

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Application of High-Density SNP Genotyping Array in Citrus Germplasm Characterization and
Genetic Dissection of Traits

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

by

Yoko Hiraoka

June 2020

Dissertation Committee:

Dr. Mikeal L. Roose, Chairperson

Dr. Timothy Close

Dr. Daniel Koenig

Dr. Brandon Gaut

The Dissertation of Yoko Hiraoka is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENT

I would like to thank my mentor, Professor Mikeal L. Roose, for teaching me how to do science and how to think critically about science with patience. You taught me that there are always learning opportunities. You also taught me to be patient, and not give up when there are challenges in front of me by giving me numerous opportunities and time to get through those challenges at my own pace. You have instilled in me a drive to never give up until the data makes sense and that there are no longer results that I can not explain. Most importantly, thank you for providing an environment that encourages self-learning, growth, and patience. Your encouragements for the past seven years is something I will never forget and will stay with me throughout my future endeavors. Thank you for being a great mentor!

I would like to thank Dr. Claire Federici, whom I interacted with on a daily basis during my Ph.D. studies, for her patience and kindness. You taught me many experimental and molecular techniques, but also guided me through life as a student and a mother. You helped ease my anxieties numerous times through our daily conversations. I can recall all the moments you told me that I would be okay. I will miss these daily chats with you. I would like to thank past and present members of the Roose lab for all the learning experiences. I want to thank Dr. Sergio Ferrante, who I shared the office with, for our daily scientific chat. You taught me so many things about being a scientist and about the career which I will about to start. I would like to thank Dr. Brandon Le for being an incredible mentor to guide me through various stages of my Ph.D. study. I would not have gotten this far if you were not there as my bioinformatics mentor. I hope our late night science talks and brain-storming sessions will continue.

I would like to thank Emi Ouchi for her support. I would not have been who I am today without her believing in me. I would also like to thank Brent Oakes for believing in me.

I would like to acknowledge financial support from the USDA NIFA and Citrus Research Board for the projects in my dissertation study.

Lastly, I am extremely thankful to Dr. Elia Scudiero for his continual support. You have been patiently guiding me through my life as a Ph.D. student and a mother and I could not have gotten through any of it without you.

To my son Ren, I can not describe how thankful I am that you came to me during the most important time of my life. You provided me strength and pride to get through my Ph.D. study. Because of you, I have now accomplished two of the most meaningful things I have done in my life. Thank you.

I dedicate my dissertation to my father, Dr. Yutaka Hiraoka

my mother Chiyoko

my brother Satoshi

my son Ren Efrem

and my partner Dr. Elia Scudiero

I am who I am today because of your patience, understanding, and trust

ABSTRACT OF THE DISSERTATION

Application of High-Density SNP Genotyping Array in Citrus Germplasm Characterization and Genetic Dissection of Traits

by

Yoko Hiraoka

Doctor of Philosophy, Graduate Program in Plant Biology
University of California, Riverside, June 2020
Dr. Mikeal L. Roose, Chairperson

High throughput genotyping technologies, such as SNP genotyping arrays that have emerged in recent years, have revolutionized the way we can analyze the origin and genetic control of traits in crops and have proven to be high-throughput and cost-effective tools. I developed Affymetrix™ Axiom® Citrus HD Genotyping Array and Affymetrix™ Axiom® Citrus Genotyping Array with 1.4 M and 58 K autosomal SNPs. DNA from 924 citrus accessions was hybridized to one or both of these arrays and yielded an extensive genomic inventory of the Citrus Variety Collection (CVC). High concordance was generally observed between DNA sequence and array data, and a simulation study suggested that errors in the range observed in more distant taxa have little to no effect on analyses such as principal component analysis (PCA), phylogeny and admixture analysis. This data was also exploited in a systematic study of genetic diversity, including phylogeny, admixture, and patterns of introgression at the chromosomal level. The study identified non-admixed accessions belonging to citrus progenitor species, and revealed the

scale of diversity found in the CVC. Putative ancestors were identified for the identified admixed accessions and many species named by Tanaka were found to be hybrids. The study also suggested *papeda*'s polyphyly, in addition to *Citrus hongheensis* being a possible progenitor species to *Citrus maxima*. Lastly, we analyzed nucellar embryony (apomixis), an important trait in citrus because it significantly influences strategies for both scion and rootstock breeding. Some of the seedlings arising from a seed of a genotype with nucellar embryony are genetically identical to the seed parent. To better understand the genetic basis of nucellar embryony, a genome-wide association study (GWAS) was conducted separately in *Citrus* and *Poncirus* for two distinct but related traits, nucellar embryony and polyembryony. In this study, an average of 46 germinated seedlings from each of 209 accessions were genotyped with KASP SNP markers to determine whether they originated from a nucellar or a zygotic embryo. Using a high density SNP array for citrus to genotype the seed source trees, the GWAS analysis identified a major dominant locus in *Citrus* and a separate locus in *Poncirus*.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iv
DEDICATION	vi
ABSTRACT OF THE DISSERTATION	vii
TABLE OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF SUPPLEMENTAL TABLES	xiv
LIST OF SUPPLEMENTAL FIGURES	xx
GENERAL INTRODUCTION	1
References cited	21
Tables	30
CHAPTER 1: Development and Assessment of SNP Genotyping Arrays for Citrus and Its Close Relatives	
Background	33
Introduction	34
Materials and Methods	42
Results	50

Discussion	59
References cited	69
Tables and Figures	76
Supplemental Tables and Figures	96

CHAPTER 2: Citrus Ancestry: Diversity Study and Admixture Inference in Diverse Accessions of Citrus with High Density SNP Genotyping Array Data

Abstract	161
Introduction	162
Materials and Methods	165
Results	170
Discussion	175
References cited	203
Tables and Figures	210
Supplemental Tables and Figures	307

CHAPTER 3: Distribution and Inheritance of Nucellar Embryony Loci in Citrus and Its Close Relatives

Abstract	322
Introduction	323
Materials and Methods	326

Results	330
Discussion	332
References cited	338
Tables and Figures	343
Supplemental Tables and Figures	360
GENERAL CONCLUSION	375

LIST OF TABLES

General Introduction

Table 1 Statistics for the sequenced <i>Citrus</i> genomes	30
Table 2 List of known polyembryonic and monoembryonic cultivars	31
Table 3 Summary of citrus diversity study	32

Chapter 1

Table 1 Accessions selected to be in Citrus variant discovery panel that were sequenced at UCR IIGB	76
Table 2 Citrus variant discovery panel from public database or collaborators	78
Table 3 Accession types of 196 and 871 accessions analyzed with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array	79
Table 4 Chromosomal distribution of SNPs represented on Axiom® Citrus HD Genotyping Array	80
Table 5 Summary of SNP classes represented on Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array	81
Table 6 Chromosomal distribution of SNPs represented on Axiom® Citrus Genotyping Array	82

Chapter 2

Table 1 924 accessions genotyped with Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array	210
Table 2 Accessions in CVC that were represented in the 382 UCR diversity set	233
Table 3 85 accessions used in maximum likelihood tree	244

Table 4 Diversity statistics	247
Table 5 ADMIXTURE analysis with UCR 382 diversity set	248
Table 6 Proportion of genomic composition for F1 hybrid	262
Table 7 ID for the accessions in UCR382 diversity panel used in phylogeny network analysis	263
Chapter 3	
Table 1 71 Accessions in Citrus GWAS panel that were screened for nucellar Embryony	343
Table 2 Contingency Chi-square test assessing environmental effects on nucellar embryony and polyembryony in GWAS panel	345
Table 3 90 accessions in Citrus GWAS panel	346
Table 4 88 accessions in Poncirus GWAS panel	349
Table 5 Average, mean and standard deviation for estimated percent zygotic seedlings and percent polyembryony for GWAS panels	352
Table 6 Peak SNPs for Citrus NE identified by GWAS LML analysis with threshold corrected with Bonferroni	353
Table 7 Citrus nucellar embryony candidate genes	354
Table 8 Peak SNPs for <i>Poncirus</i> NE QTL identified by GWAS LML analysis with threshold corrected with FDR 0.01	355

LIST OF FIGURES

Chapter 1

- Figure 1 Work Flow for citrus variant discovery panel selection, SNP detection, filtering, validation, and final selection of SNPs for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array 83
- Figure 2 Proportion of markers on the Axiom® Citrus HD Genotyping Array (a) and Axiom® Citrus Genotyping Array 84
- Figure 3 Genome-wide SNP distribution of Axiom® Citrus HD Genotyping Array, Axiom® Citrus Genotyping Array and heterozygous marker distribution of selected accessions 85
- Figure 4 Comparison of percent heterozygosity estimated with sequence data and Axiom® Citrus HD Genotyping Array data 87
- Figure 5 Genotyping call comparison between sequencing data and Axiom® Citrus HD Genotyping Array data 88
- Figure 6 Reproducibility of genotyping data between Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array was assessed using 28 accessions that were analyzed with both arrays 89
- Figure 7 PCA of citrus variant discovery panel genotyped with Axiom® Citrus HD Genotyping Array 90
- Figure 8 PCA of 871 diploid accessions in CVC genotyped with Axiom® Citrus Genotyping Array 91
- Figure 9 Maximum likelihood phylogeny trees of 26 accessions in Citrus and close relatives in the citrus variant discovery panel 93
- Figure 10 Maximum likelihood phylogeny tree of 87 accessions in Citrus and close relatives with Axiom® Citrus Genotyping Array data 94

Chapter 2

Figure 1 PCA of 382 accessions in UCR diversity panel	273
Figure 2 Cross-validation error plot for admixture analysis with the 382 UCR diversity Panel	276
Figure 3 Admixture analysis with the 382 UCR diversity panel	277
Figure 4 Maximum likelihood phylogeny tree of 85 accessions in Citrus and close relatives	279
Figure 5 Phylogeny network analysis of 382 UCR diversity set	280
Figure 6. Graphical genotyping of Koster CRC 3958 showing the pattern of introgression between pummelo and mandarin at the chromosome level	282
Figure 7 Graphical genotyping of Ellendale CRC 3960 showing the pattern of introgression between pummelo and mandarin at the chromosome level	283
Figure 8 Graphical genotyping of <i>C. amblycarpa</i> CRC 2485 showing the pattern of introgression between pummelo and mandarin at the chromosome level	284
Figure 9 Graphical genotyping of <i>C. leiocarpa</i> CRC 3147 showing the pattern of introgression between pummelo and mandarin at the chromosome level	285
Figure 10 Graphical genotyping of <i>C. keraji</i> CRC 3144 showing the pattern of introgression between pummelo and mandarin at the chromosome level	286
Figure 11 Graphical genotyping of <i>C. tachibana</i> CRC 3150 showing the pattern of introgression between pummelo and mandarin at the chromosome level	287
Figure 12 Graphical genotyping of <i>C. benikoji</i> CRC 3149 showing the pattern of introgression between pummelo and mandarin at the chromosome level	288
Figure 13 Graphical genotyping of <i>C. keraji</i> CRC 3144 showing the pattern of introgression between pummelo and mandarin at the chromosome level	289
Figure 14 Graphical genotyping of <i>C. depressa</i> CRC 2448 showing the	

pattern of introgression between pummelo and mandarin at the chromosome level	290
Figure 15 Graphical genotyping of Cleopatra CRC 3844 showing the pattern of introgression between pummelo and mandarin at the chromosome level	291
Figure 16 Graphical genotyping of <i>C. sunki</i> CRC 3143 showing the pattern of introgression between pummelo and mandarin at the chromosome level	292
Figure 17 Graphical genotyping of candidate type-1 mandarins	293
Figure 18 Graphical genotyping of candidate type-2 mandarins	294
Figure 19 Graphical genotyping of Djerok CRC 2456	295
Figure 20 Graphical genotyping of Bengal CRC 3055	296
Figure 21 Graphical genotyping of <i>C. pennivesiculata</i> CRC 2434, <i>C. excelsa</i> CRC 2316, Soh Long CRC 3199, <i>C. vulgaris</i> CRC 760, India CRC 661, and <i>C. macrophylla</i> CRC 2446	297
Figure 22 Graphical genotyping for Siamese CRC2421	299
Figure 23 Graphical genotyping for Chandler type pummelo accessions	300
Figure 24 Graphical genotyping for Washington navel CRC 1241A	302
Figure 25 Graphical genotyping for Ambersweet CRC 4002	303
Figure 26 Graphical genotyping for Indian CRC 3175 and <i>C. maderaspatana</i> CRC 3225	304
Figure 27. Graphical genotyping for <i>C. intermedia</i> CRC 3474 and Orogold CRC 693	305
Figure 28 Graphical genotyping for Orlando CRC 1588 and Smooth flat seville CRC 4079	306

Chapter 3

Figure 1 PCA for GWAS panels	356
Figure 2 Manhattan plot of <i>Citrus</i> GWAS for nucellar embryony	358
Figure 3 Manhattan plot of <i>Poncirus</i> GWAS for polyembryony	359

LIST OF SUPPLEMENTAL TABLES

Chapter 1

Table 1 Percent heterozygosity at variable bases of accessions in citrus variant discovery panel 2 Accessions in Axiom® Citrus HD Genotyping Array validation panel	114
Table 2 Accessions in Axiom® Citrus HD Genotyping Array validation panel	116
Table 3 Families used in mendelian trio analysis to assess Mendelian inheritance From parents to off-spring for SNP validation	123
Table 4 Mendelian inheritance error calculated for SNP validation with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array	124
Table 5 Summary table of percent average SNP marker heterozygosity in various groups with Axiom® Citrus HD Genotyping Array	127
Table 6 Summary table of percent average SNP marker heterozygosity in various groups with Axiom® Citrus Genotyping Array	128
Table 7 196 unique accessions from CVC for PCA analysis with data from Axiom® Citrus HD Genotyping Array	129
Table 8 871 unique accessions from CVC for PCA analysis with Axiom® Citrus Genotyping Array data	135
Table 9 New classification assignment according to the marker analysis performed during this study	160

Chapter 2

Table 1 Pairwise Fst values for 7 populations studied in the citrus diversity study	307
Table 2. Sample ID and Sample names for accessions in citrus variant discovery panel used for network analysis	308

Table 3. Delta and Qresidual scores for network analysis of 36 accessions	309
Table 4 Delta and Qresidual scores for network analysis of the 382 UCR diversity set	310
Chapter 3	
Table 1 209 individuals in CVC and F x F segregating population screened for nucellar embryony and polyembryony	360
Table 2 KASP SNP estimated % nucellar embryony for 90 accessions in Citrus GWAS panel	367
Table 3 Percent polyembryony for 88 accessions in Poncirus GWAS panel	370
Table 4 Average, median, and standard deviation of KASP estimated percent zygotic by Groups	373

LIST OF SUPPLEMENTAL FIGURES

Chapter 1

Figure 1 Venn diagram of 5.3 M (SNPs and INDELs) variants of initial candidate marker loci for designing Axiom® Citrus HD Genotyping Array	96
Figure 2 Venn diagram of variant sites used as QC probes (20,000) for Axiom® Citrus HD Genotyping Array	97
Figure 3 Venn diagram of 1.4M variants for Axiom® Citrus HD Genotyping Array	98
Figure 4 Distribution of SNPs represented on Axiom® Citrus HD Genotyping Array on each of nine chromosomes	99
Figure 5 Venn diagram of 57,933 variants for Axiom® Citrus Genotyping Array	102
Figure 6 Distribution of SNPs represented on Axiom® Citrus Genotyping Array on each of nine chromosomes	103
Figure 7 Topological comparison of two trees generated with data with Axiom® Citrus Genotyping Array and sequence data of 26 accessions in citrus variant discovery panel	106
Figure 8 Percentage distribution of markers represented on Axiom® Citrus HD Genotyping Array	107
Figure 9 Percentage distribution of markers represented on Axiom® Citrus Genotyping Array	108
Figure 10 Genotyping error simulation of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo Citrange CRC 2863 for PCA	109
Figure 11 Genotyping error simulation of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo citrange CRC 2863 for Neighbor-joining trees	110

Figure 12 Genotyping error simulation analysis of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo citrange CRC 2863 for admixture	111
---	-----

Figure 13 Neighbor-joining phylogeny tree of 29 accessions in <i>Citrus</i> and close relatives in the citrus variant discovery panel	113
---	-----

Chapter 2

Figure 1 CVC924 PCA	320
---------------------	-----

Figure2 Network analysis of 36 accessions from citrus variant discovery panel	321
---	-----

Chapter 3

Figure 1 Frequency distribution of nucellar embryony in <i>Citrus</i> GWAS panel	374
--	-----

Figure 2 Frequency distribution of polyembryony in <i>Poncirus</i> GWAS panel	374
---	-----

GENERAL INTRODUCTION

Citrus, a member of the family Rutaceae, is one of the most widely cultivated and economically valued perennial fruit tree crops in the world. It is cultivated in more than 100 countries [1] in Mediterranean and subtropical climatic regions of the world [1], with its estimated production at 146 million tons [2]. Citrus fruits are rich in various functional compounds and are nutritionally beneficial for human health [3, 4], in addition to being highly valued for their fragrance, medicinal properties, and in some cases, as ornamental fruits. The lineage from ancestral *Citrus* to many of the modern cultivated *Citrus* types has been studied to show a complex history of admixture [5–7], and domestication is thought to have started in Southeast Asia before dispersing to the rest of the world [8]. Earliest record of citrus dates back to 2205 BC in China where fruits of mandarins and pummelos were described to be only available for the imperial court [9], highlighting the long history of human and citrus interaction.

Citrus production

The global citrus production is at 146 million tons (MTs) which exceeds that of other popular commodity fruits such as watermelons (118 MTs), bananas (113 MTs), apples (83 MTs), and grapes (74 MTs) [2]. More than half of the global citrus production is by orange (56%), followed by mandarins, tangerines, and clementines (17%), lemons and limes (11%), grapefruits and pummelos (6%), and other citrus fruits (10%) such as kumquats, and citrons [2]. Major citrus producers include China (mandarins, tangerines and clementines), Brazil (oranges), and the United States [2]. In the United States, citrus production is valued at 3.35 billion, and more than half is accounted for by California (51%), followed by Florida (44%), and Texas and Arizona (5%) [10]. Citrus was reportedly introduced to the United States by Spanish and Portuguese missionaries [11], with

introduction to Florida and California in 1565 and 1769, respectively [9]. The citrus industry in the U.S.A. then grew with the expansion of railroads during the late 1800s, followed by arrival of commercial scale production and global trade, making California and Florida the leading citrus producing states [12]. Citrus production in Florida is predominantly for juice and processed citrus as opposed to California where the fresh fruit market is of primary importance.

California citrus industry

The history of Riverside, California, is deeply intertwined with that of California's citrus industry. Although the exact date is unknown, the path to the great citrus industry in California is thought to have begun around 1769 when the first mission moved from Baja California to San Diego and citrus seeds or plants were assumed to have been brought along [9]. Since the establishment of the first orange grove at San Gabriel Mission in 1804, citrus gradually spread throughout Los Angeles using seeds and trees from San Gabriel Mission [13]. By 1834, seeds from San Gabriel Mission led to California's first commercial citrus production established in downtown Los Angeles by William Wolfskill [9, 14], whose endeavor soon coincided with the west coast's 'Gold Rush'. The economic prosperity following California's Gold Rush encouraged expansion of railroads [15, 16], which became the biggest turning point for California's citrus production. By the time the city of Riverside was founded in 1870, the Old Mission Road between Los Angeles and San Bernardino was filled with citrus groves changing the landscape of California [12].

The rise of the orange empire, Riverside, began with the introduction of two Washington navel trees to Riverside in 1873 by Eliza Tibbetts [17], from which millions of Washington navel oranges were propagated. These two Washington navel trees were sent from the USDA, which they originally acquired as budded trees from Brazil [18]. One of the two original Washington navel trees was transplanted from the Tibbetts homestead in 1902 and is still located at the corner of

Magnolia and Arlington Avenues (Riverside, CA). With rapid expansion of citrus production in Riverside and the demand by growers, the Citrus Experiment Station, predecessor to the University of California Riverside, was established in 1907, in order to aid growers and industry with problems related to fruit production including diseases management, pest control, water management, nutrition deficiencies, post-harvest practice, salinity and occasional frosts. In addition, the Citrus Variety Collection (CVC) was established in 1910 in order to conserve and evaluate *Citrus* and its relatives, to maintain and provide a resource of genetic diversity for research, and to improve the understanding of genetics and diversity in citrus (CVC archives). The collection now maintains an *ex situ* collection of over 1000 unique accessions of *Citrus* and related genera collected from all over the world. Since the establishment of Citrus Experiment Station, UCR has become the main public citrus breeding institution in California involved in variety improvement and releases of varieties including navel oranges, Valencia oranges, lemons, grapefruit, mandarins, tangelos, tangors and rootstock varieties.

Citrus distribution

Citrus has been domesticated and cultivated for several thousand years [19] resulting in modern varieties resulting from centuries of natural and human-directed hybridization, mutation, and selection. The exact lineage from ancestral Citrus to many of the modern cultivated Citrus types is still debated, however, domestication is thought to have started in Southeast Asia and dispersed to the rest of the world [6, 8] leading to interspecific hybridization and introgression of ancestral progenitors into cultivated *Citrus* varieties. In addition, how and when citrus was distributed from where they were indigenous to the rest of the world, and whether or not dispersal occurred via seeds or plants could have also contributed and influenced phenotypic and genotypic variations among the citrus varieties that we observe today.

