM334 contig. Where positions of known SNPs varied among the pepper genomes, a consensus gene order was obtained by comparing the four available genome assemblies including tomato, and deferring to the better-assembled tomato gene order where no consensus could be found. This syntenic consensus was used for additional marker development.

With markers thus designed, a maximum-likelihood genetic map was determined considering all possible marker orders via the ripple function implemented in IciMapping v4.2, with distances determined by the Kosambi method (Kosambi, 1944, Meng *et al.,* 2015). Inclusive composite interval mapping (ICIM) was used to associate fruit shape and weight, which includes compound genotype effects across multiple QTL within its interval mapping model (Meng *et al.,* 2015). Once two separate QTL were identified, allele effects of each were estimated in a linear model, and two-way ANOVA was performed on their effects on fruit shape to identify any interactions between the two loci.

Since the UCD10X genome appeared most consistent with the genetic map, an expansive genomic region in tomato was designated as stretching from the top of tomato chromosome 3 to the *fs3.1* marker ortholog furthest along the chromosome. All genes from this region were used as queries in tBLASTn against UCD10X chromosome 3. MCScanX was then used to identify and visualize collinear regions of the two chromosomes based on tBLASTn results with e-values under 1e-5, gap-adjusted BLOSUM62 (Henikoff and Henikoff, 1992) scores above 50 and collinear blocks defined by at least five consecutive genes. A reciprocal BLAST was also performed using default-parameter AUGUSTUS-predicted (Stanke and Morgenstern, 2005) genes in the marker region from the UCD10X reference genome as query for validation and to obtain annotation data from tomato orthologs.

Histological evaluations of F_3 ovaries

Flower buds at anthesis were collected from each plant in a single F_3 family (19P1). Five buds per plant were bisected along the proximal-distal plane. The halves were fixed in a 3:1 solution of 95% ethanol:acetic anhydride. The bisected buds were made permeable by treatment in 80% ethanol at 80 ºC for 10 minutes, then gradually rehydrated by three 10-minute incubations in 50%, then 30%, then 0% ethanol. Tissues were cleared in two steps: a 24-hour incubation in 0.2N NaOH and 1% SDS followed by a 72-hour incubation in ClearSee (Kurihara e*t al.,* 2015). Finally, tissues were washed in water, stained with 0.25% Calcofluor White M2R (Sigma-Aldrich) and 0.001% 0.2 N NaOH for 30 minutes at room temperature, and imaged by confocal microscopy by excitation at 405 nm with an emission band of 410-550 nm. Sample preparation and microscopy was performed by Dr. Zhengkun Qiu.

Images taken on the same confocal microscopy plane were merged to form pericarp-spanning composite images from the proximal to distal end of the anthesis ovaries. These images were processed in ImageJ to measure average cell area, length, width, and the number of epidermal pericarp cells from the distal to the proximal end of the ovary. Average cell measurements were taken from a 100x100 μm square at mid-height in the middle of the pericarp of each ovary, avoiding vascular bundles. All cells that where at least half the cell's area within the square's boundaries were counted to obtain the denominator and calculate average cell area. Positions were recorded for each cell within the square, and average length and width was thus calculated for cells by each column and row.

Real-time PCR assays

Approximately 1000 ng of total RNA extracted from ~2 mm sized floral buds from 19P1 F_3 homozygous progeny was reverse transcribed using High Capacity cDNA Reverse Transcription Kit (Life Technologies, USA). Each plant was treated as a biological replicate for a total of 9 plants that were homozygous for one or the other parent. The qRT-PCR was performed using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, USA). The PCR amplification of *CaTRM25* cDNA consisted of an initial incubation at 95°C for 30 s, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. PCR products were analyzed using Bio-Rad CFX connect (Bio-Rad, USA). The gene *CA10g11620/Capana10g001439* published previously was used as a control to normalize expression (Cheng, et al. 2017, Livak and Schmittgen 2001), and all analyses were performed with three technical replicates. The 2-ΔCt method was used to calculate the relative expression of the candidate gene (Livak and Schmittgen 2001). The RNA extraction and expression studies were conducted by Dr. Zhengkun Qiu in the lab.

Exploration of TRM25 diversity in pepper landrace accessions

Sanger sequencing was performed to identify *CaTRM25* polymorphisms in a subset of Mexican landraces including the parents of the F_2 population (Taitano *et al.,* 2019). Representatives (as detailed in *Plant Materials*, above) of Tusta/Taviche population 1 (TusTav1), Chile de Agua (CAg), and Costeño were sequenced. Parental lines Chigole (a Forest category member) and Chile de Arbol (a TusTav1 member) of the F_2 parental seed lots were also sequenced. SNPs occurring in these sequences were identified, and assessed for nonsynonymous vs. synonymous substitution rates via the maximum likelihood estimation method (Yang and Bielawski, 2000).

To characterize pepper genetic diversity across the *TRM* family and compare it with those affecting organ shape in other species, TRMs from the CM334 genome were examined via phylogenetic analysis. First, TRMs in pepper were identified by BLASTp search using all tomato TRM proteins bearing the OVATE-interacting M8 domain, with a permissive e-value threshold of 0.001. MAST was performed using the known TRM motifs, with cumulative e-values and position p-values under 1e-10 and 1e-4, respectively. Matching proteins that contained both the M2 (TRM-defining) and M8 (OFP-interaction) domains were aligned by CLUSTALW. RAxML was used to build a GAMMA BLOSUM62 maximum-likelihood consensus tree from 100 bootstrap replicates.

To assess allelic diversity in Mexican landraces as well as the F_2 population, additional dCAPS markers (Table 3.S1) were developed based on the *CaTRM25* sequence at this locus and used to genotype the full collection of Mexican landrace peppers.

Results

Synteny-informed fruit shape QTL mapping

The fruit weight and shape phenotypes showed a skewed distribution towards the round and small fruited Chigole parent in the F_2 population (Figure 3.1A, B). These distributions differed significantly from normality based on the Shapiro-Wilk test, $p < 0.05$. Both shape and weight traits displayed high levels of broad sense heritability and the range of both phenotypes in the F_2 plants were within that of the parental phenotypes. Furthermore, fruit weight and shape were significantly correlated (Figure 3.1C), demonstrating that the more elongated fruits were larger as well.

Of the previously mapped pepper fruit shape QTLs, *ft3.1* near the *CaELONGATA3* gene was associated with pepper shape at p=4.2e-8, explaining 20.7% (PVE) of the fruit shape variance in this F_2 population (Figure 3.S1A) based on single marker analysis. On the other hand, *CaWUSCHEL, CaOVATE* and *fs10.1* were not associated with shape in this population. We also sought to associate the same markers with fruit weight, since *lc* underlying *SlWUSCHEL* in tomato is often associated with weight in addition to shape. Two SNPs within the fruit weight locus *fw2.1* at *CaWUSCHEL* and *CaOVATE*, and to a lesser extent *ft3.1,* were significantly associated with fruit weight (Table 3.1, Figure 3.S1B). However, this result might also reflect linkage to a fruit weight gene and that the actual gene tested was not directly associated with weight.

To determine whether the association for shape on chromosome 3 was valid, we developed and mapped additional markers in the region (Table 3.S1). The pepper genome sequences around *ft3.1* (18EP290, Table 3.S1) were in agreement in the Zunla-1 and CM334 genomes (Figure 3.S2), yielding markers 18EP813, 18EP497, 18EP549 and 18EP555. The single marker association of 18EP813 and 18EP497 in the F_2 population resulted in the surprising result that both were more associated with fruit shape than *ft3.1* (Figure 3.S1A). To develop and map more markers past the most significantly associated marker 18EP497, we could no longer rely on the pepper reference genomes since collinearity had fallen apart due to a possible inversion and the sequences were highly fragmentary (Figure 3.S2). Therefore, we used GBS SNPs from a previous study (Taitano et al, 2018) and BLAST to align a 100-bp region flanking the SNP in the CM334 against tomato, yielding these additional markers: 18EP836, 18EP842, 18EP944, 18EP950, 18EP1080, 18EP1083, 18EP1098. In an effort to obtain the order of the markers in this region, a genetic map was created with the markers (Figure 3.2A). Using the inclusive composite interval mapping (ICIM) method, we identified another fruit shape QTL, *fs3.1*, to a 3.98 cM fruit shape QTL interval flanked by markers 18EP948 and 18EP836, featuring an ICIM LOD of 9.96 and PVE of 34.5% (Figure 3.2A) with an allele effect leading to an average elongation increase of 0.85 (unitless: measured by FSI) for a Chile de Arbol parental allele. In both *fs3.1* and *ft3.1*, the allele from the long-fruited Chile de Arbol parent was associated with elongated fruits in the F2. However, *fs3.1* had a larger effect on fruit shape than *ft3.1*, which had an allele effect half as large (0.39 as measured by FSI) in addition to its LOD of 4.68 (Figure 3.2A) and PVE of 20.7%. In a two-way ANOVA interaction test, no significant interaction was found between *fs3.1* and *ft3.1* (p=0.13698).

To compare the genetic marker order with the physical map, the UCD10X sequence was aligned to the tomato genome sequence as this genome build appeared to provide the most contiguous sequence. Compared to tomato, the genomic region in pepper showed a large expansion as well as several local inversions (Figure 3.2B). The extended syntenic region spanning over 3 Mb in tomato identified 364 of 633 (57%) tomato genes as collinear with pepper orthologs (Supplemental File 3.S1). BLAST against tomato revealed that the AUGUSTUS-predicted open reading frames in the pepper *fs3.1* interval corresponded with 204 annotated genes (Table 3.S2). The list of genes included a gene encoding a GA20 oxidase, a gene encoding a TCP transcription factor, and genes encoding a paralog of SUN and an ARF5 transcription factor involved in auxin signaling among many others. However, the most interesting gene at the *fs3.1* locus was predicted to be an ortholog of tomato *TRM25*. A new marker, 18EP1221, was developed based on a SNP in the coding region of the pepper *CaTRM25* and genotyped in the F_2 population (Figure 3.2C, D). The data showed that more elongated fruits carried the Chile de Arbol allele whereas the less elongated fruits carried the Chigole allele. When this marker was included in the ICIM analysis, the highest association with fruit shape was found with marker 18EP1221 which corresponded to *CaTRM25* (Figure 3.2A).

Progeny tests of F_3 families

To confirm the *fs3.1* locus was segregating in the next generation as well as to narrow down the locus, progeny tests of F_3 families were conducted. Selected F_2 plants harbored a recombination where one end of the locus was homozygous for one allele or the other while the other end of the locus was heterozygous (Table 3.2). For example, family 18P14 was fixed for the Chigole allele at marker 18EP842 and 18EP549 whereas the other markers were heterozygous. The F_3 seedlings from this plant were screened to be either homozygous Chigole or homozygous Chile de Arbol and fruit shape was evaluated. Fruit shape segregated significantly, indicating that the gene was located to the left of marker 18EP842. Combined, the progeny testing results demonstrated that the fruit shape gene mapped between 18EP836 and 18EP948. In addition, no recombinants were found between the *fs3.1* and marker 18EP1221 that corresponded to *CaTRM25* (Table 3.2). These results supported the notion that *CaTRM25* was a likely candidate gene for fruit shape in this pepper population.

Developmental and expression analyses of fs3.1

To determine whether the fruit shape differences develop during floral or during fruit development, we evaluated the shape of the ovaries at anthesis-stage flowers in one segregating F_3 family. At anthesis, ovary length was significantly enlarged in Chile de Arbol homozygotes, whereas width was not changed (Figure 3.3A, B). This resulted in a much more elongated ovary at anthesis in the accessions carrying the Chile de Arbol allele of *fs3.1* (Figure 3.3C). For the shape at the distal end of the ovary, we found a similar pattern as for the whole ovary, demonstrating that the tip of the fruit is more tapered with the Chile de Arbol allele (Figure 3.3 D-F). A trend of a triangular ovary was observed albeit that this was not significant (Figure 3.3G). No differences in shape were observed for distal end blockiness (Figure 3.3H) and distal end angle at 10% (Figure 3.3I). We also tracked the fruit growth after pollination in the segregating F_3 family. Monitoring the length and width of the developing fruits revealed a trend of accelerated fruit shape differences beginning around four days post anthesis (dpa) and maximizing at 9 dpa (Figure 3.3J). Combined the data support the notion that fruit shape mediated by *fs3.1* took place both during floral and fruit development.

Next, the differences in tissue growth were examined at the histological level to determine the cellular basis of growth. Progeny homozygous for the Chile de Arbol parental allele at *fs3.1* neither exhibit changes in cell shape nor cell size, but showed significantly more cells in the proximal-distal direction of the ovary wall than those homozygous for the Chigole allele (Figure 3.4A-C). Therefore, the cellular patterns that underlie the ovary shape yielded differences primarily in cell number.

We wanted to know if *CaTRM25* was differentially expressed in the segregating F_3 population since this could define the molecular basis of fruit shape controlled by the gene. First, we determined where *CaTRM25* was the highest expressed using publicly available expression data. The highest expression was found in 3.5 mm young flower buds and again 0-3 days after pollination (Figure 3.5A) (Liu *et al.,* 2017). Real time quantitative PCR was performed using ~2 mm sized floral buds which was presumed to overlap developmentally with the larger buds in cultivated pepper. The results showed a significant difference in *CaTRM25* expression in the 19P1 F_3 progeny based on the genotype of the accessions (Figure 3.5B). The Chigole allele was 3.36X higher expressed than the Chile de Arbol allele and the result implied that differences in fruit shape could be due to higher expression of *CaTRM25* resulting in a rounder fruit.

Exploration of CaTRM25 diversity in landrace peppers

In the coding region of *CaTRM25*, we found SNPs that translated into four missense mutations resulting in amino acid changes and one amino acid deletion in the parents of the mapping population. The changes were at position 340 from an Isoleucine to Threonine ($340_{I\rightarrow T}$), a deletion of a Serine residue at position 389, a $428_{L\rightarrow F}$, a $718_{S\rightarrow P}$ and a $723_{A\rightarrow T}$ change. Only the latter mutation mapped in the critically important M2 domain that defines the family and is required for the interaction with TONNEAU1 (Figure 3.6, Figure 3.S3). However, the amino acid change from Alanine to Threonine appeared to be a conservative replacement. Therefore, this or any of the other amino acid changes, which were also conservative, were not likely to affect the function of the protein. The sequencing of the allele in another accession of the same original seed lot showed the same results. In addition to four missense mutations and one amino acid deletion, we found two synonymous SNPs, two promoter SNPs, four 5'UTR SNPs, and four intron SNPs in and near *CaTRM25* among the parental lines (Figure 3.S3, Figure 3.6, Supplemental File 3.S2). We also sequenced the *CaTRM25* coding region in two Chile de Agua, two Costeño and two TusTav1 landraces (see Taitano *et al.* 2019 for clustering information, Figure 3.S3, Supplemental File 3.S2). Most of the missense and deletion mutations were only segregating among Chigole and Chile de Arbol except for $718_{S\rightarrow P}$ that is segregating in Costeño subpopulation. In addition, Chile de Agua and one TusTav1 accession featured a change from Serine to Cysteine at position 671. Maximum likelihood analysis of nonsynonymous vs. synonymous substitution rates returned a value of $\partial N/\partial S = 0.9469$. In the eight landraces combined, we found five missense, one amino acid deletion, and only two synonymous mutations corresponding to amino acids 583 and 519 and a total of 16 SNPs in noncoding regions (Figure 3.6).

The SNP underlying the *CaTRM25* $718_{S\rightarrow P}$ transition (19EP764) in Chile de Arbol conferring long fruit was also found in one Costeño (CRo) accession (Figure 3.S3). Therefore, we sought to determine if that allele was tightly associated with shape in additional pepper landraces described previously (Taitano *et al.,* 2019) (Figure 3.7A, B, Table 3.S3). About a third of the two largest *C. annuum* subpopulations, Chile de Agua and Costeño carried the Chile de Arbol *CaTRM25* SNP 718 allele. This allele was also present in the TusTav1 subpopulation. In other subpopulations of *C. annuum*, the Chile de Arbol allele was present at low frequency or not at all (Figure 3.7A), and thus no marker-trait association could be performed (Figure 3.7B). In the Chile de Agua subpopulation, the allele was not associated with fruit shape whereas in the Costeño and TusTav1, the allele was associated with fruit shape (Figure 3.7B). Combined, these results implied that the change in amino acid was not sufficient to explain differences in fruit shape as it was not associated with fruit shape in the Chile de Agua. Interestingly, the Chile de Arbol allele that conferred an elongated shape was rare in the Tusta-Taviche1 subpopulation (Figure 3.7) suggesting that the allele may have been introgressed from another subpopulation such as Costeño. The Chile de Arbol $340_{I\rightarrow T}$ allele was rare and limited to the Costeño and TusTav1 subpopulations (Figure 3.7C). With the caveat that the TusTav1 shape association is based strongly on observation of fruit shape from a single datapoint, the Chile de Arbol allele was significantly associated with shape in both the Costeño and TusTav1 subpopulations (Figure 3.7D).

C. annuum M8-containing TRM phylogeny

MAST search of the CM334 pepper genome for TRMs yielded 12 proteins bearing both the M2 and M8 domains (Figure 3.S4). The M2 domain defined the family as it is known to interact with TONNEAU1 (Drevensek *et al.,* 2012). The M8 domain defined the type of TRM that is known to interact with OVATE and other OFPs (Wu et al, 2018). Except for outgroup CA04g20490, which lacked the M8 domain, most of these pepper TRMs showed a one-to-one relationship with a single tomato TRM.

Discussion

CaTRM25 is a candidate gene for pepper fs3.1

The pepper fruit shape locus *fs3.1* was mapped to a region spanning most of chromosome 3 in 2016 (Han *et al.,* 2016). In this study, we confirmed the QTL and narrowed it down to a 27 Mb region spanning 4 cM. We obtained additional recombinants in the F_2 population. However, poor germination of several seed lots precluded a narrower fine mapping of *fs3.1*. In all, several candidate genes in the region were of note. The TEOSINTE-LIKE1, CYCLOIDEA, and PROLIFERATING CELL FACTOR1 (TCP) transcription factor family (*CA03g16800*) is part of a large family with a high degree of functional redundancy for controlling cell division in developing tissues (reviewed by Martín-Trillo *et al.,* 2010). Genes in the GA biosynthesis and auxin signaling pathway also are likely to control fruit shape (Serrani *et al.,* 2007a, Serrani *et al.,* 2007b, Bünger-Kibler *et al.,* 1982, Srivastava and Handa, 2005, De Jong *et al.,* 2009). Lastly, members of the SUN/IQD family are known to control organ shape in tomato and other vegetables (Xiao *et al*., 2008, Pan *et al.,* 2017, Dou *et al.,* 2018). However, the most interesting gene from this study was an ortholog of tomato TRM25. This is because the highest associated marker at the locus corresponds to *CaTRM25*. A recent publication also showed that a marker in the same gene was associated with fruit shape in a pepper GWAS population (Colonna *et al.,* 2019).

In tomato or any other species, the role of *TRM25* is not well understood. Most of the *TRMs* associated with shape are found in the Arabidopsis *TRM1-5* subclade (Lee *et al.,* 2006, Wu *et al*., 2018, Zhou *et al.,* 2015). CRISPR-Cas9 induced or natural nulls in this subclade of *TRMs* result in shorter organs whereas overexpression results in longer organs (Lee *et al.,* 2006, Wu *et al.,* 2018). In pepper, the expression results showed that higher expression led to a rounder fruit, a finding that contrasts with that of the *TRM1-5* mutants. While this may suggest that *CaTRM25* is not a likely candidate, recently studies from our lab have shown that different *TRM* nulls affect tomato fruit shape differently. CRISPR-Cas9 induced nulls in tomato *TRM19* in combination with *TRM17/20a* showed an elongated fruit which is opposite of that found in nulls in the *TRM1-5* clade (Zhang *et al*., in preparation). These results imply that fruit shape mediated by the TRMs could result in either rounder or longer fruit, depending on the specific TRM.

The pepper genome is large (3.5 Gb) due in part to a large expansion of transposons such as Gypsy and Copia retroelements (Kim *et al.,* 2014, Hulse-Kemp *et al.,* 2018). This is one of the reasons that the genome contains regions of poor assembly, hampering the discovery of genes that underlie important traits. The syntenically-informed analysis aided in the mapping of the large *fs3.1* QTL (Chaim *et al.,* 2001, Han *et al.,* 2016, Hill *et al.,* 2017) into what appears to be a single fruit-shape QTL and a strong candidate gene. The strongest-associated marker 18EP1221 located in the *CaTRM25* CDS was in a region of the pepper chromosome 3 syntenic with tomato chromosome 3 in the first of two major inversions relative to tomato (Rinaldi *et al.,* 2016). By using the syntenic approach in conjunction with all three published pepper reference genomes (Qin *et al.,* 2014, Kim *et al.,* 2014, Hulse-Kemp *et al.,* 2018), we were able to span these and other assembly discrepancies to narrow *fs3.1* to a 27 Mb region.

The *ft3.1* QTL on chromosome 3 showed a significant association with fruit weight in the F_2 population possibly due to pleiotropy at the locus, or linkage to another QTL. For example, the ortholog of tomato *fw3.2* maps near *ft3.1* in pepper and it is associated with weight in this species (Chakrabarti et al., 2013). While some fruit shape loci also affect fruit weight (Sun *et al.,* 2015), a commonly observed phenotype for mutants in TRM-OFP pathway genes such as *SlOFP20, SlTRM5,* or *OVATE* is a change in the plane of cell division leading to changes in fruit shape without differences in fruit weight (Wu *et al.,* 2018). For *fs3.1* on *CaTRM25*, there was no association with fruit weight in the F_2 and F_3 progeny tests, which is consistent with other findings that imply that TRMs only affect shape.

At the developmental and histological level, the changes in shape appear to be driven by changes in cell number occurring before anthesis during floral development. After anthesis, fruit growth continues to expand more along the proximo-distal axis in the lines carrying the Chile de Arbol allele of *fs3.1*. This developmental pattern mirrors the expression pattern of *CaTRM25*, which shows high expression early in floral development before anthesis, followed by a second increase of expression after pollination in the developing pericarp and placenta (Liu *et al.,* 2017). We might expect a disruption in *CaTRM25* expression to have noticeable effects at the stages where that gene is highest-expressed and most active.

Exploration of TRM diversity in landrace peppers

Null mutations in TRMs have been found to cause phenotypic changes in the formation of the PPB and the plane of cell division (Schaefer *et al.,* 2017), and may ultimately lead to changes in organ shape (Wu *et al.,* 2018). Therefore, the presence of *CaTRM25* at *fs3.1* is intriguing albeit that the QTL is still too large to positively identify this gene to underlie the QTL. In an effort to identify potential haplotypes and the underlying SNP that is causal to the effect on shape, we catalogued the genetic variation in a diversity population from near the center of pepper domestication (Taitano *et al.,* 2019). Variation in *CaTRM25* among the landrace population largely fell along lines predictable by their phylogenetic relationship to each other: mutant alleles were shared between more closely related plants (Figure 3.S3). For example, no alleles differed between the two plants from the highly homozygous Chile de Agua landrace. Furthermore, the two Chile de Agua also shared a substitution with the 14Tus179-1 Tusta accession found close to the Chile de Agua in prior population genetic analysis (Taitano *et al.,* 2019). Furthermore, internal controls recapitulated all homozygous mutant genotypes found in either of the two parents, and a heterozygous SNP from Chigole. Selection for blocky, round peppers in the elite Chile de Agua pepper varieties (Bosland and Votava, 2012) could act on background loci silencing the chromosome 3 fruit shape QTL, which may account for the difference seen between Chile de Agua and Costeño in the 19EP764-shape marker-trait association (Figure 3.7).

∂N/∂S analysis returned a ratio of 0.9469, very close to the value of 1 expected under neutral selection (Yang and Bielawski, 2000). However, this approach was limited by having closely related conspecific samples with relatively few SNPs within the coding region. Such populations can exhibit relatively lower ∂N/∂S values even under positive selection (Kryazhimskiy and Plotkin, 2008). Thus, the possibility remains that this locus is under diversifying selection rather than neutrality.

We also assessed pepper TRM diversity by cataloguing the TRM protein family in pepper and comparing against tomato TRMs. Protein tree analysis recapitulated in pepper the main clusters of M8-

bearing tomato TRMs (Figure 3.S4), including the TRM1-5 protein subclade (Lee *et al.,* 2006, Wu *et al.,* 2018). There were three M8-bearing tomato TRMs lacking orthologs in pepper: SlTRM30/34a, SlTRM13/14/15/33b, and SlTRM6/7/8b. In all three cases, the nearest tomato paralog is more closely related to its pepper ortholog. This suggests that these three proteins originated prior to the pepper-tomato speciation event. These proteins may have been disrupted during the rapid proliferation of transposons in the pepper genome (Kim *et al.,* 2014), or by the several genome rearrangements in pepper relative to tomato (Rinaldi *et al.,* 2016). However, all of the strongest OVATE-interacting TRMs in tomato - TRM3/4, TRM5, TRM26, TRM17/20a, and TRM19 (Wu *et al.,* 2018) - had clear one-to-one orthologs in pepper. Furthermore, all 10 of the M8-bearing pepper TRMs had orthologs in tomato. Together, these patterns are consistent with conservation of this developmental mechanism between the two species.

Conclusion

We report a narrowed fruit-shape QTL, *fs3.1*, on *C. annuum* chromosome 3, including a promising candidate gene that would explain multiple aspects of the phenotype, and which is part of a well-conserved fruit shape regulatory network. We examine the M8-bearing members of the TRM family in pepper, and find that they closely parallel those in tomato. Further research could include transgenic complementation tests and further fine-mapping of *fs3.1* with a focus around *TRM25* in pepper.

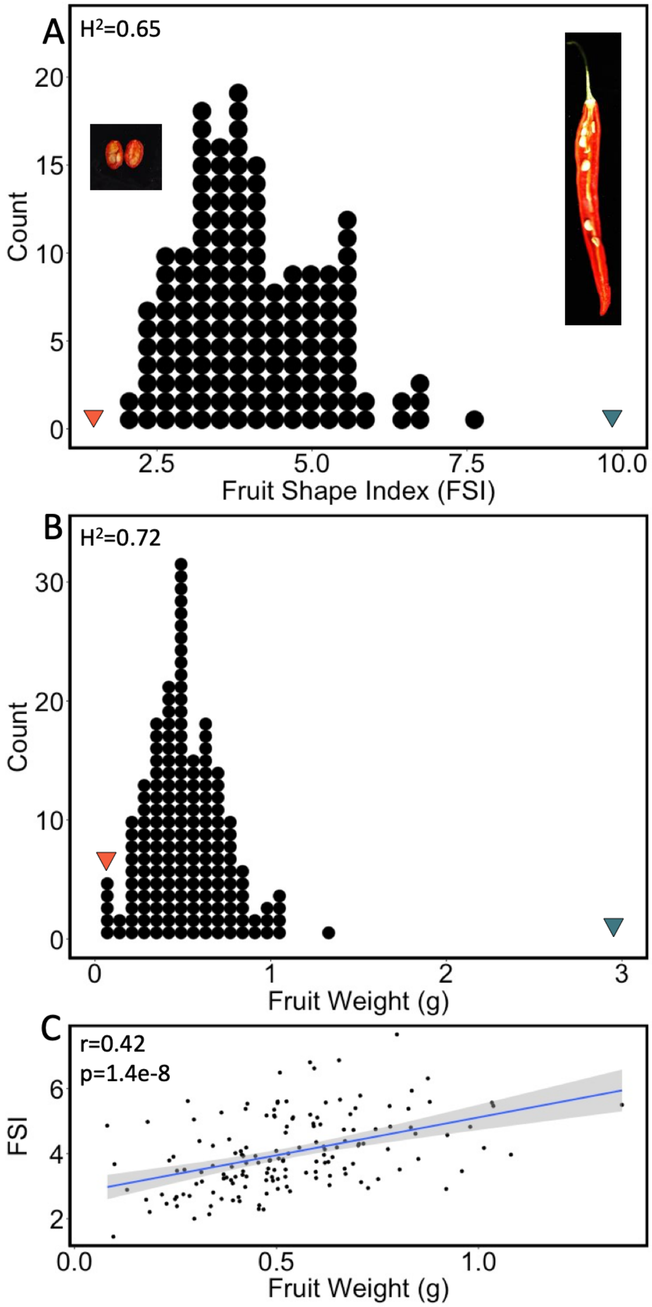

Figure 3.1. Distribution of fruit weight and shape in F₂. Histograms depicting the distribution of fruit shape (**A**) and fruit weight (**B**). Scanned pepper images of Chigole (left) and Chile de Arbol (right) included. Parent plant phenotypes depicted as colored triangles in each histogram, with orange triangles depicting the rounder and small-fruited Chigole (14P15-1) and teal triangles depicting the elongated and large-fruited Chile de Arbol (14P9-1) parental phenotypes. Broad-sense heritability estimates are displayed in the top-left corners of each trait's histogram. **C)** Scatterplot depicting a correlation between fruit weight (x-axis) and shape (y-axis) in the F_2. Line of best fit indicated in blue, with 95% confidence intervals shaded in gray and correlation statistics displayed at the top left corner.

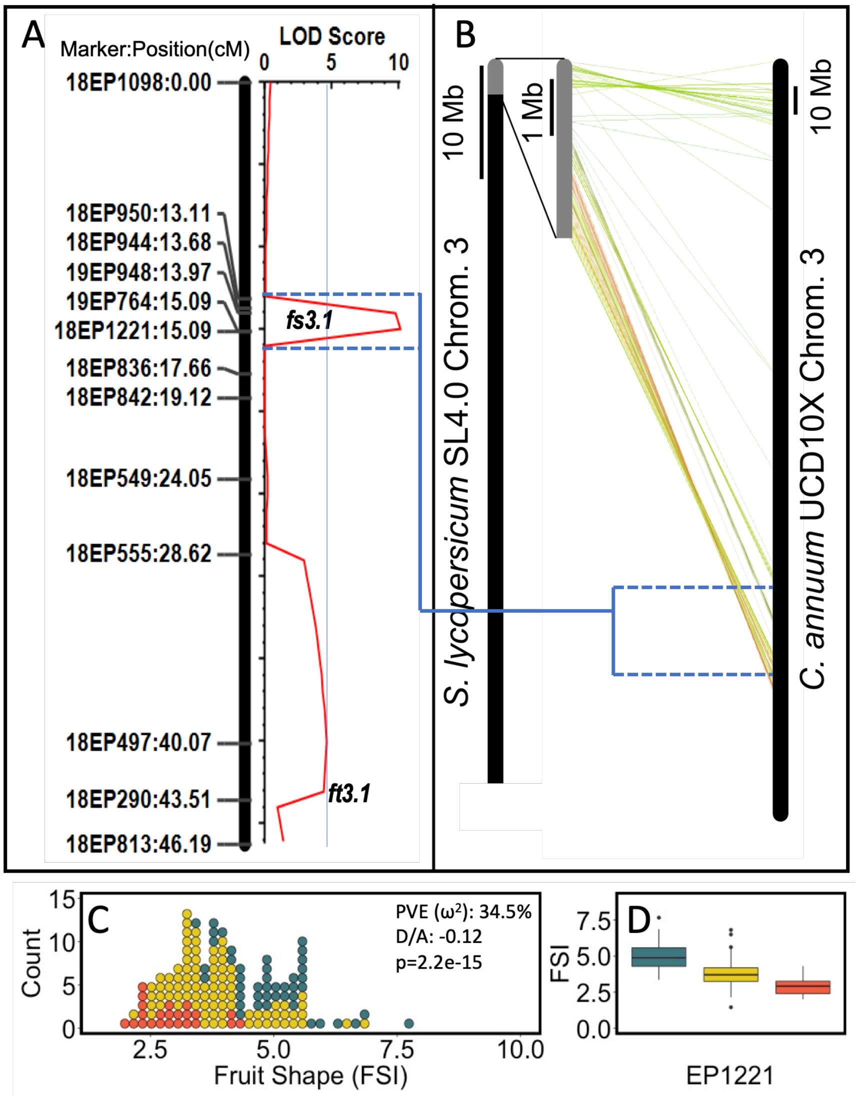

Figure 3.2. Synteny-informed mapping of chromosome 3 pepper fruit-shape QTL. A) Genetic map of the top of chromosome 3 and the association with fruit shape**.** The position of the markers is shown to the left of the chromosome diagram. The location of the previously mapped shape QTLs *ft3.1* and narrowed *fs3.1* are shown. The LOD scores for fruit shape index in the F$_2$ population are shown with the red line. **B)** Synteny between pepper UCD-10X-F$_1$ and tomato Heinz 1706 chromosome 3. Collinear blocks found by MCScanX in pepper for each tomato gene in the marker region shown by lines connecting the two chromosomes, with default MCScanX shading (randomly selected hues from green to orange) to differentiate synteny lines from one another. The blue dotted lines indicate the same interval on the genetic (A) and physical (B) map. **C)** Fruit-shape histogram as in Figure 3.1A, with each circle colored by its individual genotype. Colors correspond to Chigole-allele homozygous (orange), heterozygous (yellow), and Chile de Arbol-allele homozygous (teal) genotypes. The percent phenotypic variance explained, dominance effects (D/A), and marker-trait association are shown in the right corner. **D)** Boxplot depicting fruit shape index for each genetic group. Colors correspond to the Chile de Arbol-allele homozygous (teal), heterozygous (yellow), and Chigole-allele homozygous (orange) genotypes.