Citron (*Citrus medica* L.) was reported to be the first citrus to reach Europe [20] as indicated by the existence of historical references to citrons in early European literature, while other citrus reached Europe by the Arabian and Crusaders. Citrons were originally thought to be indigenous in Media, an ancient country of northwestern Iran, and then spread into Persia and hence the term medica [9]. However, contrary to previous beliefs, it is now widely accepted that citron originated in India, near the foothills of the Himalayas [21]. The exact path by which citron spread from its native habitat to Media is still not known. One hypothesis is that the Jewish people brought citrons from India to Palestine during their Babylonian exile around 500 B.C. [22] while another claims that the spread of citron occurred only after Alexander the Great passed through Greek settlements of Persia during 300 B.C. [9]. The most credible evidence to suggest that citron had already been spread from its native habitat by 300 B.C. is seen in the accurate description of citron by Theophrastus, a Greek botanist who supposedly never left Greece. By the middle of the first century of A.D., citron was introduced to Italy based on archeological, and literary evidence, such as sculptures and mosaics [23]. By the 4th century, it is thought that Ligurians took citrons from southern Europe including Sicily, Sardinia, and Naples and spread them to the West via Venetians trade.

While it remains uncertain, lemon (*Citrus limon* [L.] Burm. f.) is also thought to have originated in Southeast Asia, specifically, India. The importance of lemon in Indian culture and religions can be seen in the historic records, and many wild lemons grow in the region [24]. In addition, citron being native to the same area, supports the likelihood that lemon is from India as lemon is a natural hybrid of sour orange and citron. When the introduction of lemon to Europe occurred was argued by several citriculture scholars. Gallesio claimed that lemons were not known by Romans while Tolkowsky claimed otherwise [9]. Tolkowsky's argument was based on the fact that Romans were known to pay for and value highly exotic products, and there is a mosaic

depicting fruit bearing lemon trees in a Roman villa presumably built during the 2nd century A.D. [24] which is also supported by the recent findings from Italy showing that the lemon was already known to the Romans prior to the arrival of the Arabians in 7th century [25]. One evidence to support the former claims is that the Arabians continued to spread lemons into the rest of the world after the collapse of Roman empire. The earliest record of lemon in China is 1175 A.D. and it was called “li-mung” which is “limun”, the name for lemons used by the Arabians [9]. In Egypt, the historic records referred to lemon as “el-limun” [9], also indicating that the lemons were grown by 700 A.D. [26] and brought there by the Arabians as they continued to spread lemons into northern Africa, and south of Europe. The Crusaders brought lemons into other parts of Europe including to Liguria, other parts of Italy, and France as the earliest record of lemons in European literature only appears after the period of the Crusades. By the 13th century, lemons were commonly grown in these regions. From the 15th century on, the Spanish and the Portuguese continued to spread lemons into the different subtropical regions of the world [9].

Sour oranges (*Citrus aurantium* L.) are thought to have originated in Southeast Asia as with many other Citrus species [23]. The debate over the Roman’s familiarity with lemon described above goes the same for sour oranges. It is believed that sour orange trees were brought from India in 912 A.D. by the Arabians and first planted in Oman, then spread to Iraq, Syria, Palestine, Egypt first and later to western Asia, northern Africa and southern Europe including Sicily, Sardinia, and Spain [9]. The Crusaders spread sour orange into the other parts of Europe as was the case for other citrus.

Sweet orange (*Citrus sinensis* L.) is thought to have originated southeast Asia. The exact path by which sweet orange was brought to Europe is still debated. However, there are numerous written records indicating that sweet oranges had become established and commercially important in southern Europe by the 16th century [9]. Although it was argued by citriculture historians,

Tolkowsky [24] claimed that sweet oranges were grown already in the Mediterranean regions of Europe including Italy, Spain and Portugal prior to Vasco da Gama's voyage in 1497 A.D. based on the existence of a letter sent by the French king Louis XI in 1483 to the governor of Languedoc requesting to send him citrons, and sweet oranges along with other fruits. The desire of the Europeans for Eastern exotic goods was stimulated by the Crusaders, leading to the establishment of trade routes including Genoese and Venitian trade routes [25]. Opening of these two trade routes led the Arabians and those involved in trade with Asia to search for rare and novel products for commerce. Therefore, it is assumed that during this time sweet oranges were brought to Europe via Genoese trade prior to Vasco da Gama's voyage of discovery [9]. The Portuguese did not bring the first sweet oranges to Europe, however they ignited the popularization of sweet oranges by bringing back more superior varieties from China, which became known as Portugal orange [27].

Limes are diverse in terms of the tree and fruit characteristics and include several types such as small fruited Mexican lime, *Citrus aurantifolia* (Christm.), Swingle and the triploid seedless *Citrus latifolia* type [23]. The first reference to the lime appeared in 13th century literature. Based on the description, citriculture historians believed that the fruit mentioned in the literature was of the lime of Naples, leading them to think that the Arabians must have been also involved in the spread of lime from India to Persia, Palestine, Egypt and Europe as with many other citrus [9].

Mandarin (*Citrus reticulata* Blanco) is a native to China, Japan and India [28] its cultivation history of over 4000 year [23]. Unlike other citrus discussed so far, mandarin was not brought to Europe during the same period other citrus reached Europe, instead, it remained mainly cultivated in the area of origin [25] including China and Japan, leading to many mandarin varieties that are endogenous to China and Japan. The early record showed that there were several distinct varieties of mandarins already existing in 1178 A.D. [9]. During the 16th century in Japan, there was a rapid expansion of cultivation of mandarin in the regions of Higo, Kishu, and Suruga in

Japan. It was reported that most cultivated mandarins at the time were a small fruited Kishu mandarin type that originated in China. Since the development of the larger fruited Satsuma mandarins (*Citrus unshiu* Marc.) presumably in Kagoshima province, different selections of *C. unshiu* were made and gradually spread throughout Japan. Interestingly, mandarin was the last major citrus fruit to reach the West. Tolkowsky stated that the first mandarin was brought to England from China in 1805 then spread to Malta, Sicily and to Italy thereafter [23, 25]. In addition, it was reported that Owari mikan was the first mandarin exported to the United States in 1878 from Kagoshima province, then called Satsuma province [29].

Pummelo (*Citrus grandis* L.), also known as shaddock, is thought to have originated in the Malayan and East Indian archipelagos then spread to China, India and to Persia, Palestine and to Europe [9]. The path of pummelos to Europe followed a similar path as the sweet orange and sour orange. The earliest record of pummelo mentioned that it was growing in Palestine in 1187. Based on the description and illustration of pummelo, it presumably reached Europe by the 12th century, also by the Arabians. The introduction of pummelo to Japan, was in the 1770s where they were named Buntan or Zabon, and mainly grew in south west region of Japan [30].

Papeda is a subgenus of *Citrus* and comprises several different species including *Citrus ichangensis*, *Citrus micrantha*, *Citrus latipes*, *Citrus celebica*, *Citrus macroptera*, and *Citrus hystrix*. Papeda is one of the most confounding taxonomic groups in *Citrus* and species of papeda supposedly originated in different parts of the world ranging from west, central and southwest China (*C. ichangensis*), northeast India and northern Myanmar (*C. latipes*), India, southern Philippines (*C. micrantha*) and northern Sulawesi and southern Philippines (*C. celebica*) and Indonesia, Malaysia, Papua New Guinea, and Thailand (*C. macroptera*) [8, 23, 31, 32].

Fortunella (kumquat) is a subgenus of *Citrus* and also comprises several different species including *Fortunella margarita* (Lour.) Swingle, *Fortunella japonica* (Thumb.) Swingle,

Fortunella polyandra, and *Fortunella hindsii* (Champ. ex Benth.) Swingle. *Fortunella* species are mostly grown in China, southern Japan, and Taiwan and are thought to be native to the south to southeast part of China [9]. The earliest record of *Fortunella* in China was in 1178 in the first monograph of citrus fruits ever published in any country [26]. *Fortunella* species did not reach Europe until much later compared to the other citrus until the late 19th century while the record shows that it was introduced to Japan by the 18th century [9].

Microcitrus and *Eremocitrus* are thought to be native to Australia and have been isolated for millions of years from other citrus that are thought to have originated from Asia [33]. In addition, it is thought that they escaped competition with *Citrus* due to the geographical isolation and the fact that they are native to the region where *Citrus* could not grow well [19].

Poncirus (trifoliate orange), is a related genus to *Citrus* that freely hybridizes with *Citrus*. Interestingly, there is little morphological variation among *Poncirus* accessions in the CVC, in comparison to *Citrus* despite thousands of years of history of cultivation in China and in Japan since the 8th century and is considered the most stable species with a long history of cultivation [9].

After initially reaching Europe. *Citrus* finally reached the Americas when Christopher Columbus brought citrus seeds into America in 1493 [34]. By the time Columbus reached America, sweet orange fruits were already valued highly in the Mediterranean countries. It is thought that during the settlement in the New World on the island of Haiti, the gardens colonists developed included trees of sweet orange, sour orange, lemon, citron and lime which were the citrus that Spanish and Portuguese would have been already familiar with [9]. It was also during this time that citrus was brought to many Pacific islanders. The earliest record of citrus in continental America, from 1518, mentions the sowing of oranges seeds, thereby establishing the first planting of citrus by seeds in 1518. By 1554, oranges are thought to have become abundant in Central America and

it was around this time that citrus was first introduced to Brazil. By the 16th century, citrus, including oranges, lemons, limes and citrons had spread throughout the Americas [9].

Interestingly, opening routes to the Americas and the subsequent dispersal of *Citrus* into the New World had significant influence on the variations in *Citrus* that we can observe today. For example, grapefruit is thought to have originated as a chance seedling on the island of Barbados in the Caribbean in 1873 when grapefruit was referred to as the forbidden fruit *Citrus paradisi*. The Island of Barbados was English and Dutch speaking territories where citrus trees were cultivated including sweet orange and pummelos by 1684 [34] and it was there that the hybridization between the shaddock and sweet orange was thought to have occurred leading to the development of grapefruit [35].

The earliest record of citrus introduction to Southern Africa is dated 1654, after citrus was already widely distributed to most of the tropical and subtropical regions of the Eastern hemisphere. The record shows that sweet orange trees were brought to South Africa to be planted in Cape Town's governor's private garden. By 1661, oranges, lemons, pummelos and citrons had already been introduced to Southern Africa [9]. It is reported that Seville sour orange and Mazoe lemon (rough lemon) were growing wild by the river banks of Mazoe river then [36].

Orange was reportedly introduced to Australia in 1788 when colonists settled in New South Wales with the mission to introduce plants and seeds after acquiring trees of orange, lime, and lemon during the stop in Brazil prior to the arrival in Australia.

In modern days, the distribution of citrus is achieved via seeds or budwood, and in most countries the budwood is cleaned of any pathogens prior to the distribution. Nearly all citrus is propagated by grafting, to reduce the long juvenility and to ensure true to type propagation. The grafting has a long history dating back to the Greeks and the Romans [37] and is said to have been practiced in China before A.D.[38]. By the 16th centuries, grafting was supposedly a common

practice (unpublished Citrus rootstocks: their characters and reactions W.P. Bitters 1986). However, very little is known about the history of citrus grafting, such as when and how it began. If the assumption that citrus did not reach Europe and the West before the Arabic era is correct, it can be concluded that grafting of citron on other citrus species could not have been practiced around the Mediterranean before the Arabic era [37].

Many commercially important citrus including navel oranges, Valencia oranges, Eureka lemons, Lisbon lemons, and clementine are clonal groups, composed of cultivars that have arisen via accumulated somatic mutations, sport mutation or nucellar variants, and which are propagated vegetatively by grafting. This profoundly affects the pattern of genetic diversity within these groups.

Citrus Taxonomy

Citrus phylogeny and taxonomy are notoriously controversial and confounding. Phylogeny and taxonomy of Citrus have been studied based on morphology and geography prior to the arrival of molecular markers. There are two widely accepted classical Citrus taxonomic systems, Swingle and Reece, and Tanaka, who each classified Citrus into 16 species in two subgenera (*Citrus* and *Papeda*) [19] and up to 162 species [39]. This discrepancy exists mainly in *Citrus*, and stems from how much divergence within a recognized species should be allowed, and more specifically, whether natural hybrids should be considered as true species. Various molecular markers such as simple sequence repeats (SSR), restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), sequence characterized amplified region (SCAR), sequence related amplified polymorphism (SRAP), mitochondrial DNA, chloroplast DNA, chromosomal variability, and nuclear gene sequences have been used in phylogenetic investigations into Citrus to try to better understand the origin and relationships of diverse accessions [40–47]. Analysis of

markers [41, 48] and sequencing indicate four major ancestral Citrus groups, citron (*C. medica*), pummelo (*C. maxima*), mandarin (*C. reticulata*), and papaya (*C. micrantha*) [6], while Scora and Barrett and Rhodes [31, 49] only recognized the first three groups as ancestors of cultivated citrus. Further complicating phylogenetic studies of *Citrus* is more than a thousand years of domestication that led to hybridization and introgression of ancestral progenitors into the cultivated *Citrus* varieties.

Citrus genomes

The arrival of the genomic era in plants came with construction of the first plant reference genome of *Arabidopsis thaliana* in 2000 [50]. Since then, rapid advances were made in sequencing technologies and available bioinformatics tools that enabled the construction of several citrus reference genomes. *Citrus* is generally a diploid with a chromosome number of $2n=18$ and in most species its genome is highly heterozygous resulting from self-incompatibility, interspecific sexual hybridization and occasional intergeneric hybridization [51].

The first publicly available citrus reference genome *C. sinensis* was constructed using shotgun sequencing with Illumina paired-end tag sequencing technology from a double haploid *C. sinensis* cv Valencia. The size of the *C. sinensis* genome was estimated to be 320.5 Mb. There are 29,455 coding genes with average sequence length of 1.3 kb and 5.8 exons [52]. *C. clementina* v1.0 [5] was assembled using shotgun sequencing with Sanger sequencing technology from a selected haploid plant derived from *C. clementina* cv. Clemenules [53]. The size of the *C. clementina* v1.0 genome was estimated to be 301.4 Mb and synteny analysis showed highly conserved relationships with other eudicots. There are 24,533 protein coding genes with an average sequence length of 2.5 kb and 5.3 exons [5]. *C. sinensis* v 1.1 genome, constructed using a combination of Roche 454 pyrosequencing and Sanger sequencing technologies from diploid ‘Ridge Pineapple’, became

publicly available along with *C. clementina* v1.0 [5]. The size of the *C. sinensis* v 1.1 genome was estimated to be 319.2 Mb. There are 25,379 protein coding genes with an average sequence length of 2.1 kb and 4.9 exons [5]. A *C. grandis* (*C. maxima*) reference genome was assembled using Pacific Biosciences and Illumina sequencing technologies from a haploid *C. grandis*. The size of *C. grandis* genome was estimated to be 346,302 Mb [54]. There are approximately 30,123 coding genes with average sequence length of 1.1 kb [54]. A reference genome of Satsuma mandarin (*C. unshiu* Marc.) was constructed using Illumina sequencing technology and Pacific Biosciences from a diploid ‘Miyagawa Wase’. The size of the *C. unshiu* genome is estimated to be 359.7 Mb with 29,024 protein coding sequences with an average sequence length of 1.2 kb and 4.1 exons [55]. Additionally, *C. medica*, *C. ichangensis*, and *Atalantia buxifolia* (*Citrus* relative) genomes are available [54] which were constructed using Illumina sequencing technology from diploid accessions (Table 2).

The availability of these reference genomes rapidly facilitated high-resolution genomic studies in citrus, though it still lagged behind most other major crops. Recently, a detailed analysis of the admixed genome of cultivated sweet orange (*C. sinensis*), a known hybrid that originated from ancestral progenitors, mandarin (*C. reticulata* Blanco) and pummelo (*C. maxima*), has been reported [5, 52]. In addition, analysis of resequencing data from various accessions revealed that some cultivated mandarins (*C. reticulata*) are admixed with pummelo from repeated introgression [5, 56, 57], contrary to previously accepted views on the ancestry of these cultivars. Lemon, which is the third largest citrus group, was long thought to be a natural hybrid of sour orange and citron [31, 49], a hypothesis confirmed by extensive sequence data [6, 58]. Lime, also a member of the large “acid citrus” group, was confirmed as a natural hybrid of citron and one of the papeda species *C. micrantha* [6, 56]. Additionally, sequence analysis revealed *C. tachibana* to be more a divergent

type-1 mandarin (no introgression with pummelo) and that *C. ichangensis* is divergent from *C. micrantha*, suggesting polyphyletic nature of papeda [6].

Citrus genotyping methods

Progress in high-resolution genomic studies of citrus germplasm resources was hindered by the lack of high-throughput genomics tools available and appropriate to analyze complex citrus populations. One of the alternatives available to whole genome sequencing is high throughput genotyping technologies such as high-density genome-wide single nucleotide polymorphism (SNP) genotyping arrays. Such high-density SNP genotyping arrays give high fidelity data with sufficient resolution and power to conduct high-resolution genomic studies [59–63].

Several arrays have been previously developed for *Citrus* and were applied to gene expression studies, phylogeny studies, construction of genetic maps, trait mapping and diversity studies. The Affymetrix™ GeneChip® citrus genome microarray was developed by the laboratories of Dr. Roose and Dr. Close at the University of California Riverside for gene expression studies. Total of 33,879 *Citrus* transcripts, based on EST sequences from *Citrus* and its hybrids, are targeted by 30,171 probes represented on the array (Affymetrix™, Thermo Fisher Scientific, Santa Clara, CA). Other gene expression arrays, 7K and 20K expression arrays were developed which allowed the expression study of 6,875 and 21,081 putative citrus genes based on EST sequences, respectively [64, 65]. A 1,457 GoldenGate SNP array was developed from SNPs mined from BAC end sequences of *C. clementina* and was applied to a citrus diversity study by genotyping 96 accessions [66]. The 384 marker SNP array, *CitSGA-1*, was developed from SNPs mined from the SNP discovery panel consisting of 8 accessions representing the genetic diversity of citrus important for breeding in Japan and used to genotype 191 accessions [67]. The 756 GoldenGate SNP array was developed from a SNP discovery panel consisting of 15 accessions

important in breeding in Japan and used to genotype 120 citrus accessions [68]. Alternative to the sequencing and the array approaches is genotyping by sequencing (GBS) which utilizes restriction enzyme technologies to generate a representation (subset) of the genome that is then sequenced [69]. The GBS approach has been applied for admixture and phylogeny analyses of 55 mandarin and pummelo hybrids (did not include non-admixed type-1 mandarins nor pummelo), using 9,995 SNPs and 1,178 Indels with missing data of 12.3% [57].

Genome-wide association mapping (GWAS) was conducted using the GBS approach for fruit quality traits such as fruit weight, skin color, surface texture, peelability, pulp firmness, segment firmness, sugar content, and acid content using 2,309 SNPs and 110 accessions including citrus landraces, modern cultivars and breeding lines [70]. GBS was also applied to infer 15,946 diagnostic single nucleotide polymorphism (DSNPs) for ancestral taxa mined from a panel of 29 accessions, and phylogenomic karyotypes showing genome wide introgression pattern were inferred in 53 accessions (did not include non-admixed type-1 mandarins) [7]. A variant of GBS, RAD-sequencing, was used to create high resolution maps and identify QTLs affecting HLB tolerance [71].

Citrus apomixis

Apomixis is defined as asexual reproduction by seed, allows clonal propagation through generations, and is found in more than 400 plant species [72]. Progenies produced by apomictic plants are genetically identical to the maternal plant. Such a trait is advantageous in plant breeding and for seed production for its ability to allow continuous propagation via seeds of genotypes fixed for desired traits. There are two distinctive modes of apomixis: (1) gametophytic apomixis and (2) sporophytic apomixis [73]. Sexual reproduction is a highly regulated process [74], and processes involved in apomixis are expected to follow similar developmental and gene expression patterns as

sexual reproduction [54]. Relevant processes involved in sexual reproduction include female gametogenesis and double fertilization events within the ovule [74, 75], hence, the generation and fusion of reduced gametes, resulting in recombination of parental genetic materials.

Gametophytic apomixis, to the contrary, involves the formation of an unreduced embryo sac (female gametophyte) arising from an unreduced embryo sac initial and parthenogenetic development of their egg cells, thereby skipping the process of meiosis all together [72]. There are two components to gametophytic apomixis; (1) diplospory and (2) apospory. Diplospory involves the unreduced female gametophyte arising from a megaspore mother cell by mitosis or modified meiosis leading to an unreduced nucleus [76]. Apospory involves the unreduced female gametophyte arising from somatic cells of the ovule, including the archesporium, nucellus, or inner integument [77].