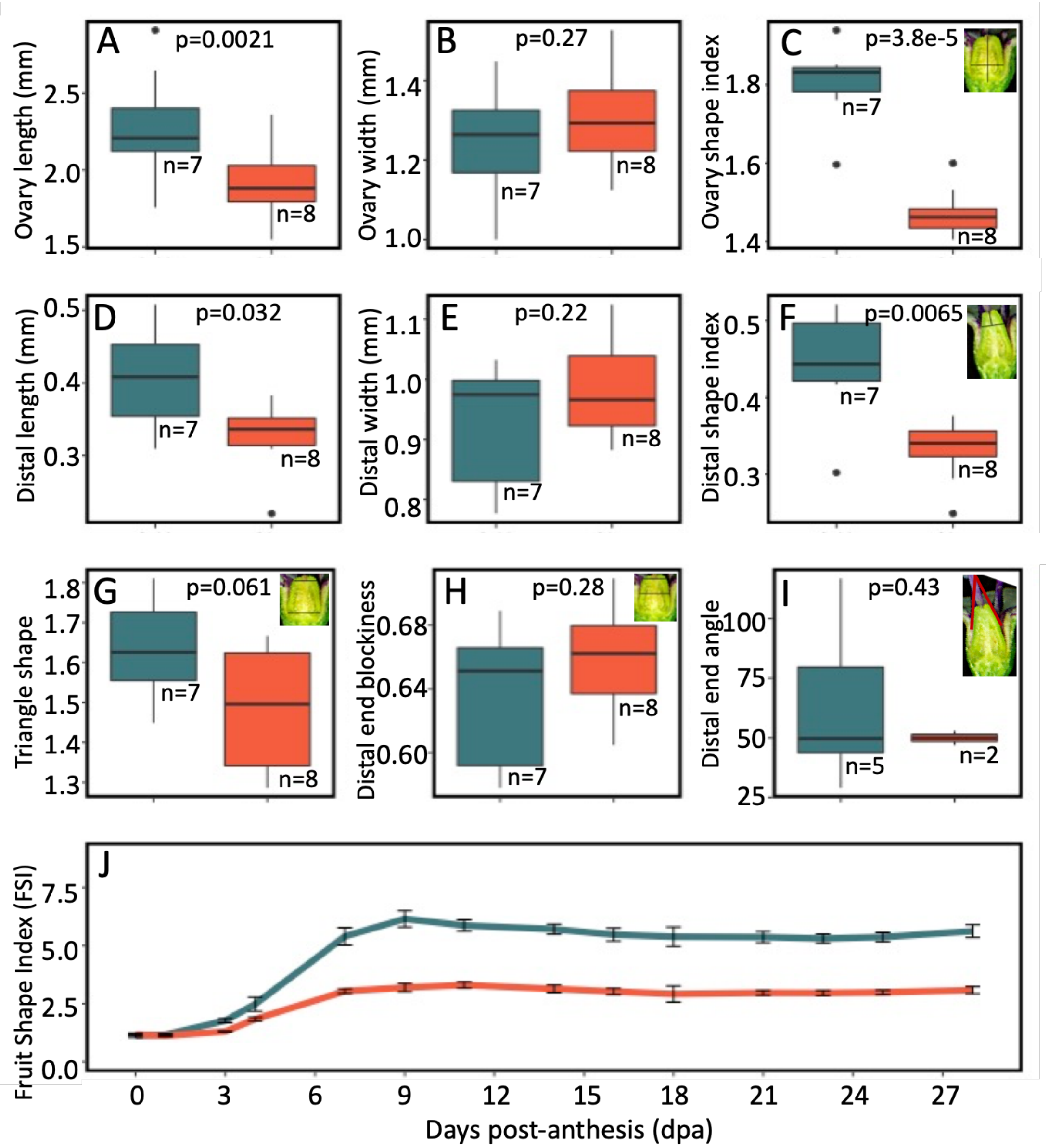

Figure 3.3. Fruit shape, controlled by *fs3.1,* of the ovary. Homozygotes of each genotype (marker 18EP1221) in F_3 population 19P1 were evaluated. **A)** Ovary length from base to distal tip. **B)** Ovary maximum width. **C)** Ovary shape index, as ratio of length:width. Inset depicts both ovary length and width measurements as black lines. **D)** Distal length from end of locules to distal tip. **E)** Distal width at distal end of locules. **F)** Distal end shape index, measured as ratio of distal length:width. Inset depicts distal tip length and width measurements as black lines. **G)** Triangle shape index, measured as ratio of widths at 10% and 90% ovary length. Example width lines depicted as black lines in inset. **H)** Distal end blockiness, measured as ratio of widths (depicted as black lines in inset) at 50% and 90% ovary length. **I)** Distal end angle macro, measured as the angle formed by the tangents (depicted as red lines in inset) 10% along the pericarp in either direction of the distal tip. A-I, P-values obtained from Bonferroni-corrected t-tests and degrees of freedom are shown at top center. **J)** Shape index of developing fruits in population 19P1 from anthesis until end of shape elongation. Colors correspond to Chile de Arbol-allele homozygous (teal) and Chigole-allele homozygous (orange) genotypes.

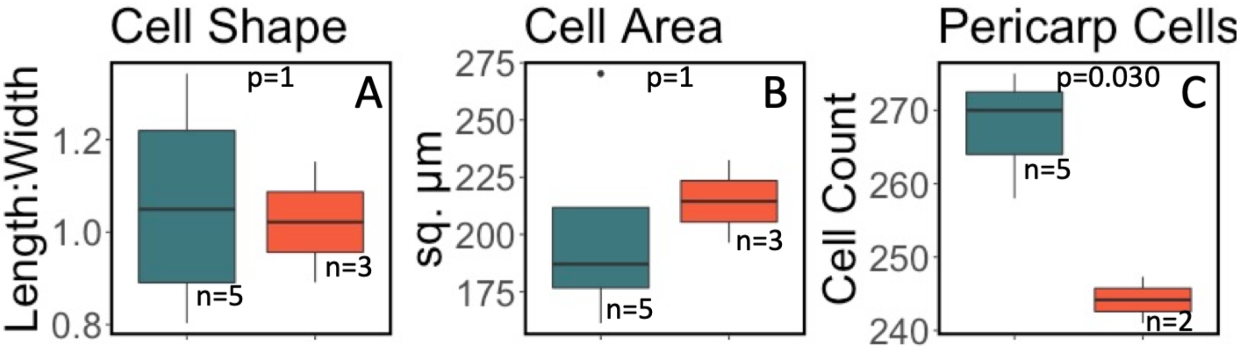

Figure 3.4. Histological differences in the ovary walls controlled by *fs3.1*. A) Cell shape index as measured by length (proximal-distal):width (medial-lateral) ratio. **B)** Cell area, averaged for 100x100µm squares in the center of each pericarp at mid-height of the ovary. **C)** Cell count along the pericarp, from pericarp base to distal tip. Colors correspond to Chile de Arbol-allele homozygous (teal) and Chigole-allele homozygous (orange) genotypes.

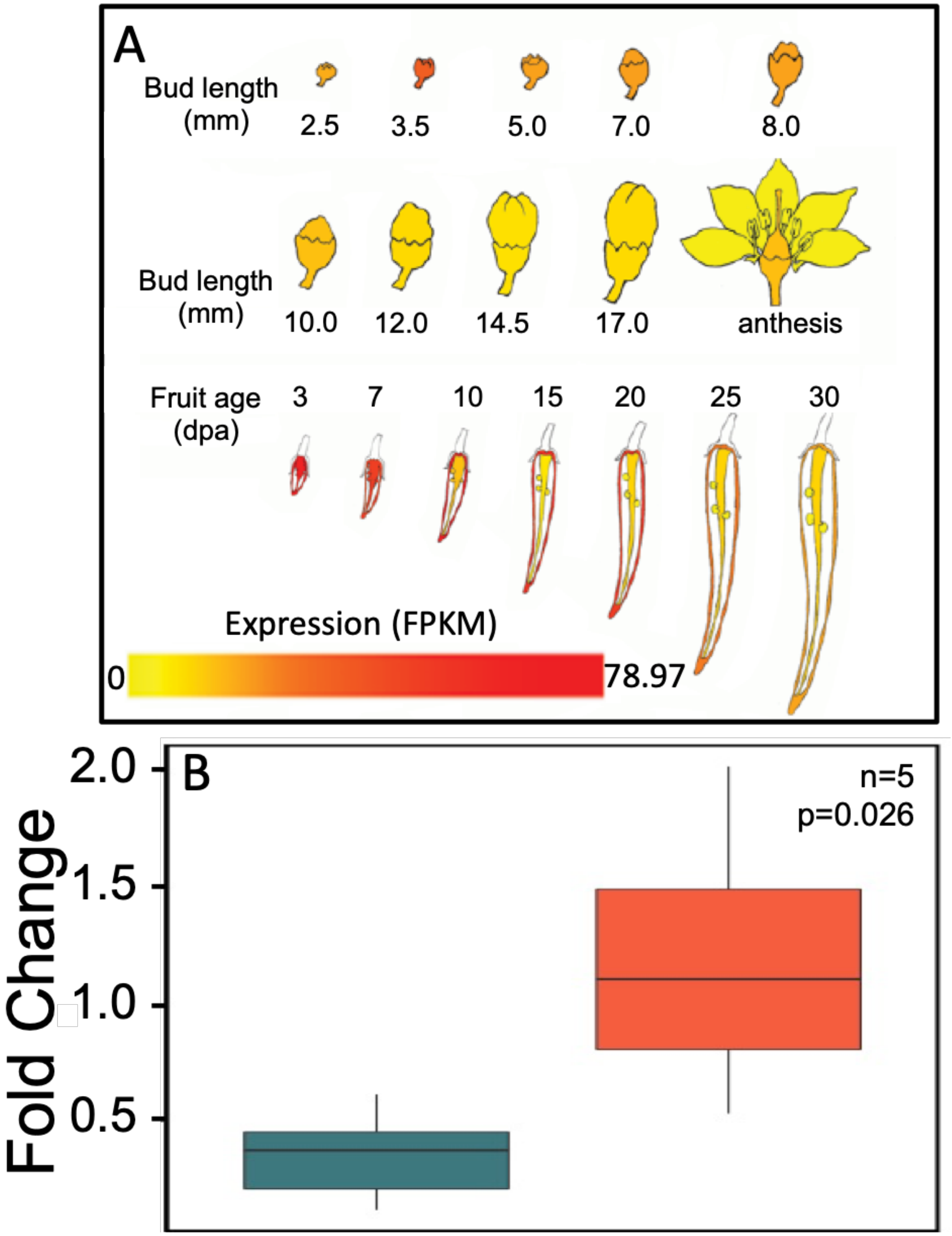

Figure 3.5. Expression of *CaTRM25*. A) *CaTRM25* expression pattern during floral and fruit development. Figure produced using the PepperHub expression database (Liu *et al.*, 2017) depicting the expression of *CaTRM25 (CA03g16080/Capana03g002426)*. Expression depicted as a color gradient from yellow (0 FPKM) to red (78.97 FPKM). **B)** Expression of *CaTRM25* in young ~2 mm flower buds. Expression of *CaTRM25* in buds collected from pepper plants of the 19P1 progeny population as a proportion of reference gene *CA10g11620/Capana10g001439*. Colors correspond to Chile de Arbol-allele homozygous (teal) and Chigole-allele homozygous (orange) genotypes.

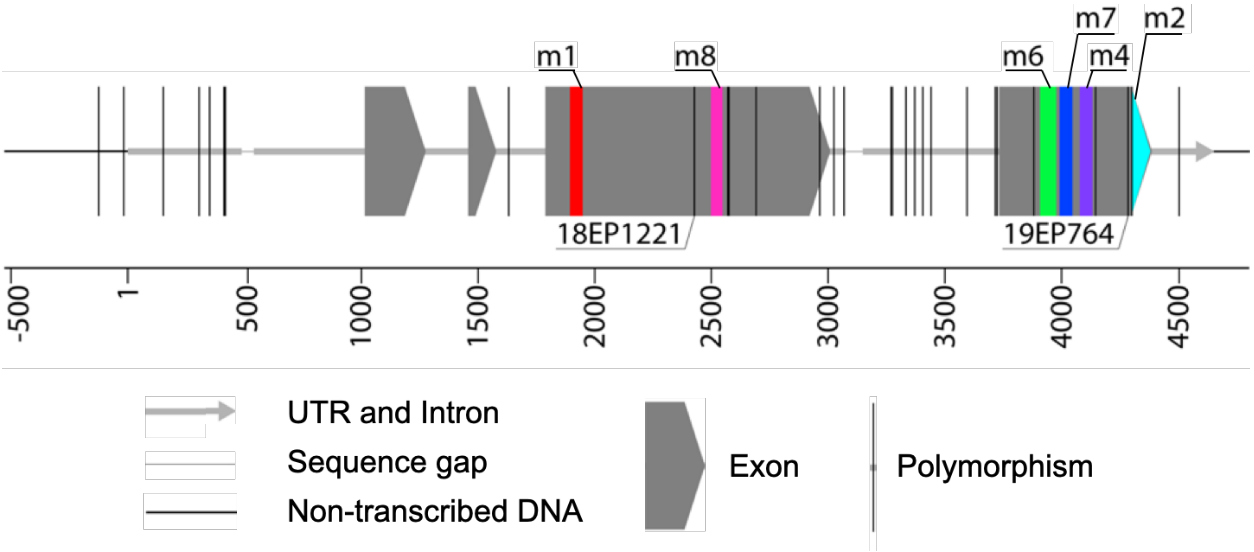

Figure 3.6. Diagram of *CaTRM25* nucleotide polymorphisms. Diagram depicting gene features found in 8 chile pepper landrace accessions including the two parents of the mapping population. Numbering begins with 1 as the transcription start site predicted by AUGUSTUS. The transcribed region is also shown as a light gray arrow pointing from 5' to 3'. Exons are shown as large dark gray arrows. Unknown bases (gaps in the CM334 genome assembly and our sequence data) are represented by thin grey horizontal line segments. Nucleotides encoding the known protein motifs are colored according to the legend above the diagram. A FASTA file with the complete nucleotide alignment is found in Supplemental File 3.S2.

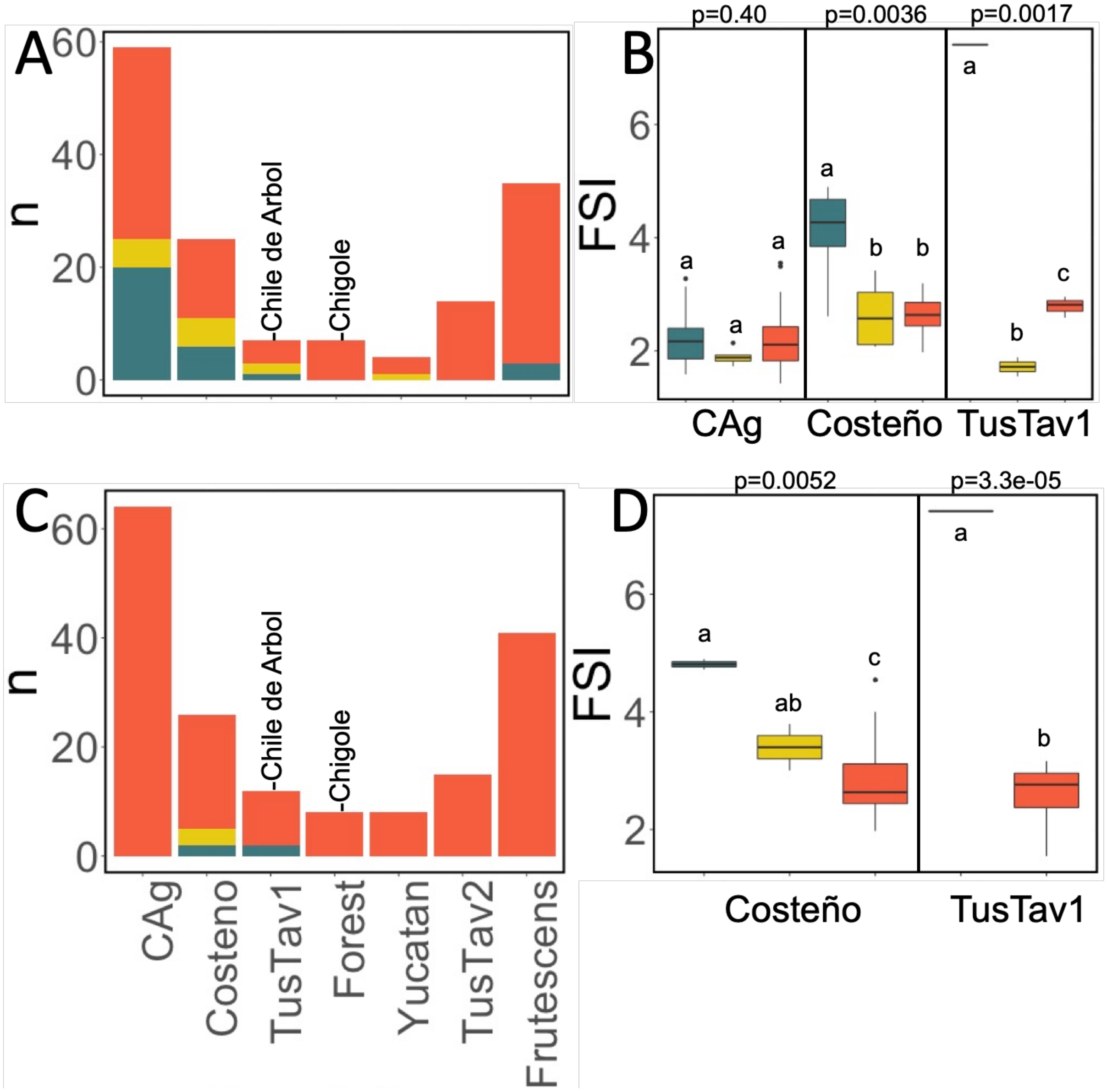

Figure 3.7. Pepper landrace genotypes of $718_{S\rightarrow P}$ and $340_{I\rightarrow T}$ in CaTRM25. A) Stacked barplot showing the distribution of the $718_{S\rightarrow P}$ alleles for each Mexican landrace subpopulation with categories following Taitano *et al.,* 2019. Parent populations labeled above subpopulation bar. Colors correspond to Chile de Arbol-allele homozygous (teal) and Chigole-allele homozygous (orange) genotypes. **B)** Boxplots displaying fruit-shape index for the three landrace subpopulations that carry sufficient number of plants with both $718_{S\rightarrow P}$ alleles. FDR-adjusted p-values displayed above each subpopulation. **C)** Stacked barplot showing the distribution of the $340_{I\rightarrow T}$ alleles as in (A). **D)** Boxplots displaying fruit-shape index for the two landraces with both alleles of the $340_{I\rightarrow T}$ substitution as in (B).

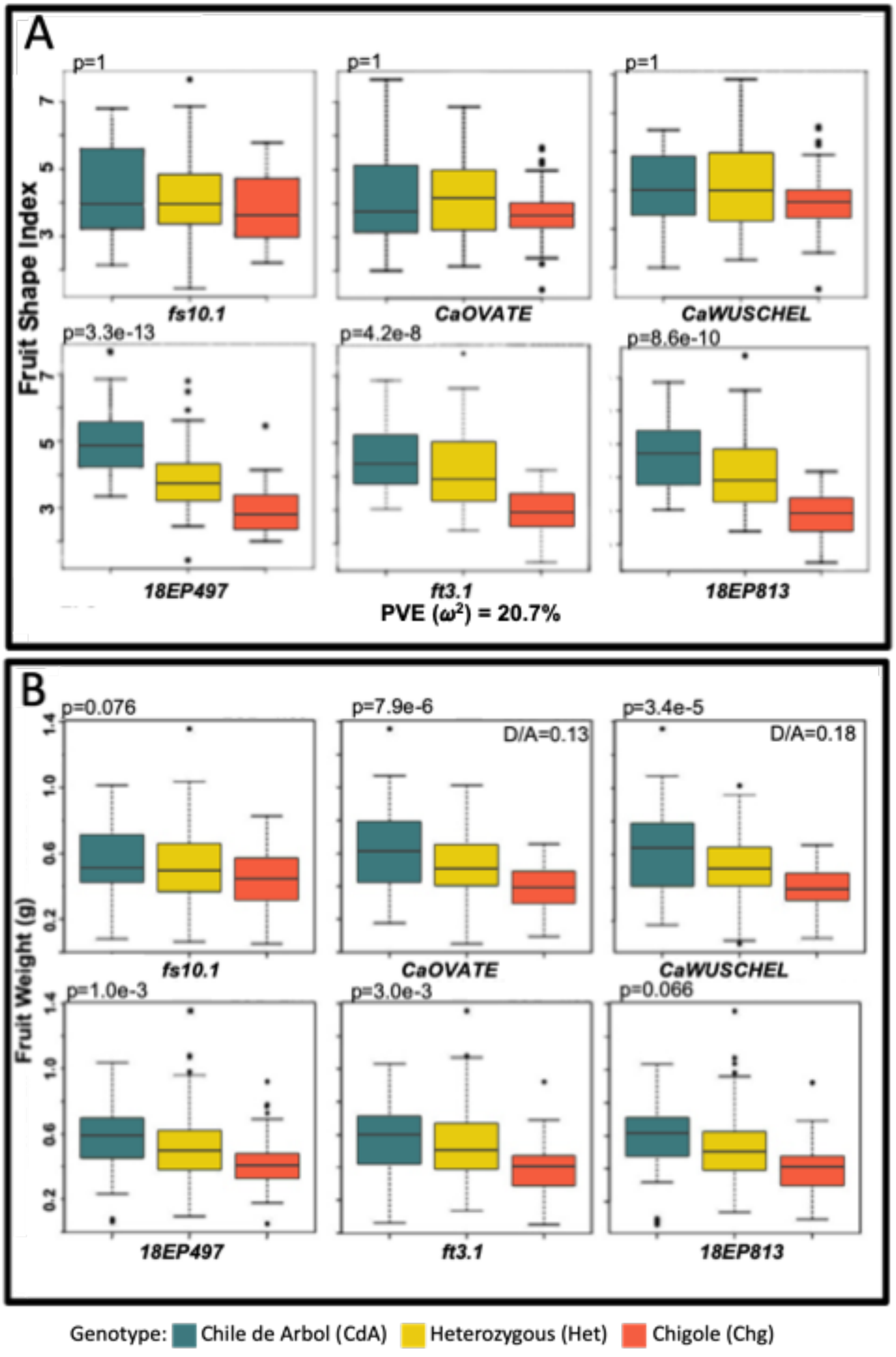

Figure 3.S1. Marker-trait association with known pepper fruit shape QTLs in the F$_2$ population. A) Marker association with fruit shape index. **B)** Marker association with fruit weight.

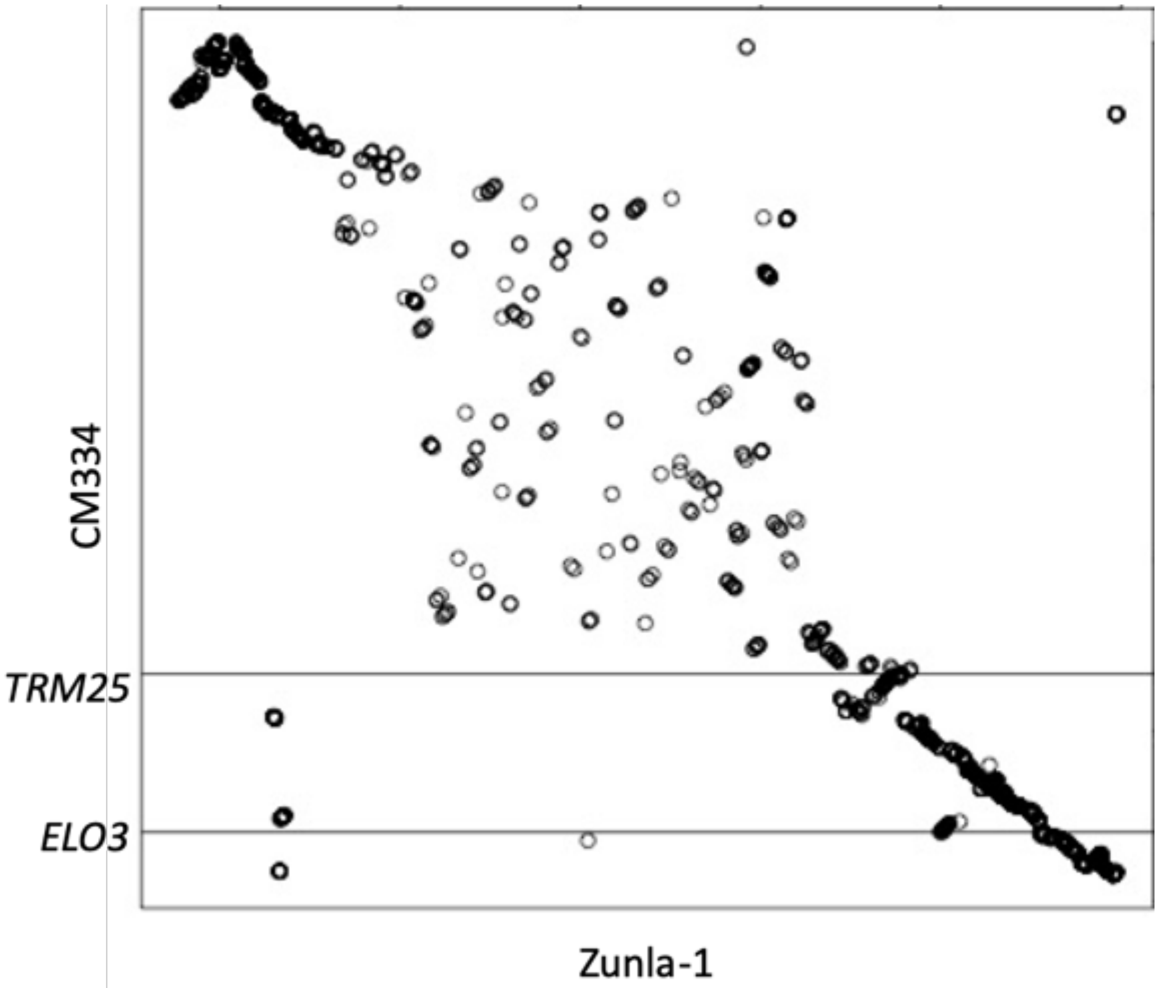

Figure 3.S2. Dot-plot of gene positions on CM334 vs. Zunla-1 chromosome 3 assemblies. Gene positions denoted by the circles plotted with CM334 positions on the y-axis and Zunla-1 positions on the x-axis, oriented with the top of chromosome 3 in the top left corner. Position of the candidate genes corresponding to *ft3.1*, *CaELONGATA3* (Chunthawodtiporn *et al.,* 2018) and *fs3.1*, *CaTRM25* (Han *et al.,* 2016) are shown. The genome sequence information is from previous studies (Qin *et al.*, 2014, Kim *et al.*, 2014).

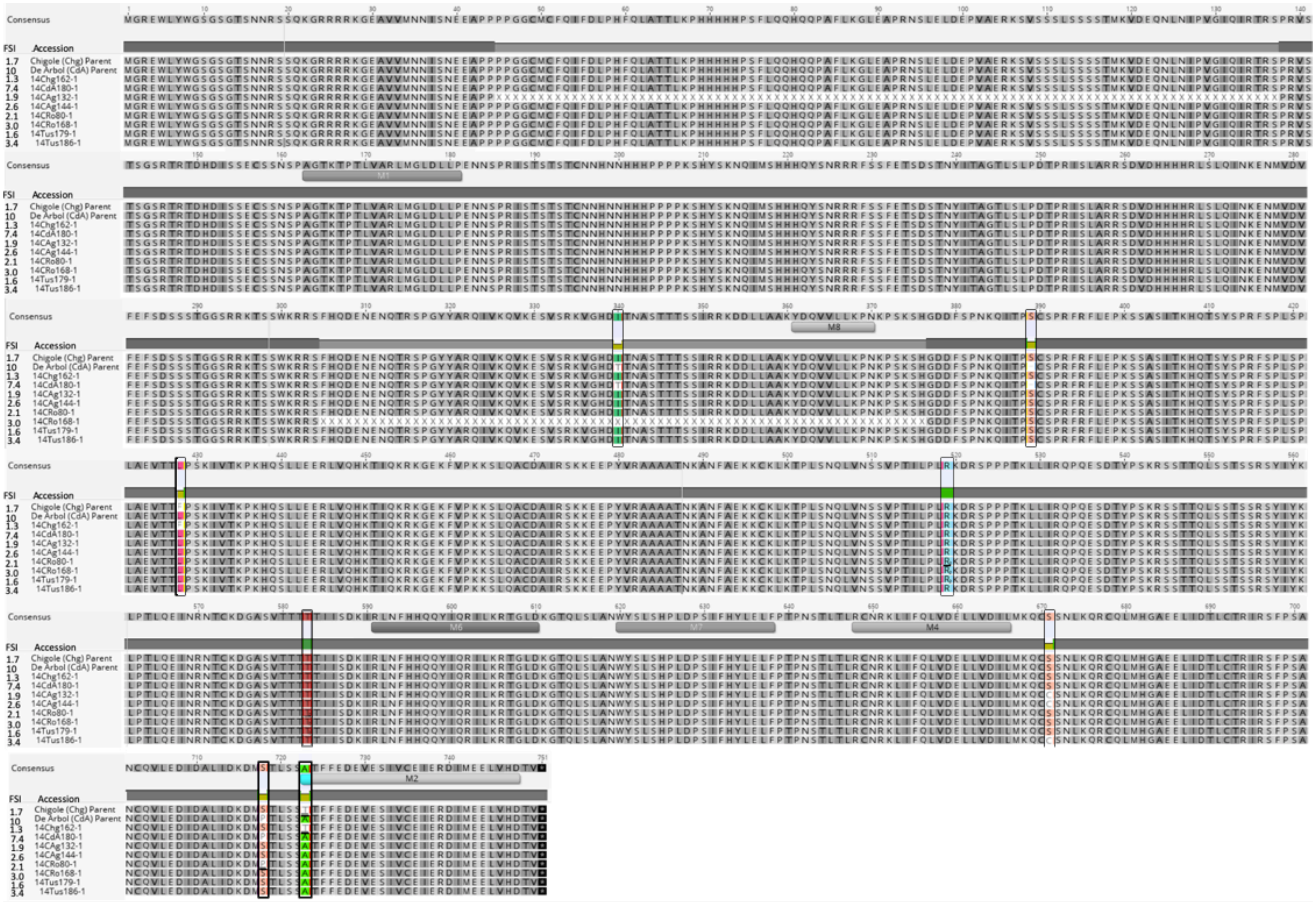

Figure 3.S3. ClustalW-aligned sequences of TRM25 in pepper landraces. *CaTRM25* from two each of Chigole (Chg), Chile de Arbol (CdA), Costeño (CRo), Chile de Agua (CAg), and two additional TusTav1 (Tus) were sequenced and translated from sequencing data in Supplemental File 3.S2. Amino acids are indicated by 1-letter code, with unknowns marked as X per IUPAC standards (IUPAC, 1984). Strings of X represent gaps in sequence data. Heterozygotes are represented by underlined amino acids, and synonymous mutations are represented by a subscript numeral 2.

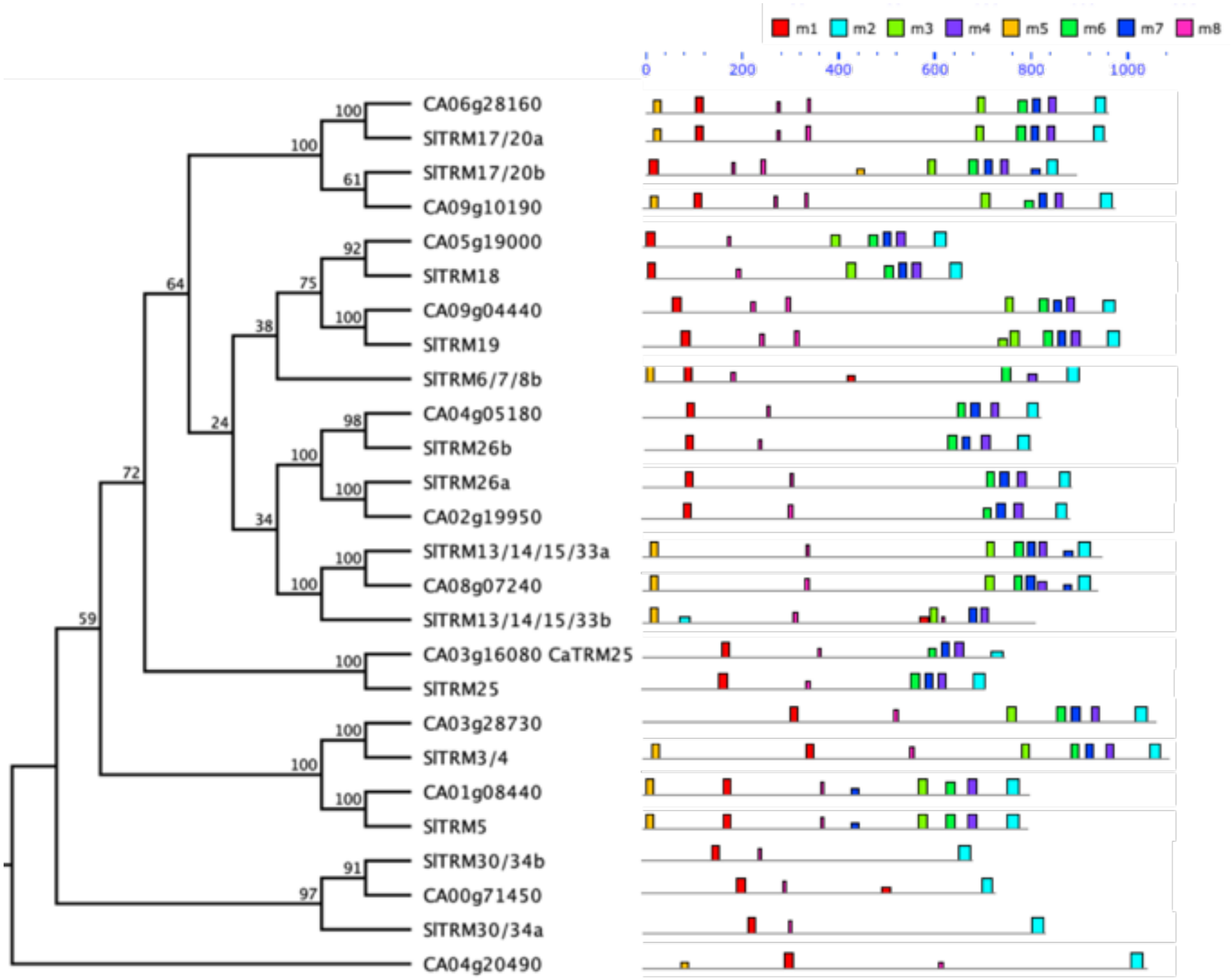

Figure 3.S4. Phylogenetic tree of a subset of CaTRM and SlTRM proteins. Phylogeny constructed from ClustalW-aligned TRMs that carry the M2 and M8 domains. Tomato TRM labels begin with "Sl", pepper TRM labels begin with "Ca". Percent bootstrap values shown at branch nodes. Protein structure, including TRM motifs (numbered as per Wu *et al.,* 2018) shown by diagram, with the height of each colored box denoting to the strength of the MAST hit for the corresponding motif, indicated by the legend.

Table 3.1. Fruit weight and shape marker-trait association. Marker-trait association performed on 178 F2 plants and several known fruit shape QTL markers.

Locus	**Gene marker**	**Chrom.**	**Trait**	**p-value**
fs10.1	*TG63*	10	weight	0.076
			shape	1
fw2.1	*CaOVATE*	2	weight	7.90E-06
			shape	1
fw2.1	*CaWUSCHEL*	2	weight	3.40E-05
			shape	1
18EP497		3	weight	1.00E-03
			shape	3.30E-13
ft3.1	*CaELONGATA*	3	weight	3.00E-03
			shape	4.20E-08
18EP813		3	weight	0.066
			shape	8.60E-10
fs3.1	*CaTRM25*	3	weight	0.43
			shape	1.20E-14

Table 3.2. F_3 progeny test results. Mean fruit shape index and fruit weight per genotype for homozygotes of each F3 family with number of plants shown in parentheses. Families segregating for 18EP1221 marker on CaTRM25 exclusively and without exception had genotypes differing significantly for fruit shape in bold. Het = heterozygous; Chg = chigole allele, CdA = Chile de Arbol allele. *p-values adjusted for multiple testing

Marker	**18EP950**	**18EP948**	**18EP 1221**	**18EP836**	**18EP842**	**18EP549**	**Mean FSI**		***p-value**	**Mean Weight (g)**		***p-value**
QTL			*TRM25*									
Family	0 cM	0.86 cM	2.03 cM	4.6 cM	6.06 cM	10.99 cM	CdA	Chg		CdA	Chg	
18P14	**Het**	**Het**	**Het**	**Het**	**Chg**	**Chg**	**4.94 (4)**	**3.27 (8)**	**0.019**	0.434 (3)	0.407 (8)	0.76
18P15	**Het**	**Het**	**Het**	**Het**	**Het**	**CdA**	**4.02 (3)**	**2.56 (2)**	**0.007**	0.557 (2)	0.400 (3)	0.07
18P16	**Het**	**Het**	**Het**	**Het**	**Het**	**CdA**	**5.9 (9)**	**3.44 (6)**	**2e-5**	0.684 (4)	0.918 (3)	0.17
18P17	Het	Het	Chg	Chg	Chg	Chg	2.22 (4)	2.17 (2)	0.757	0.295 (3)	0.360 (2)	0.77
18P19	Chg	Chg	Chg	Het	Het	Het	3.39 (6)	3.56 (5)	0.624	0.700 (5)	0.625 (5)	0.32
18P21	Chg	Chg	Chg	Het	Het	Het	3.5 (5)	3.48 (9)	0.859	0.564 (5)	0.559 (9)	0.91
18P25	CdA	CdA	CdA	Het	Het	Het	5.12 (7)	4.9 (4)	0.785	0.525 (7)	0.462 (4)	0.70
18P26	**Chg**	**Chg**	**Het**	**Het**	**Het**	**Het**	**4.79 (5)**	**3.14 (2)**	**2e-4**	0.643 (5)	0.723 (2)	0.83

Table 3.S1. Table of PCR primers. †Physical map positions in two genomes displayed for markers on the chromosome 3 QTL. *Primers used in Sanger sequencing are labeled with an ID matching the Sanger sequence data (Supplemental File 3.S2).

KASP Markers

Marker/gene name	*Marker ID*	*Sequence common primer*	*Sequence allele-specific primer1*	*Sequence allele-specific primer2*	*UCD10X ch3 Position*	*CM334 ch3 Position*
18EP1080	18EP1080	TTCAAAACGAAGCTCTAGGTGTCA	GAAGGTGACCAAGTTCATGCTCCCCTTCCCATCTCTTTTCG	GAAGGTCGGAGTCAACGGATTTCCCCTTCCCATCTCTTTTCA	19243792	18955782
18EP1083	18EP1083	CAACCCCATCAGGTCTCGTT	GAAGGTGACCAAGTTCATGCTGCTCGTACCAAGTTACTCTGTCCAC	GAAGGTCGGAGTCAACGGATTAGCTCGTACCAAGTTACTCTGTCCAT	19665009	19380563
18EP1098	18EP1098	CTTCGTTCCAGCCACACTGA	GAAGGTGACCAAGTTCATGCTAAAGAAGGAAGAGAGATGGAAGATTGT	GAAGGTCGGAGTCAACGGATTAAGAAGGAAGAGAGATGGAAGATTGA	27133905	23922523
Cafs3.1	18EP950	ATGGAGATCAGAAGGGCTCG	GAAGGTGACCAAGTTCATGCTTGCAATTTTCATTCGTCCAGTTACT	GAAGGTCGGAGTCAACGGATTGCAATTTTCATTCGTCCAGTTACA	199157999	193721492

18EP944	18EP944	TGCTTCTGCATTTGGACTGTCT	GAAGGTGACCAAGTTCATGCTTCCTCCGAAAGCACTTAATAACAAT	GAAGGTCGGAGTCAACGGATTTCCTCCGAAAGCACTTAATAACAAG	187008123	197064742
18EP948	18EP948	CCATTGTTCTTGTATGCGGGA	GAAGGTGACCAAGTTCATGCTTGCAATTTTCATTCGTCCAGTTACT	GAAGGTCGGAGTCAACGGATTGCAATTTTCATTCGTCCAGTTACA	199157931	193721424
CaTRM25	18EP1221	TGATGGTTTGTTTGGTTTGAGC	GAAGGTCGGAGTCAACGGATTCAGGAAAGTTGGCCATGACAC	GAAGGTGACCAAGTTCATGCTCAGGAAAGTTGGCCATGACAT	201972259	183386127
18EP836	18EP836	GAGCCATGGCCATATCCAGT	GAAGGTGACCAAGTTCATGCTGCTGCGACGTTTCTTGAGGTA	GAAGGTCGGAGTCAACGGATTCTGCGACGTTTCTTGAGGTG	213519984	201352601
18EP842	18EP842	GGTGCACTAACGCTCCCACT	GAAGGTGACCAAGTTCATGCTTGCGTACAAATACACCCTTGTGA	GAAGGTCGGAGTCAACGGATTGCGTACAAATACACCCTTGTGG	214753255	204524143
18EP549	18EP549	AGCATTGGAGGCTCTTCAGG	GAAGGTGACCAAGTTCATGCTCGACGGGCTGGGTCTCTT	GAAGGTCGGAGTCAACGGATTACGACGGGCTGGGTCTCTA	226312727	211375286

18EP555	18EP555	GAAAGCCATTCCTGTTGTAATCG	GAAGGTGACCAAGTTCATGCTTTGAGCATCTTTGTCAAGTCCAA	GAAGGTCGGAGTCAACGGATTCTTGAGCATCTTTGTCAAGTCCAT	230594194	216362841
18EP497	18EP497	TTCCAGCTGAAGTGGAGCAA	GAAGGTGACCAAGTTCATGCTATGTGAGGTTGAGTAACAGGAAAGTGT	GAAGGTCGGAGTCAACGGATTGTGAGGTTGAGTAACAGGAAAGTGG	238542312	225114417
Caft3.1	18EP290	TCATCCTTTACGCCAGACATGAG	GAAGGTGACCAAGTTCATGCTTCTAGATTGAGGGATCGCGGTT	GAAGGTCGGAGTCAACGGATTTCTAGATTGAGGGATCGCGGTA	UCD10X scf80788 pos49783	229456971
18EP813	18EP813	AATCCACTGCTCGCAAACAAC	GAAGGTGACCAAGTTCATGCTGATATTGCGACTATTTGTCCTGTTTCT	GAAGGTCGGAGTCAACGGATTATTGCGACTATTTGTCCTGTTTCC	243903101	232118790
Cafs10.1	17EP286	GTTGCGAACTTGAGAGGGTGT	GAGGTGACCAAGTTCATGCTGTAACAGTGTTGATCCTAACTTAAAAGAGTATG	GAAGGTCGGAGTCAACGGATTCAACAGTGTTGATCCTAACTTAAAAGAGTATA		
CaWUS	18EP370	CTCCAACCCATGTGAAATTTTG	GAAGGTGACCAAGTTCATGCACCAGGTGGTAGCTCAAACGA	GAAGGTCGGAGTCAACGGATTACCAGGTGGTAGCTCAAACGG		

CaOVATE	18EP373	GCCATAAAGCCGCCAACTACT	GAAGGTGACCAAGTTCATGCCAGCTGAAGACGTGCTCATTCTAAT	GAAGGTCGGAGTCAACGGATTCAGCTGAAGACGTGCTCATTCTAAC

qRT-PCR primers

Gene	*Forward Primer ID*	*Forward Primer*	*Reverse Primer*
CaTRM 25	19EP568	GTACTCAAACAGAAGGCGATTC	CTGACCTCCTGGCTAAAGATAT
CA10g1 1620	20EP166	ACATCAGGAATTGGATCCGT	GATGTAGGAACAGCCCGTTT

Sequencing primers

Primer Pair IDs*	*Forward Primer ID*	*Forward Primer*	*Reverse Primer*
af(fwd) - ar(rev)	19EP536	AAACTCCAAATATAATTCATAGACAAA	CGGTTCTACCCACTCACCTC
bf - br	19EP538	TCCAAATTGATGTTTCGTCTTG	GCTCGGTTCTACCCACTCAC
cf - cr	19EP540	GAGCTCAGTTCGTTGGGAAA	GTATAGTGGGAGGCGCAAAA

df - dr	19EP542	GCCATGTTGCTTGGACTTTT	TCCTCTGTCCTGTCCTGTCC
ef - er	19EP544	GGACAGGACAGGACAGAGGA	GAGGAGGTGGAGCTTCTTCA
ff - fr	19EP546	AACGCTCTAATTAAAACACACACA	AAACACATGCAGCCACCAG
gf - gr	19EP548	AAGAAAAGGAGAAGCAGTTGTCA	TTTGAATGCCCACCTGAAAT
hf - hr	19EP550	AAGGAGAAGCAGTTGTCATGAAT	ATTTGAATGCCCACCTGAAA
18EP1190 - 18EP1192	18EP1190	CGTTGTCGTCATCATCAACC	TGGGCATTCAAATCAGAACA
xf - xr	18EP1191	AGTACCCCGGACTTCTCGTT	CTGCCAACGGAGATAATGGT
if - ir	19EP676	TCAAGAACCCCATTAAATGAAAG	AGGGCTCGTTTGATACAAGG
kf - kr	19EP680	AGCTGCTAAATATGATCAAGTTGTCT	GAACATAACGATTAAGCAAGGAAAA

lf - lr	19EP682	AAACCGAGGAGTTCAAATGTT	GTTTGGGGTGGGAAAAAGTT
mf - mr	19EP684	TGTTGAACATAAGGAGGCAAGA	TCATCAAGATGTCAACGAGGA
nf - nr	19EP686	AGACTCAATTTTCATCATCAGCA	CAAATTCTCGATTTACCGCTTC
of - or	19EP688	AAAGGTACTCAATTGTCCTTAGCC	CATCTGCATCGACATGAACA

Table 3.S2. Table of orthologous gene annotations found for ORFs within *fs3.1*. List of gene annotations, from AUGUSTUS-predicted ORFs from the UCD10X assembly within the narrowed fs3.1 QTL, BLASTed against the SL4.0 genome, with SL4.0 gene IDs, positions, and annotations.

SL4.0 Gene ID	*SL4.0 Chrom*	*SL4.0 Start*	*SL4.0 End*	*Strand*	*Annotation*
Solyc03g005610	SL4.0ch03	448192	450667	-	Polyubiquitin (AHRD V3.3 *** A0A2P6TPU3_CHLSO)
Solyc03g006800	SL4.0ch03	1353615	1354595	+	TCP transcription factor 21
Solyc03g006810	SL4.0ch03	1357942	1359324	+	Peroxidase (AHRD V3.3 *** K4BEA4_SOLLC);Dbxref=InterPro:IPR002016,Pfam:PF00141;Ontology_term=GO:0004601,GO:0006979,GO:0020037,GO:0055114
Solyc03g006820	SL4.0ch03	1359981	1363744	-	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (AHRD V3.3 *** A0A2U1M9G2_ARTAN)
Solyc03g006840	SL4.0ch03	1391066	1393915	+	*TRM25*
Solyc03g006850	SL4.0ch03	1404608	1409513	-	junctophilin-like protein (AHRD V3.3 *** AT5G51840.1)
Solyc03g006860	SL4.0ch03	1409663	1414589	-	fructokinase;Dbxref=InterPro:IPR011611,Pfam:PF00294
Solyc03g006870	SL4.0ch03	1419978	1427817	+	Phosphoglucomutase protein (AHRD V3.3 *** A0A2U1KZG8_ARTAN)
Solyc03g006880	SL4.0ch03	1445090	1446880	+	gibberellin 20-oxidase-1
Solyc03g006890	SL4.0ch03	1459736	1461802	+	Protein kinase domain (AHRD V3.3 *** A0A200PMN1_9MAGN);Dbxref=InterPro:IPR000719,Pfam:PF00069;Ontology_term=GO:0004672,GO:0005524,GO:0006468
Solyc03g006900	SL4.0ch03	1476468	1479760	+	Protein kinase superfamily protein (AHRD V3.3 *** AT5G51800.1)

Solyc03g006960	SL4.0ch03	1554820	1557816	-	Protein phosphatase 2C (AHRD V3.3 *** A0A2U1PQC9_ARTAN);Dbxref=InterPro:IPR001932,Pfam:PF00481;Ontology_term=GO:0003824
Solyc03g006970	SL4.0ch03	1563229	1565556	-	serine protease SBT2;Dbxref=InterPro:IPR000209,Pfam:PF00082;Ontology_term=GO:0004252,GO:0006508
Solyc03g006980	SL4.0ch03	1575284	1578681	+	Alpha-L-fucosidase 1 (AHRD V3.3 *** A0A2G2X6A5_CAPBA);Dbxref=InterPro:IPR000933,Pfam:PF01120;Ontology_term=GO:0004560,GO:0005975
Solyc03g006990	SL4.0ch03	1580689	1582305	-	Pentatricopeptide repeat (AHRD V3.3 *** A0A200PMS4_9MAGN)
Solyc03g007000	SL4.0ch03	1582910	1588130	-	Peptidase M48 (AHRD V3.3 *** A0A200PMQ6_9MAGN)
Solyc03g007010	SL4.0ch03	1588607	1593861	-	DEAD-box ATP-dependent RNA helicase 7 (AHRD V3.3 *** A0A2G3CVX4_CAPCH)
Solyc03g007030	SL4.0ch03	1603306	1605907	-	CDGSH iron-sulfur domain-containing protein NEET (AHRD V3.3 *** A0A1U8HGA0_CAPAN)
Solyc03g007040	SL4.0ch03	1608523	1616753	+	K(+) efflux antiporter 5 (AHRD V3.3 *** A0A2G2X685_CAPBA);Dbxref=InterPro:IPR006153,Pfam:PF00999;Ontology_term=GO:0006812,GO:0015299,GO:0016021,GO:0055085
Solyc03g007050	SL4.0ch03	1644931	1651125	+	LRR receptor-like kinase family protein (AHRD V3.3 *** A0A072VP04_MEDTR)
Solyc03g007060	SL4.0ch03	1651432	1658197	-	Cell cycle regulated microtubule associated protein (AHRD V3.3 *** A0A1P8B7D6_ARATH)
Solyc03g007070	SL4.0ch03	1663228	1669454	+	1-aminocyclopropane-1-carboxylate synthase (AHRD V3.3 *** Q6U5H3_ROSHC);Dbxref=InterPro:IPR004839,Pfam:PF00155;Ontology_term=GO:0009058,GO:0030170
Solyc03g007170	SL4.0ch03	1775626	1780036	-	Peptidylprolyl isomerase (AHRD V3.3 *** A0A2G2X6E0_CAPBA)

Solyc03g007180	SL4.0ch03	1787580	1789847	+	hypothetical protein (AHRD V3.3 *** AT1G62870.1)
Solyc03g007190	SL4.0ch03	1796318	1801468	-	Hypersensitive-induced response protein 1 (AHRD V3.3 *** A0A1U8G9S8_CAPAN);Dbxref=InterPro:IPR001107,Pfam:PF01145
Solyc03g007890	SL4.0ch03	2435796	2439035	-	class 2 small heat shock protein Le-HSP17.6;Dbxref=InterPro:IPR001404,Pfam:PF00183;Ontology_term=GO:0005524,GO:0006457,GO:0051082
Solyc03g025800	SL4.0ch03	3187679	3188363	+	Heavy metal-associated isoprenylated plant protein (AHRD V3.3 *-* A0A2R6RWA7_ACTCH)
Solyc03g025810	SL4.0ch03	3194654	3197122	-	Low-temperature-induced 65 kDa protein (AHRD V3.3 *-* LTI65_ARATH)
Solyc03g025820	SL4.0ch03	3198336	3208315	+	shortage in chiasmata 1 (AHRD V3.3 *** AT5G52290.1)
Solyc03g025830	SL4.0ch03	3209483	3215674	-	Myosin heavy chain-related protein (AHRD V3.3 *** Q9C8T4_ARATH)
Solyc03g025840	SL4.0ch03	3225756	3228440	+	Cytochrome b561/ferric reductase transmembrane protein family (AHRD V3.3 *** A0A178UZ09_ARATH);Dbxref=InterPro:IPR006593,Pfam:PF03188
Solyc03g025850	SL4.0ch03	3234434	3236874	+	remorin 1;Dbxref=InterPro:IPR005516,Pfam:PF03763
Solyc03g025860	SL4.0ch03	3241399	3243551	+	25.3 kDa vesicle transport protein (AHRD V3.3 *** A0A2U1P9C4_ARTAN)
Solyc03g025870	SL4.0ch03	3243677	3246532	-	R2R3MYB transcription factor 19
Solyc03g025880	SL4.0ch03	3263743	3275700	-	Cleavage and polyadenylation specificity factor subunit 5 (AHRD V3.3 *** A0A1U8G7I4_CAPAN);Dbxref=InterPro:IPR016706,Pfam:PF13869;Ontology_term=GO:0003729,GO:0005849,GO:0006378
Solyc03g025890	SL4.0ch03	3278281	3298185	+	DNA mismatch repair protein MutS (AHRD V3.3 *** A0A200QUI6_9MAGN)
Solyc03g0	SL4.0ch03	3300895	3306926	-	Flap endonuclease GEN-like 2 (AHRD V3.3 *** A0A2G2X6K0_CAPBA)

25900					
Solyc03g025910	SL4.0ch03	3308354	3313469	-	transmembrane protein (AHRD V3.3 *** AT4G00585.1)
Solyc03g025920	SL4.0ch03	3316239	3322423	+	F-box/LRR-repeat protein (AHRD V3.3 *** A0A1U8G7I8_CAPAN)
Solyc03g025940	SL4.0ch03	3356056	3366257	+	F-box/LRR-repeat protein (AHRD V3.3 *** A0A1U8G7I8_CAPAN)
Solyc03g025950	SL4.0ch03	3370632	3373657	+	Membrane steroid-binding protein (AHRD V3.3 *** G5DXK7_SILLA)
Solyc03g025970	SL4.0ch03	3381260	3387208	+	Methyl-CpG binding domain protein (AHRD V3.3 *-* A0A2K3NPJ5_TRIPR)
Solyc03g025980	SL4.0ch03	3388820	3394274	+	Protein phosphatase inhibitor 2-like (AHRD V3.3 *** A0A2K3NNM1_TRIPR);Dbxref=InterPro:IPR007062,Pfam:PF04979;Ontology_term=GO:0004864,GO:0009966,GO:0043666
Solyc03g026000	SL4.0ch03	3398513	3400556	-	bromodomain-containing protein DDB_G0271118-like (AHRD V3.3 *** XP_021765153.1)
Solyc03g026010	SL4.0ch03	3407169	3408545	+	transmembrane protein 161AB protein (AHRD V3.3 *** AT5G52180.1)
Solyc03g026020	SL4.0ch03	3416861	3418590	+	Heat shock transcription factor (AHRD V3.3 *** E1U3C8_CICAR)
Solyc03g026030	SL4.0ch03	3421144	3422517	-	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (AHRD V3.3 *** Q9M144_ARATH)
Solyc03g026040	SL4.0ch03	3425975	3430043	-	Leucine-rich repeat receptor protein kinase (AHRD V3.3 *** A0A2R6RQU4_ACTCH)
Solyc03g026050	SL4.0ch03	3442049	3443000	+	Terminal flower 1 (AHRD V3.3 *** F7J0V7_GENTR);Dbxref=InterPro:IPR008914,Pfam:PF01161
Solyc03g026060	SL4.0ch03	3445262	3451905	-	Ankyrin repeat-containing protein (AHRD V3.3 *** A0A2U1MVN7_ARTAN)
Solyc03g026070	SL4.0ch03	3461959	3466620	-	Homeobox-leucine zipper protein ROC5 (AHRD V3.3 *** A0A2G2X6G6_CAPBA)

Solyc03g026080	SL4.0ch03	3492797	3495348	+	Rhomboid-like protein (AHRD V3.3 *** A0A2G2X6Z4_CAPBA)
Solyc03g026090	SL4.0ch03	3495610	3496311	+	Non-specific lipid-transfer protein (AHRD V3.3 *** A0A164YT40_DAUCA)
Solyc03g026100	SL4.0ch03	3498287	3499483	-	tRNA-specific adenosine deaminase TAD3 (AHRD V3.3 *** TAD3_ARATH)
Solyc03g026110	SL4.0ch03	3504661	3508609	-	SUN-like protein 8
Solyc03g026115	SL4.0ch03	3511442	3512421	-	Stamen-specific protein FIL1 (AHRD V3.3 *** A0A2G2ZYZ1_CAPAN)
Solyc03g026120	SL4.0ch03	3515207	3518214	-	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (AHRD V3.3 *** A0A2U1NM08_ARTAN);Dbxref=InterPro:IPR004159,Pfam:PF03141;Ontology_term=GO:0008168
Solyc03g026140	SL4.0ch03	3532390	3534106	-	Cytochrome P450 (AHRD V3.3 *** A0A2U1KE50_ARTAN);Dbxref=InterPro:IPR001128,Pfam:PF00067;Ontology_term=GO:0005506,GO:0016705,GO:0020037,GO:0055114
Solyc03g026150	SL4.0ch03	3541115	3548288	+	RING/U-box superfamily protein (AHRD V3.3 *-* A0A2U1NQN4_ARTAN)
Solyc03g026170	SL4.0ch03	3561604	3565474	+	F-box family protein (AHRD V3.3 *** A0A2U1QK66_ARTAN)
Solyc03g026190	SL4.0ch03	3570257	3575804	-	Transcriptional corepressor SEUSS (AHRD V3.3 *-* A0A2I0WPA8_9ASPA)
Solyc03g026200	SL4.0ch03	3586933	3592441	+	protein COFACTOR ASSEMBLY OF COMPLEX C SUBUNIT B CCB2, chloroplastic (AHRD V3.3 *** A0A2I4GK97_9ROSI);Dbxref=InterPro:IPR021325,Pfam:PF11152
Solyc03g026210	SL4.0ch03	3592715	3596845	-	Dihydrodipicolinate reductase-like protein (AHRD V3.3 *** A0A0F7GZT6_9ROSI)

Solyc03g026220	SL4.0ch03	3614792	3618046	+	BAG family molecular chaperone regulator 1 (AHRD V3.3 *** A0A2G2ZYV9_CAPAN)
Solyc03g026230	SL4.0ch03	3652866	3654449	-	Protein DETOXIFICATION (AHRD V3.3 *** A0A1Q3B1C8_CEPFO)
Solyc03g026240	SL4.0ch03	3662711	3667616	-	Serine/arginine-rich-splicing factor RS40 (AHRD V3.3 *** A0A0M8KTC1_NICAT)
Solyc03g026250	SL4.0ch03	3673085	3676092	-	Two-component response regulator (AHRD V3.3 *-* A0A0K9QUN4_SPIOL)
Solyc03g026260	SL4.0ch03	3689190	3692210	-	TraB family protein (AHRD V3.3 *** A0A2U1PQE1_ARTAN);Dbxref=InterPro:IPR002816,Pfam:PF01963
Solyc03g026270	SL4.0ch03	3699231	3699848	+	CRT binding factor 3 (AHRD V3.3 *** B3TPP7_SOLPI)
Solyc03g026280	SL4.0ch03	3702349	3703245	+	C-repeat binding factor 1
Solyc03g026290	SL4.0ch03	3707585	3716068	-	Per1-like (AHRD V3.3 *** A0A200PW77_9MAGN);Dbxref=InterPro:IPR007217,Pfam:PF04080
Solyc03g026300	SL4.0ch03	3718652	3724663	-	Per1-like (AHRD V3.3 *** A0A200PW77_9MAGN);Dbxref=InterPro:IPR007217,Pfam:PF04080
Solyc03g026310	SL4.0ch03	3729882	3740631	+	ABC transporter (AHRD V3.3 *** A0A200R0A6_9MAGN)
Solyc03g026330	SL4.0ch03	3755248	3755967	-	ATP-dependent Clp protease ATP-binding subunit (AHRD V3.3 *-* AT5G62140.1)
Solyc03g026340	SL4.0ch03	3760278	3764142	-	Calcium-dependent protein kinase 4 (AHRD V3.3 *** A0A2G2ZYX8_CAPAN);Dbxref=InterPro:IPR000719,Pfam:PF00069;Ontology_term=GO:0004672,GO:0005524,GO:0006468
Solyc03g026350	SL4.0ch03	3770590	3778731	+	Zinc finger transcription factor 24
Solyc03g026360	SL4.0ch03	3782767	3783123	+	Peptidoglycan-binding LysM domain-containing protein (AHRD V3.3 *** A0A2U1MKY7_ARTAN)
Solyc03g0	SL4.0ch03	3892513	3894288	+	molybdenum cofactor sulfurase (AHRD V3.3 *** A0A2I4FGR8_9ROSI)

31420					
Solyc03g031430	SL4.0ch03	3899081	3899809	+	urease accessory protein F;Dbxref=InterPro:IPR002639,Pfam:PF01730;Ontology_term=GO:0006807,GO:0016151
Solyc03g034207	SL4.0ch03	5762757	5764044	-	K-box region protein (DUF1985) (AHRD V3.3 -** AT1G31150.2)
Solyc03g070470	SL4.0ch03	13179475	13185907	+	ARF guanine-nucleotide exchange factor GNL2 (AHRD V3.3 *** A0A1U8EYH6_CAPAN)
Solyc03g078330	SL4.0ch03	45102298	45103725	-	3-ketoacyl-CoA synthase (AHRD V3.3 *** A0A2G2VZG2_CAPBA);Dbxref=InterPro:IPR013601,Pfam:PF08392;Ontology_term=GO:0006633,GO:0016020,GO:0016747
Solyc03g078360	SL4.0ch03	45257359	45259737	+	Receptor-like protein kinase (AHRD V3.3 *** A0A2U1PPA0_ARTAN)
Solyc03g078370	SL4.0ch03	45260411	45262780	+	Receptor-like protein kinase (AHRD V3.3 *** A0A2U1PPA0_ARTAN)
Solyc03g078623	SL4.0ch03	45770134	45771247	+	UBN2_3 domain-containing protein (AHRD V3.3 *-* A0A1Q3AQH1_CEPFO)
Solyc03g082860	SL4.0ch03	47321635	47324948	+	Histone-lysine N-methyltransferase ATXR6-like protein (AHRD V3.3 *** A0A2K3P9S3_TRIPR)
Solyc03g083130	SL4.0ch03	47527947	47542762	+	Protein DEFECTIVE IN MERISTEM SILENCING 3 (AHRD V3.3 *** A0A1J3I4J6_NOCCA)
Solyc03g083490	SL4.0ch03	47993147	47993957	+	DUF1442 domain-containing protein (AHRD V3.3 *** A0A1Q3BAQ3_CEPFO);Dbxref=InterPro:IPR009902,Pfam:PF07279
Solyc03g083500	SL4.0ch03	48001926	48002811	+	DUF1442 domain-containing protein (AHRD V3.3 *** A0A1Q3BAQ3_CEPFO);Dbxref=InterPro:IPR009902,Pfam:PF07279
Solyc03g083520	SL4.0ch03	48040751	48045215	+	Calmodulin (AHRD V3.3 *** B6T4U8_MAIZE)
Solyc03g083540	SL4.0ch03	48058586	48065725	+	Chloride conductance regulatory protein ICln (AHRD V3.3 *** A0A2G3CVQ4_CAPCH);Dbxref=InterPro:IPR039924,Pfam:PF03517

Solyc03g083550	SL4.0ch03	48066469	48067689	-	UDP-glucuronate 4-epimerase 4 (AHRD V3.3 *** A0A1J3F7T2_NOCCA);Dbxref=InterPro:IPR001509,Pfam:PF01370;Ontology_term=GO:0003824,GO:0050662
Solyc03g083560	SL4.0ch03	48088322	48089371	+	Protein EXORDIUM-like 7 (AHRD V3.3 *** A0A2G2X692_CAPBA);Dbxref=InterPro:IPR006766,Pfam:PF04674
Solyc03g083580	SL4.0ch03	48107954	48121274	-	Peptidase_M3 domain-containing protein (AHRD V3.3 *** A0A1Q3BAD5_CEPFO);Dbxref=InterPro:IPR001567,Pfam:PF01432;Ontology_term=GO:0004222,GO:0006508
Solyc03g083590	SL4.0ch03	48122236	48133054	+	Protein kinase domain (AHRD V3.3 *** A0A200RA50_9MAGN)
Solyc03g083610	SL4.0ch03	48141092	48144626	+	UMP-CMP kinase (AHRD V3.3 *** A0A2G2ZYL4_CAPAN),Pfam:PF00406
Solyc03g083650	SL4.0ch03	48177406	48180886	-	Glycosyltransferase (AHRD V3.3 *** U5NH37_CATRO)
Solyc03g083710	SL4.0ch03	48226903	48227490	-	Plant invertase/pectin methylesterase inhibitor superfamily protein (AHRD V3.3 *** A0A2U1LKC0_ARTAN);Dbxref=InterPro:IPR006501,Pfam:PF04043;Ontology_term=GO:0004857
Solyc03g083720	SL4.0ch03	48230029	48230658	-	Plant invertase/pectin methylesterase inhibitor superfamily protein, putative (AHRD V3.3 *** A0A061G4U8_THECC);Dbxref=InterPro:IPR006501,Pfam:PF04043;Ontology_term=GO:0004857
Solyc03g083770	SL4.0ch03	48279855	48280457	+	Plant invertase/pectin methylesterase inhibitor superfamily protein, putative (AHRD V3.3 *** A0A061G4U8_THECC);Dbxref=InterPro:IPR006501,Pfam:PF04043;Ontology_term=GO:0004857

Solyc03g083820	SL4.0ch03	48294863	48297090	-	Endoglucanase (AHRD V3.3 *** A0A2G3CVU8_CAPCH);Dbxref=InterPro:IPR001701,Pfam:PF00759;Ontology_term=GO:0004553,GO:0005975
Solyc03g083840	SL4.0ch03	48334113	48336839	+	Pectinesterase (AHRD V3.3 *** A0A1U8G949_CAPAN);Dbxref=InterPro:IPR000070,Pfam:PF01095;Ontology_term=GO:0030599,GO:0042545
Solyc03g083880	SL4.0ch03	48394009	48396858	+	NAC domain-containing protein (AHRD V3.3 *** A0A2U1LJ98_ARTAN)
Solyc03g083900	SL4.0ch03	48403467	48407475	-	monocopper oxidase-like protein SKU5 (AHRD V3.3 *** A0A2I4GW38_9ROSI)
Solyc03g083910	SL4.0ch03	48429643	48433972	-	sucrose accumulator
Solyc03g083960	SL4.0ch03	48480502	48483514	-	Trehalose 6-phosphate phosphatase (AHRD V3.3 *** A0A1U8G0B4_CAPAN);Dbxref=InterPro:IPR003337,Pfam:PF02358;Ontology_term=GO:0003824,GO:0005992
Solyc03g083970	SL4.0ch03	48506932	48509768	-	BAG family molecular chaperone regulator 7-like (AHRD V3.3 *** A0A2I4H718_9ROSI)
Solyc03g083980	SL4.0ch03	48511608	48512168	-	Unknown protein
Solyc03g093140	SL4.0ch03	48876374	48879286	-	Glycerol-3-phosphate transporter 1-like protein (AHRD V3.3 *** A0A2K3PR18_TRIPR);Dbxref=InterPro:IPR011701,Pfam:PF07690;Ontology_term=GO:0016021,GO:0055085
Solyc03g093150	SL4.0ch03	48909548	48912490	-	Pentatricopeptide repeat-containing protein (AHRD V3.3 *** A0A2U1NTS8_ARTAN)
Solyc03g093160	SL4.0ch03	48919621	48925891	-	LUC7-related (AHRD V3.3 *** A0A200QTT4_9MAGN);Dbxref=InterPro:IPR004882,Pfam:PF03194;Ontology_term=GO:0003729,GO:0005685,GO:0006376