While most cases of apomixis are considered gametophytic [78, 79], citrus apomixis is a nucellar embryony, a type of sporophytic apomixis. Sporophytic apomixis involves somatic embryos arising directly from the maternal nucellar tissue surrounding the embryo sac [80], and is a unique and common phenomenon in some citrus groups.

Citrus apomixis research

The term nucellar embryony or adventitious embryony, is often interchangeably used with polyembryony in citrus, however, nucellar embryony and polyembryony are two distinct but related traits. Nucellar embryony is a rare trait in plants, best characterized in some citrus and mango (*Mangifera indica* L.) [81]. Many popular citrus cultivars are known to be polyembryonic, except for clementine, kinokuni mandarin, and pummelo [82–87] (Table 3). Interestingly, monoembryonic cultivar ‘Kiyomi’ is a mandarin derived from two polyembryonic cultivars [88]. The first detailed in vitro culture study of development of nucellar embryos in Valencia orange (*C.*

sinensis) identified the presence of nucellar initial cells at the binucleate stage of megagametogenesis [89, 90], much earlier than it was previously reported [80].

Considerable variation among citrus genotypes in the relative proportions of zygotic and nucellar embryos has long been known, and to show that environmental factors influence the trait [91, 92]. Originally these were largely based on using morphological characters to distinguish zygotic from nucellar seedlings. With the development of various types of molecular markers, isozymes and later DNA markers have been used to determine the relative proportions among seedling populations [83–85]. For rootstock breeding, nucellar embryony is often a desired trait, because it allows most citrus rootstocks to be seed propagated but remains mostly true to type genotypes. Prior to the arrival of molecular markers, the roguing was most often practiced where off-type seedlings are removed based on visual assessment [93]. This method is still employed to an extent, however, nursery escape of the zygotic seedlings has been reported [94]. For scion breeding, in contrast to rootstock breeding, nucellar embryony hinders the breeding effort as some genotypes with this trait produce few to no hybrid seeds, making it difficult for them to be used as seed parents.

Several hypotheses have been proposed during the last several decades to explain the characteristics and the genetics underlying the inheritance of nucellar embryony. One study using type and the number of embryos of the individuals in the segregating populations, reported that polyembryony in both *Citrus* and *Poncirus* is a single dominant gene controlled trait [95], which was later questioned by an another study suggesting that a single dominant gene could not explain the inheritance of the trait in both *Citrus* and *Poncirus* [96]. These findings lead to the two dominant loci control hypotheses in the crosses of *Citrus* and *Poncirus* [92, 97]. The first study with molecular markers analyzed a segregating population of *C. volkameriana* $\times$ *P. trifoliata*, and reported 6 QTL that influence nucellar embryony in *Citrus* [98]. A study of % polyembryonic

seed using both bulk segregant and QTL analysis with AFLP markers identified a major locus essential for polyembryony in a population created by a cross of Chandler pummelo x *P. trifoliata* [99]. This study showed that the major dominant locus controlling polyembryony, and by inference most polyembryony, is heterozygous in *P. trifoliata* and absent in pummelo, and a significant segregation distortion was observed [99]. A study conducted with two segregating populations created by a cross between *C. reshni* (Cleopatra mandarin) and *P. trifoliata* and by self-pollinating selected monoembryonic hybrids from the *C. reshni* and *P. trifoliata* population reported 6 QTL and a segregation distortion [100]. These later reports implied that the presence of nucellar embryony is under a single dominant locus control while the proportion of polyembryonic seeds a given genotype can produce is determined by multiple loci, and that the genetic control in *C. reticulata* (*C. reshni*) and *P. trifoliata* is likely different. Construction of a genetic and physical map aided investigating the inheritance of nucellar embryony in *C. reticulata* by studying segregating populations involving monoembryonic and polyembryonic cultivars using cleaved amplified polymorphic sequence (CAPS) [101], and expressed sequence tags (EST), leading to the identification of a 380 kb genomic region controlling nucellar embryony [82]. In addition, there have been numerous transcriptomic studies targeting nucellar embryony. The first transcriptome study investigating *Citrus* nucellar embryony looked into the differentially expressed genes (DEG) during nucellar embryo initiation at pre- and post-anthesis stage along with the expression patterns of genes in the nucellar embryony locus in both monoembryonic ‘Temple’ and polyembryonic ‘Vaniglia Sanguigno’ and found the stress signaling genes such as JSP and WRKY are likely to be involved [102]. Subsequently, a subtractive suppression hybridization (SSH) based microarray was used to study the expression profiles of DEG between the ovules of monoembryonic pummelo hybrids *C. tamurana* Tanaka, Hyuuganatsu and *C. hassaku* Tanaka, Hassaku and polyembryonic

sour orange hybrids *C. aurantium* L., Daidai and *C. natsudaoidai* Hayata, Natsudaoidai further characterizing ESTs [103].

After the release of the *C. sinensis* [52] and *C. clementina* [5] reference genomes, the possible role of miRNA was been reported from studies using sweet orange, sour orange, navel orange, mandarin, Satsuma and grapefruit [104]. In addition, transcriptional and post-transcriptional regulators were investigated by studying DEG between the ovules of monoembryonic clementine and polyembryonic Ponkan (*C. reticulata*) using mRNA-seq and sRNA-seq approaches, suggesting the possible role of miRNA in nucellar embryogenesis in *C. reticulata* [105]. Additional genomic studies enabled identification of the first candidate gene responsible for nucellar embryony in *C. reticulata* on chromosome 4 (between 24.54 and 24.62 Mb) on the *C. maxima* reference genome [54] within the previously identified polyembryony locus [82]. The reported candidate gene *CitRWP* was highly expressed in the ovules of polyembryonic cultivars and a miniature inverted repeat transposable element (MITE) was found to be inserted in the promoter of *CitRWP* in polyembryonic cultivars [54, 106]. This recent finding supported the standing hypothesis that one major dominant locus controls the induction of nucellar embryony in *Citrus*. Additionally, their study implicated the involvement of similar genes or pathways in both nucellar and sexual embryogenesis [54].

Significance of the study

High-resolution genomic studies as achieved in most major crop species lagged in citrus due to its complex genetics and lack of available and appropriate genomic tools. The high-throughput and cost-effective tools to analyze complex citrus populations involving a large number of individuals were not readily available. The initial investment for the development of high-density genome wide SNP genotyping arrays can be large, due to the amount of time and knowledge required for SNP

detection and array design. However, once developed, a SNP array genotyping approach is often more efficient and cost effective, in comparison to the whole genome sequencing [107, 108] or Genotyping by Sequencing (GBS) approaches. Citrus genomes are highly heterozygous and imputation of missing calls associated with the GBS approach demands well-established complex pipelines. Therefore, the computational requirements necessary for downstream analysis with data from whole genome sequencing and GBS approach makes SNP array genotyping often more efficient, cost effective and high throughput. Development of high-density genome-wide SNP genotyping arrays for citrus and close relatives, with SNPs that have good coverage of diverse genomes and target genes will provide citrus breeders with high throughput genetic marker screening systems. Such SNP arrays also provide breeders a platform with a set of high-density makers that can be used among different mapping populations and germplasms. These tools will significantly reduce the time that citrus breeders take to characterize germplasm, breeding populations and mapping populations to identify and locate genes associated with traits of interest.

In addition, phylogeny and diversity studies in citrus previously reported used a relatively small number of accessions or focused on specific taxonomic groups. The high-density genome wide SNP arrays for citrus provides enough resolution, throughput and efficiency to perform a range of high resolution genomic studies including phylogeny studies, diversity studies, chromosome level introgression analysis, genome-wide association studies (GWAS), genomic selection (GS) and breeding applications. Diverse citrus germplasm is available from the University of California Riverside Citrus Variety Collection (CVC), however, the majority of the collection have never been characterized with high-density marker systems. Genotyping the entire collection with high-density genome wide SNP arrays will enable development of an extensive inventory of the genetic profile of the accessions in the CVC. This information will help to understand the

history of admixture and ancestry, and aid in updating sometimes questionable passport records of many accessions added to the CVC during more than 100 years of history.

Objective of the study

The major objective of this study was to design and develop high-density SNP genotyping arrays, Axiom® Citrus HD Genotyping Array HD Genotyping Array and Axiom® Citrus Genotyping Array and assess array performance to test their utility and application in genomic and genetic studies in *Citrus* and close relatives. The design criteria for the SNP genotyping arrays included accurately calling SNPs in diverse accessions of *Citrus* and its close relatives that sexually hybridize with *Citrus*, providing high genome coverage, inclusion of chloroplast SNPs and robustness. Every single accession in the CVC was to be genotyped to yield an extensive genomic inventory. The SNP genotyping array data from CVC was to be used to analyze phylogeny, characterize diversity and determine admixture, including chromosomal level introgression analysis, to reveal the complex history of admixture and ancestry in *Citrus*. Lastly, GWAS was to be conducted in order to study the distribution and inheritance of nucellar embryony loci in *Citrus* and its close relative *Poncirus trifoliata*. A GWAS panel from the CVC and segregating populations was to be genotyped with the SNP arrays. The study also aimed to compare analyses conducted using data produced with high-density SNP genotyping arrays for citrus which can provide resolution, and information content to infer phylogeny, admixture and chromosomal level introgression in complex citrus population. High-density SNP genotyping arrays, Axiom® Citrus HD Genotyping Array HD Genotyping Array and Axiom® Citrus Genotyping Array were intended to be made available publicly, allowing facile comparison of the composition of germplasm collections, and providing breeders and geneticists with a valuable new tool for genomic analysis.

Reference cites

1. Fao. FAO statistical yearbook: World Food and Agriculture. FAO; 2013.
2. Stat F. Food and agriculture organization statistical database. Retrieved February. 2017;27:2017.
3. United Nations, editor. Food and Agriculture Organization of the United Nations (FAO). In: Yearbook of the United Nations 1995. UN; 1995. p. 1469–72. doi:10.18356/d6ce0776-en.
4. Tripoli E, Guardia ML, Giammanco S, Majo DD, Giammanco M. Citrus flavonoids: Molecular structure, biological activity and nutritional properties: A review. Food Chem. 2007;104:466–79. doi:10.1016/j.foodchem.2006.11.054.
5. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotechnol. 2014;32:656–62. doi:10.1038/nbt.2906.
6. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of Citrus. Nature. 2018;554:311–6. doi:10.1038/nature25447.
7. Ahmed D, Comte A, Curk F, Costantino G, Luro F, Dereeper A, et al. Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: a case study in diploid and polyploid citrus. Ann Bot. 2019;123:1231–51. doi:10.1093/aob/mcz029.
8. Swingle, T W. The botany of Citrus and its wild relatives. The citrus industry. 1967;:190–430. <https://ci.nii.ac.jp/naid/10006100992/>. Accessed 14 May 2018.
9. Reuther W, Webber HJ, Batchelor LD. The citrus industry volume I. University of California. 1967.
10. Statistics A. United States Department of Agriculture, National Agricultural Statistical Service. Washington DC: United States government printing office, 512 p. 2019.
11. Liu Y, Heying E, Tanumihardjo SA. History, Global Distribution, and Nutritional Importance of Citrus Fruits. Compr Rev Food Sci Food Saf. 2012;11:530–45. doi:10.1111/j.1541-4337.2012.00201.x.
12. Patterson TC. From Acorns to Warehouses: Historical Political Economy of Southern California's Inland Empire. Left Coast Press; 2014. <https://play.google.com/store/books/details?id=BgU0BgAAQBAJ>.
13. Zierer CM. The Citrus Fruit Industry of the Los Angeles Basin. Econ Geogr. 1934;10:53–73. doi:10.2307/140629.
14. Tobey R, Wetherell C. The citrus industry and the revolution of corporate capitalism in Southern California, 1887-1944. California History. 1995;74:6–21. doi:10.2307/25177466.

15. White R. The gold rush: Consequences and contingencies. *California history*. 1998;77:42–55. <http://history.msu.edu/hst321/files/2010/07/white.pdf>.
16. Jung MA. Capitalism comes to the diggings: From gold rush adventure to corporate enterprise. *california history*. 1998;77:52–77. doi:10.2307/25462508.
17. Weathers L, Lawton H, Others. History of the citrus research center. *Calif Agric*. 1982;36:0–0. <http://calag.ucanr.edu/archive/?article=ca.v036n11pCR-2>.
18. Sackman DC. *Orange Empire: California and the Fruits of Eden*. University of California Press; 2005. <https://play.google.com/store/books/details?id=Ef8kDQAAQBAJ>.
19. Swingle, T W. The botany of Citrus and its wild relatives in the orange subfamily. *The citrus industry*. 1943;1:128–474. <https://ci.nii.ac.jp/naid/10018851984/>. Accessed 13 Apr 2020.
20. Ramana KVR, Govindarajan VS, Ranganna S, Kefford JF. Citrus fruits — varieties, chemistry, technology, and quality evaluation. Part I: Varieties, production, handling, and storage. *C R C Critical Reviews in Food Science and Nutrition*. 1981;15:353–431. doi:10.1080/10408398109527321.
21. Lim TK. Citrus medica. In: Lim TK, editor. *Edible Medicinal And Non-Medicinal Plants: Volume 4, Fruits*. Dordrecht: Springer Netherlands; 2012. p. 682–9. doi:10.1007/978-94-007-4053-2_84.
22. Gallesio G. *Traité du citrus*. 1811. doi:10.5962/bhl.title.50431.
23. Nicolosi E. Origin and taxonomy. *Citrus genetics, breeding and biotechnology*. :19–43. doi:10.1079/9780851990194.0019.
24. Tolkowsky S. *Hesperides. A history of the culture and use of citrus fruits*. Hesperides A history of the culture and use of citrus. 1938. <https://www.cabdirect.org/cabdirect/abstract/19391601818>.
25. Langgut D. The Citrus Route Revealed: From Southeast Asia into the Mediterranean. *HortScience*. 2017;52:814–22. <https://journals.ashs.org/hortsci/view/journals/hortsci/52/6/article-p814.xml>.
26. Morton J. *Fruits of warm climate* (1987). Published by Julia F. Morton, 20534 SW 92 Ct. Miami, FL. 33189. ISBN: 0-9610184-1-0, Distributed by Creative Resource Systems. Inc Box. 1987;890.
27. Ramon-Laca L. The introduction of cultivated citrus to Europe via Northern Africa and the Iberian Peninsula. *Econ Bot*. 2003;57:502–14. doi:10.1663/0013-0001(2003)057[0502:TIOCCT]2.0.CO;2.
28. Hirai M, Kozaki I, Kajiura I. Isozyme analysis and phylogenic relationship of citrus. *Japanese Journal of Breeding*. 1986;36:377–89. https://www.jstage.jst.go.jp/article/jsbbs1951/36/4/36_4_377/_article/-char/ja/.

29. Saunt J, Others. Citrus varieties of the world. An illustrated guide. Sinclair International Ltd.; 1990. <https://www.cabdirect.org/cabdirect/abstract/19900300000>.
30. Sawamura M, Kuriyama T. Quantitative determination of volatile constituents in the pummelo (*Citrus grandis* Osbeck forma Tosa-buntan). J Agric Food Chem. 1988;36:567–9. doi:10.1021/jf00081a040.
31. Scora RW. On the history and origin of Citrus. Bull Torrey Bot Club. 1975;;369–75. <https://www.jstor.org/stable/2484763>.
32. Mukhopadhyay S. Citrus: production, post harvest, disease and pest management. CRC Press; 2010. <https://play.google.com/store/books/details?id=aUhZDwAAQBAJ>.
33. Pang X-M, Hu C-G, Deng X-X. Phylogenetic relationships within *Citrus* and its related genera as inferred from AFLP markers. Genet Resour Crop Evol. 2007;54:429–36. doi:10.1007/s10722-006-0005-5.
34. Scora RW, Kumamoto J, Soost RK, Nauer EM. Contribution to the origin of the grapefruit, *Citrus paradisi* (Rutaceae). Syst Bot. 1982;7:170–7. doi:10.2307/2418325.
35. Gmitter, F. G. Jr. Origin, evolution, and breeding of the grapefruit. Plant Breed Rev. 1995;13. <http://agris.fao.org/agris-search/search.do?recordID=US9620156>.
36. J. H. H. Citrus growing in South Africa. Nature. 1927;119:776. doi:10.1038/119776c0.
37. Nicolosi E, La Malfa S, El-Otmani M, Negbi M, Goldschmidt EE. The search for the authentic citron (*Citrus medica* L.): historic and genetic analysis. HortScience. 2005;40:1963–8. <https://journals.ashs.org/hortsci/previewpdf/journals/hortsci/40/7/article-p1963.xml>.
38. Schenk W. Der heutige Stand der Affinitätsforschung in der Rebenveredlung. Weinberg u Keller. 1962;9:391–402.
39. Tanaka T, Others. Species problem in Citrus. 1954. <http://agris.fao.org/agris-search/search.do?recordID=US201300690891>.
40. Federici CT, Fang DQ, Scora RW, Roose ML. Phylogenetic relationships within the genus *Citrus* (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. Theor Appl Genet. 1998;96:812–22. doi:10.1007/s001220050807.
41. Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor Appl Genet. 2000;100:1155–66. doi:10.1007/s001220051419.
42. Yamamoto M, Tominaga S. High chromosomal variability of mandarins (*Citrus* spp.) revealed by CMA banding. Euphytica. 2003;129:267. doi:10.1023/A:1022284908424.
43. Barkley NA, Roose ML, Krueger RR, Federici CT. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). Theor Appl Genet. 2006;112:1519–31. doi:10.1007/s00122-006-0255-9.

44. Penjor T, Anai T, Nagano Y, Matsumoto R, Yamamoto M. Phylogenetic relationships of *Citrus* and its relatives based on *rbcl* gene sequences. *Tree Genet Genomes*. 2010;6:931–9. doi:10.1007/s11295-010-0302-1.
45. Uzun A, Yesiloglu T, Aka-Kacar Y, Tuzcu O, Gulsen O. Genetic diversity and relationships within *Citrus* and related genera based on sequence related amplified polymorphism markers (SRAPs). *Sci Hortic* . 2009;121:306–12. doi:10.1016/j.scienta.2009.02.018.
46. Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, et al. Phylogenetic relationships of citrus and its relatives based on *matK* gene sequences. *PLoS One*. 2013;8:e62574. doi:10.1371/journal.pone.0062574.
47. Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Mol Biol Evol*. 2015;32:2015–35. doi:10.1093/molbev/msv082.
48. Dugrand-Judek A, Olry A, Hehn A, Costantino G, Ollitrault P, Froelicher Y, et al. The distribution of coumarins and furanocoumarins in *Citrus* species closely matches *Citrus* phylogeny and reflects the organization of biosynthetic pathways. *PLoS One*. 2015;10:e0142757. doi:10.1371/journal.pone.0142757.
49. Barrett HC, Rhodes AM. A numerical taxonomic study of affinity relationships in cultivated *Citrus* and its Close Relatives. *Syst Bot*. 1976;1:105–36. doi:10.2307/2418763.
50. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000;408:796–815. doi:10.1038/35048692.
51. Khan IA. *Citrus Genetics, Breeding and Biotechnology*. CABI; 2007. <https://market.android.com/details?id=book-KaJ0-7XuQLEC>.
52. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al. The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet*. 2013;45:59–66. doi:10.1038/ng.2472.
53. Aleza P, Juárez J, Hernández M, Pina JA, Ollitrault P, Navarro L. Recovery and characterization of a *Citrus clementina* Hort. ex Tan.'Clemenules' haploid plant selected to establish the reference whole *Citrus* genome sequence. *BMC Plant Biol*. 2009;9:110. <https://bmcpplantbiol.biomedcentral.com/articles/10.1186/1471-2229-9-110>.
54. Wang X, Xu Y, Zhang S, Cao L, Huang Y, Cheng J, et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat Genet*. 2017;49:765–72. doi:10.1038/ng.3839.
55. Shimizu T, Tanizawa Y, Mochizuki T, Nagasaki H, Yoshioka T, Toyoda A, et al. Draft sequencing of the heterozygous diploid genome of Satsuma (*Citrus unshiu* Marc.) using a hybrid assembly approach. *Front Genet*. 2017;8:180. doi:10.3389/fgene.2017.00180.
56. Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet J-P, Garcia-Lor A, et al. Nuclear species-diagnostic SNP markers mined from 454 amplicon sequencing reveal admixture genomic

structure of modern citrus varieties. PLoS One. 2015;10:e0125628. doi:10.1371/journal.pone.0125628.

57. Oueslati A, Salhi-Hannachi A, Luro F, Vignes H, Mournet P, Ollitrault P. Genotyping by sequencing reveals the interspecific *C. maxima* / *C. reticulata* admixture along the genomes of modern citrus varieties of mandarins, tangors, tangelos, orangelos and grapefruits. PLoS One. 2017;12:e0185618. doi:10.1371/journal.pone.0185618.

58. Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P. Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. Ann Bot. 2016;117:565–83. doi:10.1093/aob/mcw005.

59. Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, Gasic K, et al. Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. PLoS One. 2012;7:e35668. doi:10.1371/journal.pone.0035668.

60. Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, et al. Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus × domestica* Borkh). PLoS One. 2014;9:e110377. doi:10.1371/journal.pone.0110377.

61. Chen H, Xie W, He H, Yu H, Chen W, Li J, et al. A high-density SNP genotyping array for rice biology and molecular breeding. Mol Plant. 2014;7:541–53. doi:10.1093/mp/sst135.

62. Unterseer S, Bauer E, Haberer G, Seidel M, Knaak C, Ouzunova M, et al. A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. BMC Genomics. 2014;15:823. doi:10.1186/1471-2164-15-823.

63. Bassil NV, Davis TM, Zhang H, Ficklin S, Mittmann M, Webster T, et al. Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry *Fragaria × ananassa*. BMC Genomics. 2015;16:155. doi:10.1186/s12864-015-1310-1.

64. Martinez-Godoy MA, Mauri N, Juarez J, Marques MC, Santiago J, Forment J, et al. A genome-wide 20 K citrus microarray for gene expression analysis. BMC Genomics. 2008;9:318. doi:10.1186/1471-2164-9-318.

65. Forment J, Gadea J, Huerta L, Abizanda L, Agusti J, Alamar S, et al. Development of a citrus genome-wide EST collection and cDNA microarray as resources for genomic studies. Plant Mol Biol. 2005;57:375–91. doi:10.1007/s11103-004-7926-1.

66. Ollitrault P, Terol J, Garcia-Lor A, Bérard A, Chauveau A, Froelicher Y, et al. SNP mining in *C. clementina* BAC end sequences; transferability in the *Citrus* genus (Rutaceae), phylogenetic inferences and perspectives for genetic mapping. BMC Genomics. 2012;13:13. doi:10.1186/1471-2164-13-13.

67. Fujii H, Shimada T, Nonaka K, Kita M, Kuniga T, Endo T, et al. High-throughput genotyping in citrus accessions using an SNP genotyping array. Tree Genet Genomes. 2013;9:145–53. doi:10.1007/s11295-012-0542-3.

68. Shimizu T, Kaminuma E, Nonaka K, Yoshioka T, Goto S, Matsumoto T, et al. A genomic approach to selecting robust and versatile SNP sets from next-generation sequencing data for genome-wide association study in citrus cultivars. In: III International Symposium on Citrus Biotechnology 1135. 2014. p. 23–32. https://www.actahort.org/books/1135/1135_4.htm.
69. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011;6:e19379. doi:10.1371/journal.pone.0019379.
70. Imai A, Nonaka K, Kuniga T, Yoshioka T, Hayashi T. Genome-wide association mapping of fruit-quality traits using genotyping-by-sequencing approach in citrus landraces, modern cultivars, and breeding lines in Japan. Tree Genet Genomes. 2018;14:24. doi:10.1007/s11295-018-1238-0.
71. Huang M, Roose ML, Yu Q, Du D, Yu Y, Zhang Y, et al. Construction of high-density genetic maps and detection of QTLs associated with Huanglongbing tolerance in *Citrus*. Front Plant Sci. 2018;9:1694. doi:10.3389/fpls.2018.01694.
72. Nogler GA. Gametophytic Apomixis. In: Johri BM, editor. Embryology of angiosperms. Berlin, Heidelberg: Springer Berlin Heidelberg; 1984. p. 475–518. doi:10.1007/978-3-642-69302-1_10.
73. Koltunow AM, Johnson SD, Bicknell RA. Apomixis is not developmentally conserved in related, genetically characterized *Hieracium* plants of varying ploidy. Sex Plant Reprod. 2000;12:253–66. doi:10.1007/s004970050193.
74. Yang WC, Sundaresan V. Genetics of gametophyte biogenesis in *Arabidopsis*. Curr Opin Plant Biol. 2000;3:53–7. doi:10.1016/s1369-5266(99)00037-0.
75. Drews GN, Lee D, Christensen CA. Genetic analysis of female gametophyte development and function. Plant Cell. 1998;10:5–17. doi:10.1105/tpc.10.1.5.
76. Asker S. Gametophytic apomixis: elements and genetic regulation. Hereditas. 1980;93:277–93. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1601-5223.1980.tb01367.x>.
77. Talent N. Evolution of gametophytic apomixis in flowering plants: an alternative model from Maloid Rosaceae. Theory Biosci. 2009;128:121–38. doi:10.1007/s12064-009-0061-4.
78. Bicknell RA, Koltunow AM. Understanding apomixis: recent advances and remaining conundrums. Plant Cell. 2004;16 Suppl:S228–45. doi:10.1105/tpc.017921.
79. Spillane C, Steimer A, Grossniklaus U. Apomixis in agriculture: the quest for clonal seeds. Sex Plant Reprod. 2001;14:179–87. doi:10.1007/s00497-001-0117-1.
80. Kobayashi S, Ikeda I, Nakatani M. Studies on Nucellar Embryogenesis in Citrus. Journal of the Japanese Society for Horticultural Science. 1979;48:179–85. doi:10.2503/jjshs.48.179.
81. Mukherjee SK. Origin of mango (*Mangifera indica*). Economic Botany. 1972;26:260–4. doi:10.1007/bf02861039.

82. Nakano M, Shimada T, Endo T, Fujii H, Nesumi H, Kita M, et al. Characterization of genomic sequence showing strong association with polyembryony among diverse *Citrus* species and cultivars, and its synteny with *Vitis* and *Populus*. *Plant Sci.* 2012;183:131–42. doi:10.1016/j.plantsci.2011.08.002.
83. Moore GA, Castle WS. Morphological and isozymic analysis of open-pollinated citrus rootstock populations. *J Hered.* 1988;79:59–63. doi:10.1093/oxfordjournals.jhered.a110448.
84. Ruiz C, Paz Breto M, Asíns MJ. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. *Euphytica.* 2000;112:89–94. doi:10.1023/A:1003992719598.
85. Xiang C, Roose ML. Frequency and characteristics of nucellar and zygotic seedlings in 12 citrus rootstocks. *Sci Hortic.* 1988;37:47–59. doi:10.1016/0304-4238(88)90150-1.
86. Ueno, I. Embryo number of various varieties of *Citrus* and its relatives. *Bull Hort Res Sta Japan*, ser B. 1967;7:11–22. <https://ci.nii.ac.jp/naid/10024030075/>. Accessed 19 Apr 2020.
87. Zhang S, Liang M, Wang N, Xu Q, Deng X, Chai L. Reproduction in woody perennial *Citrus*: an update on nucellar embryony and self-incompatibility. *Plant Reprod.* 2018;31:43–57. doi:10.1007/s00497-018-0327-4.
88. Nakano M, Shimizu T, Kuniga T, Nesumi H, Omura M. Mapping and haplotyping of the flanking region of the polyembryony locus in *Citrus unshiu* Marcow. *J Japan Soc Hort Sci.* 2008;77:109–14. https://www.jstage.jst.go.jp/article/jjshs1/77/2/77_2_109/_article/-char/ja/.
89. Koltunow AM. Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. *Plant Cell.* 1993;5:1425–37. doi:10.1105/tpc.5.10.1425.
90. Koltunow AM, Soltys K, Nito N, McClure S. Anther, ovule, seed, and nucellar embryo development in *Citrus sinensis* cv. Valencia. *Canadian Journal of Botany.* 1995;73:1567–82. doi:10.1139/b95-170.
91. Cameron JW, Frost HB. Genetics, breeding, and nucellar embryony. *The Citrus Industry* (eds W. Reuther, HJ Webber & LD Batchelor), vol. II. 1968; 325-370. Univ. Calif. Division Agric. Sci.
92. Cameron JW, Soost RK. Sexual and nucellar embryony in F1 hybrids and advanced crosses of *Citrus* with *Poncirus*. *J Amer Soc Hort Sci.* 1979. 104:408-410. <https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCALAGROLINEINRA8010051248>.
93. Webber H, Others. Variations in citrus seedlings and their relation to rootstock selection. *Hilgardia.* 1932;7:1–79. <http://hilgardia.ucanr.edu/Abstract/?a=hilg.v07n01p001>.
94. Anderson CM, Castle WS, Moore GA. Isozymic identification of zygotic seedlings in Swingle citrumelo *Citrus paradisi* × *Poncirus trifoliata* nursery and field populations. *Journal of the American Society for Horticultural Science.* 1991;116:322–6. doi:10.21273/jashs.116.2.322.

95. Parlevliet JE, Cameron JW. Evidence on the inheritance of nucellar embryony in *Citrus*. Proc Am Soc Hortic Sci. 1959;74:252–60. <https://ci.nii.ac.jp/naid/10020830116/>. Accessed 20 Apr 2020.
96. Iwamasa, M. Inheritance of nucellar embryony in citrus. Bull Hort Res Stn. 1967;7:1–10. <https://ci.nii.ac.jp/naid/10006289656/>. Accessed 20 Apr 2020.
97. Hong Q-B, Xiang S-Q, Chen K-L, Chen L-G. Two complementary dominant genes controlling apomixis in genus *Citrus* and *Poncirus*. Yi chuan xue bao= Acta genetica Sinica. 2001;28:1062–7. <https://europepmc.org/abstract/med/11725641>.
98. García R, Asíns MJ, Forner J, Carbonell EA. Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. Theor Appl Genet. 1999;99:511–8. doi:10.1007/s001220051264.
99. Kepiro JL, Roose ML. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*. Tree Genet Genomes. 2010;6:1–11. doi:10.1007/s11295-009-0223-z.
100. Raga V, Bernet GP, Carbonell EA, Asins MJ. Segregation and linkage analyses in two complex populations derived from the citrus rootstock Cleopatra mandarin. Inheritance of seed reproductive traits. Tree Genet Genomes. 2012;8:1061–71. doi:10.1007/s11295-012-0486-7.
101. Nakano M, Shimizu T, Fujii H, Shimada T, Endo T, Nesumi H, et al. Marker enrichment and construction of haplotype-specific BAC contigs for the polyembryony genomic region in *Citrus*. Breed Sci. 2008;58:375–83. https://www.jstage.jst.go.jp/article/jsbbs/58/4/58_4_375/_article/-char/ja/.
102. Kumar V, Malik SK, Pal D, Srinivasan R, Bhat SR. Comparative transcriptome analysis of ovules reveals stress related genes associated with nucellar polyembryony in citrus. Tree Genet Genomes. 2014;10:449–64. doi:10.1007/s11295-013-0690-0.
103. Nakano M, Kigoshi K, Shimizu T, Endo T, Shimada T, Fujii H, et al. Characterization of genes associated with polyembryony and in vitro somatic embryogenesis in *Citrus*. Tree Genet Genomes. 2013;9:795–803. doi:10.1007/s11295-013-0598-8.
104. Wu XM, Kou SJ, Liu YL, Fang YN, Xu Q. Genomewide analysis of small RNA s in nonembryogenic and embryogenic tissues of citrus: micro RNA-and si RNA-mediated transcript cleavage involved in somatic embryogenesis. Plant Biotechnol. 2015. <https://onlinelibrary.wiley.com/doi/abs/10.1111/pbi.12317>.
105. Long J-M, Liu Z, Wu X-M, Fang Y-N, Jia H-H, Xie Z-Z, et al. Genome-scale mRNA and small RNA transcriptomic insights into initiation of citrus apomixis. J Exp Bot. 2016;67:5743–56. doi:10.1093/jxb/erw338.
106. Shimada T, Endo T, Fujii H, Nakano M, Sugiyama A, Daido G, et al. MITE insertion-dependent expression of CitRKD1 with a RWP-RK domain regulates somatic embryogenesis in citrus nucellar tissues. BMC Plant Biol. 2018;18:166. doi:10.1186/s12870-018-1369-3.

107. Viquez-Zamora M, Vosman B, van de Geest H, Bovy A, Visser RGF, Finkers R, et al. Tomato breeding in the genomics era: insights from a SNP array. *BMC Genomics*. 2013;14:354. doi:10.1186/1471-2164-14-354.
108. Yu H, Xie W, Li J, Zhou F, Zhang Q. A whole-genome SNP array (RICE6K) for genomic breeding in rice. *Plant Biotechnol J*. 2014;12:28–37. doi:10.1111/pbi.12113.

Table 1. Statistics for the sequenced *Citrus* genomes

Genome	Ploidy	Platform	Total size of assemblyed scaffolds (Mb)	Number of genes	Mean gene length (kb)	Number of exons	Reference
<i>C. clementina</i> v 1.0	haploid	Sanger	301	24,533	2.5	5.3	Wu et al. 2014
<i>C. sinensis</i>	double haploid	Illumina GA11	301	29,455	1.3	5.8	Xu et al. 2013
<i>C. sinensis</i> v 1.1	diploid	454/Sanger	319	25,379	2.1	4.9	Wu et al. 2014
<i>C. grandis</i> (<i>C. maxima</i>)	haploid	PACBi RSII/Illumina	302	30,123	1.1	NA	Wang et al. 2017
<i>C. medica</i>	diploid	PACBi RSII/Illumina	404	32,579	1.8	NA	Wang et al. 2017
<i>C. ichangensis</i>	diploid	PACBi RSII/Illumina	357	32,067	1.6	NA	Wang et al. 2017
<i>Atalantia buxifolia</i>	diploid	PACBi RSII/Illumina	316	28,420	2.2	NA	Wang et al. 2017
<i>C. unshiu</i>	diploid	Illumina HiSeq 2000/PacBio RSII	360	29,024	1.2	4.1	Shimizu et al. 2017

Table.2. List of known polyembryonic and monoembryonic cultivars

Polyembryonic	Monoembryonic
sweet orange (<i>C. sinensis</i>)	clementine (<i>C. clementina</i>)
grapefruits (<i>C. paradisi</i>)	Kinokuni mandarin (<i>C. kinokuni</i>)
lemon (<i>C. limon</i>)	citron (<i>C. medica</i>)
Satsuma (<i>C. unshiu</i>)	pummelo (<i>C. grandis</i>)
Ponkan (<i>C. reticulata</i>)	<i>C. mangshanensis</i>
some mandarins (<i>C. reticulata</i>)	Papeda
most trifoliate orange (<i>P. trifoliata</i>)	

Table 3. Summary of citrus diversity study

Study	Marker type	Number of markers	Types analyzed	Number of accessions
Barkley et al. 2006	SSR	24	<i>Citrus</i>	370
Uzun et al. 2009	SRAP	21	Relatives <i>Citrus</i>	86
Baig et al. 2009	RAPD	25	Relatives <i>Citrus</i> <i>Poncirus hybrid with Citrus</i>	18
Amar et al. 2011	SSR SRAP CAPS-SNP	61 33 24	<i>Citrus</i> Relatives	24
Uzun et al. 2011	SSR SRAP	13 21	Lemon Citron Rough lemon Citron	56
Biswas et al. 2011	AFLP	11	<i>Citrus</i>	34
	SSAP SAMPL SSR	8 7 7		
Luro et al. 2012	SSR Chloroplast SSR	22 6		24
Garcia-Lor et al. 2011	SSR	50	<i>Citrus</i>	90
Ollitrault et al. 2012	Indel SNP	12 506	<i>Citrus</i>	54
Queslati et al. 2017	SNP (GBS)	30,289	<i>Citrus</i>	55
Wu et al. 2018	Indel (GBS) Whole genome sequence	8,794	<i>Citrus</i>	60
Ahmed et al. 2019	SNP (GBS)	43,598	Relatives <i>Citrus</i>	48
This study	SNPs (array)	51,296	<i>Citrus</i> Relative	924

CHAPTER 1

Development and Assessment of Axiom® Citrus 15AX and Axiom® Citrus 56AX

Background

Cultivated *Citrus* is one of the most important fruit crops in the world. Rapid advancements in technologies provide various tools to analyze fruit crop genomes to better understand genetic diversity, relationships, and to aid breeding. One such tool is a genome-wide single nucleotide polymorphism (SNP) genotyping array which offers highly multiplexed assays at a relatively low cost per data point. Given the complex ancestry of *Citrus*, such technology can provide a powerful and high throughput tool for *Citrus* genomic studies and significantly reduce the time breeders take to screen and characterize germplasm and to identify the genes responsible for traits of interest.

Results

I report development and validation of 1.4M SNP Axiom® Citrus HD Genotyping Array (Citrus 15AX 1 and Citrus 15AX 2) and 58K SNP Axiom® Citrus Genotyping Arrays for *Citrus* and close relatives. SNPs represented were chosen from a citrus variant discovery panel consisting of 41 diverse whole-genome re-sequenced accessions of *Citrus* and close relatives, including nine progenitor citrus species. They are three Australian species (*Microcitrus australasica*, *M. australis* and *Eremocitrus glauca*) and six Asian species (*C. reticulata*, *C. maxima*, *C. medica*, *C. micrantha*/*C. hystrix*, *Fortunella*, and *Poncirus trifoliata*) SNPs chosen mainly target putative genic regions of the genome, and are accurately called in both *Citrus* and its closely related genera, while providing good genome coverage. The Axiom® Citrus HD Genotyping Array HD Genotyping Array and Axiom® Citrus Genotyping Array include 1002 and 500 chloroplast SNPs,

respectively. These arrays were used to genotype and analyze diversity in 871 accessions from a diverse germplasm collection. Relationships among variant discovery panel accessions inferred with the array were similar to those inferred from sequence data. Reproducibility of the arrays was nearly 100% with the vast majority of SNPs classified as the most stringent class of markers, “*Poly High Resolution*” (PHR) polymorphisms.

Conclusion

Both the Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array are publicly available and provide citrus geneticists with high throughput genetic marker screening systems to perform a range of high resolution genomic studies such as genome-wide association studies (GWAS), genomic selection (GS), germplasm management and other breeding applications.

Keywords

Citrus, Genotyping array, Plant breeding, Germplasm, Single nucleotide polymorphism

Introduction

Citrus, a member of family Rutaceae, is one of the most widely cultivated and economically valued fruit tree crops in the world. It is cultivated in more than 100 countries [1] and its production is estimated at 124 million tons [2], mostly in a belt that spans Mediterranean and subtropical climatic regions of the world [1]. Citrus fruits are rich in various functional compounds such as vitamins, limonoids, and antioxidants, and are nutritionally beneficial for human health [3, 4]. The exact

lineage from ancestral *Citrus* to many of the modern cultivated *Citrus* types has not been resolved, however, domestication is thought to have started in Southeast Asia and dispersed to the rest of the world [5] leading to interspecific hybridization and introgression of ancestral progenitors into cultivated *Citrus* varieties.

The arrival of the genomic era revealed that domestication-related selection by humans has tremendously shaped modern crop genomes as reported in apple, grape, peach, tomato, soy bean, rice and maize [6–18]. *Citrus* has been domesticated and cultivated for several thousand years [5] and modern varieties are thought to have resulted from similar alteration of the genome composition guided by human selection. Recently, the admixed genome of cultivated sweet orange (*C. sinensis*) that originated from ancestral progenitors, mandarin (*C. reticulata* Blanco) and pummelo (*C. maxima*), has been reported [19, 20]. In addition, it was revealed that some cultivated mandarins (*C. reticulata*) are admixed with pummelo from repeated introgression [20–22], contrary to previously accepted views on the ancestry of these cultivars. Many species of *Citrus* are self-incompatible and hybridize sexually interspecifically and even intergenerically occasionally [23], resulting in highly heterozygous genomes [19]. Further complicating genomic studies of *Citrus* are the many clonal varieties derived by mutations such as accumulated somatic mutations, sport mutations or nucellar variants [24, 25]. In fact, most commercially important citrus species, often referred to as secondary species, such as sweet orange, many mandarins, lemon (*C. limon*), lime (*C. aurantifolia*) and grapefruit (*C. paradisi*), are either natural or artificial interspecific hybrid species derived by hybridization among four major ancestral *Citrus* taxa, citron (*C. medica*), pummelo, mandarin, and papeda (*C. micrantha*) which then further diversify by mutation and, in some cases hybridization [26].

Recent advances in sequencing and genomic technologies enabled the construction of reference genomes of *C. sinensis* [19], *C. clementina* (or *C. reticulata*) [20] and *C. maxima* [27],

facilitating high-resolution genomic studies and characterization of germplasm. *Citrus* is generally a diploid and highly heterozygous, with a chromosome number of $2n=18$ and an estimated genome size of 360 Mb [19]. There are approximately 29,000 coding genes with average sequence length of 1.3 kb and 5.8 exons [19]. However, availability and utilization of genomic resources for citrus have lagged behind most of the major crops due to lack of high throughput genomic tools available and appropriate to analyze complex citrus populations.

Due to high heterozygosity and long juvenility, citrus production relies on clonal vegetative propagation to ensure uniform propagation of the desired genotype. Nearly all citrus is propagated by grafting onto seedling rootstocks which are fairly uniform because rootstock varieties are selected to have nucellar embryony, a type of apomixis which results in seedlings identical the seed parent. Therefore, citrus breeders must consider two separate strategies in breeding new varieties of rootstocks and scions. Furthermore, it is very difficult to develop inbred lines in most citrus because of self-incompatibility and severe inbreeding depression [25]. The long generation time and large space requirement of citrus make conventional genetic studies and breeding very costly and laborious. Typically, a citrus breeding cycle takes 5-10 years and further selection processes average more than 30 years until the release of a new elite variety [23]. However, in the genomic era, the available bioinformatics and genomic tools can assist citrus breeders to understand the genetic mechanisms underlying horticultural traits and significantly accelerate release of new varieties

Understanding the relationships between genomic variations and phenotypic variation associated with various horticultural traits, is the key for breeding of novel and improved citrus varieties. Molecular markers such as Simple Sequence Repeats (SSRs) and Amplified Fragment Length Polymorphism (AFLP) have previously been popular marker systems often utilized in citrus breeding and diversity studies for their large information content [28–31]. However, some previous

marker systems were sequence anonymous, low-throughput, or did not provide enough density across the genome for high resolution genomic studies in multiple populations. In recent years, high density and genome-wide SNP markers have been easily mined from sequence data, and have served as genetic markers for GWAS [32], population structure analysis [33], admixture analysis [34] in various populations, linkage disequilibrium (LD) analysis [32] and haplotype construction [20], as well as structural variant (SV)/copy number variation (CNV) and Loss of Heterozygosity (LOH) detection [35] in mutationally derived populations.