Solyc03g093180	SL4.0ch03	48941259	48941933	+	Peroxisomal membrane protein 11-4 (AHRD V3.3 *** A0A2G3CVP3_CAPCH);Dbxref=InterPro:IPR008733,Pfam:PF05648;Ontology_term=GO:0005779,GO:0016559
Solyc03g093200	SL4.0ch03	48949050	48955415	+	Cell number regulator 6 (AHRD V3.3 *** A0A2G2X665_CAPBA);Dbxref=InterPro:IPR006461,Pfam:PF04749
Solyc03g093210	SL4.0ch03	48958868	48964403	+	Nucleic acid-binding, OB-fold protein (AHRD V3.3 *** A0A2U1P2W4_ARTAN)
Solyc03g093220	SL4.0ch03	48966372	48969389	-	transmembrane protein (AHRD V3.3 *** AT3G42150.3)
Solyc03g093230	SL4.0ch03	48983364	48984678	+	tonoplast intrinsic protein 5.1;Dbxref=InterPro:IPR000425,Pfam:PF00230;Ontology_term=GO:0015267,GO:0016020,GO:0055085
Solyc03g093240	SL4.0ch03	48986176	48992380	+	nitric oxide synthase 1
Solyc03g093270	SL4.0ch03	49011703	49013733	+	NAD(P)-linked oxidoreductase superfamily protein (AHRD V3.3 *** A0A2U1L511_ARTAN);Dbxref=InterPro:IPR023210,Pfam:PF00248
Solyc03g093290	SL4.0ch03	49026067	49027370	+	Aldo/keto reductase family protein (AHRD V3.3 *** Q9FJK0_ARATH)
Solyc03g096305	SL4.0ch03	52947156	52951712	-	4-hydroxy-3-methylbut-2-enyl diphosphate reductase, chloroplastic (AHRD V3.3 *-* A0A2G2W0B0_CAPBA)
Solyc03g096385	SL4.0ch03	53027606	53028379	-	UBN2_2 domain-containing protein (AHRD V3.3 *** A0A1Q3AQU5_CEPFO),Pfam:PF14223
Solyc03g110880	SL4.0ch03	56013746	56030347	+	RNA polymerase 4 second largest subunit (RPD2)
Solyc03g111500	SL4.0ch03	56512646	56519803	-	B3 domain-containing protein (AHRD V3.3 *** A0A1Q3B6Y6_CEPFO)
Solyc03g117520	SL4.0ch03	61108642	61110498	+	Protein FAR1-RELATED SEQUENCE 6 (AHRD V3.3 *** A0A2G2WQF9_CAPBA)
Solyc00g500003	SL4.0ch00	311496	382066	-	MP domain-containing protein (AHRD V3.3 *-* A0A1Q3D0H5_CEPFO)

Solyc00g500147	SL4.0ch00	2564107	2566997	-	Protein Ycf2 (AHRD V3.3 *-* A0A0U1ZJJ5_9SOLN)
Solyc00g500167	SL4.0ch00	3011082	3012921	+	Cytochrome c biogenesis FC (AHRD V3.3 *-* A0A0C5APX0_HYONI)
Solyc01g021700	SL4.0ch01	32625286	32627233	+	Photosystem II CP43 reaction center protein (AHRD V3.3 *-* A0A2G2WUM4_CAPBA)
Solyc01g057740	SL4.0ch01	56780600	56781448	-	Ribonuclease H protein (AHRD V3.3 *-* A0A2I0V757_9ASPA)
Solyc01g090510	SL4.0ch01	76414827	76416964	+	Pentatricopeptide repeat (AHRD V3.3 *** A0A200QT61_9MAGN)
Solyc01g091230	SL4.0ch01	77150274	77153523	+	Receptor-like kinase (AHRD V3.3 *** A0A191UPW4_CASSA)
Solyc01g091380	SL4.0ch01	77253959	77258166	-	65-kDa microtubule-associated protein 9 (AHRD V3.3 *** A0A2G2XNT4_CAPBA),Pfam:PF03999
Solyc01g094080	SL4.0ch01	77917367	77920041	-	Cytochrome P450 (AHRD V3.3 *** A0A200PMP8_9MAGN);Dbxref=InterPro:IPR001128,Pfam:PF00067;Ontology_term=GO:0005506,GO:0016705,GO:0020037,GO:0055114
Solyc01g095300	SL4.0ch01	78878133	78887955	+	Glycosyl transferase, family 1 (AHRD V3.3 *** A0A2U1P6H1_ARTAN)
Solyc01g095310	SL4.0ch01	78888385	78889293	+	Acidic endochitinase (AHRD V3.3 *** A0A1U8DT03_CAPAN);Dbxref=InterPro:IPR001223,Pfam:PF00704;Ontology_term=GO:0005975
Solyc01g096950	SL4.0ch01	80212674	80220690	+	WD40 repeat-containing protein (AHRD V3.3 *** A0A2U1MAJ5_ARTAN)
Solyc01g098980	SL4.0ch01	81591928	81605694	-	MAP kinase kinase kinase 8
Solyc01g101180	SL4.0ch01	83295219	83297506	+	Terpene synthase (AHRD V3.3 *** G5CV44_SOLLC)
Solyc01g109370	SL4.0ch01	88691454	88694668	+	Acyltransferase domain-containing protein (AHRD V3.3 *** A0A1Q3CKP3_CEPFO)

Solyc01g111040	SL4.0ch01	89802074	89805107	+	Calcium-binding EF-hand family protein (AHRD V3.3 *** A0A2U1MN65_ARTAN)
Solyc02g044005	SL4.0ch02	238504	248699	+	RNA-directed DNA polymerase (AHRD V3.3 *-* A0A2I0VXC8_9ASPA)
Solyc02g071260	SL4.0ch02	38758222	38763749	+	Phytochrome (AHRD V3.3 *** A0A2G2XE21_CAPBA)
Solyc02g081700	SL4.0ch02	43546831	43547577	+	Proteasome subunit alpha type (AHRD V3.3 *** A0A1U8FNG0_CAPAN);Dbxref=InterPro:IPR001353,Pfam:PF00227;Ontology_term=GO:0004298,GO:0005839,GO:0051603
Solyc02g082180	SL4.0ch02	43873572	43879221	+	DNA helicase (AHRD V3.3 *** A0A2G2XFH5_CAPBA)
Solyc02g085270	SL4.0ch02	46238998	46239936	+	RING/U-box superfamily protein (AHRD V3.3 *-* A0A1P8AQV3_ARATH)
Solyc02g085700	SL4.0ch02	46535275	46536357	+	geranylgeranyl pyrophosphate synthetase;Dbxref=InterPro:IPR000092,Pfam:PF00348;Ontology_term=GO:0008299
Solyc02g088580	SL4.0ch02	48681372	48683489	+	Plant calmodulin-binding protein-like protein (AHRD V3.3 *-* AT5G39380.5)
Solyc02g093380	SL4.0ch02	52319776	52321725	-	Pentatricopeptide repeat-containing protein (AHRD V3.3 *** A0A2U1PB47_ARTAN)
Solyc02g094430	SL4.0ch02	53045963	53054518	-	Acyltransferase-like protein, chloroplastic (AHRD V3.3 *** A0A2G3AD32_CAPAN)
Solyc04g005650	SL4.0ch04	437647	438699	-	Peroxisomal nicotinamide adenine dinucleotide carrier (AHRD V3.3 *** A0A2G2ZHF7_CAPAN)
Solyc04g025755	SL4.0ch04	20460976	20473545	+	Exocyst complex component EXO84A (AHRD V3.3 *-* A0A2G2VT08_CAPBA)
Solyc04g048905	SL4.0ch04	37838928	37841874	-	Aminotransferase-like, plant mobile domain family protein (AHRD V3.3 *-* F4I9K2_ARATH);Dbxref=InterPro:IPR019557,Pfam:PF10536

Solyc04g049105	SL4.0ch04	38927433	38927799	+	LOW QUALITY PROTEIN: putative transporter svop-1 (AHRD V3.3 --* XP_021743117.1)
Solyc04g049380	SL4.0ch04	40991204	41012342	+	Protein kinase superfamily protein (AHRD V3.3 *-* O80962_ARATH)
Solyc04g063380	SL4.0ch04	54857058	54860371	+	F-box protein SKP2A (AHRD V3.3 *** A0A1U8F5B3_CAPAN)
Solyc04g071380	SL4.0ch04	56174912	56175710	+	ubiquitin-like-specific protease ESD4 (AHRD V3.3 -** XP_021808599.1);Dbxref=InterPro:IPR003653,Pfam:PF02902;Ontology_term=GO:0006508,GO:0008234
Solyc04g076990	SL4.0ch04	59936680	59939027	+	Receptor-like protein kinase (AHRD V3.3 *** A0A2R6RNH0_ACTCH);Dbxref=InterPro:IPR000719,Pfam:PF00069;Ontology_term=GO:0004672,GO:0005524,GO:0006468
Solyc04g081240	SL4.0ch04	63197184	63203418	-	Auxin Response Factor 5
Solyc04g082490	SL4.0ch04	64095472	64112070	-	Niemann-Pick C1 protein-like (AHRD V3.3 *** A0A2I4DPI3_9ROSI)
Solyc05g010817	SL4.0ch05	5082562	5088488	-	Aminotransferase-like, plant mobile domain family protein (AHRD V3.3 *-* A0A0K9NPT3_ZOSMR)
Solyc05g014970	SL4.0ch05	9232063	9244844	-	Urease (AHRD V3.3 *** Q93WI8_SOLTU)
Solyc05g023895	SL4.0ch05	30129434	30136091	-	Gag/pol polyprotein (AHRD V3.3 *-* Q6WAY3_PEA)
Solyc05g026520	SL4.0ch05	41858114	41860382	+	Serine/threonine-protein phosphatase 7 long form-like protein (AHRD V3.3 *-* A0A1J3DHF8_NOCCA)
Solyc05g050640	SL4.0ch05	60170467	60171024	-	B3 domain protein (AHRD V3.3 -** AT5G26805.1)
Solyc05g055090	SL4.0ch05	64256968	64275666	-	Ubiquitin carboxyl-terminal hydrolase 12 (AHRD V3.3 *** A0A2G2XX24_CAPAN)
Solyc06g009530	SL4.0ch06	3507057	3534738	-	Carbohydrate-binding-like fold (AHRD V3.3 *** A0A061DSB9_THECC)
Solyc06g033970	SL4.0ch06	21214252	21214701	+	Ribonuclease H domain (AHRD V3.3 *-* A0A200QTT8_9MAGN)

Solyc06g034370	SL4.0ch06	21922312	21922890	-	Plant invertase/pectin methylesterase inhibitor superfamily protein (AHRD V3.3 *** A0A061G797_THECC);Dbxref=InterPro:IPR006501,Pfam:PF04043;Ontology_term=GO:0004857
Solyc06g043065	SL4.0ch06	27615768	27618065	-	PttA' (AHRD V3.3 *-* O22369_PETHY);Dbxref=InterPro:IPR004252,Pfam:PF03004
Solyc06g062295	SL4.0ch06	36960007	36961568	-	UBN2_3 domain-containing protein (AHRD V3.3 *** A0A1Q3AR18_CEPFO)
Solyc06g082380	SL4.0ch06	45791441	45793429	-	protein LURP-one-related 7-like (AHRD V3.3 *** A0A2I4F9W5_9ROSI);Dbxref=InterPro:IPR007612,Pfam:PF04525
Solyc07g008320	SL4.0ch07	3054506	3072790	-	Calcium-transporting ATPase (AHRD V3.3 *** G3LY22_SOLLC)
Solyc07g019640	SL4.0ch07	12698098	12699654	-	G-type lectin S-receptor-like serine/threonine-protein kinase (AHRD V3.3 *-* A0A2G3D6E9_CAPCH)
Solyc07g041660	SL4.0ch07	53680979	53692448	-	DENN domain-containing protein (AHRD V3.3 *** A0A1Q3BXR1_CEPFO)
Solyc07g042910	SL4.0ch07	56277686	56314327	-	transmembrane protein (AHRD V3.3 *** AT5G58100.2)
Solyc07g053280	SL4.0ch07	61675820	61681031	-	Ketol-acid reductoisomerase (AHRD V3.3 *** K4CFM5_SOLLC)
Solyc07g053380	SL4.0ch07	61760134	61762395	-	Pentatricopeptide repeat (AHRD V3.3 *** A0A200PYA7_9MAGN)
Solyc08g014618	SL4.0ch08	4790183	4791797	+	UB-like protease 1B (AHRD V3.3 --* AT4G00690.1);Dbxref=InterPro:IPR003653,Pfam:PF02902;Ontology_term=GO:0006508,GO:0008234
Solyc08g044337	SL4.0ch08	21207222	21208345	+	UB-like protease 1B (AHRD V3.3 --* AT4G00690.1);Dbxref=InterPro:IPR003653,Pfam:PF02902;Ontology_term=GO:0006508,GO:0008234

Solyc08g061245	SL4.0ch08	46396121	46397834	-	UBN2_3 domain-containing protein (AHRD V3.3 *** A0A1Q3AQH1_CEPFO)
Solyc08g077460	SL4.0ch08	59486975	59489375	-	SBP (S-ribonuclease binding protein) family protein (AHRD V3.3 *** Q8LCK5_ARATH)
Solyc08g080360	SL4.0ch08	61770913	61773282	-	F-box protein PP2-A15 (AHRD V3.3 *** A0A2G2Y8Y0_CAPAN);Dbxref=InterPro:IPR025886,Pfam:PF14299
Solyc08g081050	SL4.0ch08	62289421	62297810	+	Bromo adjacent homology (BAH) domain (AHRD V3.3 *-* A0A200QRT9_9MAGN)
Solyc08g082040	SL4.0ch08	63073209	63079130	+	Extra-large G-like protein, putative (DUF3133) (AHRD V3.3 *** A0A1P8B1E9_ARATH)
Solyc09g007830	SL4.0ch09	1376408	1380774	-	Cytokinin riboside 5'-monophosphate phosphoribohydrolase (AHRD V3.3 *** A0A1U8E671_CAPAN);Dbxref=InterPro:IPR031100,Pfam:PF03641
Solyc09g014642	SL4.0ch09	6452321	6454274	+	Unknown protein
Solyc09g065900	SL4.0ch09	60291959	60299093	-	Glutathione reductase (AHRD V3.3 *** B8PWQ9_SOLLC);Dbxref=InterPro:IPR023753,Pfam:PF07992;Ontology_term=GO:0016491,GO:0055114
Solyc09g072730	SL4.0ch09	61499535	61500532	-	Zinc finger, GRF-type (AHRD V3.3 *-* A0A2U1LVQ4_ARTAN)
Solyc09g075730	SL4.0ch09	63563550	63568352	+	Leucine-rich repeat protein kinase family protein (AHRD V3.3 *** A0A0K9NJV0_ZOSMR);Dbxref=InterPro:IPR024788,Pfam:PF12819
Solyc10g011790	SL4.0ch10	3813337	3818128	-	BTB/POZ domain-containing protein (AHRD V3.3 *** A0A1U8E4A8_CAPAN)
Solyc10g062340	SL4.0ch10	25066982	25074066	-	RNA binding protein (AHRD V3.3 *-* A0A2K3MUE2_TRIPR)
Solyc10g044470	SL4.0ch10	25485051	25496803	+	Chloride channel protein (AHRD V3.3 *** K4CZH2_SOLLC);Dbxref=InterPro:IPR001807,Pfam:PF00654;Ontology_term=GO:0005247,GO:0006821,GO:0016020,GO:0055085

Solyc10g047865	SL4.0ch10	41881820	41882747	-	L-ascorbate oxidase-like protein (AHRD V3.3 *-* A0A2G2Y5H1_CAPAN)
Solyc10g055685	SL4.0ch10	56266311	56268740	-	Plant mobile domain protein family (AHRD V3.3 *-* A0A1D6NQG9_MAIZE)
Solyc11g018720	SL4.0ch11	9399321	9404120	+	ALBINO3-like protein 2, chloroplastic (AHRD V3.3 *-* A0A1U8G0A5_CAPAN)
Solyc11g018740	SL4.0ch11	9428168	9437359	+	ALBINO3-like protein 2, chloroplastic (AHRD V3.3 *** A0A1U8G0A5_CAPAN)
Solyc11g021160	SL4.0ch11	13440475	13441314	+	ATP synthase subunit beta, chloroplastic (AHRD V3.3 *-* A0A2G2WCU2_CAPBA)
Solyc11g022400	SL4.0ch11	13804649	13807618	-	C2 calcium/lipid-binding plant phosphoribosyltransferase family protein (AHRD V3.3 *** A0A061FS12_THECC)
Solyc11g069250	SL4.0ch11	51948248	51950105	-	Fasciclin-like arabinogalactan protein 12 (AHRD V3.3 *** A0A1J6K653_NICAT);Dbxref=InterPro:IPR000782,Pfam:PF02469
Solyc11g072260	SL4.0ch11	53533858	53536945	-	40S ribosomal protein S13 (AHRD V3.3 *** A0A2I0WN36_9ASPA)
Solyc12g008720	SL4.0ch12	2137806	2139812	+	60S ribosomal protein L31 (AHRD V3.3 *** A0A1U8EAP7_CAPAN);Dbxref=InterPro:IPR000054,Pfam:PF01198;Ontology_term=GO:0003735,GO:0005840,GO:0006412
Solyc12g036870	SL4.0ch12	47172279	47173328	+	Myb/SANT-like DNA-binding domain protein (AHRD V3.3 *** Q8GYR6_ARATH)
Solyc12g099200	SL4.0ch12	65963482	65965261	+	Invertase inhibitor (AHRD V3.3 *** O49908_TOBAC);Dbxref=InterPro:IPR006501,Pfam:PF04043;Ontology_term=GO:0004857
Solyc12g100190	SL4.0ch12	66541343	66548062	+	Pleiotropic drug resistance protein (AHRD V3.3 *** C8CA13_CUCSA)

Table 3.S3. All Mexican landrace pepper population accessions. With associated genotypes, fruit-shape, and pedigree information (Taitano *et al.,* 2019). Full landrace and population data included instead of prefix description (e.g. 14CAg1-1 becomes 1-1). *Genotypes are scored as follows: 1 = homozygous for the Chigole parent allele, 2 = heterozygous, 3 = homozygous for the Chile de Arbol parent allele

ID	*Location*	*Sample ID*	*Species*	*Region*	*Village*	*Local name*	*Growth env. of accession:*	Pop.	FSI	19EP764*	18EP1221*
1-1	Santa Cruz Nexila	CanStaCruz 20	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.2	3	1
75-1	Añil	CanAnil10	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Costeno	3	1	1
75-2	Añil	CanAnil10	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Costeno	2.7	1	1
76-1	Añil	CanAnil2	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	3.4	2	1
76-2	Añil	CanAnil2	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno		1	1
77-1	Añil	CanAnil3	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	2.6	1	1
77-2	Añil	CanAnil3	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	3	2	1

79-1	Añil	CanAnil5	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	4.5	3	1
79-2	Añil	CanAnil5	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	2.1	2	1
80-1	Añil	CanAnil6	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	2.1	2	1
81-1	Añil	CanAnil7	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Costeno	2	1	1
81-2	Añil	CanAnil7	C. annuum	COAST	Añil, Pinotepa Nacional	COSTEÑO AMARILLO	Plantation	Costeno		1	1
84-1	GpeVictoria	CanGpe1	C. anuum	COAST	Guadalupe Victoria, Pinotepa Nacional	COSTEÑO ROJO	Plantation	Costeno	2.6	2	1
88-1	Huaxpaltepec	CanHuax2	C. frutescens	COAST	Huaxpaltepec	PIQUIN	Backyard	*frutescens*	1.3	1	1
91-1	Huaxpaltepec	CfrHuax5	C. frutescens	COAST	Huaxpaltepec	PIQUIN	Backyard	*frutescens*	1.3	1	1
92-1	Huaxpaltepec	CfrHuax6	C. frutescens	COAST	Huaxpaltepec	PIQUIN	Backyard	*frutescens*	2.1	1	1
93-1	JuanDiegal	CanJuan1	C. annuum	COAST	Juan Diegal, Pochutla	TUSTA	Milpa	TusTav2	1.4	1	1
93-2	JuanDiegal	CanJuan1	C. annuum	COAST	Juan Diegal, Pochutla	TUSTA	Milpa	TusTav2	1	1	1

98-1	Lagartero	CanLag4	C. annuum	COAST	Lagartero, Pinotepa Nacional	COSTEÑO ROJO	Milpa	Costeno	2.4	1	1
98-2	Lagartero	CanLag4	C. annuum	COAST	Lagartero, Pinotepa Nacional	COSTEÑO ROJO	Milpa	Costeno	3.2	1	1
100-2	Lagunilla	CfrLaguna1	C. frutescens	COAST	Lagunilla, Pochutla	CHILE BOLITA	Backyard	*frutescens*	1.2	1	1
102-1	Los Reyes	CfrReyes2	C. frutescens	COAST	Los Reyes, Pochutla		Backyard	*frutescens*	2.2	1	1
103-1	Los Reyes	CfrReyes3	C. frutescens	COAST	Los Reyes, Pochutla		Backyard	*frutescens*	1.8	1	1
103-2	Los Reyes	CfrReyes3	C. frutescens	COAST	Los Reyes, Pochutla		Backyard	*frutescens*	2.2	1	1
104-1	Los Reyes	CfrReyes4	C. frutesce-ns	COAST	Los Reyes, Pochutla		Backyard	*frutescens*	1.1	1	1
104-2	Los Reyes	CfrReyes4	C. frutesce-ns	COAST	Los Reyes, Pochutla		Backyard	*frutescens*	1	1	1
105-1	Los Reyes	CanReyes1	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.3	1	1
105-2	Los Reyes	CanReyes1	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.5	1	1
106-1	Los Reyes	CanReyes2	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.3	1	1
106-2	Los Reyes	CanReyes2	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.2	1	1

107-1	Los Reyes	CanReyes3	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.4	1	1
107-2	Los Reyes	CanReyes3	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Backyard	TusTav2	1.2	1	1
108-1	Los Reyes	CanReyes4	C. annuum	COAST	Los Reyes, Pochutla	TUSTA	Milpa	TusTav2	1.3	1	1
109-1	Paso de Los Indios	CfrPind1	C. frutescens	COAST	Paso de los Indios, Pochutla	MIRASOL	Backyard	*frutescens*	3.1	1	1
109-2	Paso de Los Indios	CfrPind1	C. frutescens	COAST	Paso de los Indios, Pochutla	MIRASOL	Backyard	*frutescens*	3.6	1	1
112-1	Paso de Los Indios	CanPind1	C. annuum	COAST	Paso de los Indios, Pochutla	TUSTA	Backyard	*frutescens*	1.2	1	1
112-2	Paso de Los Indios	CanPind2	C. annuum	COAST	Paso de los Indios, Pochutla	TUSTA	Backyard	*frutescens*		1	1
116-1	Polvorín	CanPol9	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Milpa	Forest		1	1
116-2	Polvorín	CanPol9	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Milpa	Forest		1	1
117-1	Polvorín	CanPol10	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Forest	Forest		1	1
117-2	Polvorín	CanPol10	C. annuum	COAST	POLVORÍN, POCHUTLA	CHILE DE MONTE	Forest	Forest		1	1
118-1	Río Seco	CanRS1	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	*frutescens*	1.3	1	1

118-2	Río Seco	CanRS1	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	*frutescens*	1.4	1	1
119-1	Río Seco	CanRS2	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	*frutescens*	1.2	1	1
119-2	Río Seco	CanRS2	C. annuum	COAST	Río Seco, San Pedro Huamelula	CHIGOLE	Backyard	*frutescens*	1.2	1	1
120-1	Rosedal	CanRose1	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO	Plantation	Costeno	2.5	1	1
120-2	Rosedal	CanRose1	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO	Plantation	Costeno		1	1
121-1	Rosedal	CanRose2	C. annuum	COAST	Rosedal, close to Tomatal	COSTEÑO ROJO FLAQUITO	Plantation	Costeno		1	2
124-1	Saachilac	CanSaa1	C. annuum	COAST	Saachilac, Santiago Astata	MAREÑO	Backyard	*frutescens*	4.1	1	1
124-2	Saachilac	CanSaa1	C. annuum	COAST	Saachilac, Santiago Astata	MAREÑO	Backyard	*frutescens*	4.3	1	1
125-1	Saachilac	CanSaa2	C. annuum	COAST	Saachilac, Santiago Astata	SOLTERITO	Backyard	*frutescens*		1	1
125-2	Saachilac	CanSaa2	C. annuum	COAST	Saachilac, Santiago Astata	SOLTERITO	Backyard	*frutescens*	4.1	1	1

126-1	Saachilac	CanSaa3	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	*frutescens*	3.4	1	1
126-2	Saachilac	CanSaa3	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	*frutescens*	3.5	1	1
127-1	Saachilac	CanSaa5	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	Forest	1.4	1	1
127-2	Saachilac	CanSaa5	C. annuum	COAST	Saachilac, Santiago Astata	CHILE DE MONTE	Forest	Forest			1
128-1	Coatecas Altas	CanCoat2	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.7	1	1
128-2	Coatecas Altas	CanCoat2	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.8		1
129-1	Coatecas Altas	CanCoat4	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.6	1	1
130-1	Coatecas Altas	CanCoat5	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.8	1	1
130-2	Coatecas Altas	CanCoat5	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg		1	1
131-1	Coatecas Altas	CanCoat6	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.1	3	1
131-2	Coatecas Altas	CanCoat6	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.2	3	1
132-1	Coatecas Altas	CanCoat7	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.9	1	1

132-2	Coatecas Altas	CanCoat7	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.6	1	1
133-1	Coatecas Altas	CanCoat8	C. annuum	CENTRAL VALLEYS	Coatecas Altas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2	1	1
134-1	La Labor	CanLabor1	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	2.4	1	1
134-2	La Labor	CanLabor1	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	2.2	1	1
135-1	La Labor	CanLabor2	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.8		1
135-2	La Labor	CanLabor2	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	2.4	1	1
136-1	La Labor	CanLabor3	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.9	1	1
136-2	La Labor	CanLabor3	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.9	2	1
137-1	La Labor	CanLabor4	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	2.4	1	1
137-2	La Labor	CanLabor4	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.8	1	1
138-1	La Labor	CanLabor5	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.8	3	1
138-2	La Labor	CanLabor5	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	1.8	2	1

139-1	La Labor	CanLabor6	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	3		1
139-2	La Labor	CanLabor6	C. annuum	CENTRAL VALLEYS	La Labor, San Dionisio Ocotlán	CHILE DE AGUA	Plantation	CAg	3.6	1	1
140-1	La Lobera	CanLobera10	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	1.7	1	1
140-2	La Lobera	CanLobera10	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	1.7	2	1
141-1	La Lobera	CanLobera2	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.3	1	1
141-2	La Lobera	CanLobera2	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	1.4	1	1
142-1	La Lobera	CanLobera3	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	3.1	3	1
142-2	La Lobera	CanLobera3	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	3.3	3	1
143-1	La Lobera	CanLobera4	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.7	1	1
143-2	La Lobera	CanLobera4	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2	1	1
144-1	La Lobera	CanLobera5	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.6	1	1
144-2	La Lobera	CanLobera5	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.4	1	1

145-1	La Lobera	CanLobera6	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.1	1	1
145-2	La Lobera	CanLobera6	C. annuum	CENTRAL VALLEYS	La Lobera, Santa Inés del Monte	CHILE DE AGUA	Milpa	CAg	2.3	1	1
146-1	Paraje Coatequillas	CanParaje1	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.7	3	1
147-1	Paraje Coatequillas	CanParaje2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.1	2	1
147-2	Paraje Coatequillas	CanParaje2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.1		1
148-1	Paraje Coatequillas	CanParaje4	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.8	3	1
148-2	Paraje Coatequillas	CanParaje4	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.6	3	1
149-1	Paraje Coatequillas	CanParaje5	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.4	3	1
149-2	Paraje Coatequillas	CanParaje5	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.2	3	1

150-1	Paraje Coatequillas	CanParaje6	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.1	1	1
150-2	Paraje Coatequillas	CanParaje6	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.3	1	1
151-1	Paraje Coatequillas	CanParaje7	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.8	1	1
151-2	Paraje Coatequillas	CanParaje7	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.8	1	1
152-1	Paraje Coatequillas	CanParaje9	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	2.1	1	1
152-2	Paraje Coatequillas	CanParaje9	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.9	2	1
153-1	Paraje Coatequillas	CanParaje 10	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.9	1	1
153-2	Paraje Coatequillas	CanParaje 10	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	CAg	1.4	1	1
156-1	La Tortolita	CanTorto2	C. annuum	COAST	La Tortolita, Santiago Astata	GUAJILLO	Backyard	*frutescens*	2.1	1	1

156-2	La Tortolita	CanTorto2	C. annuum	COAST	La Tortolita, Santiago Astata	GUAJILLO	Backyard	*frutescens*	1.7	1	1
157-1	Copalita	CanCopal1	C. annuum	COAST	Copalita	MIRASOL	Backyard	*frutescens*		1	1
157-2	Copalita	CanCopal1	C. annuum	COAST	Copalita	MIRASOL	Backyard	*frutescens*			1
162-1	Rosedal	CanRose2B	C. annuum	COAST	Rosedal, close to Tomatal	CHIGOLE	Backyard	Forest	1.3	1	1
162-2	Rosedal	CanRose2B	C. annuum	COAST	Rosedal, close to Tomatal	CHIGOLE	Backyard	Forest	1.5	1	1
163-1	Rosedal	CanRose11B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno		1	1
163-2	Rosedal	CanRose11B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno		1	1
164-1	Rosedal	CanRose6B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	4.9	3	3
164-2	Rosedal	CanRose6B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	4.7	3	3
165-1	Rosedal	CanRose7B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	2.6	3	1
166-1	Rosedal	CanRose8B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	4	3	1

166-2	Rosedal	CanRose8B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	3.8	3	2
168-1	Rosedal	CanRose10B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno		1	1
168-2	Rosedal	CanRose10B	C. annuum	COAST	Rosedalito, campo de "El Mojarra"	COSTEÑO ROJO	Plantation	Costeno	3		2
169-1	Santa Cruz Nexila	CanStaCruz6	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.3	3	1
169-2	Santa Cruz Nexila	CanStaCruz6	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.4	3	1
170-1	Santa Cruz Nexila	CanStaCruz11	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.1	3	1
170-2	Santa Cruz Nexila	CanStaCruz11	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	1.8	3	1
173-1	Santa Cruz Nexila	CanStaCruz14	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.1	3	1
173-2	Santa Cruz Nexila	CanStaCruz14	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	1.9	3	1

174-1	Santa Cruz Nexila	CanStaCruz15	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.6	3	1
174-2	Santa Cruz Nexila	CanStaCruz15	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	2.2	3	1
177-1	Santa Cruz Nexila	CanStaCruz19	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	3.5	1	1
177-2	Santa Cruz Nexila	CanStaCruz19	C. annuum	CENTRAL VALLEYS	Santa Cruz Nexila, Sola de Vega	CHILE DE AGUA	Plantation	CAg	3	1	1
179-1	San Pablo Coatlán	CanSPablo1	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Backyard	TusTav1	1.5	2	1
179-2	San Pablo Coatlán	CanSPablo1	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Backyard	TusTav1	1.9	2	1
180-1	San Pablo Coatlán	CanSPablo2	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	DE ÁRBOL	Backyard	TusTav1	7.4	3	3
180-2	San Pablo Coatlán	CanSPablo2	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	DE ÁRBOL	Backyard	TusTav1		3	3
181-1	San Pablo Coatlán	CanSPablo4	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav2		1	1

181-2	San Pablo Coatlán	CanSPablo4	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav2	3.8	1	1
182-1	San Pablo Coatlán	CanSPablo5	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1	3.1		1
182-2	San Pablo Coatlán	CanSPablo5	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1			
183-1	San Pablo Coatlán	CanSPablo6	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav2	2.5		
183-2	San Pablo Coatlán	CanSPablo6	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav2	2.9		
184-1	San Pablo Coatlán	CanSPablo7	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1	2.7		
184-2	San Pablo Coatlán	CanSPablo7	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1			
185-1	San Pablo Coatlán	CanSPablo10	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Milpa	TusTav1	1.3		
185-2	San Pablo Coatlán	CanSPablo10	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TUSTA	Milpa	TusTav1	1.1		

186-1	San Pablo Coatlán	CanSPablo3	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1	3.4		
186-2	San Pablo Coatlán	CanSPablo3	C. annuum	CENTRAL VALLEYS	San Pablo Coatlán, Los Coatlanes	TAVICHE	Milpa	TusTav1	2.8	1	1
187-1	San Baltazar	CanSnBalta1	C. annuum	CENTRAL VALLEYS	San Baltazar Loxicha	TUSTA		TusTav2	1.3	1	1
187-2	San Baltazar	CanSnBalta1	C. annuum	CENTRAL VALLEYS	San Baltazar Loxicha	TUSTA		TusTav2	1.3		1
188-1	Santa Lucía	CanSLucía1	C. annuum	CENTRAL VALLEYS	Santa Lucía Miahuatlán	TUSTA		TusTav2	2.2	1	1
188-2	Santa Lucía	CanSLucía1	C. annuum	CENTRAL VALLEYS	Santa Lucía Miahuatlán	TUSTA		TusTav2	1.8	1	1
189-1	Paraje Coatequillas	CanEjutla2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1	2.6	1	1
189-2	Paraje Coatequillas	CanEjutla2	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1		1	1
190-1	Paraje Coatequillas	CanEjutla1	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1	3	1	1
190-2	Paraje Coatequillas	CanEjutla1	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1	2.8		1

191-1	Paraje Coatequillas	CanEjutla3	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1	3.2		1
191-2	Paraje Coatequillas	CanEjutla3	C. annuum	CENTRAL VALLEYS	Paraje Coatequillas, Ejutla de Crespo	CHILE DE AGUA	Plantation	TusTav1	2.4		1
200-1	Acanceh	CanAcan20	C. annuum	YUCATAN	Acanceh	MAAX'IK		Yucatan			1
200-2	Acanceh	CanAcan20	C. annuum	YUCATAN	Acanceh	MAAX'IK		Yucatan		1	1
202-1	Maní	CanMani1	C. annuum	YUCATAN	Maní	PARADITO	Backyard	*frutescens*	1.2	1	1
203-1	Maní	CanMani2	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Yucatan			1
203-2	Maní	CanMani2	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Yucatan		1	1
204-1	Maní	CanMani3	C. annuum	YUCATAN	Maní	PARADITO	Backyard	*frutescens*	1.1		1
204-2	Maní	CanMani3	C. annuum	YUCATAN	Maní	PARADITO	Backyard	*frutescens*	1.2		1
206-1	Maní	CanMani7	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Yucatan		2	1
206-2	Maní	CanMani7	C. annuum	YUCATAN	Maní	MAAX'IK	Backyard	Yucatan			1
208-1	Maní	CanMani11	C. annuum	YUCATAN	Maní	PARADITO	Backyard	*frutescens*	1.1	1	1

208-2	Maní	CanMani11	C. annuum	YUCATAN	Maní	PARADITO	Backyard	*frutescens*	1.1	1	1
211-1	Maní	CanMani14	C. annuum	YUCATAN	Maní	DULCE	Backyard	*frutescens*	0.5	3	1
212-1	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	*frutescens*	1	3	1
212-2	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	*frutescens*	0.7	3	1
213-1	Dzibtzantun	CanDzib14	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	*frutescens*	3.2		1
213-2	Dzibtzantun	CanDzib14	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	*frutescens*	2.6		1
214-1	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	Yucatan			1
214-2	Dzibtzantun	CanDzib1	C. annuum	YUCATAN	Cansahcab	DULCE	Milpa	Yucatan	1.5	1	1
215-1	Abasolo	CanAbasolo 1	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	CAg	3	1	1
215-2	Abasolo	CanAbasolo 1	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	CAg	3.3		1
216-1	Abasolo	CanAbasolo 2	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	CAg	1.8	1	1
216-2	Abasolo	CanAbasolo 2	C. annuum	CENTRAL VALLEYS	Abasolo	CHILE DE AGUA	Milpa	CAg	1.6	3	1
217-1	Rancho Llano	CanRan1B	C. annuum	COAST	Rancho Llano, Tehuantepec	PAYASO	Milpa	*frutescens*		1	1

217-2	Rancho Llano	CanRan1B	C. annuum	COAST	Rancho Lllano, Tehuantepec	PAYASO	Milpa	*frutescens*		1	1
219-1	La Tortolita	CanTorto3	C. annuum	COAST	La Tortolita, Santiago Astata			*frutescens*		1	1
219-2	La Tortolita	CanTorto3	C. annuum	COAST	La Tortolita			*frutescens*	2.2		1

CHAPTER 4

STRUCTURAL VARIANT DISCOVERY USING COEXPRESSION NETWORKS OF SIX POTENTIAL FRUIT SHAPE *TONNEAU1 RECRUITING MOTIF PROTEIN GENES*[1]

---

[1] Taitano, N., Zhang, B., Visa, S., Alonge, M., and van der Knaap, E. 2020. *To be submitted to BMC Genomics.*

Abstract

Genome structural variants (SV) underlie many agronomical traits such as fruit size and shape. Such variants include *OVATE FAMILY PROTEIN (OFP) 20*, a fruit shape gene in the recently discovered *TRM-OFP* regulatory network. This study selected six TONNEAU1 RECRUITING MOTIF (TRM) proteins based on their known interactions with OFPs and preliminary evidence of fruit-shape regulation. The genes encoding these selected TRMs were subjected to a co-expression network analysis using both direct correlation and WGCNA and identified enriched GO-terms in the co-expression clusters. Next, we sought to identify SVs that might affect the expression of the same *TRMs* as well as *TRM* co-expressed genes. The co-expressed genes as well as the potential for gene expression regulation by SVs provide a useful resource for identifying natural mutations that affect the *TRM-OFP* coexpression network. These potential informative mutations could be considered in future QTL-mapping of fruit morphology traits, and serving as potential targets for CRISPR-induced mutations to assess the function of various *TRM* co-regulated genes.

1) Introduction

Variation in fruit morphology has formed the basis for much diversification in cultivated fruit and vegetable crops. Different fruit shapes and sizes form the basis for different use-types (Perry, 2007) and can facilitate processing. Tomato (*Solanum lycopersicum* L.) and its wild relative (*S. pimpinellifolium* L.) have been used as the model organism for genetic research into fruit shape. Therefore, the presence of standing variation in wild relative, semi-domesticated, and landrace tomato populations is of continued interest not only to tomato breeders as a potential source of phenotypic variation, but also to developmental biologists interested in genes controlling aspects of fruit development (Blanca *et al.,* 2015, Razifard *et al.,* 2020).

Despite the routine use of SNPs as convenient markers in all steps of gene mapping (*reviewed by* Korte and Farlow, 2013), the causative mutations underlying fruit morphology traits in natural populations have more often than not been structural variants (SVs) (van der Knaap *et al.,* 2014). SVs are functionally defined as any mutations that exceed 20 base-pairs (bp) in length, and thus are not readily identified by ordinary read-to-reference comparison methods such as the GATK pipeline (Poplin *et al.,* 2018). SNPs and other small mutations can certainly - and have indeed been found to - induce important mutations including to genes regulating fruit shape such as *WUSCHEL* (Liu *et al.,* 2011) and *OVATE* (Liu *et al.,* 2002). Nevertheless, whether in the form of a large indel in the upstream regulatory region of *OFP20* (Wu *et al.,* 2018), transposon-mediated translocation in the case of *SUN* (van der Knaap *et al.,* 2004; Xiao *et al.,* 2008), or inversion disrupting the promoter of *CLV3* (Xu *et al.,* 2015), SVs have often been found to affect fruit morphology genes and their cis regulatory regions.

For fruit shape in particular, among the most promising genes for further study are the OVATE family proteins (OFPs). OFPs and TONNEAU1 RECRUITING MOTIF proteins (TRMs) together interact in a network that appears to influence fruit shape by controlling the plane of cell division in floral meristems. Both protein families are evolutionarily conserved (Liu *et al.,* 2014), function across species, and exhibit protein-protein interactions with one another (Wu *et al.,* 2018), suggesting roles in a wider network that presents a promising avenue of developmental biology research.

The fruit-shape effects of *OVATE* were first discovered in tomato, where a null mutation leads to elongation in the proximal end resulting in pear-shaped fruit (Liu *et al.,* 2002; van der Knaap *et al.,*2014). This phenotype is driven by increased proximal-distal and decreased medial-lateral cell division in the proximal end of developing ovaries (Wu *et al.,* 2018). A 31 Kbp deletion SV in the promoter region of a second OFP gene, *OFP20*, is necessary together with the *OVATE* mutation to exhibit a strongly pear-shaped phenotype (Rodriguez *et al.,* 2013; Wu *et al.,* 2018). Organ length changes have also been reported on *OVATE*-silenced pepper fruits (Tsaballa *et al.,* 2011), Arabidopsis leaves and siliques (Hackbusch *et al.,* 2005; Wang *et al.,* 2007; Li *et al.,* 2011; Wang *et al.,* 2011; Wang *et al.,* 2019), rice grains (Schmitz *et al.,* 2015; Ma *et al.,* 2017; Zhao *et al.,* 2018), melon fruits, and potato tubers (Wu *et al.,* 2018). A potential mechanism of OFP fruit-shape regulation is given by the relocalization of cell-division orientation proteins from the cytoskeleton to cytoplasm, and *trm* mutants rescue the *ovate/ofp20* mutant phenotype (Wu *et al.,* 2018).

TRMs affect the direction of cell division as part of a TRM, TONNEAU1, and PROTEIN PHOSPHATASE 2A (TTP) complex that controls the orientation of the cytoskeletal preprophase band (PPB), which forms along what will be the metaphase plate (Schaeffer *et al.,* 2017). As TRMs relocalize TONNEAU1 and vary in cytoskeletal localization motifs, the TRM subunit of this complex likely play a key role in patterning cell division (Drevensek *et al.,* 2012). Originally found through elongated-leaf mutants in Arabidopsis (Lee *et al.,* 2006), they also affect silique shape in Arabidopsis (Drevensek *et al.,* 2012), fruit shape in tomato and cucumber (Wu *et al.,* 2018), and grain shape in rice (Wang Y. *et al.,* 2015; Wang S. *et al.,* 2015; Zhou *et al.,* 2015).

Together, the OFP-TRM network regulates the shape of fruits and other organs through cell division patterning in early floral meristems. In tobacco leaf cells, OFPs relocalize TRMs from the microtubules to cytoplasm (Wu *et al.,* 2018). TRMs, in turn, interact with the TTP (Spinner *et al.,* 2013), often bind microtubules, but vary in their ability to do so (Drevensek *et al.,* 2012). Thus, TRMs are hypothesized to control TTP cytoskeletal localization and OFPs may further fine-tune this function (Spinner *et al.,* 2013; Drevensek *et al.,* 2012; Wu *et al.,* 2018; Snouffer *et al.,* 2020). This may explain phenotypic effects of *TRM* mutations on PPB formation, the cell division plane (Schaefer *et al.,* 2017), and organ shape (Lee *et al.,* 2006; Wu *et al.,* 2018). To further investigate the role of TRMs, co-expression analyses are a possible route to identify genes that may be acting in conjunction with the target gene as their expression dynamics mirror that of the given TRM. Such genes may be controlled by the same regulatory elements and avail themselves as interaction partners. Given the large role SVs have played in domestication traits, another useful task is to explore the SV diversity in the population, and identify those most likely to affect expression of the relevant genes.

Whole genome assembly-to-assembly comparison methods (e.g. Kurtz *et al.,* 2004; Li, 2018) are ideal for SV discovery, especially when based mainly on ≳20kbp repeat-spanning fragments like bacterial artificial chromosomes or long-read sequencing. However, large-fragment whole genome sequencing remains prohibitively expensive when applied to population genetics and GWAS populations. SV-calling software packages such as Lumpy (Layer *et al.,* 2014), are capable of detecting SVs from cheaply produced Illumina high throughput sequence data. Yet even once obtained, sifting through a large number of SVs for causative mutations remains a challenge. Like SNPs, many SVs are located in regions of the genome where they produce no phenotypic effects, or produce effects outside the scope of the research area. Further, very large inversions, duplications, and translocations may simply transport intact genes together with their regulatory elements. GWAS and other forward-genetics approaches can address this issue to find major QTL in the entire genome (Yang *et al.,* 2019).

This study explores coregulation networks of a selected set of TRMs that are known to interact with OVATE and SlOFP20. The expression of the genes that encode these TRMs and their coregulated network may be involved in regulating fruit morphology as well. We focus on tomato, which is well-studied model organism for fleshy fruit development. Our first objective is to explore the genes coregulated with the selected TRMs using expression data from various studies of the SP line LA1589. Our second objective is to identify and catalogue a set of SVs in a diversity population, with a focus on SVs affecting genes in the co-expression networks of the selected TRMs. We genotyped the SVs in a morphologically and genetically diverse population of 166 tomato accessions from highly cultivated *Solanum lycopersicum* var. *lycopersicum* (SLL), semi-domesticated *S. lycopersicum* var. *cerasiforme* (SLC), and wild relative *S. pimpinellifolium* (SP) (Razifard *et al.,* 2020). The result is a set of coexpressed genes, alongside a list of their SVs, that may inform future study into mechanisms underlying fruit-shape regulation by these six TRMs.

Materials and Methods

Sequence data and plant material

RNA-seq data were combined from several previous experiments in SP accession LA1589 at various stages of development. These developmental stages are: inflorescence and floral meristems, flower buds at 2 days post floral initiation (dpi) (Chu *et al.,* 2019); buds at 4 dpi, 6 dpi, 8 dpi, 10 dpi, 13 dpi, and 16 dpi (Wang *et al.,* 2019); flowers at anthesis and fruits 10 days post anthesis (dpa), 20 dpa, and 30 dpa (Huang *et al*., 2013); and lastly columella, pericarp, and seeds from developing fruits 2 days post anthesis (dpa), 4 dpa, 6 dpa, 8 dpa, and 10 dpa (Biyao Zhang, *unpublished data*).

To call SVs, we used a population of SLL, SLC, and SP plants from South America and Mexico (Razifard et al, 2020). The plant population comprises 166 accessions: 27 SP, 111 SLC, and 18 SLL, which we further categorize into subpopulations following Razifard *et al.,* (2020). SLC is divided into subpopulations from Peru (SLC-PER: 21 accessions), Ecuador (SLC-ECU: 42 accessions), San Martin (SLC-SM: 12 accessions), Central America (SLC-CA: 15 accessions), and Mexico (SLC-MEX: 22 accessions). SP is divided into subpopulations from Peru (SP-PER: 13 accessions), Southern Ecuador (SP-SECU, four accessions), and Northern Ecuador (SP-NECU, 10 accessions). The 18 SLL accessions were not further subdivided.

Paired-end sequence data with 150 bp reads from 550-650 bp fragments was originally produced by Razifard *et al.,* (2020) from the tomato domestication population on the NextSeq Illumina platform for 300 cycles to an average depth of 18X per individual.

RNA-seq analysis of OFP-associated TRM co-expression networks

Raw counts from the 26 tissue types above were converted to TPM, following the standard formula (Conesa *et al.,* 2016). Genes with < 0.1 TPM in 90% of tissue types were removed from analysis. This gene filtering threshold falls on the permissive end in the range of standard RNA-seq procedures for low-coverage gene removal, which ranges from excluding all genes with mean TPM ≤ 2.5 (Xia *et al.,* 2019), to keeping all genes with two replicates showing ≥ 0.1 TPM in any tissue (Kagale *et al.,* 2016).

Including *TRM25*, a fruit-shape candidate gene in pepper (Colonna *et al.,* 2019*,* Taitano, *unpublished data*), there are six TRMs that interact with OVATE and have mutants exhibiting fruit-shape changes: *TRM17/20a*, *TRM26*, *TRM3/4*, *TRM5*, and *TRM19* (Wu *et al.,* 2018, Biyao Zhang, *unpublished data*). Co-expression analysis was done for each of these six core genes to determine significant Pearson's R values at $p < 0.05$: $R > 0.75$ as determined by a permutation test of 10,000 randomly drawn gene pairs. A second coexpression analysis was conducted using weighted gene correlation network analysis (WGCNA) to generate clusters from whole transcriptome data with a soft threshold value of 10. In this WGCNA analysis, *TRM* genes were assigned to co-expression networks based on hierarchical clustering analysis along with the other genes in the transcriptome. Gene ontology (GO) term enrichment analysis was conducted using R/topGO (Alexa and Rahnenfuhrer, 2019; Alexa *et al.,* 2006). FDR-corrected hypergeometric tests were performed using topGO::runTest, and GO-terms were returned at a significance level of $p < 0.05$.

SV-calling and filtering

To overcome potential ascertainment bias arising from the use of a reference genome more closely related to one subpopulation, two distant reference genomes were used: BGV006775, an SP accession from Northern Ecuador, and the Heinz 1706 SL2.50 reference genome, a North American SLL cultivar. After BWA-mem alignment to either reference genome, SVs were called from Illumina data generated for each individual in the tomato domestication population described above using Lumpy with default parameters (Layer *et al.,* 2014). The original SV calls were then filtered by evidence, requiring at least one observation of both split-read and paired-end evidence for each SV. Based on testing performed on a subset of these SVs (Lei Zhang and Sofia Visa, *unpublished data*), a maximum split-read to paired-end ratio of 1:1, a minimum total evidence equal to half the average depth of sequencing, and a maximum total evidence count of thrice average depth of sequencing calculated for samples in which the SV is called, was set to avoid false-positives caused by partial read misalignment in repetitive elements. A maximum threshold of 1 Mbp was also used to remove SVs which were spurious outliers in size and of less interest as they would span many genes, confounding analysis. Finally, redundant SVs of the same type and where both start and end positions were located within 300 bp of each other were merged recursively.

SVs were then screened based on position within 5 Kbp of genes in the TRM co-expression network, and categorized as overlapping promoters, UTRs, or protein-coding regions. SVs that deleted ≥30% of the gene, including the start codon, were flagged as likely knockouts. The resulting genes were used in GO-term analysis following the procedure outlined above for co expressed gene clusters.

The filtered Illumina-generated SVs were compared to the 100-genome project's ONT-sequenced dataset with a partially overlapping tomato population (Alonge *et al.*, submitted 2020). These were also screened against our candidate gene list, and ONT SVs affecting the same loci as our Illumina SVs were determined based on provided gene IDs, SV type, and position information.

Results

Gene Expression Clustering

After filtering for low-expressed genes in all samples and replicates, we performed hierarchical clustering based on Euclidean gene expression distances between all replicates and developmental tissue-timepoints (Figure 4.S1). This largely returned clusters based on developmental stage. Thus, the replicates in each developmental tissue-timepoint tended to cluster together, with the exception of 20dpa2 and IM&FM-3, which were then removed as outliers.

We extracted the gene expression data for the six selected *TRM*s as well as *SlOFP20* and *OVATE* (Figure 4.1). The results revealed different expression profile for each gene. The earliest-expressed genes were *OVATE* and *TRM5*, with *OVATE* especially exhibiting a markedly higher expression than any other gene in the earliest inflorescence meristem (IM) and floral meristem (FM) developmental timepoint. Until 6 dpi (Figure 4.1A-D), *TRM5* and *OVATE* remained the two highest-expressed genes of those examined, albeit with *OVATE* expression decreasing during development and both *TRM5* and *TRM25* increasing in expression until 8 dpi (Figure 4.1E). *TRM26* was consistently lower-expressed, showing higher expression only in the columella at 2 dpa (Figure 4.1J). Each of the four other genes exhibited the highest expression at its own tissue-timepoint: *OFP20* was more highly expressed than any of the other seven only at anthesis (Figure 4.1I), *TRM19* was the highest expressed in the columella and seeds 2-4 dpa (Figure 4.1J-K, M-N), *TRM3/4* in the seeds from 4-8 dpa (Figure 4.1P,S,W), and *TRM17/20a* in the pericarp after anthesis until 10 dpa (Figure 4.1L,O,Q,T,X).

The soft power threshold of 10 attained an optimal clustering model fit at R^2=0.7, after which further increases in threshold power led to sharply diminishing returns in model fit. (Figure 4.S2). WGCNA using the optimum model fit clustering returned 19 clusters of co-expressed genes (Figure 4.S3). The six OFP-interacting TRMs were distributed among four of these clusters (Figure 4.2). The *TRM17/20a* cluster consisted of 406 genes expressed at an intermediate level through most tissues, with a slight reduction in expression just before and at anthesis, and again at 30 dpa (Figure 4.2A). The green and yellow lines in the figure represented OVATE and SlOFP20 expression dynamics, respectively. The *TRM19* and *TRM26* cluster consisted of 4,353 genes and covered a wide range of expression changes, with a small expression bump in the four days immediately after anthesis (Figure 4.2B). The *TRM5* and *TRM25* cluster, which comprised 328 genes, was high expressed in early developing flowers 4-8 dpi, was low at anthesis, and then increased in expression after anthesis before tapering off (Figure 4.2C). The 415 genes of the *TRM3/4* cluster were, like *TRM3/4,* higher expressed in the seeds 6 and 8 days post anthesis. *OVATE* was located in a fifth cluster, which includes higher-expressed genes in early timepoints from the IM & FM stage to 6 dpi (Figure 4.S3). *OFP20* was extremely high-expressed at anthesis, but not elsewhere, and was assigned to a sixth WGCNA cluster with genes that featured similar gene expression dynamics (Figure 4.S3).

Hard-threshold TRM Gene Clusters

Our second co-expression network approach - correlating all genes with each of the six selected *TRMs* – yielded 2,326 genes co-expressed with a selected *TRM* at a hard-threshold $p < 0.05$ level, determined via permutation test to be $R > 0.8$ (Figure 4.S4). This yielded networks of genes that were more closely co-expressed with each of the selected *TRMs* (Figure 4.3). However, these network genes were more distinct between *TRMs*, with none of them attaining high enough co-expression with any other to appear in the same hard-threshold network. Though overall patterns remained roughly the same, notable differences included the closer correlation early in developing floral buds for *TRM3/4* (Figure 4.3A) and *TRM17/20a* (Figure 4.3B). Direct correlation identified genes more closely following the expression of *TRM25* (Figure 4.3C) and *TRM26* (Figure 4.3D), but with far sparser networks. Unlike *TRM25, TRM5* exhibited a wider range of expression in different fruit tissues after anthesis (Figure 4.3E). This reduced tissue specificity was similar to many of the genes in the WGCNA cluster (Figure 4.2C), and remained so in the direct-correlation network. The largest difference was shown in the *TRM19* network, which exhibited a clear increase in expression in columella at 2 dpa, and reduction in expression after 10 dpa (Figure 4.3F). *TRM3/4* was significantly correlated with 385 genes, *TRM5* with 31, *TRM25* with 47, *TRM26* with 22, *TRM17/20a* with 241, and *TRM19* with 1,940 (Figure 4.4). No overlap was found between any of these genes save for those that were co-expressed with *TRM25* and *TRM5*. Over two-thirds of the genes in the *TRM25*-based cluster were also present in the *TRM5*-based cluster (Supplemental File 4.S1 – see Appendix).

Of the six co-expression networks resulting from direct correlation to a selected *TRM*, only the largest *TRM19*-based cluster returned significantly enriched GO-terms (Table 4.1). The top enriched GO-terms for *TRM19* included several related to organ development and especially cell-division, including "chromosome organization", "DNA repair", "reproductive structure development", "cell cycle", and "DNA replication" (Table 4.1). GO-term analysis was also performed on WGCNA clusters bearing any of the six selected *TRMs,* with significantly enriched terms found in three out of the four WGCNA clusters. In the *TRM17/20a* WGCNA cluster, 19 GO-terms were found, most of which - such as "cutin biosynthetic process", "wax biosynthetic process", and "cuticle development" – were related to lipid biosynthesis. Additionally, several, such as "cellulose metabolic process", "beta-glucan metabolic process", and "cellular polysaccharide metabolic process" were terms related to cell wall synthesis (Table 4.2). Being the largest of the four WGCNA clusters with a selected *TRM*, the cluster containing *TRM19* and *TRM26* also had the greatest number of GO-terms: 536 in total. As such, these GO-terms covered a wide array of biological processes. However, many of these were higher-level GO-terms that apply to >2,000 annotated genes. Including only more focused GO-terms applying to ≤2,000 genes each revealed that, aside from three related to RNA processing, the top ten GO-terms all involve such relevant processes as "anatomical structure development", "organelle organization", "post-embryonic development", and "multicellular organism development" Table 4.3. The WGCNA cluster containing *TRM5* and *TRM25* was enriched mainly for GO-terms related to active cell division such as "cell cycle", "DNA metabolic process", "cell cycle process", and "DNA replication", to name the top five (Table 4.4).

Production and analysis of substantial co-expression networks for each of the six selected *TRMs* enabled the creation and testing of an informatics pipeline for examining SVs intersecting those TRMs and categorizing their effects (Figure 4.S4). Stringent quality controls selected the best-supported 8% of

Lumpy-called SVs against both the BGV006775 (Supplemental File 4.S2 – see Appendix) and SL2.50 genomes (Supplemental File 4.S3 – see Appendix). Deletions were the most commonly identified SVs (Figure 4.5A). After more stringent filtering, they remained the most abundant SV class (Figure 4.5B). Deletions were also the most strongly and consistently supported by both split-read and read pair evidence (Figure 4.S6). Concordantly, deletions, along with similarly well-supported inversions, were kept in higher proportions than breakends and duplications (Figure 4.5B). Furthermore, a lower proportion of deletions (1.1%) relative to either duplications (13.5%) or inversions (5.2%) were questionably large (>1 Mbp) (Figure 4.S7).

We called the SVs against the SLL Heinz 1706 (Figure 4.6A). The highest number of SVs were found in SP from Peru which was expected due to the relatively large phylogenetic distance of this subclade in the *S. pimpinellifolium* (Razifard et al., 2020). When aligning to the SP NECU BGV006775 assembly, the set of returned SVs were more evenly distributed among the various subpopulations (Figure 4.6B). Relatively low per-accession SV counts were found in the populations used as reference genome. The SP PER subpopulation in both cases had the highest number of private SV alleles (alleles exclusive to a single subpopulation).

Direct correlation returned 2,326 target genes co-expressed with one of the six core *OFP*-interacting *TRM* genes. Of the 58,470 SVs called against the Heinz SL2.50 reference genome, 1.4% (817) were within 5 Kbp of the target gene (Figure 4.7A; Supplemental File 4.S4). Of these, 248 (30%) deleted part of the CDS, 108 (13%) deleted at least half of the CDS or the transcription start site, and 84 (10%) deleted the entire CDS (Figure 4.7B), and 318 (39%) fell within the gene including its UTRs (Figure 4.8A).

Compared to the Illumina-based SVs (Figure 4.8A), SVs provided for comparison by Dr. Xingang Wang generated from higher-coverage (ONT) reads (Table 4.5) returned a higher number (6,084) of SVs in the target genes (Figure 4.8B). Trends were similar between the two datasets in terms of the proportion of genes intersecting the UTRs vs. the CDS (Figure 4.8), and 132 (49%) of genes intersected by an SV in the target genome were also intersected by an SV in the ONT dataset (Figure 4.9).

Several SVs intersected genes co-expressed with *TRM19* (Table 4.6). Foremost among these was a 539 Kbp duplication in the Central American SLC accession LA1712, which duplicated *TON1,* and *CELL DIVISION CYCLE 5.* Another duplication found across three genetically distant accessions (BGV006767, BGV014522, and TS-75) affected *Solyc04g010220*, which is homologous to an animal centromere-recruitment gene *CEP135* (Kim *et al.,* 2008). Also notable is a 64 Kbp deletion in the SLC-MEX accession BGV007908 deleting the entirety of *Solyc01g091380*, a microtubule-associated protein (*MAP*)-like gene and closest tomato ortholog to Arabidopsis *MAP65-9*, which is highly expressed in stamens and pollen (Smertenko *et al.,* 2008). Finally, a 5.5 Kbp SV deleted the first third of the *Solyc05g006470* CDS, annotated as an ortholog of Arabidopsis *ACTIN-RELATED PROTEIN 2/3*.

Discussion

Variation between RNA-seq experiments performed on different samples can complicate experiments searching for differential expression between samples (Conesa *et al.,* 2016). These issues present less of a problem for network-based analyses, where the primary goal is to identify genes with consistently close expression patterns within samples. Furthermore, in this study hierarchical clustering (Figure 4.S1) revealed that similar tissue-timepoints originating from different experiments (e.g. buds 2, and 4 dpi) bore gene expression patterns closer to one another than more different timepoints originating from the same experiment (e.g., buds at 10 dpi and flowers at anthesis). Nevertheless, while viewing data from different tissue-timepoints as in Figure 4.2-3 provides informative context, caution should still be used when interpreting expression changes of a gene between different experiments.

Of the coexpression networks determined by direct correlation to the six selected *TRMs, TRM19* was the largest coexpression group and the only one to return significantly enriched GO-terms. Enriched terms were largely related to cell division and related biological processes such as DNA recombination and replication, as might be expected for a pathway that orients the cell division plane (Wu *et al.,* 2018). However, a second interpretation is that these genes are expressed in actively dividing cells. Nevertheless, certain GO-terms, such as “covalent chromatin modification” and “chromatin organization” point to potentially relevant pathways for remodeling the PPB as expected in this gene network.

Our SV-calling results reinforce the notion that Lumpy is best suited for calling deletions, as a deletion returns clear split-read and paired end evidence and a clear lack of read alignment over the length of the deleted region (Layer *et al.,* 2014).

SV counts in different subpopulations demonstrated the value of calling Illumina-based SVs from multiple reference genomes. Due to the ease of calling deletions relative to other types of SVs, including insertions (Figure 4.5), SVs called against different genomes (Figure 4.6) were not merely mirror images of each other. The relative evenness in the BGV-called SVs suggests greater sensitivity in the SLL reference to detecting SVs in more distant populations, though the reverse does not seem to be true. The low per-accession SV counts in the populations used as reference genome are to be expected due to higher relatedness and lower variation between individuals in that subpopulation and the reference genome. That the SP PER subpopulation in both cases had the highest number of private SV alleles (alleles exclusive to a single subpopulation), which fits with its position as a strongly distinct subpopulation (Razifard *et al.,* 2020).

The duplication of several genes in SLC accession LA1712 is interesting, as the duplicated region includes *TON1* and the *CDC5* gene necessary for mitosis in yeast (Ohi *et al.,* 1999). The expression of these genes could be assessed in plants bearing the LA1712 duplication allele and compared to wild-type, for example in near-isogenic lines, to test whether this duplication led to increased expression of either gene. However, *TON1* duplication may not have an effect on organ development. In Arabidopsis, two copies of *TON1* (*TON1a* and *TON1b*) are present, both must be mutated to prevent preprophase band formation and induce organ shape phenotypic changes, and overexpression causes no phenotypic changes relative to wild-type (Azimzadeh *et al.,* 2008). The duplication spanning the upstream region and first half of *Solyc04g010220* was also notable as this tomato gene exhibits homology with an animal protein necessary for recruiting proteins to new centrioles in the animal centrosome in mitosis (Kim *et al.,* 2008). Though plant cells do not have centrioles or centrosomes, the same recruitment networks, including the TON1-binding TRM M2 motif, have been found in both the plant TTP complex and animal centrosomes (Drevensek *et al.,* 2012).

Finally, a 5.5 Kbp SV deleted the first third of the *Solyc05g006470* CDS including the translation start site, which may lead to a loss-of-function mutation at that locus. This gene is noteworthy for its homology with Arabidopsis *ACTIN-RELATED PROTEINS (ARP) 2* and *3*, is part of a highly conserved complex that regulates cytoskeletal organization in a range of organisms, mutations of which cause cell shape changes in Arabidopsis leaves and trichomes (Mathur *et al.,* 2003). As PPB formation also relies on actin and cytoskeletal reorganization (Azimzadeh *et al.,* 2008, Schaefer *et al.,* 2017), a knockout to an *ARP* coexpressed with *TRM19* is potentially an interesting candidate to study for this gene network.

Conclusion

This study identified genes coexpressed with six select *OFP*-interacting *TRMs*, and natural SVs in several promising candidate genes that warrant future study. Biparental mapping using the accessions bearing these SVs could focus on the SV regions to assess phenotypic effects of the SV mutant allele relative to wild-type. Expression analyses on these genes might be useful, especially where SVs intersect regulatory elements, or where gene copy number is affected. These expression analyses could either be performed on plants grown from the accession seedlots, or on near-isogenic lines designed to differ only for the SV locus. Knockouts targeting the SV regions using CRISPR may also determine the phenotypic effects of the SV allele. For genes duplicated by an SV, overexpression constructs may provide useful insights into mutant phenotypes.

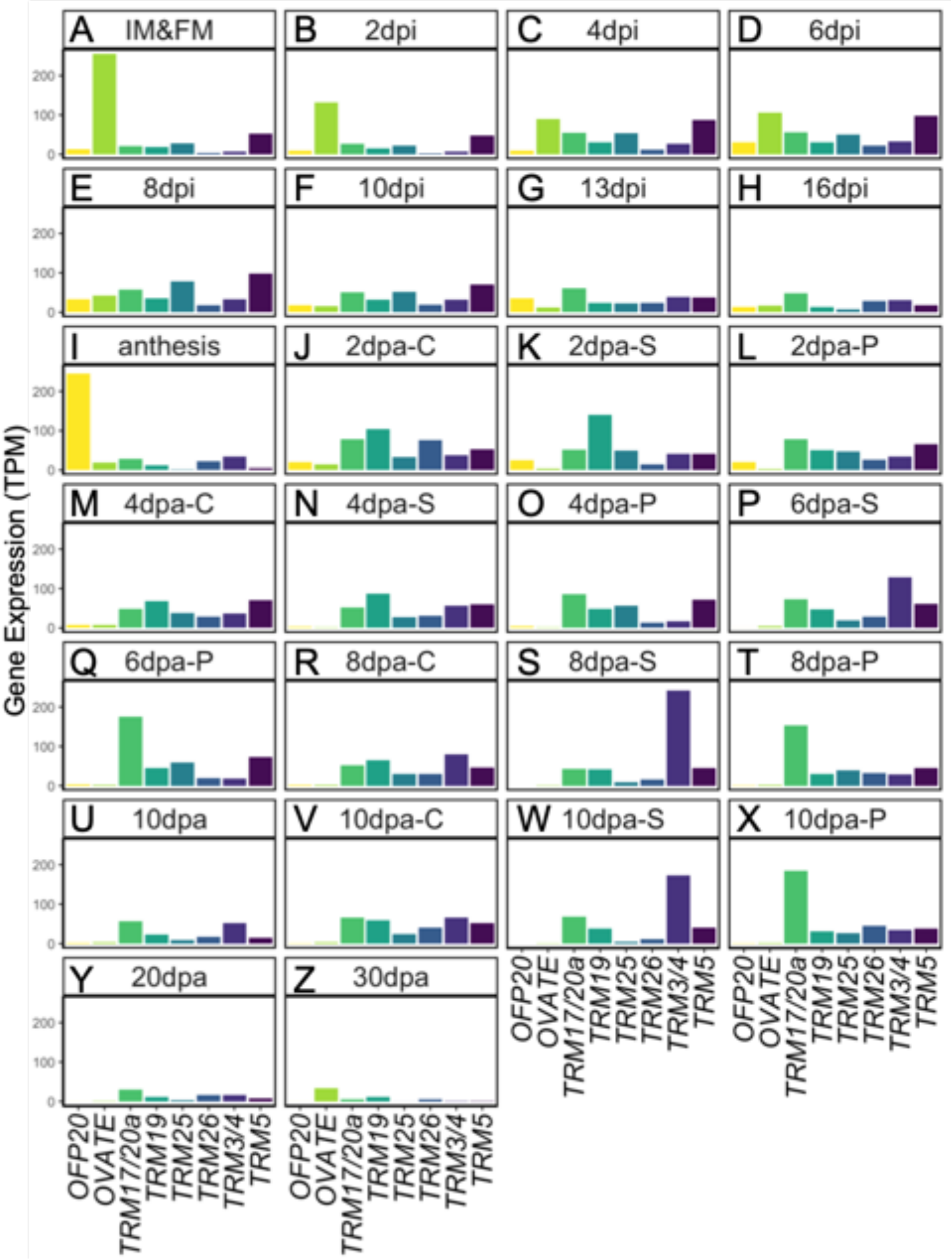

Figure 4.1. Expression bar charts. Depicts the six selected *OFP*-interacting *TRMs*, *OVATE,* and *OFP20* at each developing floral and fruit tissue-timepoint. Expression data for each tissue-timepoint is depicted in its own panel, with panels (**A-Z**) ordered chronologically from left-to-right then top-to-bottom. The height of each colored bar depicts gene expression in TPM, labeled on Y-axes, and genes are labeled along the X-axes at the bottom of the composite plot.