High throughput genotyping technologies, such as SNP genotyping arrays that have emerged in the recent years, have revolutionized the way we can analyze the origin and genetic control of traits in crops. These tools are now available to crop breeders and produce high fidelity data to perform high resolution genomic studies as demonstrated in apple, maize, rice, peach and strawberry [36–40]. The development of SNP genotyping arrays requires a tremendous amount of knowledge about the organism under investigation [41] to maximize the opportunities for SNP detection, and to ensure that selected SNPs will be suitable for the final application of the arrays. Despite these prerequisites and the large initial investment, an array-based genotyping approach is often a more efficient alternative to the whole genome sequencing approach [42, 43] and the Genotyping by Sequencing (GBS) approach in terms of computational requirements necessary for downstream analysis for non-model organisms such as *Citrus*. GBS utilizes restriction enzyme technologies to generate a representation (subset) of the genome that is then sequenced [44]. However, imputing missing calls can be computationally challenging and demands well established complex pipelines to analyze a highly heterozygous citrus genome [22]. Therefore, an array based genotyping approach could be much more high-throughput and cost-effective once it is developed.

We selected 1.4 million SNPs for the design of a larger validation array, Affymetrix™ Axiom® Citrus HD Genotyping Array and about 58,000 SNPs for the smaller and more economical array Affymetrix™ Axiom® Citrus Genotyping Array. Diverse citrus germplasm is available from the University of California Riverside Citrus Variety Collection (CVC), and contains diversity relevant to disease resistance, plant architecture, fruit characteristics, and adaptation to various environmental conditions. Accessions from the CVC were hybridized to these two arrays. Data obtained from the two arrays yielded an extensive genomic inventory of the CVC.

Development of genome-wide high density SNP genotyping arrays for citrus will provide citrus breeders with tools to rapidly characterize citrus germplasm and locate genes responsible for traits under investigation with much higher resolution. They are useful, not only for the maintenance of the germplasm, but also to facilitate quicker release of novel citrus varieties.

We demonstrated the potential applications and utility of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array by performing population structure analysis and phylogenetic analysis in diverse citrus germplasm and anticipate additional downstream analyses such as linkage mapping, GWAS, diversity analysis, CNV and LOH analysis.

Design of genome-wide SNP genotyping arrays Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

Designing the SNP genotyping arrays for citrus requires strategizing to optimize the array composition for its intended uses. The genus *Citrus* and cross-compatible taxa considered to be its close relatives need to be taken into consideration, in order for the arrays to be applicable for a wide range of genomic studies and breeding programs. For example, some relatives such as *Poncirus trifoliata* (L.) Raf. are important parents of citrus rootstocks. Design criteria include 1) ability to

call SNPs accurately in the diverse accessions of *Citrus* and close relatives that sexually hybridize with *Citrus*, 2) sufficient nuclear genome coverage, and 3) inclusion of chloroplast SNP markers. SNP marker development for genotyping arrays involves several steps: 1) selection of genotypes to be included in the citrus variant discovery panel, 2) variant discovery, 3) variant validation, and 4) final selection of variants to be represented on the array (Fig. 1). Composition of the citrus variant discovery panel defines the level and phylogenetic distribution of ascertainment bias, and the range of array utility for subsequent studies. Therefore, this step requires careful analysis of existing information. Once the citrus variant discovery panel is defined, selected individuals in the discovery panel are re-sequenced, followed by variant discovery and further filtering of the identified variants. Identified polymorphic variants should be then validated and screened using an informative set of individuals. This process also assesses and identifies the range of diversity over which each SNP variant can be genotyped with confidence. Finally, informative SNPs that meet the criteria specified in the design strategy are selected to be represented on the arrays. An Axiom® Citrus HD Genotyping Array was used as a validation array to assess performance of the SNPs represented on it and overall performance of the array itself. Data from the Axiom® Citrus HD Genotyping Array was subsequently used to select a set of SNPs to be represented on an Axiom® Citrus Genotyping Array.

Methods

Sample preparation for array hybridization

Leaf samples (young flush) of the accessions of *Citrus* and its relatives were collected in the field from the CVC (Riverside, CA) from 2014-2016 and placed on ice. Each tree from which a leaf sample was collected was tagged with a unique identifier. Collected leaf samples were then returned

to the laboratory on ice and the leaf surface was quickly and gently wiped with chloroform to remove wax prior to drying with silica packs in a sealed, labeled sample bag. Silica dried leaf samples were homogenized with FastPrep-24™ (MP Biomedicals, Santa Ana, CA) with ¼” ceramic spheres or ¼” stainless steel beads. High molecular weight genomic DNA was then isolated from homogenized dried leaf samples with a DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). The quality of DNA was assessed by electrophoresis, 260/280 and 260/230 ratios with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) and genomic DNA was quantified with a Qubit fluorometric system (Thermo Fisher Scientific, Wilmington, USA) according to manufacturer’s instructions. Genomic DNA of concentration of 25 ng/µl in 50 µl and 30 ng/µl in 25 µl for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array, respectively were prepared in Affymetrix™ bar-coded 96 well plates.

Variant selection according to *in-silico* analysis of re-sequencing data

The citrus variant discovery panel was composed of 30 accessions that were whole genome re-sequenced using an Illumina® HighSeq 2500 at the UCR Institute for Integrative Genome Biology (IIGB, Riverside, CA, USA) genomics core, and whole genome sequence data of 11 accessions taken from publicly available databases and from a cooperating laboratory (National Institute of Genetics, Japan). For whole genome sequencing, young leaf tissue was sampled from trees growing in the UCR Citrus Variety Collection, now called Givaudan Citrus Variety Collection (Riverside, CA, USA), and DNA was isolated from tissue ground with liquid nitrogen in a mortar and pestle using Qiagen DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). Sequencing libraries were prepared at IIGB using a NEBNext Ultra DNA library prep kit for Illumina (NEB), and sequenced (100 bp paired-end reads) with a target depth of at least at 30X. Variant sites were mined

using a variant discovery pipeline at the DOE Joint Genome Institute (JGI) using Phytozome *C. clementina* v1.0 reference genome (<http://www.phytozome.org/citrus/>) for nuclear genome variant discovery and by mapping reads to the sweet orange chloroplast genome [45] chloroplast variant discovery. The Illumina® paired end reads were mapped to the haploid clementine reference sequence using bwa/mem, followed by removal of duplicated reads using Picard (<http://picard.sourceforge.net/>). The Genome Analysis Toolkit (GATK) HaplotypeCaller (<https://software.broadinstitute.org/gatk/>) was used for SNP calls with the following filters:

1. map quality score ≥ 25
2. base quality score ≥ 30
3. sample genotype quality score ≥ 20
4. read depth filter for each sample: $10 \leq DP < 2 \times \text{median}$
5. allele balance filter for heterozygous SNPs: alt allele frequency from reads of a sample does not reside in the tails of a binomial distribution accounting for 5% of total area.
6. the final variant set was restricted to sites with no more than 2 alleles and fewer than half of the samples having missing genotypes.

This analysis identified 19.9M (19,907,622) autosomal variants and 2,216 chloroplast variants from the whole genome sequence data of 41 diverse accessions of *Citrus* species (33 accessions, groups 1 and 2) and close relatives (8 accessions, group 3) in the citrus variant discovery panel (Table 1, Table 2). The raw sequence variants were filtered according to the quality and their support by sequence reads. In the 19.9M autosomal variants mined in citrus variant discovery panel, the observed heterozygous calls ranged from 1.84% in Ethrog citron to 19.37% in Schaub rough lemon (See Supplementary Table 1, Additional File 1).

Variant Filtering Part 1

Variant Call Format (VCF) files of 41 accessions in the citrus variant discovery panel were generated from the Illumina® short read data and grouped into Group 1 (Citrus, 29 accessions), Group 2 (Citron, 4 accessions), and Group 3 (Citrus relatives, 8 accessions) based on the sequence divergence between groups, calculated prior to SNP filtering (data not shown). SNPs were selected according to the phylogenetic group or groups in which they were polymorphic to enable selection of a more appropriate number of SNPs for analysis of each group. Without some limitation of this type, most SNPs on the array might be useful for differentiating the divergent groups but not useful for analysis within the core group of *Citrus* accessions in which most breeding is done. SNP filtering was then performed using SnpEff (snpeff.sourceforge.net) and SnpSift (snpeff.sourceforge.net/SnpSift.htm), mainly targeting putative genic regions, which included exons, introns, 5' and 3' UTRs, 5 kb upstream, and downstream regions of a putative gene. This strategy was expected to select variants more likely to affect gene functions and having flanking sequences relatively conserved among the target taxa to reduce miscalls due to multiple SNPs. 6,931,772 autosomal variants resulted from further SNP filtering based on the alternate allele frequency (> 0.1) in group 1, 2 and 3, separately. In addition, A/T or G/C transversion variants were removed from these filtered autosomal variants as recommended by the Affymetrix™ bioinformatics team, resulting in 5,320,680 autosomal variants including 639,833 InDels (See Supplementary Figure 1, Additional File 1). Chloroplast variants were not filtered at this stage and all were added to the autosomal variants. Finally, the 5,322,896 candidate filtered variants were submitted to Affymetrix™ bioinformatics team for *in-silico* analysis to predict variant to probe conversion performance for the array. Affymetrix™ *in-silico* analysis included following filters:

1. duplicate count to assess the number of instances where 24 bases on each side of the SNP match other sequences in the list

2. poly count to assess the number of instances where 24 bases on each side of SNP match other sequences in the list but the alleles differ
3. forward/reverse P-convert to assess the expected probability of conversion based on the forward strand of the sequences
4. forward/reverse repetitive to assess whether or not the 16-mer hit count exceeds threshold
5. forward/reverse wobble count to assess the number of interfering polymorphisms within 24 bases
6. forward/reverse wobble distance to assess the distance to the nearest interfering the polymorphism
7. forward/reverse buildable to assess whether or not a probe can be built from the forward strand

A marker/forward/reverse is recommended if

1. P-convert > 0.6
2. No wobbles
3. Poly count =0

A marker/forward/reverse is not recommended if

1. Duplicate count > 0
2. P-convert < 0.4
3. Wobble distance < 21
4. Wobble count >= 3

A marker/forward/reverse is not possible

1. If a probe cannot be built to interrogate the SNP in that direction

All other markers are considered neutral

***in-silico* analysis of variants**

Upon Affymetrix™ *in-silico* analysis of 5.3M (5,320,680 autosomal and 2,215 chloroplast) candidate variants, 436,431 variants were recommended for tiling on both strands, 1,599,832 variants for tiling only on one strand, 3,285,540 neutral or not recommended variants that could be tiled on at least one strand, and 1,092 variants that couldn't be tiled because the probes for these loci could not be created. Overall 2,036,263 variants (38.25%) were recommended for tiling on at least one strand. The 639,833 InDels failed to pass Affymetrix™ *in-silico* analysis and were excluded from further filtering at this stage. In order to perform further filtering of the candidate variants with maximal efficiency, we focused on filtering from 2,036,263 variants that were recommended for tiling on at least one strand. Further filtering was performed based on P-conver value. If both forward and reverse strand tiling probes had same P-conver value, the variants were accepted to be tiled on both strands. If both forward and reverse strand tiling probes were recommended but P-conver values differed, the strand with the higher P-conver value was selected if this value was above 0.818. Finally, P-conver values above 0.69 was used to accept the rest of the variants to be tiled on either forward or reverse strand. No further filtering was performed on chloroplast variants and 1,002 Affymetrix™ recommended chloroplast variants were accepted based on the *in-silico* analysis. Affymetrix™ Axiom® genotyping array technology uses probes of 70 bases to query the targeted SNP and probes on the Axiom® Citrus HD Genotyping Array were designed to match the *C. clementina* v1.0 reference genome sequence except at the targeted SNP position base.

SNP Validation Panel

To assess if selection of a robust set of variants that work in both *Citrus* and its near relatives was achieved, a total of 277 genomic DNA samples from 232 accessions (Supplementary Table 2, Additional File 2) were selected from CVC as a validation panel and screened with Axiom® Citrus HD Genotyping Array. This process allowed performance assessment of the represented SNPs and the array, but also was essential in screening for additional potential SNP filtering parameters. In addition, we assessed the range of phylogenetic divergence over which the selected SNP probes were reliable. Trios (7 families) (See Supplementary Table 3a, Additional File 1), technical and clonal duplicate samples were included in the validation panel. SNPs and array performance were investigated with respect to 1) SNP genotyping call rates, cluster separation, and reproducibility, 2) Level of heterozygosity in the validation panel, and 3) Mendelian inheritance from parents to offspring in trios (See Supplementary Table 4a, Additional File 1)

Variant Filtering Part 2

To design an array that accurately genotypes SNPs in *Citrus* and its relatives, we performed variant filtering part 2 to enhance genotyping success in the close relatives. Very distant relative, polyploid and synthetic polyploid (mixed DNA) samples were removed from the set of 277 genomics DNA samples for this additional further analysis because the Axiom® Analysis Suite software calls SNPs only in diploids, and genotyping distant relatives was a low-priority objective.

1. SNP genotyping calls for 254 genomic DNA samples analyzed with Axiom® Citrus HD Genotyping Array were made using Axiom® Analysis Suite 1.1.1.66 by Affymetrix™ at Axiom® BestPractice workflow default setting (default DQC, QC, and % passing) to identify “*Poly High Resolution*” (PHR) probes. PHR probes are the most informative and accurate probe performance

class on the Axiom® platform based on the genotype call cluster separation, cluster variance and cluster position.

2. SNP genotyping calls of 254 genomic DNA samples were made for PHR probes loci identified in analysis 1 using Axiom® Analysis Suite at lowered DQC, QC, and % passing thresholds (=0) to allow Axiom® Analysis Suite to analyze all the samples including *Citrus* relatives which otherwise failed the analysis due to their divergence from the majority of the samples being analyzed. DQC and QC probes are used by Affymetrix™ Axiom® platform as quality controls. DQC probes are 31-mer sequences that are non-polymorphic in the variant discovery panel and QC probes are the best performing probes selected by Affymetrix™ *in silico* analysis. Among the 20,000 QC probes, the numbers of probes polymorphic in various groups were 15,539 in Group 1 only, 786 in Group 2 only, 120 in Group 3 only, 1,181 in Groups 1 and 2, 487 in Groups 1 and 3, 78 in Groups 2 and 3, and 1,809 polymorphic in all three Groups (See Supplementary Figure 2, Additional File 1).

3. R package SNPolisher (SNPolisher, v1.5.2, Affymetrix™) was used to compare the SNP genotyping calls between analysis 1 and analysis 2 at PHR probe loci in each individual analyzed. We then selected loci for which calls did not differ between analysis 1 and analysis 2. Analysis 1 identified 728,831 PHR probe loci that accurately genotyped SNPs in *Citrus* and analysis 2 identified 369,269 PHR probe loci that accurately genotyped SNPs in both *Citrus* and its relatives. In order to select high-confidence variants to be represented on the smaller and more economical array Axiom® *Citrus* Genotyping Array, we selected the best performing probes from 369,269 PHR probe loci identified in variant filtering parts 1 and 2 based on 1) the best performing probe from each pair for which probes were tiled on both strands, 2) concordance between *in silico* variant calls from the sequencing data of the variant discovery panel and SNP genotyping calls with validation array Axiom® *Citrus* HD Genotyping Array. We then randomly selected 58,433 variants

from the list and assessed distribution of the selected 58,433 SNPs along the genome and the number of heterozygous variants in selected genotypes. Finally, a total of 58,433 SNPs (57,933 autosomal SNPs, and 500 chloroplast SNPs) were chosen to be represented on Axiom® Citrus Genotyping Array.

PCA

Principal component analysis (PCA) was carried out with accessions from the CVC genotyped with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. A total of 196 accessions analyzed with the Axiom® Citrus HD Genotyping Array including most members of the variant discovery panel (Table 3) and 871 CVC accessions analyzed with either the Axiom® Citrus Genotyping Array or only the Axiom® Citrus HD Genotyping Array (193 accessions) were included in the analysis after removal of clonal and technical duplicate samples (Table 3). Only the genotyping calls at PHR with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array were included for the PCA analysis with the software TASSEL v5.2.31 (<http://www.maizegenetics.net/tassel>).

Phylogenetic Inference

Phylogenetic analyses of the citrus variant discovery panel and CVC accessions representing major *Citrus* taxa and close relatives were performed. Genotyping calls of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array and sequence data of 26 selected accessions were used to infer phylogenetic relationships among accessions in the citrus variant discovery panel. The 26 accessions were selected based on samples passing at default setting with

Axiom® Analysis Suite and no more than 20% missing calls with the sequence data. Genotyping calls of Axiom® Citrus HD Genotyping Array, and Axiom® Citrus Genotyping Array were filtered for PHR. In addition, PHR loci from Axiom® Citrus Genotyping Array were pulled from the sequence data and used for a phylogenetic analysis. The resulting phylogenetic trees generated with array data and sequence data were then compared to each other topologically with phylo.io [46]. For phylogenetic analysis of non-hybrid and non-admixed CVC accessions, 87 accessions representing major citrus taxa and close relatives were selected. The SNP genotyping dataset was filtered for PHR and all 87 accessions which had less than 10 percent heterozygosity were selected (a threshold supported by preliminary Admixture analysis, not shown). The 33 accessions belonging to the important *Citrus* relative, *P. trifoliata*, were grouped into three major clusters with little genetic variation within each cluster, based on preliminary phylogenetic analysis using 871 CVC accessions (data not shown). Only one individual representing each of the three *Poncirus* clusters was included in the analysis to reduce redundancy. The accession of *C. micrantha* var. *microcarpa* failed the initial analysis with Axiom® Analysis Suite when analyzed with default settings. Therefore, the QC call rate threshold was reduced from 97% to 95% in Axiom® Analysis Suite in order for *C. micrantha* var. *microcarpa* to be included in the phylogenetic analysis. The pipeline SNPhylo [47] was used to perform phylogenetic analysis with *Poncirus* as an outgroup genus. SNPhylo first concatenated and aligned the SNPs, then performed a maximum likelihood distance estimation without molecular clock and bootstrapping analysis [47]. For SNPhylo the parameter settings were the minor allele frequency filter was set to 0, the maximum missing data filter was set to 1, and the LD filter was set to 1 to remove redundant SNPs by SNPrelate [48]. The maximum likelihood tree by DNAML [49] with bootstrap analysis with 1000 replications by Phangorn [50] was constructed. The phylogenetic tree was then generated and visualized using FigTree v1.4.3. (<http://tree.bio.ed.ac.uk/software/figtree/>). A neighbor-joining phylogeny tree of 29

selected accessions in the citrus variant discovery panel was generated with sequence data that are not filtered for Axiom array loci. SNPhylo was used to filter 2,251,517 SNPs from the 19M SNPs in the full SNP database. For parameters for SNPhylo, the minor allele frequency filter was set to 0, the maximum missing data filter was set to 1, and the LD filter was set to 0.9. In addition, quality filtering of 5 and 5 for maximum PLCS (the percent of low coverage sample) and minimum depth coverage, respectively, was applied. A neighbor-joining tree was then generated with TASSEL v5.2.31 and visualized with Dendroscope 3 [51].

Simulation Study

A genotyping error simulation study with Axiom® Citrus Genotyping Array data on 26 selected CVC accessions including 25 accessions in the citrus variant discovery panel and Carrizo Citrange 2863 was performed for PCA, phylogeny, and ADMIXTURE analyses to understand the effect of errors in selected analyses. Genotyping errors of 0, 3, and 5 % were introduced in Flying Dragon trifoliolate 3330A and Carrizo Citrange CRC 2863 because these accessions have a higher occurrence of genotyping errors. Carrizo Citrange 2863 was not represented in the discovery panel and array data on diploid Carrizo may have a higher error rate than that of *Citrus* samples. Therefore, the genotype of Carrizo was reconstructed from haplotypes inferred from whole genome amplified single pollen grains genotyped with Axiom® Citrus Genotyping Array (data not shown). Genotype calls for haploids are more accurate than those of diploids. Genotyping data with 0 % error was generated by excluding all missing calls, and calls not identical between the array and sequence calls, resulting in total of 14,470 loci used in the simulation study. Synthetic errors were then introduced by selecting random 3 and 5 % of loci and changing heterozygous calls to homozygous and homozygous calls to heterozygous. PCA was conducted following the method described above.