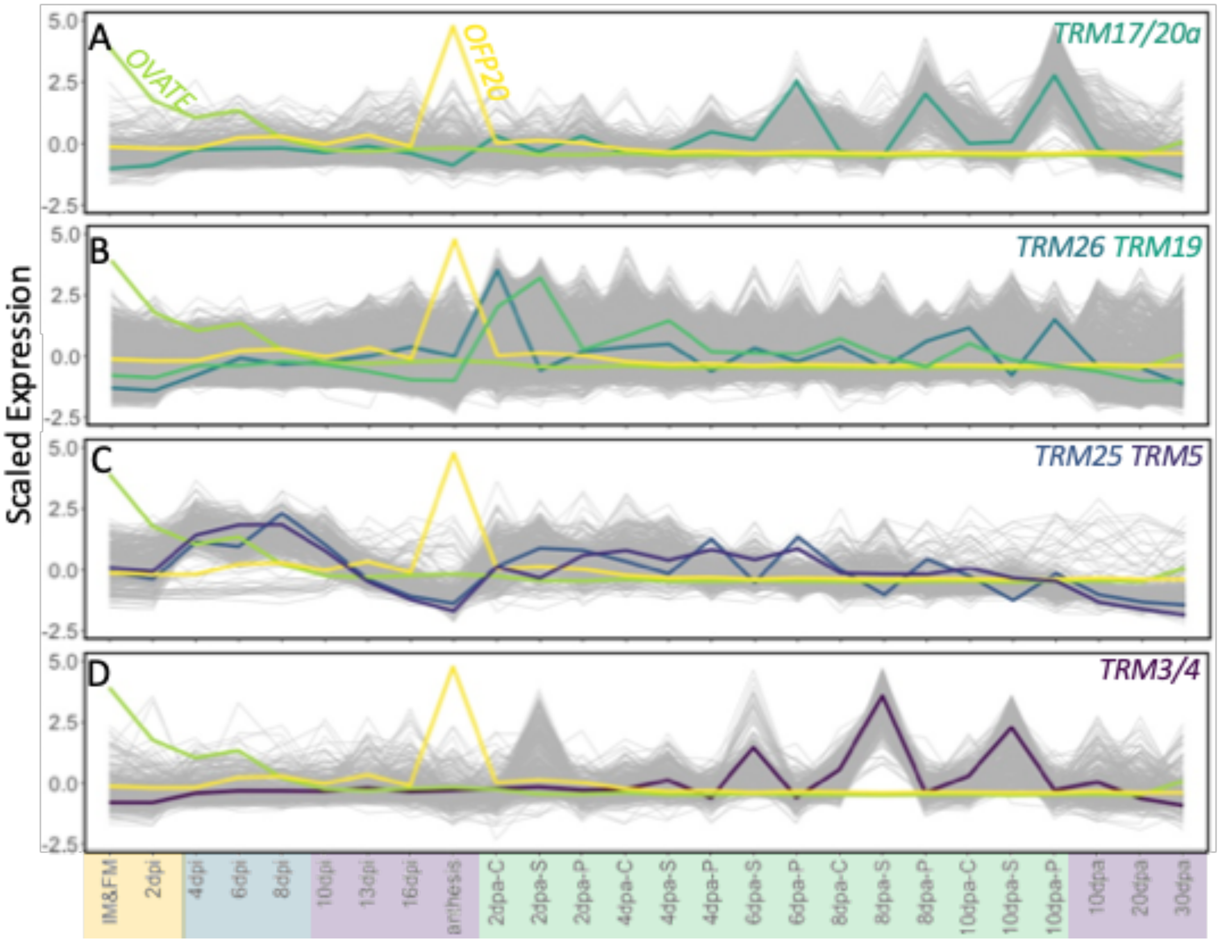

Figure 4.2. Line graphs of WGCNA coexpression networks containing select *OFP*-interacting *TRMs*. Z-score scaled expression displayed on the Y-axis, with lines connecting expression levels of the same gene at each developmental tissue-timepoint (displayed on the X-axis). *OVATE* and *OFP20* are displayed in lime green and yellow on each plot for comparison. **A)** WGCNA cluster containing *TRM17/20a* (marked in blue-green). **B)** WGCNA cluster containing *TRM26* (pale blue) and *TRM19* (green). **C)** WGCNA cluster containing *TRM25* (blue) and *TRM5* (purple). **D)** WGCNA cluster containing *TRM3/4* (magenta).

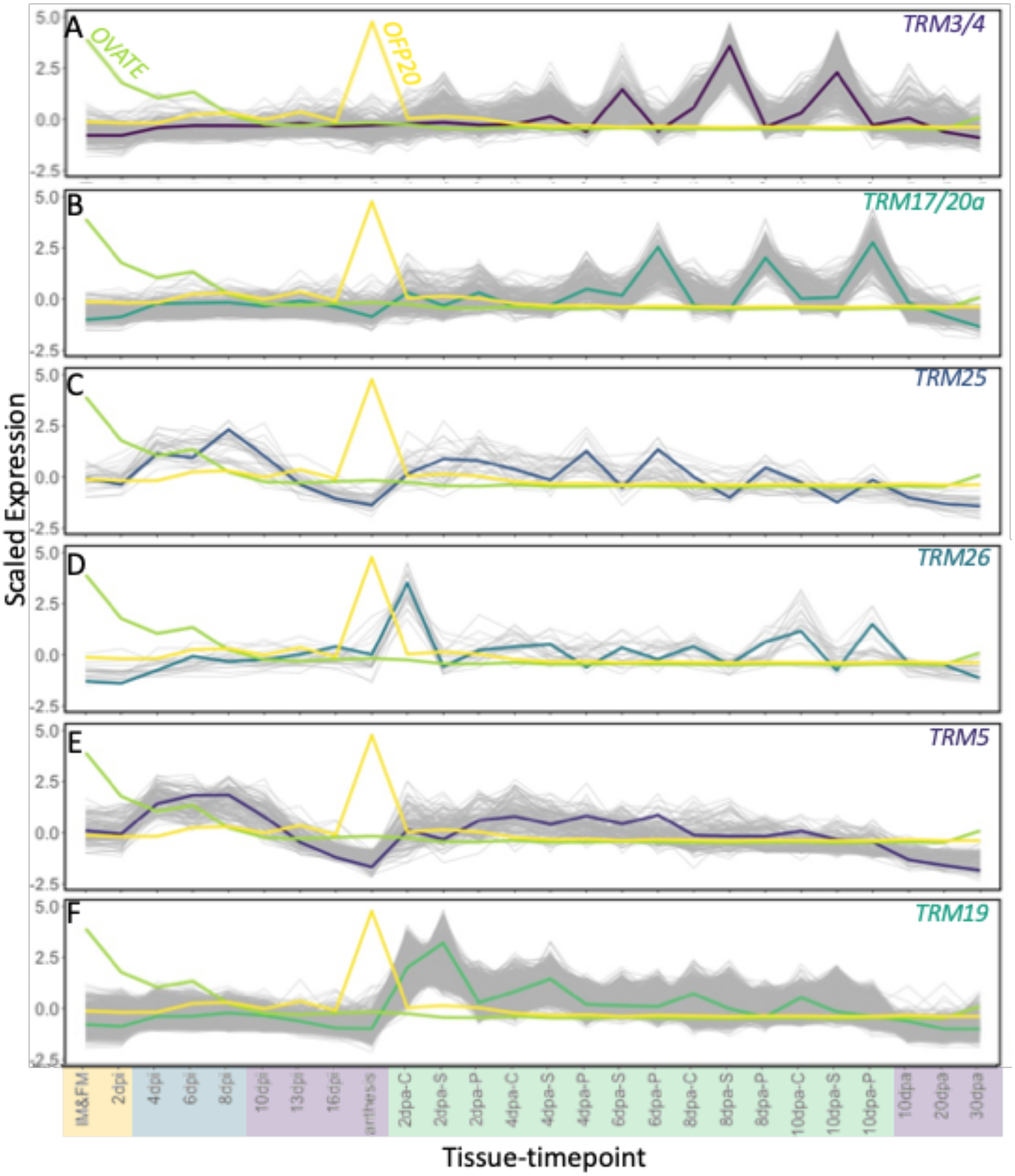

Figure 4.3. Select *OFP*-interacting *TRM* hard-threshold correlation network line-graphs. Expression scaled as in Figure 4.5 displayed on the Y-axis at each tissue-timepoint (X-axis), with expression of *OFP20* depicted in yellow, *OVATE* in lime green, one of the six selected *OFP*-interacting *TRM* genes in colored lines with its respective coexpressed genes in grey, for each panel. **A)** *TRM3/4* and coexpressed genes **B)** *TRM17/20a* and coexpressed genes **C)** *TRM25* and coexpressed genes **D)** *TRM26* and coexpressed genes **E)** *TRM5* and coexpressed genes **F)** *TRM19* and coexpressed genes.

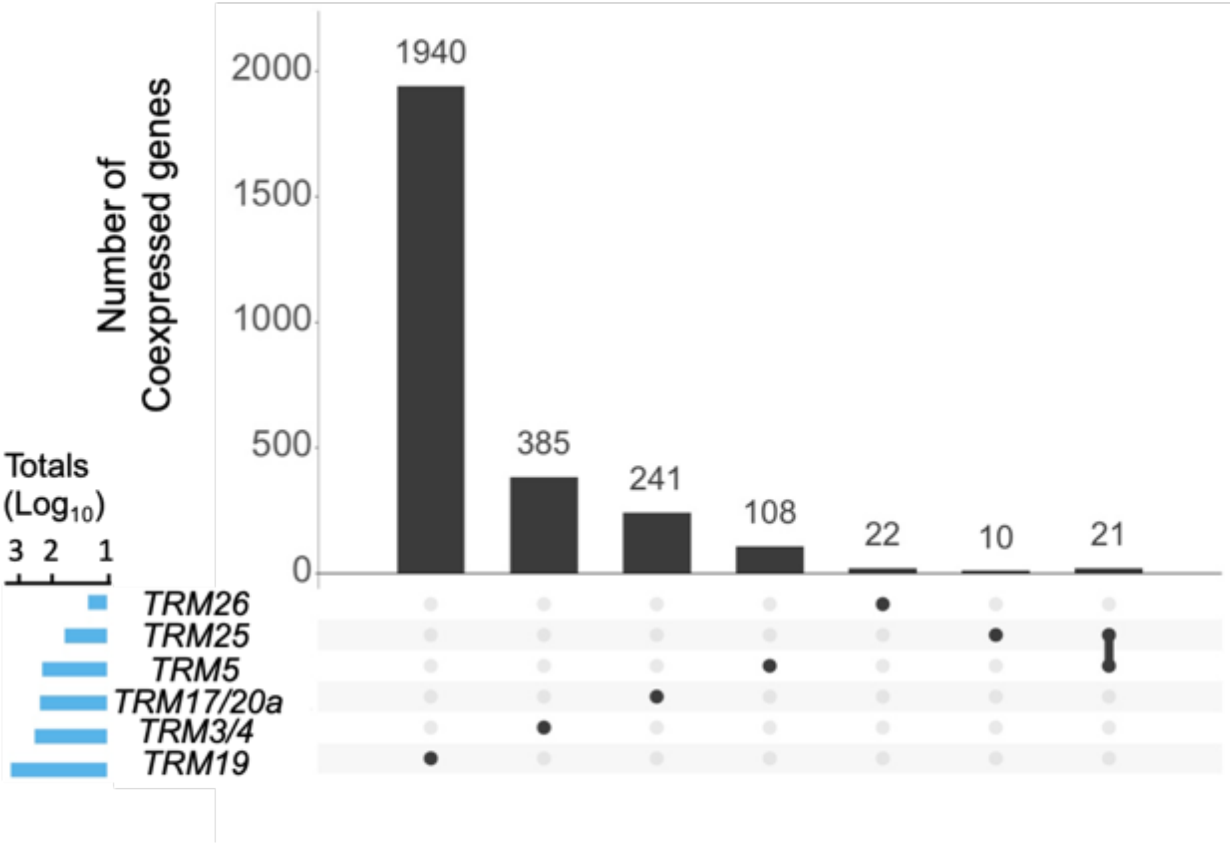

Figure 4.4. UpSet plot depicting the number of coexpressed genes. Number of genes unique to a specific *TRM*'s hard-threshold co-expression network shown by black bars (top right), in groups indicated by black circles beneath each black bar (bottom right). Circles are aligned horizontally with the core gene used for coexpression analysis. Black bars above connected circles represent the number of shared genes between core coexpression groups indicated. Blue bars (bottom left) represent total numbers of genes coexpressed with each core gene.

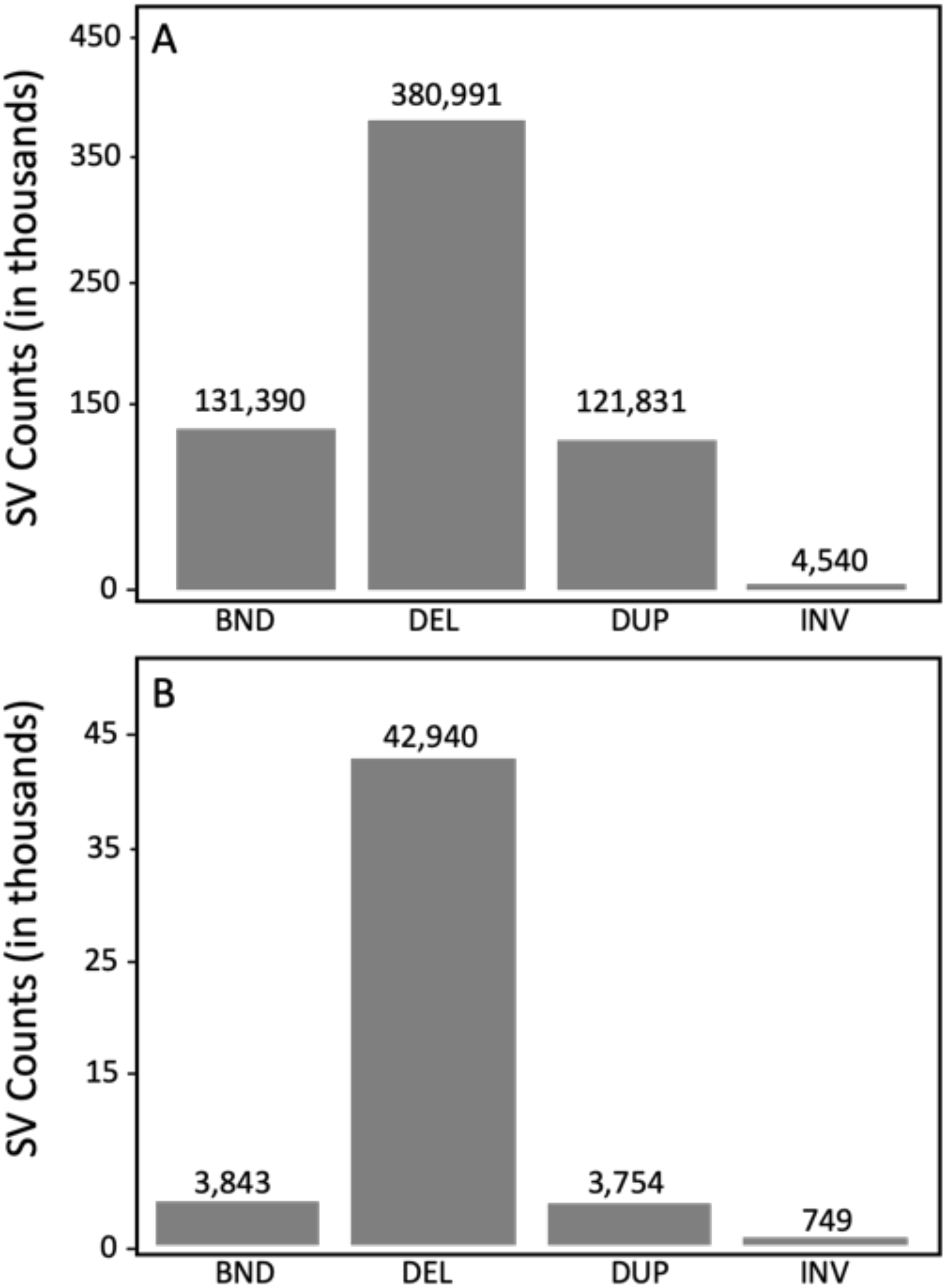

Figure 4.5. Barplots depicting structural variant counts before and after filtering. A) Number of breakend (BND), deletion (DEL), duplication (DUP), and inversion (INV) SVs shown in gray, with SV count on the Y-axis and exact figures above each bar.

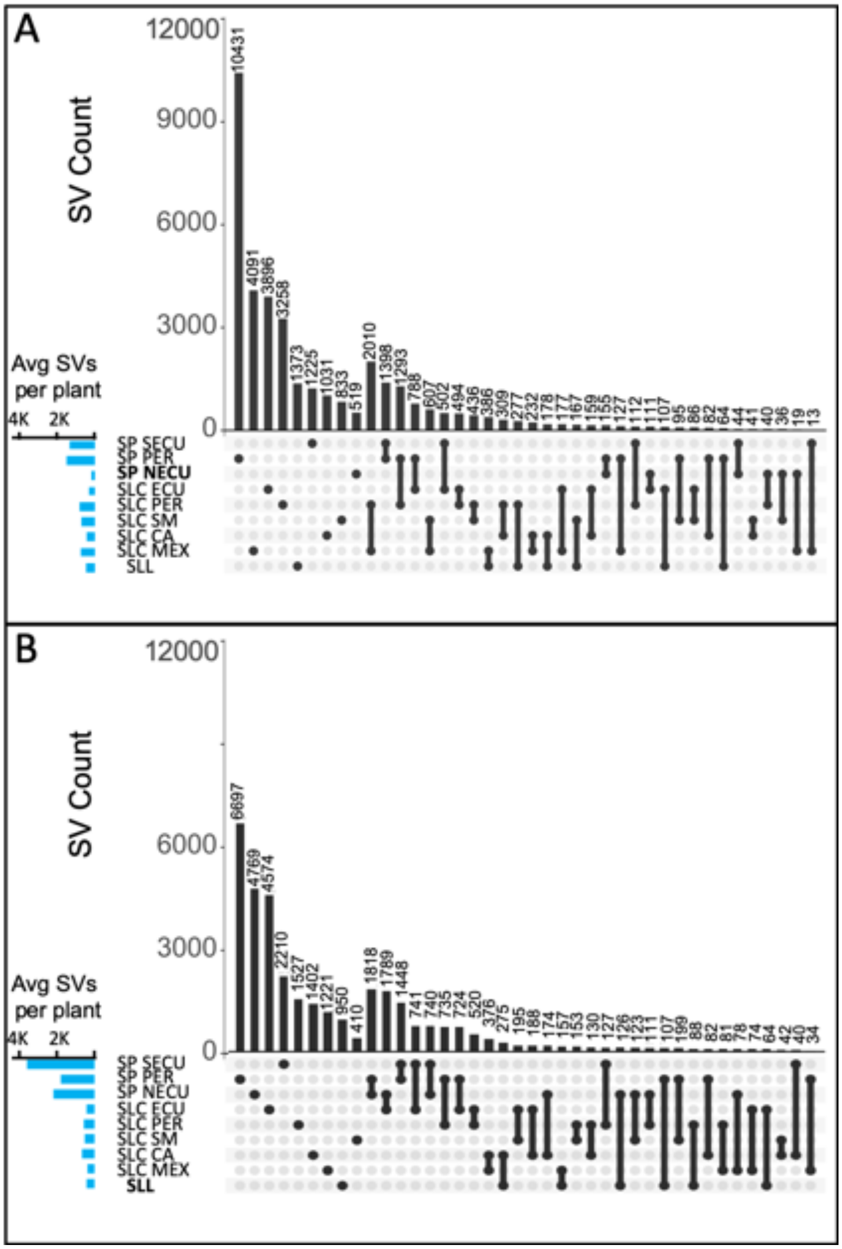

Figure 4.6. SV count by subpopulation UpSet plot. UpSet plots depicting the number of SVs unique to and shared between each subpopulation for **A)** SVs called against the BGV006775 reference genome and **B)** SVs called against the Heinz 1706 reference genome. Black bars depict the number of SVs in each group. Black circles under each black bar are aligned with the corresponding subpopulation(s), with circles connected when the black bar above represents the SVs shared between separate subpopulations. Blue bars at left depict the total SV counts per subpopulation divided by the number of accessions in that subpopulation, for an average number of SVs per accession in each subpopulation.

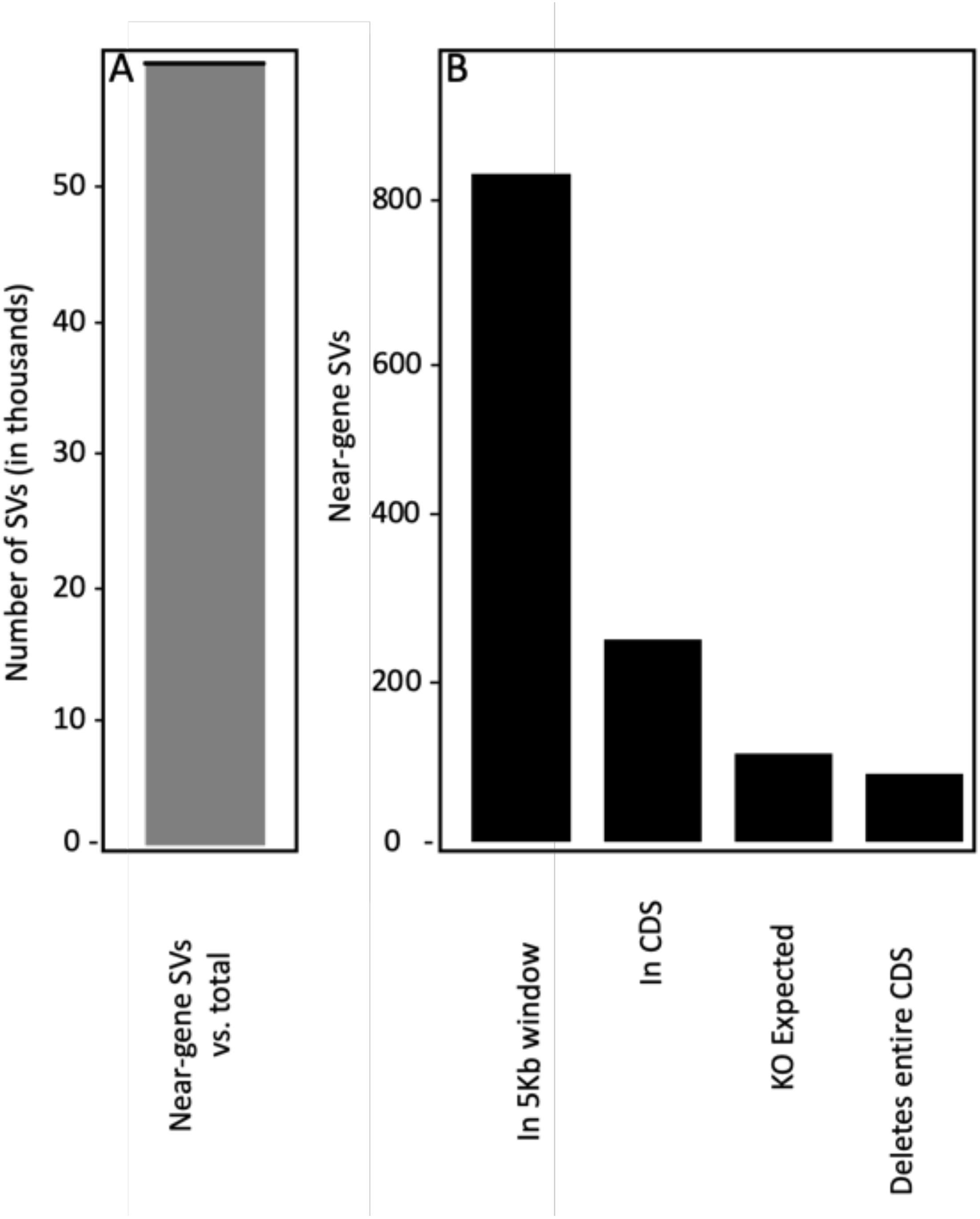

Figure 4.7. Barplots depicting SV-gene overlap patterns. A) Stacked barplot showing proportion of SVs falling within 5 Kbp of any suspect genes identified via coexpression analysis (black), and those falling outside of suspect genes (gray). **B)** Barplot characterizing SVs falling near or on genes. From left-to-right: SVs falling within 5Kbp of the protein-coding sequence (CDS) of the gene, SVs falling within the CDS, SVs deleting the transcription start site or >50% of the gene's CDS, SVs deleting the entire CDS.

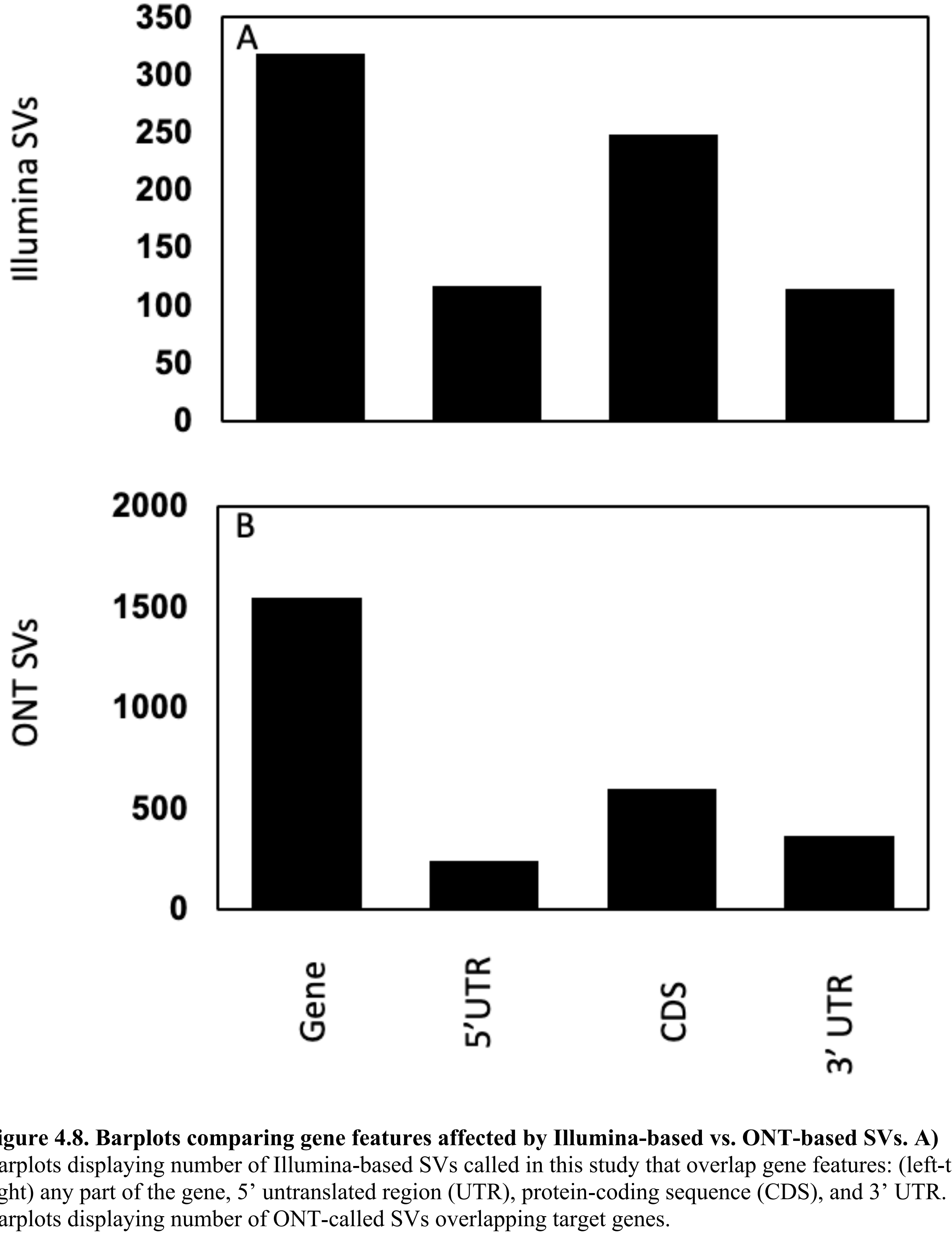

Figure 4.8. Barplots comparing gene features affected by Illumina-based vs. ONT-based SVs. A) Barplots displaying number of Illumina-based SVs called in this study that overlap gene features: (left-to-right) any part of the gene, 5' untranslated region (UTR), protein-coding sequence (CDS), and 3' UTR. **B)** Barplots displaying number of ONT-called SVs overlapping target genes.

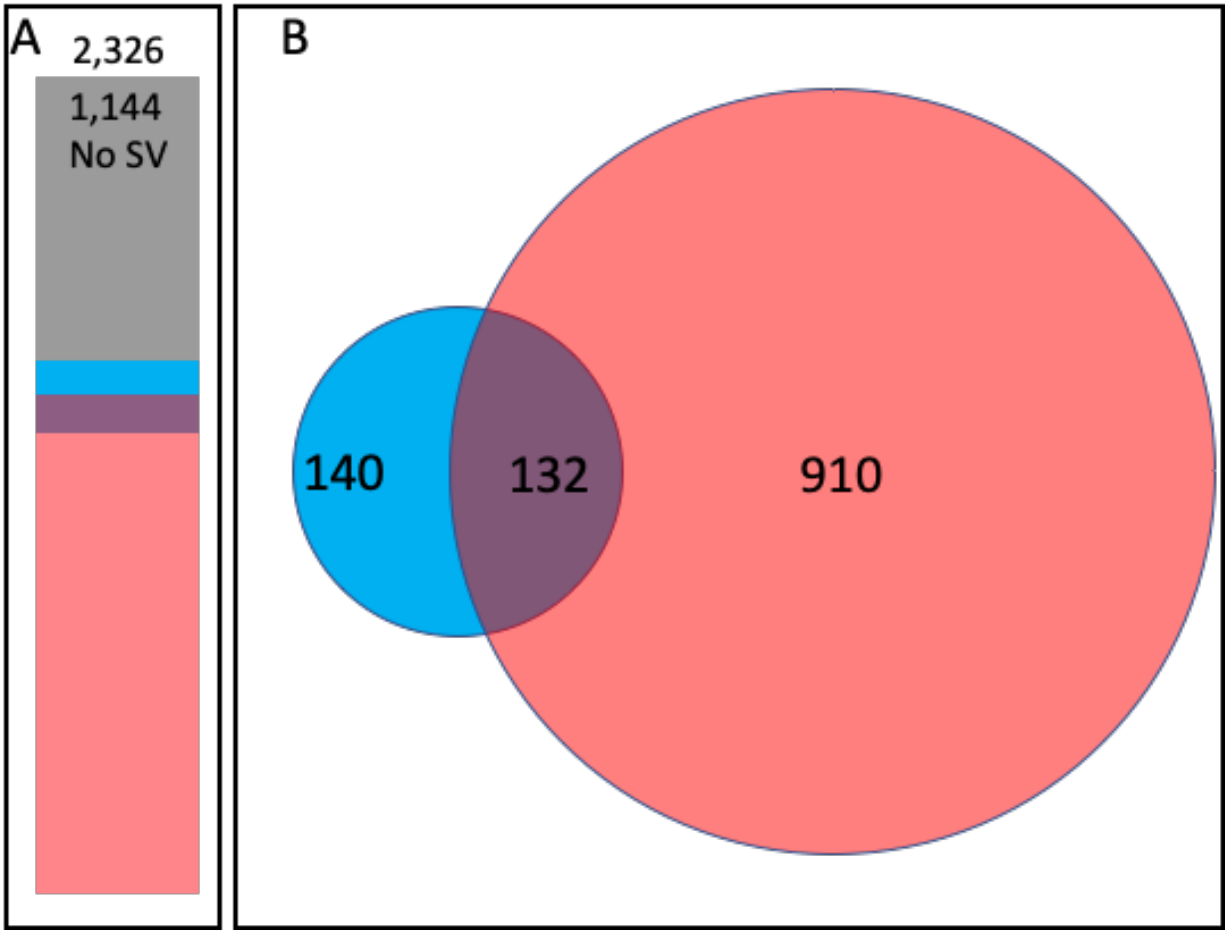

Figure 4.9. Target genes with SVs in Illumina vs ONT datasets. A) Stacked barplot displaying proportion of target genes with no SVs (in gray), Illumina-based SVs only (in blue), both Illumina and ONT-based SVs (purple), and ONT-based SVs only (red). **B)** Venn diagram depicting the overlap between genes bearing Illumina-based SVs (blue) and ONT-based SVs (red).

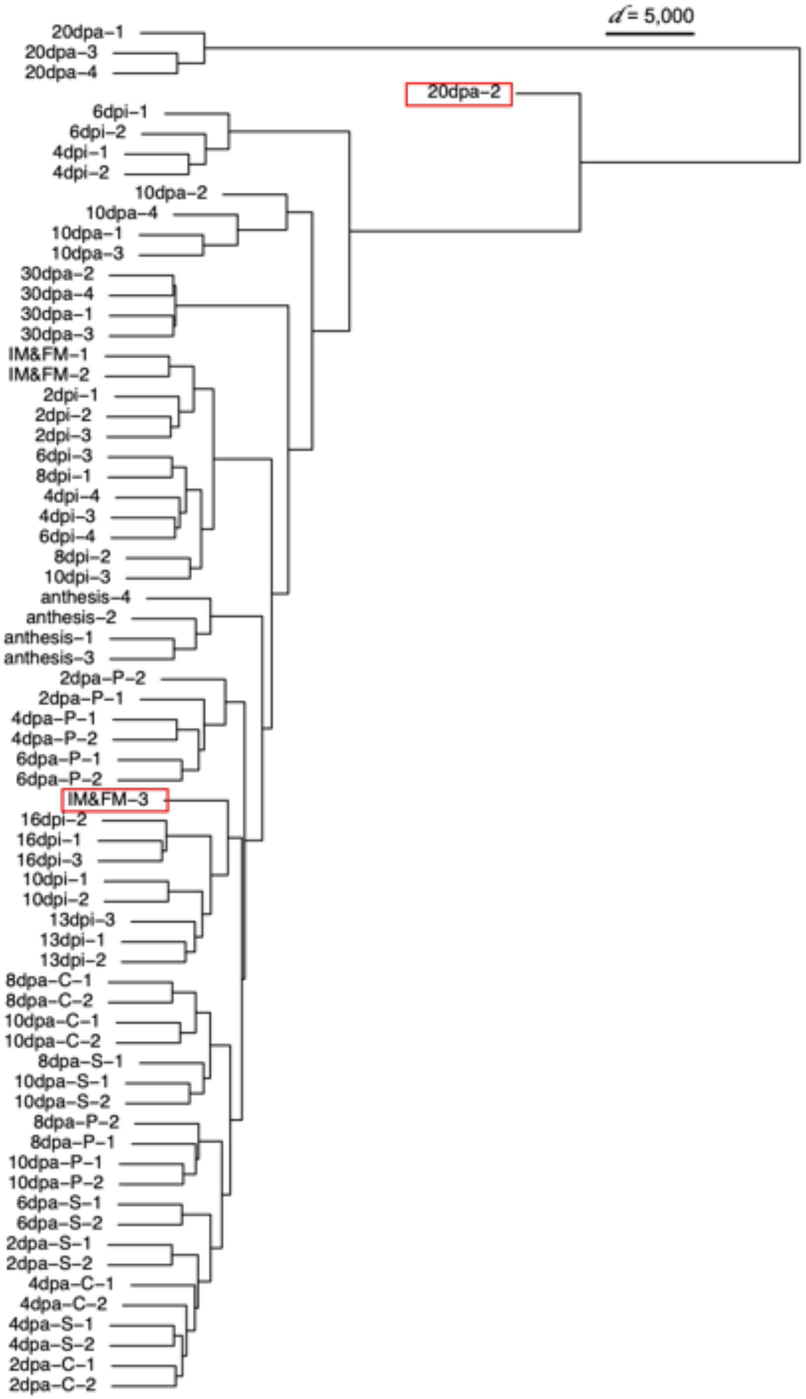

Figure 4.S1. Sample clustering by similar expression. Hierarchical clustering of samples as labeled at branch tips, calculated by Euclidean distance (*d*) between each sample based on gene expression profile across all filtered genes at that timepoint. Branch lengths depict *d* between samples as shown by the scale bar in black. Removed samples shown in red.

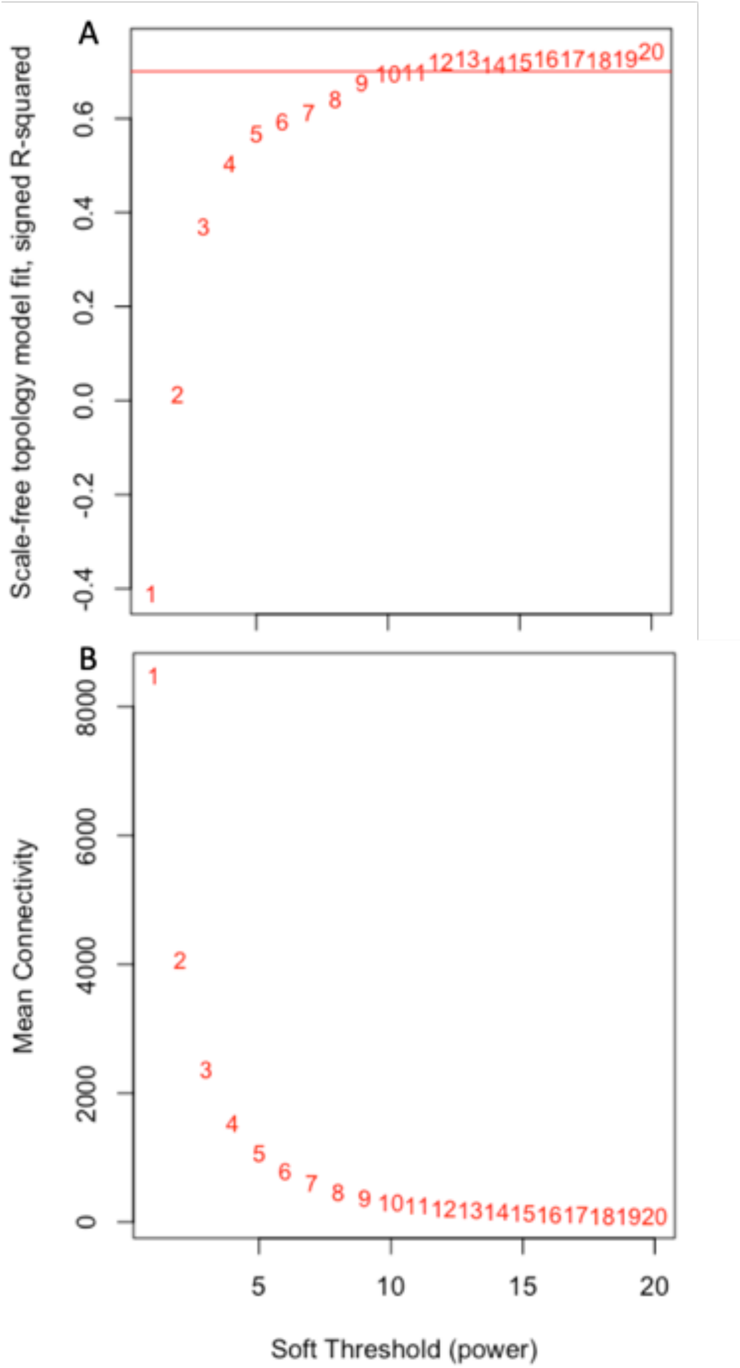

Figure 4.S2. Identification of WGCNA model soft threshold parameter. A) R-squared model fit (Y-axis) at considered soft threshold levels (X-axis) plotted as red numbers, with the R-squared level (0.7) at the selected soft threshold level (10) indicated as a horizontal red line. **B)** Mean connectivity of the model (Y-axis) at various soft threshold levels (X-axis) plotted as red numbers.

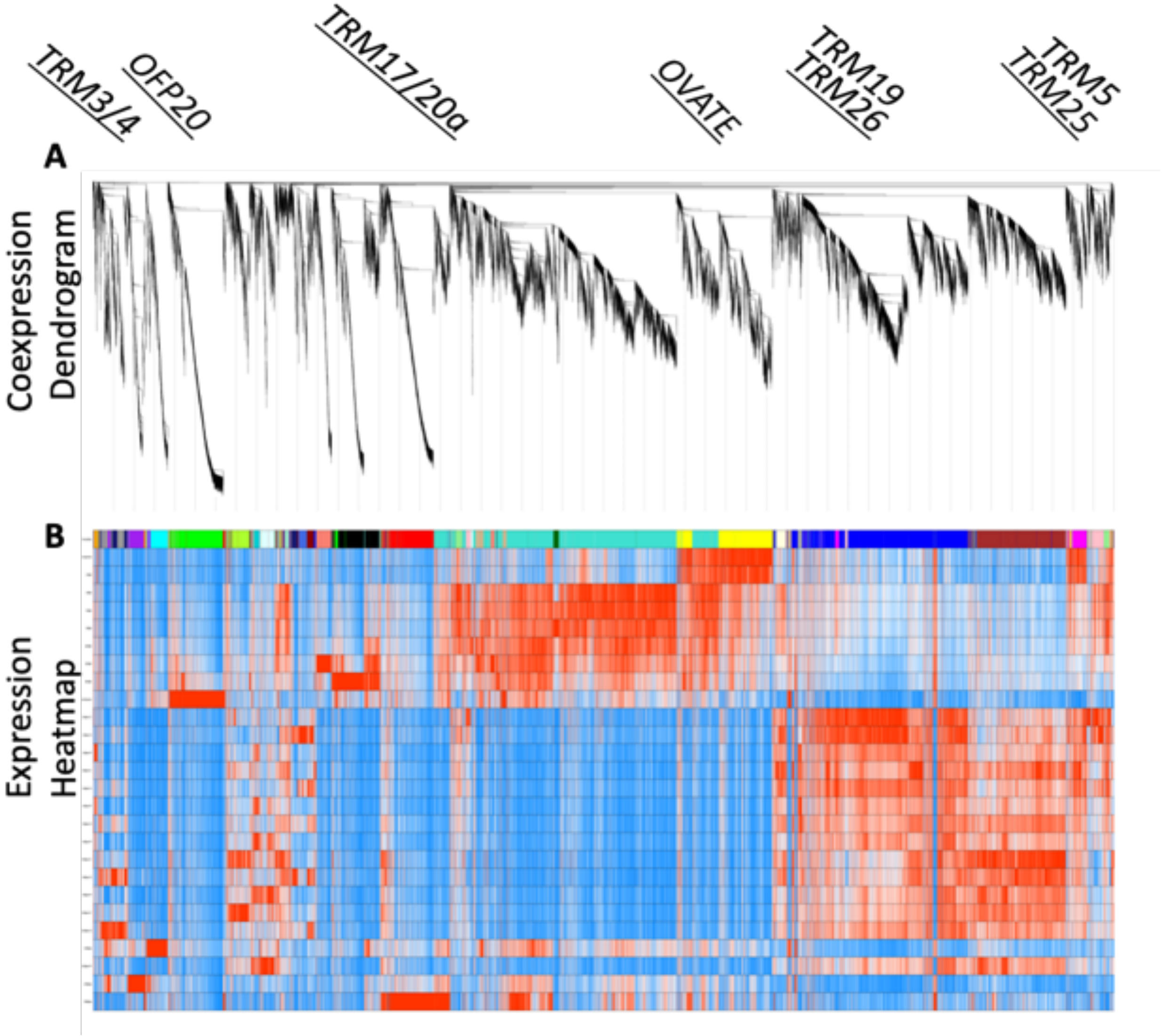

Figure 4.S3. WGCNA gene cluster and expression. A) Dendrogram depicting gene expression hierarchical clusters as calculated by WGCNA, with genes of interest labeled above their respective groups: *TRM5* and *TRM25* in cerulean, *TRM3/4* in turquoise, *TRM17/20a* in purple, *TRM19* and *TRM26* in blue, and *OVATE* in yellow colored groups. *OFP20* was designated as gray (unassigned). **B)** Expression heatmap depicting gene expression for each gene across all samples, with higher expression shown in red, and lower expression shown in blue.

Table 4.1. GO-terms significantly enriched in stringent TRM19 hard-threshold correlation coexpression cluster. ($p < 0.01$, $R > 0.8$). No significantly enriched GO-terms were found for the five other selected *TRMs*.

GO.ID	*Term*	*Annotated*	*Significant*	*Expected*	*P-value (FDR-adj.)*
GO:0006259	DNA metabolic process	569	21	3.26	1.76E-08
GO:0006974	cellular response to DNA damage stimulus	331	13	1.9	2.67E-05
GO:0010605	negative regulation of macromolecule met.	525	14	3.01	0.000450274
GO:0016070	RNA metabolic process	2805	35	16.06	0.000560162
GO:0045934	negative regulation of nucleobase-cont.	236	9	1.35	0.001643707
GO:0010467	gene expression	3454	37	19.77	0.006182469
GO:0006325	chromatin organization	325	9	1.86	0.013710256
GO:0031327	negative regulation of cellular biosynthesis	272	8	1.56	0.021642619
GO:1903046	meiotic cell cycle process	106	5	0.61	0.035310377
GO:0048608	reproductive structure development	773	13	4.43	0.042213158
GO:0090305	nucleic acid phosphodiester bond hydrolysis	242	7	1.39	0.043501017
GO:0000280	nuclear division	118	5	0.68	0.049226349

Table 4.2. GO-terms enriched in coexpression cluster for WGCNA coexpression network containing TRM17/20a.

GO.ID	*Term*	*Annotated*	*Significant*	*Expected*	*P-value (FDR-adj.)*
GO:0010143	cutin biosynthetic process	18	6	0.27	0.00085552
GO:0010025	wax biosynthetic process	18	5	0.27	0.00727192
GO:0044264	cellular polysaccharide metabolic process	238	13	3.53	0.020675067
GO:0044275	cellular carbohydrate catabolic process	70	7	1.04	0.026400813
GO:0000272	polysaccharide catabolic process	183	11	2.72	0.028142105

Table 4.3. GO-terms enriched in coexpression cluster for WGCNA coexpression network containing TRM26 and TRM19.

GO.ID	*Term*	*Annotated*	*Significant*	*Expected*	*P-value (FDR-adj.)*
GO:0016070	RNA metabolic process	2805	623	442.4	1.23E-20
GO:0015031	protein transport	443	132	69.87	5.48E-12
GO:0043414	macromolecule methylation	198	74	31.23	1.26E-11
GO:0044267	cellular protein metabolic process	3514	692	554.23	1.34E-10
GO:0006259	DNA metabolic process	569	154	89.74	2.26E-10
GO:0009790	embryo development	287	92	45.27	3.98E-10
GO:0010154	fruit development	429	117	67.66	4.59E-08
GO:0010228	vegetative to reproductive phase transit.	128	47	20.19	3.97E-07
GO:0048366	leaf development	222	66	35.01	5.75E-06
GO:0016569	covalent chromatin modification	189	57	29.81	2.25E-05
GO:0036211	protein modification process	2422	464	382	3.34E-05
GO:0010604	positive regulation of macromolecule met.	369	93	58.2	6.95E-05
GO:1902275	regulation of chromatin organization	46	21	7.26	6.95E-05
GO:0010605	negative regulation of macromolecule met.	525	122	82.8	0.00014864
GO:0043043	peptide biosynthetic process	781	170	123.18	0.000150396
GO:0090567	reproductive shoot system development	321	81	50.63	0.000259625
GO:1905037	autophagosome organization	20	12	3.15	0.000303251
GO:0034645	cellular macromolecule biosynthetic proc.	2936	540	463.07	0.000448123
GO:0090305	nucleic acid phosphodiester bond hydroly.	242	64	38.17	0.000448123
GO:0006338	chromatin remodeling	37	17	5.84	0.000504057
GO:0010015	root morphogenesis	143	41	22.55	0.00186859
GO:0044265	cellular macromolecule catabolic process	623	134	98.26	0.002068258
GO:0090558	plant epidermis development	192	49	30.28	0.006950244
GO:0006520	cellular amino acid metabolic process	367	80	57.88	0.02003625
GO:0090627	plant epidermal cell differentiation	82	23	12.93	0.037918065

Table 4.4 GO-terms enriched in coexpression cluster for WGCNA coexpression network containing TRM5 and TRM25.

GO.ID	*Term*	*Annotated*	*Significant*	*Expected*	*P-value*
GO:1903047	mitotic cell cycle process	113	11	1.25	4.54E-05
GO:0051276	chromosome organization	556	22	6.13	0.00014036
GO:0010564	regulation of cell cycle process	59	8	0.65	0.00015447
GO:0000280	nuclear division	118	9	1.3	0.00259124
GO:0090304	nucleic acid metabolic process	3287	61	36.22	0.00342208
GO:0045786	negative regulation of cell cycle	45	5	0.5	0.0224229
GO:0098813	nuclear chromosome segregation	76	6	0.84	0.02988029
GO:0036211	protein modification process	2422	45	26.69	0.03267611
GO:0000226	microtubule cytoskeleton organization	79	6	0.87	0.03290462
GO:0048367	shoot system development	557	16	6.14	0.04627212
GO:0044770	cell cycle phase transition	33	4	0.36	0.04627212

Table 4.5. SV-calling tools and statistics. Note: *Polished with paired-end Illumina data

	Varitome SVs	**Comparison SVs**
Sequencing Platform	Illumina	Oxford Nanopore*
Sequence Type	Paired-end	Long-read
Read length	150 bp	20-30 kbp
Fragment length	550-650 bp	20-30 kbp
Aligner	BWA-mem	NGMLR
SV-caller	LUMPY	SNIFFLES
Variants per sample	170-8,700	5,000-15,000
Average sequencing depth	18X	74X

Table 4.6. Select SVs and affected coexpression network genes. Rows in bold mark genes of particular interest for the TRM-OFP network.

Start	*End*	*Type*	*Affected gene*	*Accessions*	*% CDS affected*	*CDS Start affected*	*CDS end affected*	*TRM*
84992006	**85056059**	**DEL**	**Solyc01g091380**	**BGV007908**	**100**	**yes**	**yes**	**TRM19**
2900301	**3439085**	**DUP**	**Solyc04g009550**	**LA1712**	**100**	**yes**	**yes**	**TRM19**
2900301	3439085	DUP	Solyc04g009730	LA1712	100	yes	yes	TRM19
2900301	**3439085**	**DUP**	**Solyc04g009950**	**LA1712**	**100**	**yes**	**yes**	**TRM19**
2900301	3439085	DUP	Solyc04g009980	LA1712	100	yes	yes	TRM17/20a
2900301	3439085	DUP	Solyc04g009990	LA1712	100	yes	yes	TRM19
3504462	**3547002**	**DUP**	**Solyc04g010220**	**BGV006767, BGV014522, TS-75**	**58**	**yes**	**no**	**TRM19**
1095529	**1101076**	**DEL**	**Solyc05g006470**	**BGV015380, BGV015382, PAS014479, TS-411, TS-420**	**35**	**yes**	**no**	**TRM19**

CHAPTER 5

CONCLUSION

Summary

This purpose of this research was to better understand the molecular underpinnings of fruit morphological diversity in two diverse fruit biological systems: *Capsicum spp.* peppers and *Solanum spp.* tomatoes. It quantified and categorized the relationships between landrace chile peppers from southern Mexico, near the domestication hearth of *C. annuum*, and compared these to a diverse global population of peppers from seed companies and hobbyists. Two of these accessions, varying widely in fruit shape, were then used as parents to narrow and identify candidate genes for a fruit shape QTL on chromosome 3 that includes a candidate gene homologous to the tomato *OVATE*-interacting *TRM25,* implying parallels between fruit shape control in pepper and tomato. In tomato, this work leverages a higher degree of sequencing, histological, and protein-protein interaction evidence to explore a promising fruit-shape regulatory pathway. It identifies coexpression networks for several genes identified as components of that pathway, and catalogues structural variants in a basal tomato population.

The first study explored a novel collection of chile peppers, which was mostly focused on diverse Mexican landraces that had distinctly different uses. Chapter 2 explored the genetic structure of a novel collection of chile pepper landraces and named pepper types. It showed that, with few but notable exceptions, traditionally named pepper types did indeed form genetically distinct populations. Genetic diversity was found to vary by the cultivation systems that landrace populations were primarily grown in. In the Tusta population, evidence of *C. annuum* – *C. frutescens* interspecific hybridization was found at the root of the population. In scans of genetic diversity across the genomes of each population, we identified signals of selection acting on certain pepper subpopulations with unique phenotypes. A high degree of outcrossing was found in the global population relative to the landrace populations, possibly a result of extensive crosses done by pepper breeders and enthusiasts worldwide, and the breakdown of landrace distinctions when removed from their traditional contexts. This work points toward avenues for further research into abiotic stress, disease resistance, and ancient pepper landrace hybridization.

Chapter 3 reports a refined fruit shape QTL on chromosome 3, and identifies promising candidate genes, including a TRM that would explain much of the observed mutant phenotype. A SNP in the TRM gene within *fs3.1* was genotyped in the Mexican landrace population. Finally, the work also explores pepper TRMs with the M8 OFP-binding domain, where close parallels are observed between the pepper and tomato TRM families.

Chapter 4 explores the coexpression networks of six *TRMs* whose protein products interact with OFPs implicated in fruit shape regulation. We searched genes in those networks for SVs varying in a tomato population also taken from near that species' center of genetic diversity, and identifies key candidate genes. These SVs could serve as landmarks for fruit morphology QTL-mappers to be aware of when seeking to identify candidate genes. Affected gene expression analyses of SV-mutant accessions might be useful, especially where the SV leads to a change in gene copy number or disrupts a promoter region. Eventually, CRISPR mutants, overexpression constructs, or gene silencing could be employed to confirm phenotypic effects suspected to be caused by one of these SVs.

Opportunities for Future Research

These studies present promising opportunities for future research. More detailed population genetics, together with an exploration or modeling of ancient trade routes in the mountainous regions of Oaxaca, could help us to better understand the forces that contributed to landrace pepper diversity in the first study. The QTL in the second study can also be further narrowed, and complementation tests, transgenic manipulation, and the creation of near isogenic lines might help to further explore the fruit shape phenotype of *TRM25* or other important candidate genes. The third study provided a series of SVs that could be explored by biparental populations to map the loci identified as disrupting *TRM-OFP* coexpression network genes when such loci are also associated with phenotypes. The genes in our network can be assessed for altered expression in our population, especially where SVs are identified nearby.

References

Chapter 1 References

Bosland, P. W. 1996. Capsicums: Innovative Uses of an Ancient Crop. Progress in New Crops. Arlington, VA, ASHS Press**:** 479-487.

Bosland, P. W. and E. J. Votava 2012. Peppers: vegetable and spice capsicums. New York, USA, CABI Publishing.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss and E. S. Buckler 2007. "TASSEL: software for association mapping of complex traits in diverse samples." Bioinformatics **23**(19): 2633-2635.

Brown, C. H., C. R. Clement, P. Epps, E. Luedeling and S. Wichmann 2013. "The paleobiolinguistics of domesticated chili pepper (Capsicum spp.)." Ethnobiology Letters **4**: 1-11.

Castellón-Martínez, É., J. L. Chávez-Servia, J. C. Carrillo-Rodríguez and A. M. Vera-Guzman 2012. "Preferencias de consumo de chiles (Capsicum annuum L.) nativos en los valles centrales de Oaxaca, México." Revista fitotecnia mexicana **35**(SPE.5): 27-35.

Darwin, C. 1868. The variation of animals and plants under domestication, O. Judd.

Drevensek, S., M. Goussot, Y. Duroc, A. Christodoulidou, S. Steyaert, E. Schaefer, E. Duvernois, O. Grandjean, M. Vantard and D. Bouchez 2012. "The Arabidopsis TRM1–TON1 interaction reveals a recruitment network common to plant cortical microtubule arrays and eukaryotic centrosomes." The Plant Cell **24**(1): 178-191.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler and S. E. Mitchell. 2011. "A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species." PLoS ONE **6**(5): e19379.

Fernandez-Pozo, N., Y. Zheng, S. I. Snyder, P. Nicolas, Y. Shinozaki, Z. Fei, C. Catala, J. J. Giovannoni, J. K. Rose and L. A. Mueller. 2017. "The Tomato Expression Atlas." Bioinformatics **33**(15): 2397-2398.

González-Jara, P., A. Moreno-Letelier, A. Fraile, D. Piñero and F. García-Arenal. 2011. "Impact of Human Management on the Genetic Variation of Wild Pepper, Capsicum annuum var. glabriusculum." PLoS ONE **6**(12): e28715.

Hernández-Verdugo, S., R. Luna-Reyes and K. Oyama. 2001. "Genetic Structure and Differentiation of Wild and Domesticated Populations of Capsicum Annuum (solanaceae) from Mexico." Plant Systematics and Evolution **226**(3-4): 129-142.

Hill, T. A., H. Ashrafi, S. Reyes-Chin-Wo, J. Yao, K. Stoffel, M.-J. Truco, A. Kozik, R. W. Michelmore and A. Van Deynze. 2013. "Characterization of Capsicum annuum genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene Pepper GeneChip." PloS one **8**(2): e56200.

Hulse-Kemp, A. M., H. Ashrafi, J. Plieske, J. Lemm, K. Stoffel, T. Hill, H. Luerssen, C. L. Pethiyagoda, C. T. Lawley and M. W. Ganal. 2016. "A HapMap leads to a Capsicum annuum SNP infinium array: a new tool for pepper breeding." Horticulture research **3**.

Kim, H.-J., K.-H. Baek, S.-W. Lee, J. Kim, B.-W. Lee, H.-S. Cho, W. T. Kim, D. Choi and C.-G. Hur. 2008. "Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (Capsicum annuum) transcriptome." BMC Plant Biology **8**(1): 101.

Kim, S., M. Park, S.-I. Yeom, Y.-M. Kim, J. M. Lee, H.-A. Lee, E. Seo, J. Choi, K. Cheong and K.-T. Kim. 2014. "Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species." Nature genetics **46**(3): 270-278.

Kraft, K. H., C. H. Brown, G. P. Nabhan, E. Luedeling, J. Luna Ruiz Jde, G. Coppens d'Eeckenbrugge, R. J. Hijmans and P. Gepts. 2014. "Multiple lines of evidence for the origin of domesticated chili pepper, Capsicum annuum, in Mexico." Proc Natl Acad Sci U S A **111**(17): 6165-6170.

Kraft, K. H., J. de Jesús Luna-Ruíz and P. Gepts. 2013. "A new collection of wild populations of Capsicum in Mexico and the southern United States." Genetic resources and crop evolution **60**(1): 225-232.

Lee, Y. K., G.-T. Kim, I.-J. Kim, J. Park, S.-S. Kwak, G. Choi and W.-I. Chung. 2006. "LONGIFOLIA1 and LONGIFOLIA2, two homologous genes, regulate longitudinal cell elongation in Arabidopsis." Development **133**(21): 4305-4314.

Liu, D., W. Sun, Y. Yuan, N. Zhang, A. Hayward, Y. Liu and Y. Wang. 2014. "Phylogenetic Analyses Provide the First Insights into the Evolution of OVATE Family Proteins in Land Plants." Annals of Botany **113**: 1219-1233.

Liu, J., J. V. Eck, B. Cong and S. D. Tanksley. 2002. "A New Class of Regulatory Genes Underlying the Cause of Pear-shaped Tomato Fruit." Proceedings of the National Academy of Sciences **99**(20): 13302-13306.

Maggioni, L., R. v. Bothmer, G. Poulsen and E. Lipman. 2017. "Domestication, Diversity and use of *Brassica oleraceae* L., Based on Ancient Greek and Latin Texts." Genet. Resour. Crop Evol.

McLeod, M., S. I. Guttman and W. H. Eshbaugh. 1982. "Early evolution of chili peppers (Capsicum)." Economic Botany **36**(4): 361-368.

Naegele, R. P., J. Mitchell and M. K. Hausbeck (2016). "Genetic Diversity, Population Structure, and Heritability of Fruit Traits in *Capsicum annuum*." PLoS ONE **11**(7): e0156969.

Perry, L., R. Dickau, S. Zarrillo, I. Holst, D. M. Pearsall, D. R. Piperno, M. J. Berman, R. G. Cooke, K. Rademaker, A. J. Ranere, J. S. Raymond, D. H. Sandweiss, F. Scaramelli, K. Tarble and J. A. Zeidler. 2007. "Starch Fossils and the Domestication and Dispersal of Chili Peppers (Capsicum spp. L.) in the Americas." Science **315**(5814): 986-988.

Perry, L. and K. V. Flannery (2007). "Precolumbian use of chili peppers in the Valley of Oaxaca, Mexico." Proceedings of the National Academy of Sciences **104**(29): 11905-11909.

Pickersgill, B. 1988. "The genus Capsicum: a multidisciplinary approach to the taxonomy of cultivated and wild plants." Biologisches Zentralblatt **107**(4): 381-389.

Popovsky-Sarid, S., Y. Borovsky, A. Faigenboim, E. P. Parsons, G. T. Lohrey, S. Alkalai-Tuvia, E. Fallik, M. A. Jenks and I. Paran. 2017. "Genetic and biochemical analysis reveals linked QTLs determining natural variation for fruit post-harvest water loss in pepper (Capsicum)." Theor Appl Genet **130**(2): 445-459.

Pritchard, J. K., M. Stephens and P. Donnelly. 2000. "Inference of population structure using multilocus genotype data." Genetics **155**.

Qin, C., C. Yu, Y. Shen, X. Fang, L. Chen, J. Min, J. Cheng, S. Zhao, M. Xu and Y. Luo 2014. "Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization." Proceedings of the National Academy of Sciences **111**(14): 5135-5140.

Schaefer, E., K. Belcram, M. Uyttewaal, Y. Duroc, M. Goussot, D. Legland, E. Laruelle, M.-L. de Tauzia-Moreau, M. Pastuglia and D. Bouchez 2017. "The preprophase band of microtubules controls the robustness of division orientation in plants." Science **356**(6334): 186-189.

Smith, B. D. 1997. "Reconsidering the Ocampo Caves and the Era of Incipient Cultivation in Mesoamerica." Latin American Antiquity **8**(4): 342-383.

Spinner, L., A. Gadeyne, K. Belcram, M. Goussot, M. Moison, Y. Duroc, D. Eeckhout, N. De Winne, E. Schaefer and E. Van De Slijke 2013. "A protein phosphatase 2A complex spatially controls plant cell division." Nature communications **4**: 1863.

Takagi, H., A. Abe, K. Yoshida, S. Kosugi, S. Natsume, C. Mitsuoka, A. Uemura, H. Utsushi, M. Tamiru and S. Takuno 2013. "QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations." The Plant Journal **74**(1): 174-183.

Weir, B. S. and C. C. Cockerham (1984). "Estimating F-Statistics for the Analysis of Population Structure." Evolution **38**(6): 1358-1370.

Wu, S. 2015. The Roles of *OVATE* and Other Elongation Genes in Regulating Proximal-Distal Patterning of Tomato Fruit. Doctor of Philosophy Dissertation, Ohio State University.

Zheng, X. 2012. "SNPRelate: parrallel computing toolset for genome-wide association studies." R package version **95**.

Zhou, Y., J. Miao, H. Gu, X. Peng, M. Leburu, F. Yuan, H. Gu, Y. Gao, Y. Tao, J. Zhu, Z. Gong, C. Yi, M. Gu, Z. Yang and G. Liang (201. "Natural Variations in *SLG7* Regulate Grain Shape in Rice." Genetics **201**: 1591-1599.

Chapter 2 References

Aguilar-Meléndez, A., Morrell, P. L., Roose, M. L., and Kim, S. C. 2009. Genetic diversity and structure in semiwild and domesticated chiles (Capsicum annuum; Solanaceae) from Mexico. American Journal of Botany, 96(6), 1190-1202.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Cambridge, UK: Babraham Institute

Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Diez, M. J., and Nuez, F. 2012. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. *PLoS ONE,* **7**, e48198.

Bosland, P. W., and Votava, E. J. 2012. *Peppers: vegetable and spice capsicums* (2 ed.). New York, USA: CABI Publishing.

Brown, C. H. 2015. Paleobiolinguistics of new world crops and the Otomanguean language family. *Ethnobiology Letters,* **6**, 189-191.

Brown, C. H., Clement, C. R., Epps, P., Luedeling, E., and Wichmann, S. 2013. The paleobiolinguistics of domesticated chili pepper (Capsicum spp.). *Ethnobiology Letters,* **4**, 1-11.

Brush, S. B., Taylor, J. E., and Bellon, M. R. 1992. Technology adoption and biological diversity in Andean potato agriculture. *Journal of Development Economics*, *39*(2), 365-387.

Cook, O. F. 1921. Milpa agriculture: a primitive tropical system: US Government Printing Office.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Sherry, S. T. 2011. The variant call format and VCFtools. *Bioinformatics,* **27**, 2156-2158.

Darwin, C. 1868. *The variation of animals and plants under domestication* (Vol. 2). O. Judd.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol, 4.

Fick, S., and Hijmans, R. 2017. Worldclim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas. *International Journal of Climatology*.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., and Buckler, E. S. 2014. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE,* **9**, e90346.

González-Jara, P., Moreno-Letelier, A., Fraile, A., Piñero, D., and García-Arenal, F. 2011. Impact of Human Management on the Genetic Variation of Wild Pepper, Capsicum annuum var. glabriusculum. *PLoS ONE,* **6**, e28715. doi:10.1371/journal.pone.0028715

Han, K., Jeong, H. J., Yang, H. B., Kang, S. M., Kwon, J. K., Kim, S., and Kang, B. C. 2016. An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (Capsicum annuum). *DNA Research*, *23*(2), 81-91.

Harlan, J. R. 1998. *The living fields: our agricultural heritage*. Cambridge University Press.

Hasegawa, M., Kishino, H., and Yano, T. 1985. Dating the Human-ape Split by a Molecular Clock of Mitochondrial DNA. *Journal of Molecular Evolution,* **22**, 160-174.

Hill, T. A., Chunthawodtiporn, J., Ashrafi, H., Stoffel, K., Weir, A., and Van Deynze, A. 2017. Regions Underlying Population Structure and the Genomics of Organ Size Determination in Capsicum annuum. *The Plant Genome*. doi:10.3835/plantgenome2017.03.0026

Hill, T. A., Ashrafi, H., Reyes-Chin-Wo, S., Yao, J., Stoffel, K., Truco, M.-J., and Van Deynze, A. 2013. Characterization of Capsicum annuum Genetic Diversity and Population Structure Based on Parallel Polymorphism Discovery with a 30K Unigene Pepper GeneChip. *PLoS ONE,* **8**, e56200. doi:10.1371/journal.pone.0056200

Hulse-Kemp, A. M., Ashrafi, H., Plieske, J., Lemm, J., Stoffel, K., Hill, T., and Ganal, M. W. 2016. A HapMap leads to a Capsicum annuum SNP infinium array: a new tool for pepper breeding. *Horticulture research,* **3**.

IPGRI. 1995 Descriptors for Capsicum (Capsicum spp.). Rome, Italy: International Plant Genetic Resources Institute

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

Kantar, M. B., Anderson, J. E., Lucht, S. A., Mercer, K., Bernau, V., Case, K. A., and Hastings, J. C. (2016). Vitamin variation in Capsicum Spp. provides opportunities to improve nutritional value of human diets. *PloS one*, ***11***, e0161464.