Phylogeny was inferred and a neighbor-joining tree was generated using TASSEL v5.2.31 (<http://www.maizegenetics.net/tassel>). Statistical comparison of trees was performed R package APE [52] and pipeline ETE3 [53]. Admixture was performed using software Admixture v1.3 with 10-fold cross validation [54].

Results

Axiom® Citrus HD Genotyping Array

The Axiom® Citrus HD Genotyping Array is comprised of Axiom® Citrus HD Genotyping Array AX1 and Axiom® Citrus HD Genotyping Array AX2. A total of 1,404,442 probes targeting 1,296,733 autosomal SNPs and 1002 chloroplast SNPs, were selected to be represented after variant filtering part 1, followed by *in-silico* analysis of candidate variants by Affymetrix™ bioinformatics team. SNP loci represented on the Axiom® Citrus HD Genotyping Array included 854,128 SNPs from (polymorphic in) Group 1, 36,070 SNPs from Group 2, 285,521 SNPs from Group 3, 35,218 SNPs shared by Groups 1 and 2, 22,402 SNPs shared by Groups 1 and 3, 3,517 SNPs shared by Groups 2 and 3 and finally 59,877 SNPs shared by all three groups. (See Supplementary Figure 3, Additional File 1) The chromosomal distribution of the SNPs represented on Axiom® Citrus HD Genotyping Array (Table 4a, See Supplementary Figure 4, Additional File 1) and heterozygous SNPs in selected genotypes were assessed to ensure adequate genome coverage. The Axiom® Citrus HD Genotyping Array AX1 contains a total of 702,244 probes targeting 648,969 SNPs, of which 7,882 are duplicate probes tiled on both strands targeting 3,941 SNPs. 254 genomic DNA samples, of which 232 were unique, were analyzed on Axiom® Citrus HD Genotyping Array AX1, of which 220 samples (86.61%) passed and 34 samples failed. 362,281 SNPs were categorized as PHR (51.59%) and passed Affymetrix™ SNP filtering metrics. For these SNPs, the average cluster

call rate (CR), average technical reproducibility (defined as analysis of the same DNA sample) and average clonal reproducibility (defined as analysis of DNA samples from separate trees of the same accession) were 99.52%, 99.86% and 99.80% respectively. Axiom® Citrus HD Genotyping Array AX2 contains a total of 702,198 probes targeting 648,766 SNPs and of which 8,090 are duplicate probes tiled on both strands targeting 4,045 SNPs. Of 254 genomic DNA samples that were analyzed on Axiom® Citrus HD Genotyping Array AX2, 224 samples (88.19%) passed and 30 samples failed. 366,550 SNPs were categorized as PHR (52.2%) and passed Affymetrix™ SNP filtering metrics. Average cluster CR, average technical reproducibility and average clonal reproducibility were 99.56%, 99.80% and 99.97% respectively. 240 samples without accessions belonging to group 3 were analyzed and 353,278 SNPs were categorized as PHR (50.31%) on Axiom® Citrus HD Genotyping Array AX1 and 336,851 SNPs were categorized as PHR (47.97%) on Axiom® Citrus HD Genotyping Array AX2.

PHR probes on Axiom® Citrus HD Genotyping Array AX1 and Axiom® Citrus HD Genotyping Array AX2

A total of 728,831 probes (50.55%) targeting 609,890 SNPs were categorized as PHR and passed Affymetrix™ SNP filtering metrics (Table 5, Fig. 2a). The QC CR of the 34 samples that were categorized as “failed” on Axiom® Citrus HD Genotyping Array AX1 and/or Array AX2 with Axiom® Analysis Suite Default settings ranged from 67.46 % to 96.99%. Out of 34 samples that failed, 20 samples were distant relatives, and 10 samples were *Poncirus* with average QC CR of 94.18% and 95.46%, respectively. Three out of 34 failed samples had less than 80% QC CR and were distant *Citrus* relatives, *Murraya paniculata*, *Aegle marmelos*, and *Balsamocitrus daweei*.

Five failed samples were synthetic polyploid hybrids which were constructed by mixing DNA samples of two genotypes which included some more distant relatives.

Axiom® Citrus Genotyping Array

A total of 58,433 probes targeting 57,933 autosomal SNPs and 500 chloroplast SNPs were selected to be represented on the smaller Axiom® Citrus Genotyping Array after variant filtering part 2. SNP loci represented on the Axiom® Citrus Genotyping Array included 37,377 SNPs from (polymorphic in) Group 1, 2,157 SNPs from Group 2, 8,788 SNPs from Group 3, 2,961 SNPs shared by Groups 1 and 2, 1,454 SNPs shared by Groups 1 and 3, 198 SNPs shared by Groups 2 and 3 and finally 4,998 SNPs shared by all three groups (See Supplementary Figure 5 Additional File 1).

PHR probes on Axiom® Citrus Genotyping Array

A total of 1507 genomic DNA samples were analyzed on Axiom® Citrus Genotyping Array, 1494 samples (99.14%) passed and 101 samples failed. 51,296 probes (87.79%) targeting 51,296 SNPs were categorized as PHR and passed Affymetrix™ SNP filtering metrics (Table 5, Fig. 2b). For 46 out of 50 CVC samples that failed on Axiom® Citrus Genotyping Array with Axiom® Suite Default settings, cluster CR ranged from 58.91% to 96.9%. Samples of four more distant citrus relatives, *Murraya paniculata*, *Clausena anisata*, *Clausena harmandiana*, and *Glycosmis pentaphylla* had dQC values ranging from 0.657 to 0.787 and no Cluster CR was calculated. The 19 failed samples having less than 80% Cluster CR were all distant *Citrus* relatives. Average cluster CR, average

technical reproducibility and average clonal reproducibility were 99.60%, 99.75% and 99.79% respectively.

Chromosome distribution and functional annotation of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array markers

Distributions of SNP markers along the Clementine 1.0 reference genome were assessed to determine whether there was adequate genome coverage for downstream analyses. SNP densities on the Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array were 3.9SNPs/kb and 0.16 SNPs/kb, respectively (Table 5). The SNP distribution among chromosomes was roughly proportional to the genome length of each chromosome (Table 4a, Table 6a). The putative targets of probes represented on Axiom® Citrus HD Genotyping Array were assessed to show that 630,821 (49%) probes target 23,434 genes out of 24,533 Phytozome annotated genes and on the Axiom® Citrus Genotyping Array, 38,289 (66%) probes target 14,355 genes, as they were designed (Table 4b, Table 6b). Distribution of the SNPs on Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array along the genome was assessed by plotting the SNPs against the Phytozome *C. clementina* v1.0 reference genome (Fig. 3 and See Supplementary Figure 6, Additional File 1). The SNPs are well distributed throughout the genome with the largest marker gaps at 454,807 bases (chromosome 5) for Axiom® Citrus HD Genotyping Array and 953,089 bases (chromosome 5) for Axiom® Citrus Genotyping Array. In addition, SNPs on Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array were plotted against gene density of Phytozome *C. clementina* v1.0 reference genome to show that the regions of high gene density correspond well to the regions of high SNP density (Fig. 3).

Heterozygosity and reproducibility of genotyping calls between sequencing data and Axiom® Citrus HD Genotyping Array

Percent average heterozygosity in the sequence data was estimated among the variants retained in the 19M VCF file generated from sequence data, and separately for the array data (Fig. 4). Highest estimated percent heterozygosity with both sequence data and array data was observed with Schaub rough lemon (JAM) and lowest estimated percent heterozygosity was observed with Ethrog citron (ETH). Lower estimated percent heterozygosity was observed with array data in comparison to sequence data in *C. hystrix* (HYS), Flying Dragon trifoliolate (FDT), Pomeroy trifoliolate (PON), Yunnanese citron (YUN), *M. australasica* (AFL), *E. glauca* (ADL), *M. australis* (ARL), *C. micrantha* (MIC), Buddha's Hand citron (BUD) and Ethrog citron (ETH) due to SNP selection processes for the array design. Hybrid groups such lemons (*C. limon*), limes (*C. aurantifolia*), and sweet oranges (*C. sinensis*) had higher heterozygosity with array data than the sequence data and more distant groups such as papeda, trifoliolate orange, *Microcitrus*, and *Eremocitrus* had lower heterozygosity with array data. The r^2 value between estimated percent heterozygosity with sequence data and array data was 0.75. Heterozygosity in Mexican lime was underestimated with the array data compared with sequence data and other accessions (Fig. 4), probably because papeda-specific alleles are somewhat under-represented on the array. Average percent marker heterozygosity in selected taxa was estimated with SNP genotyping data of Axiom® Citrus HD Genotyping Array (See Supplementary Table 5, Additional File 1) and Axiom® Citrus Genotyping Array (See Supplementary Table 6, Additional File 1). The highest average percent heterozygosity was observed in hybrid groups, such as sweet oranges (*C. sinensis*), and lemons (*C. limon*), limes (*C. aurantifolia*), and lowest in citron (*C. medica*) and the *Citrus* relative trifoliolate orange (*Poncirus trifoliata*) with data from both arrays. Estimates of percent marker heterozygosity of 1.4% and 2.6% in citron and trifoliolate orange respectively with Axiom® Citrus Genotyping Array data are

relatively low, however, in the case of trifoliate orange, there are still more than 4,500 (Axiom® Citrus HD Genotyping Array) and 1,000 (Axiom® Citrus Genotyping Array) heterozygous SNPs that can be expected to be useful in mapping traits in these taxa. In addition, average percent heterozygosity estimated using genotyping data from Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array were proportionally comparable to those measured from the sequencing data. (See Supplementary Table 1, Additional File 1). Application of the Axiom® Citrus Genotyping Array to citrus mapping can be estimated from the chromosomal distribution of heterozygous SNPs in various accessions. The distributions of heterozygous markers in accessions with high (sweet orange; 18,159 heterozygous markers), medium (Chandler pummelo; 4,554 heterozygous markers) and low (Pomeroy trifoliate; 1,293 heterozygous markers) heterozygosity (Fig. 3) show that, even in *Poncirus*, maps with excellent coverage are predicted. Concordance between Axiom® Citrus Genotyping Array SNP data and sequence data of Flying Dragon trifoliate 3330A and Pomeroy trifoliate 1717 are 98.66 % and 98.63 %, respectively. Among the 632 loci at which calls did not match, 82.1 % to 79.4 % of calls were homozygous sites with sequence data, wrongly called as heterozygous with Axiom® Citrus Genotyping Array. Further investigation of the cluster plots with Axiom® Analysis Suite showed that hybridization signals were significantly lower at these loci.

Genotyping accuracy and reproducibility of genotyping calls of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

The genotyping accuracy of the Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Arrays was assessed by comparing genotype calls for 635,809 PHR loci in 17 accessions and 51,296 PHR loci in 28 accessions from the citrus variant discovery panel that were

also genotyped with these arrays. The pairwise concordance of a genotype between the two data sets was defined as the fraction of the SNP genotype calls for a given SNP that are in agreement with each other. “NoCall” was excluded from the calculation. The concordance between sequence data and Axiom® Citrus HD Genotyping Array data ranged from 94.88% in Mexican lime (MXL) to 99.57% in Sun Chu Sha mandarin (SCM) and average concordance was 98.09% with a median of 98.93% (Fig. 5). Reproducibility of genotyping data between Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array was assessed using 36 accessions that were analyzed with both arrays. Concordance of the genotyping data on the same accession analyzed by the both arrays ranged from 99.14% for *C. amblycarpa* CRC 2485 to 99.845% for Royal grapefruit CRC 248, and the average concordance was 99.79% with median 99.87% (Fig. 6).

Percent identity of accessions in groups composed of clonal accessions.

Percent identity of accessions in groups composed of clonal accessions was assessed by comparison of genotype calls at 51,296 PHR loci within each group. Groups assessed included Navel orange (*C. sinensis*, 67 accessions), Valencia orange (*C. sinensis*, 25 accessions), Satsuma mandarin (*C. unshiu*, 45 accessions), Eureka and Lisbon lemon (24 accessions) and Clementine (*C. clementina*, 14 accessions) and percent identity in each group ranged from 97.94 to 100%, 99.60 to 100%, 99.08 to 100%, 99.08 to 99.98%, and 99.77 to 100%, respectively. Percent identity of accessions between Navel oranges and Valencia oranges ranged from 98.12 to 99.95 %.

PCA

PCA was carried out with both Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array genotyping data. PCA with Citrus HD Genotyping Array analyzed genotyping calls at 728,831 PHR loci of 196 diverse genomic DNA samples from the CVC (See Supplementary Table 7, Additional File 1). The first three principal components (PCs), PC1, PC2 and PC3 explained 62.2% (35.6 %, 22.5%, and 4.1% respectively) of the total variation in the dataset. The representatives in the citrus variant discovery panel appear at three of the corners of the 196 unique accessions plotted (Fig. 7). PCA was also carried out with genotyping calls at 51,296 PHR loci for combined data of both arrays from 871 diploid CVC accessions (See Supplementary Table 8, Additional File 1. Fig. 8). The first three PCs explained 59.7% (30.1%, 18.5% and 11.1 % respectively) of the total variation in the dataset. PC4 explained 6.11 % of the variation in the dataset. PCA carried out with these 871 accessions show that accessions in groups constituted of clonal accessions, *C. sinensis* (171 accessions), *C. unshiu* (45 accessions), *C. clementina* (14 accessions), *Citrus limon* (*L.*) *Burm. f.* (Eureka and Lisbon lemons: 24 accessions) clustered tightly together, showing almost no variation within each group (Fig. 8d). All three PCA analyses showed that three classical *Citrus* ancestral taxa [55, 56], mandarins (*C. reticulata*), pummelos (*C. maxima*), and citrons (*C. medica*) form discrete, well-separated clusters. In addition, PC3 in all three PCA analyses is mostly explained by inclusion of accessions of *Poncirus* (red), its hybrids with *Citrus* (gray), kumquat (*Fortunella*) (pink) and papeda (yellow) in the data analyzed.

Phylogenetic Inference

Phylogeny was inferred for the variant discovery panel and selected CVC accessions. Analysis of the citrus variant discovery panel was performed separately for Axiom® Citrus Genotyping Array

data and corresponding sequence data (Fig. 9). 42,987 and 32,754 LD filtered SNPs were included for phylogenetic tree construction with Axiom® Citrus Genotyping Array (Fig. 9a) and sequence data (Fig. 9b), respectively. Sequence data included fewer SNPs than array data for this analysis due to more missing calls. The bootstrap values of above 90% are shown as bold lines. Both phylogenetic trees are supported with high bootstrap values, with 89% and 83% of the above 90% bootstrap values being 100% in the trees generated with array data and sequence data, respectively. Topological comparison of the trees performed with Phylo.io [46] showed that the trees displayed mostly similar relationships (dark blue) except for some branching patterns (light blue) (See Supplementary Figure 7, Additional File 1). Analysis of 87 CVC accessions was performed using 41,626 LD-filtered SNPs from Axiom® Citrus Genotyping Array data. The bootstrap values of above 90% are shown in bold and the phylogenetic tree is supported with high bootstrap values, 89% of the above 90% bootstrap values corresponding to 100% bootstrap support (Fig. 10). These 87 accessions were considered non-hybrid and non-admixed accessions and belong to major *Citrus* taxa, *C. medica*, *C. maxima*, *C. reticulata*, and *Citrus* relatives, *Fortunella*, papeda, and outgroup genus *P. trifoliata*. The citrus cluster is well separated from outgroup *Poncirus* with 100 percent bootstrap support and the five *Citrus* and *Fortunella* taxa form discrete clusters with complete monophyly and 100% bootstrap support. The accessions belonging to mandarin, pummelo and citron are confined to three of the major clusters, and those belonging to *C. reticulata* appear to be more divergent within the cluster based on their branch length in comparison to other clusters in the analysis. Accessions belonging to papeda, *Fortunella* and *P. trifoliata* are also in separate clusters. Interestingly, *C. hongheensis* CRC 3797, which is often classified in the papeda group, did not cluster with the other papeda accessions, instead, it nested with the pummelo cluster.

Discussion

High density SNP genotyping arrays are a powerful method for characterizing elite plant germplasm and breeding populations [42, 57–59]. The SNP array technology can provide adequate throughput and efficiency, issues which have limited previous studies of highly heterozygous organisms such as *Citrus*. SNP array technology takes advantage of SNPs which are abundant, distributed across the genome, co-dominant, and likely to have high identity by descent rate. Therefore high density SNP arrays are useful not only to construct high density genetic maps [60, 61], but also for phylogenetic studies [62], association studies [63, 64] and diversity studies [65] in a fairly high throughput manner. For citrus, several arrays including a 33,879 transcript assay Affymetrix™ GeneChip® citrus genome microarray (Affymetrix™, Thermo Fisher Scientific, Santa Clara, CA), 20K expression arrays [66], 1,457 Goldengate SNP array [67], 384 marker SNP array [68], and 756 Goldengate SNP array [69] have been previously developed for *Citrus*. These have been useful for many gene expression studies, and for detection of markers in phylogenetic research, genetic map construction, and trait mapping. In the genomic era, high throughput genetic marker systems are essential for high resolution genomic studies, and combined with available bioinformatics tools, they allowed us to answer both biological and practical questions. Axiom® Citrus HD Genotyping Array (3.9 SNP/kb) and Axiom® Citrus Genotyping Array (0.16 SNP/kb) are the highest density SNP genotyping arrays constructed and available for citrus so far. Further, these high density SNP arrays allowed us to compile an extensive genomic inventory of the accessions in the CVC in a very high throughput manner and perform a range of high resolution genomic studies. As discussed previously, the heterozygous *Citrus* genome makes the GBS approach computationally demanding and, depending on sequencing cost, often more expensive per sample than in inbred plants. Alternatively, high density SNP genotyping arrays will not only

provide the coverage many genomic analyses require, but also, are efficient and high throughput. After the initial investment, the time required from sample to data is relatively short.

Accuracy of array data in citrus and citrus relatives

Our study also showed that high density SNP genotyping can produce data and results that are comparable to those of whole genome sequencing for inferring all but fine-scale relationships, while providing sufficient genome coverage with heterozygous markers for mapping in many citrus populations (Fig. 3). SNPs represented on the array also targeted mostly genic regions (Axiom® Citrus HD Genotyping Array: 49%, Axiom® Citrus Genotyping Array: 66%) (Fig. 3. See Supplementary Figure 8, 9, Additional File 1) in order to increase the success of SNP genotyping in diverse accessions of citrus. Due to the lower concordance observed in more distant taxa between array data and sequence data, a simulation study was conducted to show the effect of genotyping errors in *Poncirus* and its hybrid with *Citrus* within the range observed in array data. The simulation study showed genotyping error ranging from 0 to 5 % had little to no effect on the results of PCA (See Supplementary Figure 10, Additional File 1), phylogeny (See Supplementary Figure 11, Additional File 1), and Admixture analysis (See Supplementary Figure 12, Additional File 1). Overall, data obtained from our study indicates that these arrays are most effective when applied to *Citrus* rather than its close relatives, probably due to use of sequence from Clementine in probe design.

Population structure

The population must be carefully examined to assess the appropriate marker density and to avoid spurious associations prior to performing association studies. PCA provides a multivariate tool [70] to analyze complex genomic variation data and estimate genomic divergence between populations. A large SNP dataset can be easily analyzed by PCA [71]. Many accessions in the CVC have never been characterized with high density marker systems prior to our study. PCA allows us to efficiently determine the genomic background of such taxa. The smaller, more economical array Axiom® Citrus Genotyping Array provided sufficient marker density to perform PCA and successfully separated ancestral *Citrus* [56] and more distant taxa into separate clusters. *C. reticulata*, *C. maxima*, *C. medica* and *P. trifoliata*, a *Citrus* relative often used in rootstock breeding, clustered separately explaining the most variation within the dataset (59.6%) (Fig. 8). PC1 mainly separated *C. medica* from the other accessions, PC2 mainly separated *C. maxima* from the rest of the dataset and PC3 mainly separated *P. trifoliata* from *Citrus*. Percentage variation explained by PC3 in Axiom® Citrus HD Genotyping Array (4.1%) and Axiom® Citrus Genotyping Array (11%) differed significantly (Fig. 7, 8), probably resulting from the inclusion of a greater proportion of *Poncirus* and *Poncirus* hybrid accessions in the Axiom® Citrus Genotyping Array data (4.6% for Axiom® Citrus HD Genotyping Array data and 7.5% for Axiom® Citrus Genotyping Array data). In addition, our study with Axiom® Citrus Genotyping Array found that some accessions included in the study that were previously categorized as non-hybrid accessions, according to the classification currently employed in CVC, are instead hybrids with other *Citrus*. Many cultivated accessions of *C. reticulata* have been reported to be introgressed with *C. maxima* [20]. PCA shows that the accessions identified as not introgressed in previous studies [21, 22] clustered at the tip of the *C. reticulata* clusters. Preliminary analysis with ADMIXTURE and percent heterozygosity confirmed these results (data not shown). Reclassification of many

accessions in the CVC with unknown genomic background would have been a laborious task, had they and a large diverse collection not been analyzed with such high-throughput SNP genotyping arrays, suggesting great potential utility of Axiom® Citrus Genotyping Array.

Inference of citrus phylogeny with array data

Resolving *Citrus* lineage and phylogeny has long been an interest of citrus breeders and researchers. Phylogenetic studies are useful for both in terms of genetic diversity analysis and parental line selection in breeding programs. Historically, phylogeny and taxonomy of *Citrus* have been studied based on morphology and geography. Until recently there have been two widely accepted *Citrus* taxonomic systems, Swingle and Reece, and Tanaka. Comparing these two classical *Citrus* classifications, there are 16 species in two subgenera (*Citrus* and *Papeda*) proposed by Swingle [5] and up to 162 species proposed by Tanaka [72]. The major discrepancy between the two classical systems exists mainly in *Citrus*, and stems from how much divergence within a recognized species should be allowed, and more specifically, whether natural hybrids should be considered as true species.

Over the years, various molecular markers such as simple sequence repeats (SSR), restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), sequence characterized amplified region (SCAR), sequence related amplified polymorphism (SRAP), mitochondrial DNA, chloroplast DNA, chromosomal variability, and nuclear gene sequences have been used in phylogenetic investigations into *Citrus* to try to better understand the origin and relationships of diverse accessions [30, 73–79]. Analysis of markers [74, 81] and sequencing indicate four major ancestral *Citrus* groups, citron (*C. medica*), pummelo (*C. maxima*), mandarin (*C. reticulata*), and papeda (*C. micrantha*) [85], while Scora and Barrett and Rhodes [55,

56] only recognized the first three groups as ancestors of cultivated citrus. Conventional phylogenetic trees are used to represent the evolutionary history of taxa, however, complex evolutionary scenarios that involve hybridization, horizontal gene transfer, recombination, and gene duplication and loss violate the assumptions of these models. More than a thousand years of domestication led to hybridization and introgression of ancestral progenitors into the cultivated *Citrus* varieties. Therefore, identification of appropriate accessions prior to such analyses is very important. The phylogeny analysis performed with 26 CVC accessions included known hybrids and admixed accessions. For example, the placement of Schaub rough lemon CRC 3879 in the phylogenetic trees generated with array data and sequence data differs. Rough lemon is thought to be a natural hybrid of a citron and a mandarin [55, 81]. The Structure analysis of rough lemon with 24 SSR loci (275 alleles) showed that it derived a majority of alleles from citron and the rest from mandarin [30]. Schaub rough lemon is nested with orange and mandarins with array data while it is nested with lemon, lime and citron with sequence data. An F1 hybrid may cluster with either parent group, so sampling error can contribute to variation in its placement. There is 97.8 % genotyping call concordance between Axiom® Citrus HD Genotyping Array and sequence data and much of the discrepancy between these two trees, most notably the placement of Schaub rough lemon, is likely due to the inclusion of such hybrid accessions in the analysis. Nonetheless, topological comparison showed that Axiom® Citrus Genotyping Array could provide data and information content comparable to that of sequence data in phylogenetic studies (See Supplementary Figure 7, Additional File 1, Fig. 10, See Supplementary Figure 13, Additional File 1).

More definitive and informative evidence of ancestry can be provided by determination of chromosome-length haplotypes of hybrid accessions combined with identification and analysis of the chromosome distribution of species-specific SNPs.

Inference of 87 non-admixed accessions of *Citrus* and close relatives

Previous studies of *Citrus* phylogeny often resulted in insufficient understanding of the genetic diversity that exists in *Citrus* species due to inadequate sampling of markers or accessions. The CVC provides an excellent opportunity to assess genetic diversity and population structure given its size and diversity. In addition, the use of whole genome SNP markers provided by Axiom® citrus arrays should significantly improve the resolution of phylogenetic study in *Citrus*, given their genome coverage and SNP targets. While rapid progress has been made in this field in many other plants and some crops, progress has lagged in *Citrus* due to genetic complexity. Our phylogenetic analysis using Axiom® Citrus array data of 87 non-admixed accessions in the CVC is the most extensive phylogenetic study of non-admixed *Citrus* accessions to date [74, 76, 78, 80, 81, 82–85]. Phylogeny inferred with Axiom® Citrus array data identified five monophyletic clusters representing citron group, pummelo group, mandarin group, kumquat group, and papeda group with 100% bootstrap support (Fig. 10), consistent with previous reports that are based on fewer markers or accessions [30, 79, 81, 86]. *C. tachibana* CRC 3150 is in one of the minor subclusters and most different from other mandarin accessions, as inferred previously [75, 81, 82, 83]. Study with SSRs reported *C. tachibana* as clustered with papeda accessions, but without strong bootstrap support [30]. It is likely that increased marker density was able to better resolve the relationships of *C. tachibana* in the mandarin cluster. A recent analysis of whole genome sequence data (Wu et al. 2018) concluded that *C. tachibana* should be designated a subspecies of *C. reticulata*. Yellow mandarin CRC 3895 is in another minor subcluster and also is a rather divergent mandarin type. There has been no report of marker studies performed on yellow mandarin, however, it is thought to be an older mandarin from China according to the description on USDA Germplasm Resources Information Network (<https://www.ars-grin.gov/>). All accessions of *Fortunella* and the papeda group analyzed in the phylogeny study form two clusters with very short branch lengths between

accessions, hence, these accessions are similar to one another, except for *C. hongheensis* CRC 3797. *C. hongheensis* CRC 3797, has been considered to be a papeda, but clustered as a sister group to pummelos. Previous study with RFLP and RAPD showed that *C. hongheensis* CRC 3797 clustered with pummelo [73]. Its heterozygosity is 2.4% and lower in comparison to *C. hystrix* CRC 3103 and *C. micrantha* CRC 3605 that are in the citrus variant discovery panel. Average heterozygosity of pummelo with Axiom® Citrus 56AX genotyping array is 8.2%. Sequence analysis of selected genic regions of chloroplast DNA also clustered papedas such as *C. latipes* and *C. ichangensis* with pummelo [88]. Papeda is one of the two subgenera within *Citrus* defined by Swingle [5], but other studies suggest that papeda is a polyphyletic group [79, 88]. Although *C. hongheensis* CRC 3797 clusters with pummelo, it is very divergent, indicating the possibility that *C. hongheensis* CRC 3797 could be a sister species to pummelo. However, it is also possible that the low heterozygosity observed is due to ascertainment bias introduced by the citrus variant discovery panel not including a divergent papeda, *C. ichangensis*. *C. ichangensis* belongs to subgenus *Papeda* by Swingle's classification system [5] based on the presence of acidic oil in the fruits, and morphology, however, Tanaka placed *C. ichangensis* together with all mandarin species [72]. *C. ichangensis* is thought to be a more genetically divergent type in the papeda group [74]. The possible chloroplast genome relationship between *C. ichangensis* and *C. reticulata* is supported by marker analysis of chloroplast DNA [74]. Our phylogeny analysis conducted with array data including *C. ichangensis* and 500 chloroplast SNP array data also indicates *C. ichangensis* is related to *C. reticulata*, rather than to the papeda group, or *C. hongheensis* (data not shown). It remains possible that *C. hongheensis* is a hybrid between a pummelo and a papeda type that is not represented in the citrus variant discovery panel. The array data shows that both citron and pummelo clusters have little within-group population structure. Our phylogeny analyses strongly suggest the array data could provide resolution and information content comparable to those of

sequence data in phylogenetic analyses with smaller sample sizes (Fig. 9, Fig. 10, See Supplementary Figure 13, Additional File 1). Both phylogeny trees generated with array data and sequence data cluster the parent-offspring pair Siamese Acidless pummelo CRC 2240 and Chandler pummelo CRC 3224 together. In addition, in both trees, Yunnanese citron, *C. tachibana* mandarin, and Hunan pummelo are the most divergent accessions in their species cluster. Phylogenetic trees constructed in our study with Axiom® citrus array data and previously reported trees often differed in relative branch lengths between the identified six clusters for several reasons. First of all, the SNPs chosen from the citrus variant discovery panel were all polymorphic and mainly designed to work well in *Citrus*, therefore, overestimation of divergence between different Citrus clusters could occur. Secondly, the array was designed to discard the rarest alleles and relatively few accessions of relatives and some groups such as papedas were included, perhaps leading to underestimates of divergence between these taxa and other citrus. Lastly, we performed limited LD filtering in the analyses to reduce the bias due to redundant variants. Higher LD level can lead to overestimation of distances between clusters. LD can be created as a consequence of mutation, genetic drift, selection, and limited recombination which lead to population structure due to differences in allele frequencies between groups when individuals from different genetic origins are included in the dataset [89, 90]. Although the extent of LD has not been thoroughly investigated in *Citrus* with high density marker systems, it is presumed to vary greatly depending on the population [32]. For *Citrus* species with self-incompatibility, LD is likely to be shorter than LD in the population of interspecific hybrids and admixed accessions [91]. A report of LD in citrus, that was based on analysis of 50 SSRs in 73 accessions of which 29 belong to ancestral species, 11 to secondary species and 17 to a breeding population, showed a high LD between distant markers with little decay [92]. Another report of LD in citrus based on 1,841 SNPs showed a similar trend [32]. Therefore, SNP density represented on Axiom® Citrus HD Genotyping Array and Axiom® Citrus

Genotyping Array should be sufficient for downstream applications such as linkage mapping, ancestry analysis, and, perhaps genomic selection, that do not require very high marker densities.

Conclusion

High density Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array have been developed with more than 1.4M and 58K SNPs that are well distributed over the citrus genome. This study included many accessions in taxa relatively distant from *C. clementina*. Sequence read alignment for variant discovery in 41 accessions and probe sequence design for Axiom® citrus arrays used the *C. clementina* v1.0 reference genome and this is likely the cause for the lower concordance between sequence and array data in some taxa (Fig. 5, 6). Lower concordance in the more distant taxon *Poncirus* was observed suggesting that one or more additional adjacent SNPs, including insertions and deletions, in the genomes of genotypes in more distant taxa, interfered with hybridization. Nonetheless, our study shows that the Axiom® arrays for *Citrus* and some related genera provide a tool that is very powerful, useful and increases throughput in analyzing diversity of complex citrus genomes. They also provide a standardized set of markers that can be used repeatedly for mapping traits in various different populations. The utility of these arrays in GWAS, QTL studies, and high density linkage map construction to dissect important traits such as fruit quality, nucellar embryony, and disease resistance is promising, given the high density of robust SNPs represented on the arrays. Such results should be useful to further assist marker-assisted breeding. In addition, the arrays are used in detection of recent admixture events, population structure analysis, LOH analysis, haplotype construction and investigation into parentage of CVC accessions. Although phylogenetic analysis using array data may not always yield correct interpretation of evolutionary events due to the ascertainment bias, both arrays include

high density chloroplast SNPs, 1002 and 500 SNPs, respectively and therefore, will help improve our understanding of phylogenetic relationships between notoriously confounded citrus accessions. As improved reference genomes for citrus become available, it should be possible to improve annotation of these arrays. Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array are now available from Affymetrix™.

References cited

1. Fao. FAO Statistical Yearbook: World Food and Agriculture. FAO; 2013.
2. EST: Citrus Fruit. <http://www.fao.org/economic/est/est-commodities/citrus-fruit/en/>. Accessed 31 Oct 2018
3. United Nations, editor. Food and Agriculture Organization of the United Nations (FAO). In: Yearbook of the United Nations 1995. UN; 1995. p. 1469–72. doi:10.18356/d6ce0776-en.
4. Tripoli E, Guardia ML, Giammanco S, Majo DD, Giammanco M. *Citrus* flavonoids: molecular structure, biological activity and nutritional properties: A review. Food Chem. 2007;104:466–79. doi:10.1016/j.foodchem.2006.11.054.
5. Swingle WT, Reece PC. The botany of *Citrus* and orange relatives in the orange subfamily. (2nd ed.) W. Reuther, Webber H.J., Batchelor D.L. (Eds.), The Citrus Industry, vol. 1, California Univ. Press, Berkeley. 1967;:190-340
6. Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, et al. The effects of artificial selection on the maize genome. Science. 2005;308:1310–4. doi:10.1126/science.1107891.
7. Li C, Zhou A, Sang T. Rice domestication by reducing shattering. Science. 2006;311:1936–9. doi:10.1126/science.1123604.
8. Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, et al. Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet. 2008;40:1023–8. doi:10.1038/ng.169.
9. Lam H-M, Xu X, Liu X, Chen W, Yang G, Wong F-L, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat Genet. 2010;42:1053–9. doi:10.1038/ng.715.
10. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia J-M, Cartwright RA, et al. Comparative population genomics of maize domestication and improvement. Nat Genet. 2012;44:808–11. doi:10.1038/ng.2309.
11. Koenig D, Jiménez-Gómez JM, Kimura S, Fulop D, Chitwood DH, Headland LR, et al. Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proc Natl Acad Sci U S A. 2013;110:E2655–62. doi:10.1073/pnas.1309606110.
12. Li Y-H, Zhao S-C, Ma J-X, Li D, Yan L, Li J, et al. Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. BMC Genomics. 2013;14:579. doi:10.1186/1471-2164-14-579.
13. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet. 2013;45:487–94. doi:10.1038/ng.2586.
14. Cao K, Zheng Z, Wang L, Liu X, Zhu G, Fang W, et al. Comparative population genomics reveals the domestication history of the peach, *Prunus persica*, and human influences on perennial fruit crops. Genome Biol. 2014;15:415. doi:10.1186/s13059-014-0415-1.

15. Duan N, Bai Y, Sun H, Wang N, Ma Y, Li M, et al. Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nat Commun.* 2017;8:249. doi:10.1038/s41467-017-00336-7.
16. Ma B, Liao L, Peng Q, Fang T, Zhou H, Korban SS, et al. Reduced representation genome sequencing reveals patterns of genetic diversity and selection in apple. *J Integr Plant Biol.* 2017;59:190–204. doi:10.1111/jipb.12522.
17. Migicovsky Z, Sawler J, Gardner KM, Aradhya MK, Prins BH, Schwaninger HR, et al. Patterns of genomic and phenomic diversity in wine and table grapes. *Hortic Res.* 2017;4:17035. doi:10.1038/hortres.2017.35.
18. Zhou Y, Massonnet M, Sanjak JS, Cantu D, Gaut BS. Evolutionary genomics of grape (*Vitis vinifera* ssp. *vinifera*) domestication. *Proc Natl Acad Sci U S A.* 2017;114:11715–20. doi:10.1073/pnas.1709257114.
19. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al. The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet.* 2013;45:59–66. doi:10.1038/ng.2472.
20. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat Biotechnol.* 2014;32:656–62. doi:10.1038/nbt.2906.
21. Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet J-P, Garcia-Lor A, et al. Nuclear species-diagnostic SNP markers mined from 454 amplicon sequencing reveal admixture genomic structure of modern citrus varieties. *PLoS One.* 2015;10:e0125628. doi:10.1371/journal.pone.0125628.
22. Oueslati A, Salhi-Hannachi A, Luro F, Vignes H, Mournet P, Ollitrault P. Genotyping by sequencing reveals the interspecific *C. maxima* / *C. reticulata* admixture along the genomes of modern citrus varieties of mandarins, tangors, tangelos, orangelos and grapefruits. *PLoS One.* 2017;12:e0185618. doi:10.1371/journal.pone.0185618.
23. Nicolosi EL. Origin and taxonomy. In: Khan, I. A. (ed.). *Citrus Genetics, Breeding and Biotechnology*. CAB International, Wallingford, U.K. 2007;:19-43.
24. Frost HB. Nucellar embryony and juvenile characteristics in clonal varieties of *Citrus*. *J Hered.* 1938;29:423–92. doi:10.1093/oxfordjournals.jhered.a104431.
25. Soost RL, Roose ML. Citrus. Fruit Breeding, Vol. 1: Tree and Tropical Fruits. In Janick J. and J.N.Moore (eds.), John Wiley&Sons, Inc. 1996;:257-323
26. Ahmed D, Comte A, Curk F, Costantino G, Luro F, Dereeper A, et al. Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: a case study in diploid and polyploid citrus. *Ann Bot.* 2019;123:1231–51. doi:10.1093/aob/mcz029.
27. Wang X, Xu Y, Zhang S, Cao L, Huang Y, Cheng J, et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat Genet.* 2017;49:765–72. doi:10.1038/ng.3839.

28. Fang D, Krueger RR, Roose ML. Phylogenetic relationships among selected *Citrus* germplasm accessions revealed by Inter-simple sequence repeat (ISSR) markers. *J Am Soc Hortic Sci.* 1998;123:612–7. <http://journal.ashspublications.org/content/123/4/612.full.pdf+html>. Accessed 10 Oct 2017.
29. Ruiz C, Paz Breto M, Asíns MJ. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. *Euphytica.* 2000;112:89–94. doi:10.1023/A:1003992719598.
30. Barkley NA, Roose ML, Krueger RR, Federici CT. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor Appl Genet.* 2006;112:1519–31. doi:10.1007/s00122-006-0255-9.
31. Kepiro JL, Roose ML. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*. *Tree Genet Genomes.* 2010;6:1–11. doi:10.1007/s11295-009-0223-z.
32. Minamikawa MF, Nonaka K, Kaminuma E, Kajiya-Kanegae H, Onogi A, Goto S, et al. Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. *Sci Rep.* 2017;7:4721. doi:10.1038/s41598-017-05100-x.
33. Yu H, Yang X, Guo F, Jiang X, Deng X, Xu Q. Genetic diversity and population structure of pummelo (*Citrus maxima*) germplasm in China. *Tree Genet Genomes.* 2017;13:58. doi:10.1007/s11295-017-1133-0.
34. Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature.* 2010;464:898–902. doi:10.1038/nature08837.
35. Hu J, Diao Y, Zhou Y, Lin D, Bi Y, Pang Z, et al. Loss of heterozygosity drives clonal diversity of *Phytophthora capsici* in China. *PLoS One.* 2013;8:e82691. doi:10.1371/journal.pone.0082691.
36. Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, Gasic K, et al. Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PLoS One.* 2012;7:e35668. doi:10.1371/journal.pone.0035668.
37. Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, et al. Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* × *domestica* Borkh). *PLoS One.* 2014;9:e110377. doi:10.1371/journal.pone.0110377.
38. Chen H, Xie W, He H, Yu H, Chen W, Li J, et al. A high-density SNP genotyping array for rice biology and molecular breeding. *Mol Plant.* 2014;7:541–53. doi:10.1093/mp/sst135.
39. Unterseer S, Bauer E, Haberer G, Seidel M, Knaak C, Ouzunova M, et al. A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. *BMC Genomics.* 2014;15:823. doi:10.1186/1471-2164-15-823.
40. Bassil NV, Davis TM, Zhang H, Ficklin S, Mittmann M, Webster T, et al. Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry *Fragaria* × *ananassa*. *BMC Genomics.* 2015;16:155. doi:10.1186/s12864-015-1310-1.

41. Ganai MW, Polley A, Graner E-M, Plieske J, Wieseke R, Luerssen H, et al. Large SNP arrays for genotyping in crop plants. *J Biosci.* 2012;37:821–8. doi:10.1007/s12038-012-9225-3.
42. Viquez-Zamora M, Vosman B, van de Geest H, Bovy A, Visser RGF, Finkers R, et al. Tomato breeding in the genomics era: insights from a SNP array. *BMC Genomics.* 2013;14:354. doi:10.1186/1471-2164-14-354.
43. Yu H, Xie W, Li J, Zhou F, Zhang Q. A whole-genome SNP array (RICE6K) for genomic breeding in rice. *Plant Biotechnol J.* 2014;12:28–37. doi:10.1111/pbi.12113.
44. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One.* 2011;6:e19379. doi:10.1371/journal.pone.0019379.
45. Bausher MG, Singh ND, Lee S-B, Jansen RK, Daniell H. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol.* 2006;6:21. <https://bmcpantbiol.biomedcentral.com/articles/10.1186/1471-2229-6-21>.
46. Robinson O, Dylus D, Dessimoz C. Phylo.io: Interactive viewing and comparison of large phylogenetic trees on the web. *Mol Biol Evol.* 2016;33:2163–6. doi:10.1093/molbev/msw080.
47. Lee T-H, Guo H, Wang X, Kim C, Paterson AH. SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics.* 2014;15:162. doi:10.1186/1471-2164-15-162.
48. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics.* 2012;28:3326–8. doi:10.1093/bioinformatics/bts606.
49. Felsenstein J. *PHYLIP 3.2 Manual*. University of California Herbarium, Berkeley. 1989.
50. Schliep KP. phangorn: phylogenetic analysis in R. *Bioinformatics.* 2010;27:592–3. doi:10.1093/bioinformatics/btq706.
51. Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol.* 2012;61:1061–7. doi:10.1093/sysbio/sys062.
52. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics.* 2004;20:289–90. doi:10.1093/bioinformatics/btg412.
53. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol.* 2016;33:1635–8. doi:10.1093/molbev/msw046.
54. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;19:1655–64. doi:10.1101/gr.094052.109.
55. Scora RW. On the history and origin of *Citrus*. *Bull Torrey Bot Club.* 1975;;369–75. <https://www.jstor.org/stable/2484763>.
56. Barrett HC, Rhodes AM. A numerical taxonomic study of affinity relationships in cultivated

citrus and its close relatives. *Syst Bot.* 1976;1:105–36. doi:10.2307/2418763.