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J. M., Lee, H.-A., Kim, K.-T. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature genetics,* **46**, 270-278.

Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Ruiz, J. d. J. L., Eeckenbrugge, G. C., Gepts, P. 2014. Multiple lines of evidence for the origin of domesticated chili pepper, Capsicum annuum, in Mexico. *Proceedings of the National Academy of Sciences,* **111**, 6165-6170. doi:10.1073/pnas.1308933111

Kraft, K. H., de Jesús Luna-Ruíz, J., and Gepts, P. 2013. A new collection of wild populations of Capsicum in Mexico and the southern United States. *Genetic resources and crop evolution,* **60**, 225-232.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution,* **33**, 1870-1874.

Landaverde-González, P., Quezada-Euán, J. J. G., Theodorou, P., Murray, T. E., Husemann, M., Ayala, R., and Paxton, R. J. 2017. Sweat bees on hot chillies: provision of pollination services by native bees in traditional slash-and-burn agriculture in the Yucatán Peninsula of tropical Mexico. *Journal of Applied Ecology*.

Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods, 9*, 357-359.

Loveless, M. D., and Hamrick, J. L. 1984. Ecological determinants of genetic structure in plant populations. *Annual review of ecology and systematics*, *15*(1), 65-95.

McLeod, M., Guttman, S. I., and Eshbaugh, W. H. 1982. Early evolution of chili peppers (Capsicum). *Economic Botany,* **36**, 361-368.

McDade, L. A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution, 46*, 1329-1346.

Naegele, R. P., Mitchell, J., and Hausbeck, M. K. 2016. Genetic Diversity, Population Structure, and Heritability of Fruit Traits in Capsicum annuum. *PLoS ONE,* **11**, e0156969.

Nigh, R., and Diemont, S. A. 2013. The Maya milpa: fire and the legacy of living soil. Frontiers in Ecology and the Environment, 11(s1).

Ogundiwin, E. A., Berke, T. F., Massoudi, M., Black, L. L., Huestis, G., Choi, D., and Prince, J. P. 2005. Construction of 2 intraspecific linkage maps and identification of resistance QTLs for Phytophthora capsici root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.). *Genome*, **48**, 698-711.

Perry, L., and Flannery, K. V. 2007. Precolumbian use of chili peppers in the Valley of Oaxaca, Mexico. *Proceedings of the National Academy of Sciences,* **104**, 11905-11909.

Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D. M., Piperno, D. R., . . . Zeidler, J. A. 2007. Starch fossils and the domestication and dispersal of chili peppers (Capsicum spp. L.) in the Americas. *Science,* **315**, 986-988. doi:10.1126/science.1136914

Pickersgill, B. 1988. The genus *Capsicum*: a multidisciplinary approach to the taxonomy of cultivated and wild plants. *Biologisches Zentralblatt,* **107**, 381-389.

Pressoir, G., and Berthaud, J. 2004a. Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity*, *92*(2), 88-94.

Pressoir, G., and Berthaud, J. 2004b. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity*, *92*(2), 95-101.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., and Luo, Y. 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *Proceedings of the National Academy of Sciences,* **111**, 5135-5140.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573-589.

Rao, G. U., Chaim, A. B., Borovsky, Y., and Paran, I. 2003. Mapping of yield-related QTLs in pepper in an interspecific cross of Capsicum annuum and C. frutescens. *Theoretical and Applied Genetics*, *106*(8), 1457-1466.

Raw, A. 2000. Foraging behaviour of wild bees at hot pepper flowers (Capsicum annuum) and its possible influence on cross pollination. *Annals of Botany*, *85*(4), 487-492.

Romero, M. d. L. A. 2000. The Indigenous Population of Oaxaca From the Sixteenth Century to the Present". In R. E. W. Adams and M. J. MacLeod (Eds.), *The Cambridge History of the Native Peoples of the Americas* (Vol. 2). Cambridge, United Kingdom: Cambridge University Press.

Shrestha, S. L., Luitel, B. P., and Kang, W. H. 2011. Heterosis and heterobeltiosis studies in sweet pepper (Capsicum annuum L.). *Horticulture, Environment, and Biotechnology*, **52**(3), 278-283.

Smith, B. D. 1997. Reconsidering the Ocampo Caves and the Era of Incipient Cultivation in Mesoamerica. *Latin American Antiquity,* **8**, 342-383.

Sood, S., and Kumar, N. 2010. Heterosis for Fruit Yield and Related Horticultural Traits in Bell Pepper. *International Journal of Vegetable Science*, **16**(4), 361–373. https://doi.org/10.1080/19315260.2010.490576

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312-1313.

Tamura, K., and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, ***10***(3), 512-526.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, ***28***, 2731-2739.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences, 17*(2), 57-86.

Weir, B. S., and Cockerham, C. C. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution,* **38**, 1358-1370.

UPOV. 2004. *Sweet Pepper, Hot Pepper, Paprika, Chili: Draft Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability*. Geneva: International Union for the Protection of Varieties of Plants.

Yarnes, S. C., Ashrafi, H., Reyes-Chin-Wo, S., Hill, T. A., Stoffel, K. M., and Van Deynze, A. 2012. Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific Capsicum RIL population. *Genome*, *56*(1), 61-74.

Zeven, A. C. 1998. Landraces: a review of definitions and classifications. *Euphytica*, *104*(2), 127-139.

Zimmerer, K. S., and Douches, D. S. 1991. Geographical approaches to crop conservation: the partitioning of genetic diversity in Andean potatoes. *Economic Botany*, *45*(2), 176-189.

Chapter 3 References

Ben Chaim, A., Paran, I., Grube, R.C., Jahn, M., Van Wijk, R. and Peleman, J., 2001. QTL mapping of fruit-related traits in pepper (Capsicum annuum). *Theoretical and Applied Genetics, 102*(6-7), pp.1016-1028.

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M.J., Francis, D., Causse, M., Van der Knaap, E. and Cañizares, J., 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC genomics*, *16*(1), p.257.

Bollen, A., Van Elsacker, L. and Ganzhorn, J.U., 2004. Tree dispersal strategies in the littoral forest of Sainte Luce (SE-Madagascar). *Oecologia*, *139*(4), pp.604-616.

Borovsky, Y. and Paran, I., 2011. Characterization of fs10. 1, a major QTL controlling fruit elongation in Capsicum. *Theoretical and applied genetics*, *123*(4), pp.657-665.

Bosland, P.W., Votava, E.J. and Votava, E.M., 2012. *Peppers: vegetable and spice capsicums* (Vol. 22). Cabi.

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

Bünger-Kibler S., Bangerth F., 1982. Relationship between cell number, cell size and fruit size of seeded fruits of tomato (*Lycopersicon esculentum* Mill.), and those induced parthenocarpically by the application of plant growth regulators. *Plant Growth Regul 1*:143–154

Chakrabarti, M., Zhang, N.A., Sauvage, C., Muños, S., Blanca, J., Cañizares, J., Diez, M.J., Schneider, R., Mazourek, M., McClead, J. and Causse, M., 2014. A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proceedings of the National Academy of Sciences*, *110*(42), pp.17125-17130.

Chunthawodtiporn, J., Hill, T., Stoffel, K. and Van Deynze, A., 2018. Quantitative trait loci controlling fruit size and other horticultural traits in bell pepper (Capsicum annuum). *The plant genome*, *11*(1).

Cheng Y, Pang X, Wan H, Ahammed GJ, Yu J, Yao Z, Ruan M, Ye Q, Li Z, Wang R., 2017 Identification of optimal reference genes for normalization of qPCR analysis during pepper fruit development. *Frontiers in Plant Science, 8*, p.1128

Colonna, V., D'Agostino, N., Garrison, E., Albrechtsen, A., Meisner, J., Facchiano, A., Cardi, T. and Tripodi, P., 2019. Genomic diversity and novel genome-wide association with fruit morphology in Capsicum, from 746k polymorphic sites. *Scientific reports*, *9*(1), pp.1-14.

Cuppen, E., 2007. Genotyping by allele-specific amplification (KASPar). *CSH protocols*, *2007*, pdb-prot4841.

De Jong M., Wolters-Arts M., Feron R., Mariani C., Vriezen W.H., 2009. The Solanum lycopersicum auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant J. 57*:160–170. doi: 10.1111/j.1365-313X.2008.03671.x.

Dou, J., Zhao, S., Lu, X., He, N., Zhang, L., Ali, A., Kuang, H., Liu, W., 2018. Genetic mapping reveals a candidate gene (ClFS1) for fruit shape in watermelon (*Citrullus lanatus* L.). *Theor. Appl. Genet.131*(4):947-958

Drevensek, S., Goussot, M., Duroc, Y., Christodoulidou, A., Steyaert, S., Schaefer, E., Duvernois, E., Grandjean, O., Vantard, M., Bouchez, D. and Pastuglia, M., 2012. The Arabidopsis TRM1–TON1 interaction reveals a recruitment network common to plant cortical microtubule arrays and eukaryotic centrosomes. *The Plant Cell*, *24*(1), pp.178-191.

Federer, W.T., 1961. Augmented designs with one-way elimination of heterogeneity. *Biometrics*, *17*(3), pp.447-473.

Hackbusch, J., Richter, K., Müller, J., Salamini, F. and Uhrig, J.F., 2005. A central role of Arabidopsis thaliana ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proceedings of the National Academy of Sciences*, *102*(13), pp.4908-4912.

Henikoff, S. and Henikoff, J.G., 1992. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences*, *89*(22), pp.10915-10919.

Han, K., Jeong, H.J., Yang, H.B., Kang, S.M., Kwon, J.K., Kim, S., Choi, D. and Kang, B.C., 2016. An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (Capsicum annuum). *DNA Research*, *23*(2), pp.81-91.

Hill, T.A., Chunthawodtiporn, J., Ashrafi, H., Stoffel, K., Weir, A. and Van Deynze, A., 2017. Regions underlying population structure and the genomics of organ size determination in *Capsicum annuum*. *The Plant Genome*, *10*(3).

Hori, K. and Yano, M., 2013. Genetic improvement of grain quality in Japonica rice. *Translational genomics for crop breeding*, *2*, pp.143-160.

Hulse-Kemp, A.M., Maheshwari, S., Stoffel, K., Hill, T.A., Jaffe, D., Williams, S.R., Weisenfeld, N., Ramakrishnan, S., Kumar, V., Shah, P. and Schatz, M.C., 2018. Reference quality assembly of the 3.5-Gb genome of Capsicum annuum from a single linked-read library. *Horticulture Research*, *5*(1), p.4.

IPGRI, AVRDC, and CATIE, 1995. Descriptors for *Capsicum* (*Capsicum* spp.). International Plant Genetic Resources Institute, Rome, Italy; the Asian Vegetable Research and Development Center, Taipei, Taiwan, and the Centro Agronomico Tropical de Investigacion y Esenanza, Turrialba, Costa Rica.

IUPAC-IUB Joint Commission on Biochemical Nomenclature, 1984. Nomenclature and Symbolism for Amino Acids and Peptides: Recommendations 1983. *FEBS J. 138* (1), 9–37. DOI: 10.1111/j.1432-1033.1984.tb07877.x.

Jin, B., Kim, J., Jung, J., Kim, D. and Park, Y., 2018. Characterization of IQ domain gene homologs as common candidate genes for elongated fruit shape in cucurbits. *원예과학기술지*, *36*(1), pp.85-97.

Kraus, C.O., 2019. *OVATE and Its Associates: The Interactions of Tomato OVATE with TRMs and TON1* (Doctoral dissertation, University of Georgia).

Kim, S., Park, M., Yeom, S.I., Kim, Y.M., Lee, J.M., Lee, H.A., Seo, E., Choi, J., Cheong, K., Kim, K.T. and Jung, K., 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature genetics*, *46*(3), p.270.

Kosambi, D.D., 1944. The estimation of map distances from recombination values. *Annals of Eugenics 12,* p.172-175

Lazzaro, M.D., Wu, S., Snouffer, A., Wang, Y. and Van Der Knaap, E., 2018. Plant organ shapes are regulated by protein interactions and associations with microtubules. *Frontiers in plant science*, *9*, p.1766.

Li, E., Wang, S., Liu, Y., Chen, J.G. and Douglas, C.J., 2011. OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in Arabidopsis thaliana. *The Plant Journal*, *67*(2), pp.328-341.

Liu, F., Yu, H., Deng, Y., Zheng, J., Liu, M., Ou, L., Yang, B., Dai, X., Ma, Y., Feng, S. and He, S., 2017. PepperHub, an informatics hub for the chili pepper research community. *Molecular plant*, *10*(8), pp.1129-1132.

Liu, J., Van Eck, J., Cong, B. and Tanksley, S.D., 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proceedings of the National Academy of Sciences*, *99*(20), pp.13302-13306.

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods* 25. pp.402-408

Lee, Y.K., Kim, G.T., Kim, I.J., Park, J., Kwak, S.S., Choi, G. and Chung, W.I., 2006. *LONGIFOLIA1* and *LONGIFOLIA2*, two homologous genes, regulate longitudinal cell elongation in Arabidopsis. *Development*, *133*(21), pp.4305-4314.

Ma, Y., Yang, C., He, Y., Tian, Z., Li, J. and Sunkar, R., 2017. Rice OVATE family protein 6 regulates plant development and confers resistance to drought and cold stresses. *Journal of experimental botany*, *68*(17), pp.4885-4898.

Martín-Trillo, M. and Cubas, P., 2010. TCP genes: a family snapshot ten years later. *Trends in plant science*, *15*(1), pp.31-39.

Meng, L., Li, H., Zhang, L., & Wang, J. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*, *3*(3), 269-283.

Keyhaninejad, N., Wu, S., Snouffer, A., Meulia, T., Lazzaro, M., Li, Q., van der Knaap, E. Electrostatic interactions of tomato OFPs and TRMs leads to dynamic differential subcellular localizations in tobacco leaf epidermal cells. *Submitted for publication*.

Kryazhimskiy, S. and Plotkin, J.B., 2008. The population genetics of dN/dS. *PLoS genetics*, *4*(12).

Neff, M. M., Neff, J. D., Chory, J., & Pepper, A. E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. *Plant J, 14*(3), 387-392

Nyquist, W.E. and Baker, R.J., 1991. Estimation of heritability and prediction of selection response in plant populations. *Critical reviews in plant sciences*, *10*(3), pp.235-322.

Paran, I., Knaap, E. van der, 2007. Genetic and Molecular Regulation of Fruit and Plant Domestication Traits in Tomato and Pepper. *J. Exp. Bot. 58*(14), 3841–3852. https://doi.org/10.1093/jxb/erm257

Pan, Y., Liang, X., Gao, M., Liu, H., Meng, H., Weng, Y. and Cheng, Z., 2017. Round fruit shape in WI7239 cucumber is controlled by two interacting quantitative trait loci with one putatively encoding a tomato SUN homolog. *Theoretical and applied genetics*, *130*(3), pp.573-586.

Prince, J. P., Zhang, Y., Radwanski, E. R., & Kyle, M. M. (1997). A versatile and high-yielding protocol for the preparation of genomic DNA from Capsicum spp.(pepper). *Hortscience*, *32*(5), 937-939.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., Cheng, J., Zhao, S., Xu, M., Luo, Y. and Yang, Y., 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *Proceedings of the National Academy of Sciences*, *111*(14), pp.5135-5140.

Rasband, W.S., 2012. ImageJ: Image processing and analysis in Java. *Astrophysics Source Code Library*.

Rinaldi, R., Van Deynze, A., Portis, E., Rotino, G.L., Toppino, L., Hill, T., Ashrafi, H., Barchi, L. and Lanteri, S., 2016. New insights on eggplant/tomato/pepper synteny and identification of eggplant and pepper orthologous QTL. *Frontiers in Plant Science*, *7*, p.1031.

Rodríguez, G.R., Kim, H.J. and van der Knaap, E., 2013. Mapping of two suppressors of OVATE (sov) loci in tomato. *Heredity*, *111*(3), pp.256-264.

Schaefer, E., Belcram, K., Uyttewaal, M., Duroc, Y., Goussot, M., Legland, D., Laruelle, E., de Tauzia-Moreau, M.L., Pastuglia, M. and Bouchez, D., 2017. The preprophase band of microtubules controls the robustness of division orientation in plants. *Science*, *356*(6334), pp.186-189.

Schmitz, A.J., Begcy, K., Sarath, G. and Walia, H., 2015. Rice Ovate Family Protein 2 (OFP2) alters hormonal homeostasis and vasculature development. *Plant Science*, *241*, pp.177-188.

Serrani J.C., Fos M.., Atarés A, García-Martínez J.L., 2007a. Effect of gibberellin and auxin on parthenocarpic fruit growth induction in the cv micro-tom of tomato. *J Plant Growth Regul 26*:211–221

Serrani J.C., Sanjuán R., Ruiz-Rivero O., Fos M., García-Martínez J.L., 2007b. Gibberellin regulation of fruit set and growth in tomato. *Plant Physiol 145*:246–257

Snouffer, A., Kraus, C. and van der Knaap, E., 2020. The Shape of Things to Come: Ovate Family Proteins regulate plant organ shape. *Current Opinion in Plant Biology*, *53*, pp.98-105.

Spinner, L., Gadeyne, A., Belcram, K., Goussot, M., Moison, M., Duroc, Y., Eeckhout, D., De Winne, N., Schaefer, E., Van De Slijke, E. and Persiau, G., 2013. A Protein Phosphatase 2A complex spatially controls plant cell division. *Nature Communications*, *4*, p.1863.

Srivastava A., Handa A.K., 2005. Hormonal regulation of tomato fruit development: a molecular perspective. *J Plant Growth Regul 24*:67–82

Stanke, M. and Morgenstern, B., 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic acids research*, *33*(suppl_2), pp.W465-W467.

Sun, L., Rodriguez, G.R., Clevenger, J.P., Illa-Berenguer, E., Lin, J., Blakeslee, J.J., Liu, W., Fei, Z., Wijeratne, A., Meulia, T. and van der Knaap, E., 2015. Candidate gene selection and detailed morphological evaluations of fs8. 1, a quantitative trait locus controlling tomato fruit shape. *Journal of experimental botany*, *66*(20), pp.6471-6482.

Taitano, N., Bernau, V., Jardón-Barbolla, L., Leckie, B., Mazourek, M., Mercer, K., McHale, L., Michel, A., Baumler, D., Kantar, M. and van der Knaap, E., 2019. Genome-wide genotyping of a novel Mexican Chile Pepper collection illuminates the history of landrace differentiation after Capsicum annuum L. domestication. *Evolutionary applications*, *12*(1), pp.78-92.

Tsaballa, A., Pasentsis, K., Darzentas, N. and Tsaftaris, A.S., 2011. Multiple evidence for the role of an Ovate-like gene in determining fruit shape in pepper. *BMC plant biology*, *11*(1), p.46.

van der Knaap, E., Chakrabarti, M., Chu, Y.H., Clevenger, J.P., Illa-Berenguer, E., Huang, Z., Keyhaninejad, N., Mu, Q., Sun, L., Wang, Y. and Wu, S., 2014. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Frontiers in Plant Science*, *5*, p.227.

Wang, S., Chang, Y., Guo, J. and Chen, J.G., 2007. Arabidopsis Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. *The Plant Journal*, *50*(5), pp.858-872.

Wang, S., Chang, Y., Guo, J., Zeng, Q., Ellis, B.E. and Chen, J.G., 2011. Arabidopsis ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. *PLoS One*, *6*(8).

Wang, S., Li, S., Liu, Q., Wu, K., Zhang, J., Wang, S., Wang, Y., Chen, X., Zhang, Y., Gao, C. and Wang, F., 2015. The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nature genetics*, *47*(8), pp.949-954.

Wang, X., Wang, W., Wang, J., Zhang, N., Yang, L., Cai, L., Hussain, S., Chang, Y. and Wang, S., 2019. Class III OFPs function in the ER signaling pathway to regulate plant growth and development in Arabidopsis. *Journal of Plant Interactions*, *14*(1), pp.45-53.

Wang, Y., Tang, H., DeBarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H. and Kissinger, J.C., 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids research*, *40*(7), pp.e49-e49.

Wu, S., Xiao, H., Cabrera, A., Meulia, T. and van der Knaap, E., 2011. SUN regulates vegetative and reproductive organ shape by changing cell division patterns. *Plant Physiology*, *157*(3), pp.1175-1186.

Wu, S., Zhang, B., Keyhaninejad, N., Rodríguez, G.R., Kim, H.J., Chakrabarti, M., Illa-Berenguer, E., Taitano, N.K., Gonzalo, M.J., Díaz, A. and Pan, Y., 2018. A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nature communications*, *9*(1), p.4734.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J. and Van Der Knaap, E., 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science*, *319*(5869), pp.1527-1530.

Xiao, H., Radovich, C., Welty, N., Hsu, J., Li, D., Meulia, T. and van der Knaap, E., 2009. Integration of tomato reproductive developmental landmarks and expression profiles, and the effect of SUN on fruit shape. *BMC Plant Biology*, *9*(1), p.49.

Xu, C., Liberatore, K.L., MacAlister, C.A., Huang, Z., Chu, Y.H., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M. and Van Eck, J., 2015. A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nature Genetics*, *47*(7), p.784.

Yang, Z. and Bielawski, J.P., 2000. Statistical methods for detecting molecular adaptation. *Trends in ecology & evolution*, *15*(12), pp.496-503.

Zhao, D.S., Li, Q.F., Zhang, C.Q., Zhang, C., Yang, Q.Q., Pan, L.X., Ren, X.Y., Lu, J., Gu, M.H. and Liu, Q.Q., 2018. GS9 acts as a transcriptional activator to regulate rice grain shape and appearance quality. *Nature Communications*, *9*(1), pp.1-14.

Zhou, Y., Miao, J., Gu, H., Peng, X., Leburu, M., Yuan, F., Gu, H., Gao, Y., Tao, Y., Zhu, J. and Gong, Z., 2015. Natural variations in SLG7 regulate grain shape in rice. *Genetics*, *201*(4), pp.1591-1599.

Chapter 4 References

Alexa, A., Rahnenführer, J. and Lengauer, T., 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, *22*(13), pp.1600-1607.

Alexa, A. and Rahnenfuhrer, J., 2019. Gene set enrichment analysis with topGO. Bioconductor Improv, 27.

Alonge, A., Wang, X., Benoit, M., Soyk, S., Pereira, L., Zhang, L., Suresh, H., Ramakrishnan, S., Maumus, F., Ciren, D., Levy, Y., Harel, T., Shalev-Schlosser, G., Amsellem, Z., Razifard, H., Caicedo, A. L., Tiemen, D. M., Klee, H., Kirsche, M., Aganezov, S., Ranall-Benavidez, T. R., Lemmon, Z. H., Kim, J., Robitaille, G., Kramer, M., Goodwin, S., McCombie, W. R., Hutton, S., Van Eck, J., Gillis, J., Eshed, Y., Sedlazeck, F. J., van der Knaap, E., Schatz, M. C., Lippman, Z. B., 2020. Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. *Submitted for Publication*

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M.J., Francis, D., Causse, M., Van der Knaap, E. and Cañizares, J., 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC genomics*, *16*(1), p.257.

Chu, Y.H., Jang, J.C., Huang, Z. and van der Knaap, E., 2019. Tomato locule number and fruit size controlled by natural alleles of *lc* and *fas*. *Plant direct*, *3*(7), p.e00142.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X. and Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome biology*, *17*(1), p.13.

Huang, Z., Van Houten, J., Gonzalez, G., Xiao, H. and van der Knaap, E., 2013. Genome-wide identification, phylogeny and expression analysis of SUN, OFP and YABBY gene family in tomato. *Molecular Genetics and Genomics*, *288*(3-4), pp.111-129.

Korte, A. and Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, *9*(1), p.29.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C. and Salzberg, S.L., 2004. Versatile and open software for comparing large genomes. *Genome biology*, *5*(2), p.R12.

Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), pp.3094-3100.

Liu, D., Sun, W., Yuan, Y., Zhang, N., Hayward, A., Liu, Y. and Wang, Y., 2014. Phylogenetic analyses provide the first insights into the evolution of OVATE family proteins in land plants. *Annals of botany*, *113*(7), pp.1219-1233.

Liu, J., Van Eck, J., Cong, B. and Tanksley, S.D., 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proceedings of the National Academy of Sciences*, *99*(20), pp.13302-13306.

Layer, R.M., Chiang, C., Quinlan, A.R. and Hall, I.M., 2014. LUMPY: a probabilistic framework for structural variant discovery. *Genome biology*, *15*(6), p.R84.

Mathur, J., Mathur, N., Kernebeck, B. and Hülskamp, M., 2003. Mutations in actin-related proteins 2 and 3 affect cell shape development in Arabidopsis. *The Plant Cell*, *15*(7), pp.1632-1645.

Ohi, R., McCollum, D., Hirani, B., Den Haese, G.J., Zhang, X.I.A.O.E., Burke, J.D., Turner, K. and Gould, K.L., 1994. The Schizosaccharomyces pombe cdc5+ gene encodes an essential protein with homology to c-Myb. *The EMBO Journal*, *13*(2), pp.471-483.

Razifard, H., Ramos, A., Della Valle, A.L., Bodary, C., Goetz, E., Manser, E.J., Li, X., Zhang, L., Visa, S., Tieman, D. and van der Knaap, E., 2020. Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*.

Pachter, L., 2011. Models for transcript quantification from RNA-Seq. *arXiv preprint arXiv:1104.3889*.

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D. and Shakir, K., 2018. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, p.201178.

Smertenko, A.P., Kaloriti, D., Chang, H.Y., Fiserova, J., Opatrny, Z. and Hussey, P.J., 2008. The C-terminal variable region specifies the dynamic properties of Arabidopsis microtubule-associated protein MAP65 isotypes. *The Plant Cell*, *20*(12), pp.3346-3358.

van der Knaap, E., Chakrabarti, M., Chu, Y.H., Clevenger, J.P., Illa-Berenguer, E., Huang, Z., Keyhaninejad, N., Mu, Q., Sun, L., Wang, Y. and Wu, S., 2014. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Frontiers in Plant Science*, *5*, p.227.

Wang, Y., Clevenger, J.P., Illa-Berenguer, E., Meulia, T., van der Knaap, E. and Sun, L., 2019. A comparison of sun, ovate, fs8. 1 and auxin application on tomato fruit shape and gene expression. *Plant and Cell Physiology*, *60*(5), pp.1067-1081.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J. and Van Der Knaap, E., 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *science*, *319*(5869), pp.1527-1530.

Xu, C., Liberatore, K.L., MacAlister, C.A., Huang, Z., Chu, Y.H., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M. and Van Eck, J., 2015. A cascade of arabinosyltransferases controls shoot meristem size in tomato. Nature genetics, 47(7), pp.784-792.

Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L. and Wang, Y., 2019. Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. *Nature genetics*, *51*(6), pp.1052-1059.

Chapter 5 References

Bosland, P. W. and E. J. Votava 2012. Peppers: vegetable and spice capsicums. New York, USA, CABI Publishing.

Razifard, H., Ramos, A., Della Valle, A.L., Bodary, C., Goetz, E., Manser, E.J., Li, X., Zhang, L., Visa, S., Tieman, D. and van der Knaap, E., 2020. Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*.

APPENDIX

Appendix Summary and Supplemental Data Locations

All supplemental files are stored as plaintext and can be viewed with Notepad, TextWrangler, or any other text viewer, or a freely available spreadsheet program for ease of manipulation.

Appendix 2.S1 included below. Supplemental Files 3.S1, 3.S2, and Supplemental Files 4.S1-4 all submitted electronically together with this manuscript.

Supplemental File 3.S1 contains genes in tomato region syntenic with UCD10X chromosome 3 QTL: colinear genes returned as a text file in standard MCScanX collinearity format (Wang *et al.,* 2012).

Supplemental File 3.S2 is a text file in FASTA format of nucleotide alignment for *SlTRM25*: alignment of Sanger-sequenced reads of the two mapping population parents and the Mexican landrace accessions together with the CM334 reference genome annotation of *CaTRM25*. Each entry corresponds to one Sanger sequencing run, is labeled with both the DNA accession and primer IDs (Table 3.S1) used to obtain the sequence, and begins at nt position -834 in the promoter extending to position 3,720 relative to the methionine start site.

Supplemental Files 4.S1-4 are large spreadsheets of genes and structural variants. They contain their own descriptions at the top of each file.

Appendix 2.S1: Description of TASSEL 3.0 GBS pipeline as used in this study

(Adapted for this study from Glaubitz et al. 2014)

1.) FastqToTagCount (Step 1: Filtration)
  a. Purpose: to identify and count the number of identifiable tags generated from each lane of the Illumina sequencer. The FastqToTagCount plugin does not parse barcodes and assign them to individual plants. Instead, it merely checks each sequence read for the barcode and cutsite, keeps only those reads where the first 64 nucleotides of the gDNA tag are free of Ns or additional ApeKI cutsites, and truncates the reads - removing everything after those 64 nucleotides - to avoid reading into the Illumina adapter on the other end.
  b. Inputs/parameters:
    i. FASTQ read+quality data returned by Illumina sequencer
    ii. Key file relating barcodes to wells on a plate (and thus individual plants)
  c. Outputs:
    i. A TagCount file, isolating the readable tags returned by each lane of the sequencer, not yet assigned to individual plants

2.) MergeMultipleTagCount (Step 1: data management)
  a. Purpose: to merge the same tags when found multiple times, in order to create a reference list of all unique tags in the study.
  b. Inputs/parameters:
    i. TagCount file, generated by FastqToTagCount.
    ii. Threshold for the minimum number of times the same tag sequence must be read. Tags not meeting this threshold are removed from the dataset. We set this at one, in order to avoid excluding whole tags with a single base-pair sequencing error and otherwise informative data. Instead of removing whole tags with low depth, we elected to filter out SNPs covered at low depth later in the pipeline.
  c. Output:

        i. Master tag list containing all unique gDNA tags and the number of times they appear in the entire population, which - after conversion to FASTQ - can be referenced in the TagsByTaxa file produced by SeqToTBTHDF5 (4).

3.) TagCountToFastq (Step 1: read data management and filtration)
    a. Purpose: to convert the master tag list from (2) to a format in which they can be aligned by BWA (6) and SeqToTBTHDF5 (4).
    b. Inputs/parameters
        i. Master tag list
    c. Outputs
        i. Converted master tag list

4.) SeqToTBTHDF5 (Step 2: barcode parsing to identify read source)
    a. Purpose: to parse barcodes and relate tags in the Master tag list to the corresponding individual.
    b. Inputs/parameters:
        i. Original FASTQ sequences
        ii. Master tag list returned by MergeMultipleTagCount, reformatted by TagCountToFastq
        iii. Barcode key file
        iv. Enzyme (ApeKI)
    c. Output
        i. TagsByTaxa (TBT) file stored in the Hierarchical Data Format v5 (.h5). Relates the tags contained in the master tag list to the individual plants whose genomes contain that tag sequence (as determined by the attached barcode).

5.) MergeTagsByTaxa (Step 2: barcode parsing)

- a. Purpose: merge the TBT files from each plate into a single file for the experiment. Because the same barcodes were used for each of the two plates, it is important that data from the two plates remain separate until this point.
- b. Inputs/parameters:
    - i. TBT files returned by running 1-3 on each plate.
- c. Output:
    - i. TBT file including all greenhouse plants for the whole experiment.

6.) Burrows-Wheeler Aligner (BWA) (Step 3: Calling SNPs)

- a. Purpose: align all tags to the reference genome in preparation for SNP-calling
- b. Inputs/parameters:
    - i. *C. annuum* cv CM334 reference genome (Kim et al. 2014)
    - ii. Number of CPU threads to use: specified 4 threads
    - iii. Master tag list in fastq format
    - iv. “samse” option to indicate single-end reads
- c. Output:
    - i. SAM alignment file

7.) SAMConverter (Step 3: Calling SNPs)

- a. Purpose: convert the BWA alignment file to TASSEL’s TagsOnPhysicalMap (TOPM) format.
- b. Inputs/parameters
    - i. SAM alignment file returned by BWA
- c. Output
    - i. TagsOnPhysicalMap file for SNP calling

8.) ModifyTagsByTaxa (Step 3: Calling SNPs)

- a. Purpose: “pivot” the TBT file returned by MergeTagsByTaxa so that it is in the proper orientation for use by DiscoverySNPCaller

   b. Inputs/parameters
      i. TBT file returned by MergeTagsByTaxa
      ii. –p parameter to designate pivoting as the modification
   c. Output
      i. pivoted TBT file

9.) DiscoverySNPCaller (Step 3: Calling SNPs)
   a. Purpose: call SNPs that pass a final filtration step, and mark the SNP locations on the genome. Does not yet assign genotypes to individuals.
   b. Inputs/parameters
      i. pivoted TBT file from ModifyTagsByTaxa
      ii. TOPM-formatted, reference-aligned master tags list from SAMConverter
      iii. Mexican *chiltepin* (wild chile pepper) reference genome (Qin et al. 2014)
      iv. –mnF (minimum inbreeding coefficient) parameter set to -0.1. This was because some outcrossing was expected in the population, and thus we avoided removing SNPs where this was low. This option is only set above -0.1 for highly homozygous inbred lines (Glaubitz et al. 2014).
      v. –mnMAF (minimum minor allele frequency) set at 0.01, including SNPs with minor allele frequencies of >1% to exclude sequencing errors.
      vi. –mnMAC (minimum minor allele count) set at three, including SNPs where the minor allele was read at least three times to exclude sequencing errors.
   c. Output
      i. TOPM file containing all variants in the population, aligned to the reference genome

10.) ProductionSNPCaller (Step 3: Calling SNPs)
   a. Purpose: Assign genotypes to individuals
   b. Inputs/parameters:
      i. Original fastq files returned by Illumina sequencer

    ii. TOPM file with aligned variants, returned by DiscoverySNPCaller
    iii. –e (enzyme) parameter specified as ApeKI
  c. Output

Genotype table for all individuals, in HDF5 format, ready for viewing in the TASSEL GUI, whereupon they can be exported into the desired format. In our case, we selected the Variant Call Format (VCF) for later use with VCFTools