57. Ganai MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, et al. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS One.* 2011;6:e28334. doi:10.1371/journal.pone.0028334.

58. Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, et al. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS One.* 2012;7:e31745. doi:10.1371/journal.pone.0031745.

59. Muñoz-Amatriaín M, Mirebrahim H, Xu P, Wanamaker SI, Luo M, Alhakami H, et al. Genome resources for climate-resilient cowpea, an essential crop for food security. *Plant J.* 2017;89:1042–54. doi:10.1111/tpj.13404.

60. Sim S-C, Durstewitz G, Plieske J, Wieseke R, Ganai MW, Van Deynze A, et al. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS One.* 2012;7:e40563. doi:10.1371/journal.pone.0040563.

61. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol J.* 2014;12:787–96. doi:10.1111/pbi.12183.

62. Li Y-H, Li W, Zhang C, Yang L, Chang R-Z, Gaut BS, et al. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytol.* 2010;188:242–53. doi:10.1111/j.1469-8137.2010.03344.x.

63. Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, et al. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun.* 2011;2:467. doi:10.1038/ncomms1467.

64. Wang M, Yan J, Zhao J, Song W, Zhang X, Xiao Y, et al. Genome-wide association study (GWAS) of resistance to head smut in maize. *Plant Sci.* 2012;196:125–31. doi:10.1016/j.plantsci.2012.08.004.

65. Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, et al. Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. *PLoS One.* 2010;5:e10780. doi:10.1371/journal.pone.0010780.

66. Martinez-Godoy MA, Mauri N, Juarez J, Marques MC, Santiago J, Forment J, et al. A genome-wide 20 K citrus microarray for gene expression analysis. *BMC Genomics.* 2008;9:318. doi:10.1186/1471-2164-9-318.

67. Ollitrault P, Terol J, Garcia-Lor A, Bérard A, Chauveau A, Froelicher Y, et al. SNP mining in *C. clementina* BAC end sequences; transferability in the *Citrus* genus (*Rutaceae*), phylogenetic inferences and perspectives for genetic mapping. *BMC Genomics.* 2012;13:13. doi:10.1186/1471-2164-13-13.

68. Fujii H, Shimada T, Nonaka K, Kita M, Kuniga T, Endo T, et al. High-throughput genotyping

- in citrus accessions using an SNP genotyping array. *Tree Genet Genomes*. 2013;9:145–53. doi:10.1007/s11295-012-0542-3.
69. Shimizu T, Kaminuma E, Nonaka K, Yoshioka T, Goto S, Matsumoto T, et al. A genomic approach to selecting robust and versatile SNP sets from next-generation sequencing data for genome-wide association study in citrus cultivars. In: III International Symposium on Citrus Biotechnology 1135. 2014. p. 23–32. https://www.actahort.org/books/1135/1135_4.htm.
70. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet*. 2006;2:e190. doi:10.1371/journal.pgen.0020190.
71. McVean G. A genealogical interpretation of principal components analysis. *PLoS Genet*. 2009;5:e1000686. doi:10.1371/journal.pgen.1000686.
72. Tanaka T, Others. Species problem in *Citrus*. 1954. <http://agris.fao.org/agris-search/search.do?recordID=US201300690891>.
73. Federici CT, Fang DQ, Scora RW, Roose ML. Phylogenetic relationships within the genus *Citrus* (*Rutaceae*) and related genera as revealed by RFLP and RAPD analysis. *Theor Appl Genet*. 1998;96:812–22. doi:10.1007/s001220050807.
74. Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor Appl Genet*. 2000;100:1155–66. doi:10.1007/s001220051419.
75. Yamamoto M, Tominaga S. High chromosomal variability of mandarins (*Citrus* spp.) revealed by CMA banding. *Euphytica*. 2003;129:267. doi:10.1023/A:1022284908424.
76. Penjor T, Anai T, Nagano Y, Matsumoto R, Yamamoto M. Phylogenetic relationships of *Citrus* and its relatives based on *rbcL* gene sequences. *Tree Genet Genomes*. 2010;6:931–9. doi:10.1007/s11295-010-0302-1.
77. Uzun A, Yesiloglu T, Aka-Kacar Y, Tuzcu O, Gulsen O. Genetic diversity and relationships within *Citrus* and related genera based on sequence related amplified polymorphism markers (SRAPs). *Sci Hortic*. 2009;121:306–12. doi:10.1016/j.scienta.2009.02.018.
78. Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, et al. Phylogenetic relationships of *Citrus* and its relatives based on *matK* gene sequences. *PLoS One*. 2013;8:e62574. doi:10.1371/journal.pone.0062574.
79. Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Mol Biol Evol*. 2015;32:2015–35. doi:10.1093/molbev/msv082.
80. Dugrand-Judek A, Olry A, Hehn A, Costantino G, Ollitrault P, Froelicher Y, et al. The distribution of coumarins and furanocoumarins in *Citrus* species closely matches *Citrus* phylogeny and reflects the organization of biosynthetic pathways. *PLoS One*. 2015;10:e0142757. doi:10.1371/journal.pone.0142757.
81. Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P. Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. *Ann Bot*. 2016;117:565–83.

doi:10.1093/aob/mcw005.

82. Yamamoto M, Kobayashi S, Nakamura Y, Yamada Y. Phylogenetic relationships of citrus revealed by RFLP analysis of mitochondrial and chloroplast DNA. *Japanese Journal of Breeding*. 1993;43:355–65. https://www.jstage.jst.go.jp/article/jsbbs1951/43/3/43_3_355/_article/-char/ja/.
83. de Araújo EF, de Queiroz LP, Machado MA. What is Citrus? Taxonomic implications from a study of cp-DNA evolution in the tribe *Citreae* (*Rutaceae* subfamily *Aurantioideae*). *Org Divers Evol*. 2003;3:55–62. doi:10.1078/1439-6092-00058.
84. Yingzhi L, Yunjiang C, Nengguo T, Xiuxin D. Phylogenetic analysis of mandarin landraces, wild mandarins, and related species in China using nuclear LEAFY second intron and plastid trnL-trnF sequence. *J Am Soc Hortic Sci*. 2007;132:796–806. <http://journal.ashspublications.org/content/132/6/796.abstract>.
85. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of *Citrus*. *Nature*. 2018;32:656. doi:10.1038/nature25447.
86. Garcia-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, et al. A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the “true citrus fruit trees” group (*Citrinae*, *Rutaceae*) and the origin of cultivated species. *Ann Bot*. 2012;111:1–19. <https://academic.oup.com/aob/article-abstract/111/1/1/181653>.
87. Uchoi A, Malik SK, Choudhary R, Kumar S, Rohini MR, Pal D, et al. Inferring phylogenetic relationships of Indian citron (*Citrus medica* L.) based on rbcL and matK sequences of chloroplast DNA. *Biochem Genet*. 2016;54:249–69. doi:10.1007/s10528-016-9716-2.
88. Butelli E, Garcia-Lor A, Licciardello C, Las Casas G, Hill L, Recupero GR, et al. Changes in anthocyanin production during domestication of *Citrus*. *Plant Physiol*. 2017;173:2225–42. doi:10.1104/pp.16.01701.
89. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. *Am J Hum Genet*. 2001;69:1–14. doi:10.1086/321275.
90. Slatkin M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. 2008;9:477–85. doi:10.1038/nrg2361.
91. Roose ML. Mapping and marker assisted selection in Citrus. In: Khan, I. A. (ed.). *Citrus Genetics, Breeding and Biotechnology*. CAB International, Wallingford, U.K. 2007;: 275-286
92. García-Lor A, Luro F, Navarro L, Ollitrault P. Comparative use of InDel and SSR markers in deciphering the interspecific structure of cultivated citrus genetic diversity: a perspective for genetic association studies. *Mol Genet Genomics*. 2012;287:77–94. doi:10.1007/s00438-011-0658-4.

Table 1 Accessions selected to be in Citrus variant discovery panel that were sequenced at UCR IIGB

Accessions of *Citrus* and its close relatives in citrus variant discovery panel that were re-sequenced at the UCR Institute of Integrative Genome Biology (IIGB, University of California Riverside, CA, USA) for the designing of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. These sequences are deposited at NCBI SRA: SRP095606. Group 1 includes 29 accessions in citrus group, Group 2 includes 4 accessions from citron group, and Group 3 includes 8 accessions of citrus relatives.

ID	Variety Name	Genus and Species Name	CRC	Group
ADL	Australian desert lime	<i>Eremocitrus glauca</i> (Lindl.) Swingle	3463	3
AFL	Red pulp finger lime	<i>Microcitrus australasica</i> var. <i>sanguinia</i> (F. Muell.) Swingle	1484	3
ARL	Australian round lime	<i>Microcitrus australis</i> (A. Cunn. Ex Mudie) Swingle	3673	3
BUD	Buddha's Hand citron (fingered citron)	<i>Citrus medica</i> var. <i>sarcodactylis</i> (Hoola van Nooten) Swingle	3768	2
CLP	Cleopatra mandarin	<i>Citrus reshni</i> Hort. ex Tanaka	3844	1
CPP	Cariappa CM3 pummelo	<i>Citrus maxima</i> (Burm.) Merr.	NA	1
CSM	Changsha mandarin	<i>Citrus reticulata</i> Blanco	3577	1
DNC	Dancy mandarin (Frost nucellar)	<i>Citrus reticulata</i> Blanco	3026	1
DUN	Duncan grapefruit	<i>Citrus paradisi</i> Macf.	3832	1
ETH	Ethrog citron	<i>Citrus medica</i> L.	3891	2
FDT	Flying Dragon trifoliolate	<i>Poncirus trifoliata</i> (L.) Raf.	3330A	3
FOR	Nagami kumquat	<i>Fortunella margarita</i> (Lour.) Swingle	3877	3
HUP	Hunan pummelo	<i>Citrus maxima</i> (Burm.) Merr.	1225	1
HYS	Mauritius papeda	<i>Citrus hystrix</i> D. C.	3103	3
KPP	Kao Phuang pummelo	<i>Citrus maxima</i> (Burm.) Merr.	2352	1
KSH	Mukakukishu (Kinokuni mandarin)	<i>Citrus kinokuni</i> Hort. ex Tanaka	3887	1
LSB	Frost nucellar Lisbon lemon	<i>Citrus limon</i> L. Burm. f.	3176	1
MBP	Mato Buntan pummelo	<i>Citrus maxima</i> (Burm.) Merr.	3945	1
MIC	Small-fruited papeda	<i>Citrus micrantha</i> var. <i>microcarpa</i> Wester	3605	3
MXL	Mexican lime	<i>Citrus aurantifolia</i> Swingle	1710	1
QIN	Persistent Stigma - OPS citron	<i>Citrus medica</i> L.	NA	2
PON	Pomeroy trifoliolate	<i>Poncirus trifoliata</i> (L.) Raf.	1717	3
JAM	Schaub rough lemon	<i>Citrus jambhiri</i> Lush.	3879	1
SCM	Sun Chu Sha Kat mandarin	<i>Citrus reticulata</i> Blanco	4003	1
SNK	Sunki mandarin	<i>Citrus sunki</i> Hort. ex Tanaka	3143	1
TAP	Tahitian pummelo	<i>Citrus maxima</i> (Burm.) Merr.	3806	1

TBM	Tachibana orange	<i>Citrus tachibana</i> Tanaka	3150	1
TDP	Thong Dee pummelo	<i>Citrus maxima</i> (Burm.) Merr	3927	1
TRM	Triumph grapefruit	<i>Citrus paradisi</i> Macf.	297	1
YUN	Yunnanese citron	<i>Citrus medica</i> L.	3798	2

Table 2 Citrus variant discovery panel from public database or collaborators

Accessions of *Citrus* and its close relatives in citrus variant discovery panel from which whole genome sequence data was obtained from publicly available sources or our collaborating institution for designing of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. All accessions belong to Group 1.

Genotype ID	Variety Name	Genus Species Name	Group
BAN*	Banpeiyu	<i>Citrus maxima</i> (Burm.) Merr.	1
CHP	Chandler pummelo	<i>Citrus maxima</i> (Burm.) Merr.	1
CLM	Clementine mandarin	<i>Citrus clementina</i> hort. ex Tanaka	1
KNG*	King mandarin	<i>Citrus nobilis</i> Lour.	1
LAP	Low acid pummelo	<i>Citrus maxima</i> (Burm.) Merr.	1
PKM	Ponkan mandarin	<i>Citrus reticulata</i> Blanco	1
SSO	Sour orange	<i>Citrus aurantium</i> L.	1
SWO	Sweet orange	<i>Citrus sinensis</i> (L.) Osbeck	1
UNS*	Satsuma mandarin	<i>Citrus unshiu</i> Marcovitch	1
WLM	Willowleaf mandarin	<i>Citrus reticulata</i> Blanco	1
WMM	W Murcott mandarin	<i>Citrus reticulata</i> Blanco	1

*provided by National Institute of Genetics, Japan

Table 3 Accession types of 196 and 871 accessions analyzed with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

CVC Type	Axiom® Citrus HD Genotyping Array	Axiom® Citrus Genotyping Array
Blood orange	3	17
Calamondin	1	3
Citrango	2	24
Citragelo	0	1
Citron	12	20
Citrumelo	1	8
Clementine	3	16
Eureka lemon	2	14
Grapefruit	6	34
Kumquat	2	6
Lemon	12	49
Lime	5	26
Lisbon lemon	3	13
Mandarin	35	94
Navel orange	6	73
Papeda	3	7
Pummelo	25	58
Rangpur lime	3	17
Rough lemon	6	15
Satsuma	3	45
Sour orange	5	26
Sweet orange	3	71
Tangelo	3	29
Tangor	5	18
Trifoliate	6	33
Valencia orange	4	26
Hybrids	7	75
Others	30	55

Table 4 Chromosomal distribution of SNPs represented on Axiom® Citrus HD Genotyping Array 1,297,735 SNPs (1,296,733 autosomal, 1002 chloroplast) represented on Axiom® Citrus HD Genotyping Array against *Phytozome C. clementina* v1.0 reference genome. (a) Distribution of 1,297,735 SNPs represented on Axiom® Citrus HD Genotyping Array per each of 9 chromosomes. (b) SNP targets of Axiom® Citrus HD Genotyping Array against *Phytozome C. clementina* v1.0 reference genome

a)

Linkage Group	Probe count	Size
1	137,783	28.94 Mb
2	171,732	36.38 Mb
3	232,654	51.05 Mb
4	132,515	25.65 Mb
5	152,556	43.3 Mb
6	120,276	25.61 Mb
7	115,841	21.13 Mb
8	107,769	25.11 Mb
9	125,607	31.41 Mb
Chloroplast	1,002	

b)

SNP Count	Phytozome Annotation
336	Overlap with START Codons
638	Overlap with STOP Codons
4,815	Overlap with Splice Sites
665,912	Intergenic
630,821	Gene Body
245,120	Introns
385,701	Exons
630,821	mRNA
291,849	CDS
93,852	UTR
275	Transcription Start Site

Table 5 Summary of SNP classes represented on Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

Summary count for total probes, total autosomal SNPs, total chloroplast SNPs, SNP density and PHR SNPs after analysis with Axiom® Suite default settings for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. Analysis was performed with 1507 genomic DNA samples that mostly include citrus group.

	Total Probes	Autosomal SNPs	Chloroplast SNPs	SNP Density	PHR
<i>Axiom® Citrus 15AX</i>	1,404,442	1,296,733	1,002	3.9 SNP/kb	728,831
<i>Axiom® Citrus 56AX</i>	58,433	57,933	500	0.16 SNP/kb	51,296

Table 6 Chromosomal distribution of SNPs represented on Axiom® Citrus Genotyping Array 58,433 SNPs, (57,933 autosomal, 500 chloroplast) represented on Axiom® Citrus Genotyping Array against *Phytozome C. clementina* v1.0 reference genome. (a) Distribution of 58,433 SNPs represented on Axiom® Citrus 56AX on the 9 chromosomes. (b) Annotation of SNP targets of Axiom® Citrus Genotyping Array on *Phytozome C. clementina* v1.0 reference genome.

(a)

Linkage Group	Probe Count	Size
1	6,289	28.94 Mb
2	7,773	36.38 Mb
3	10,572	51.05 Mb
4	6,259	25.65 Mb
5	6,665	43.3 Mb
6	5,446	25.61 Mb
7	4,961	21.13 Mb
8	4,608	25.11 Mb
9	5,360	31.41 Mb
Chloroplast	500	

(b)

SNP Count	Phytozome Annotation
8	Overlap with START Codons
24	Overlap with STOP Codons
236	Overlap with Splice Sites
20,144	Intergenic
38,289	Gene Body
14,787	Introns
23,502	Exons
38,289	mRNA
19,550	CDS
3,952	UTR
9	Transcription Start Site

Fig. 1 Work Flow for citrus variant discovery panel selection, SNP detection, filtering, validation, and final selection of SNPs for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

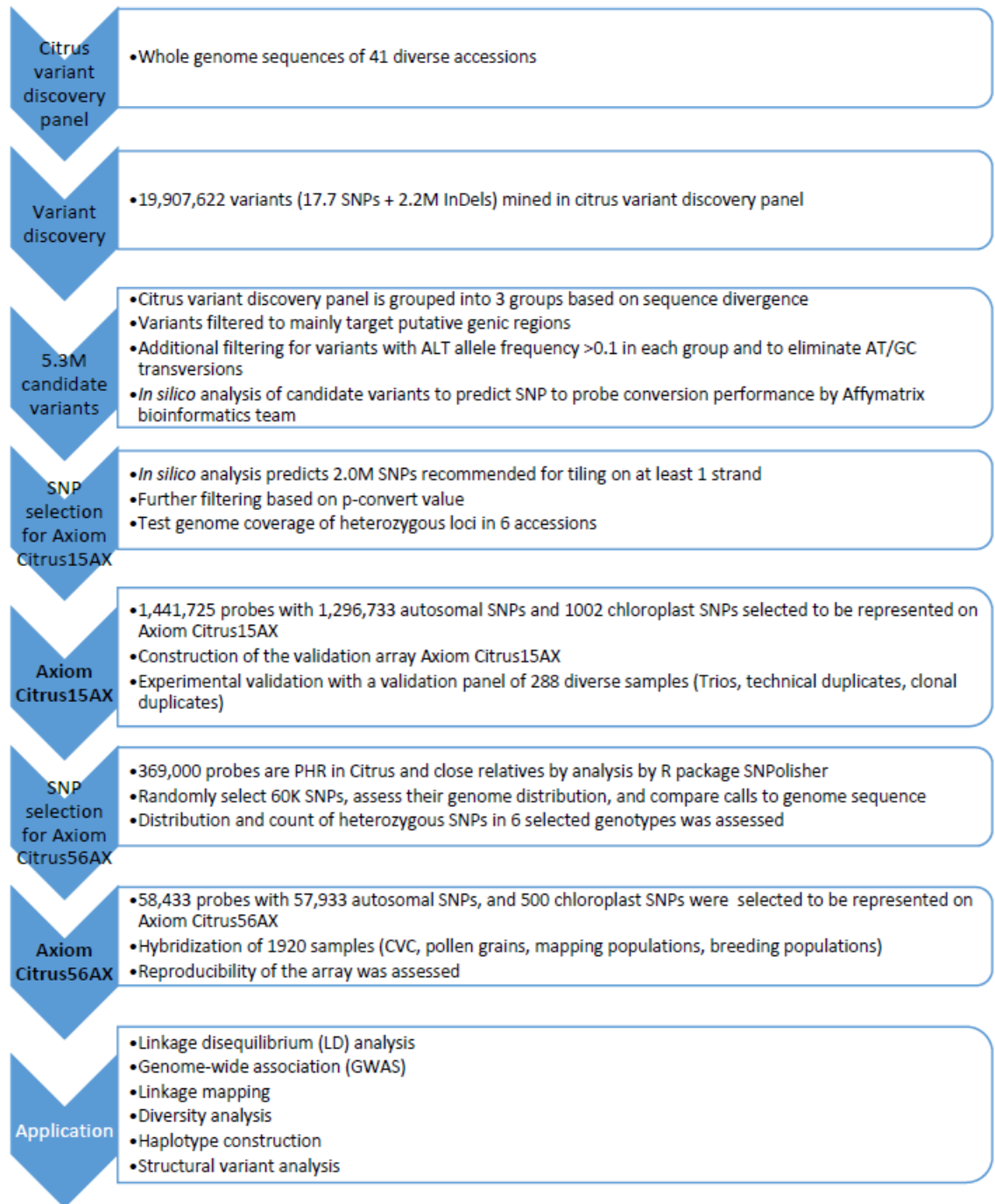


Fig. 2 Proportion of markers on the Axiom® Citrus HD Genotyping Array (a) and Axiom® Citrus Genotyping Array (b) arrays classified in 6 categories by Axiom Analysis Suite with default settings. Categories of markers are PolyHighResolution (yellow), OTV (gray), CallRateBelowThreshold (orange), Hemizygous (Red), other (green), and MonoHighResolution(blue)

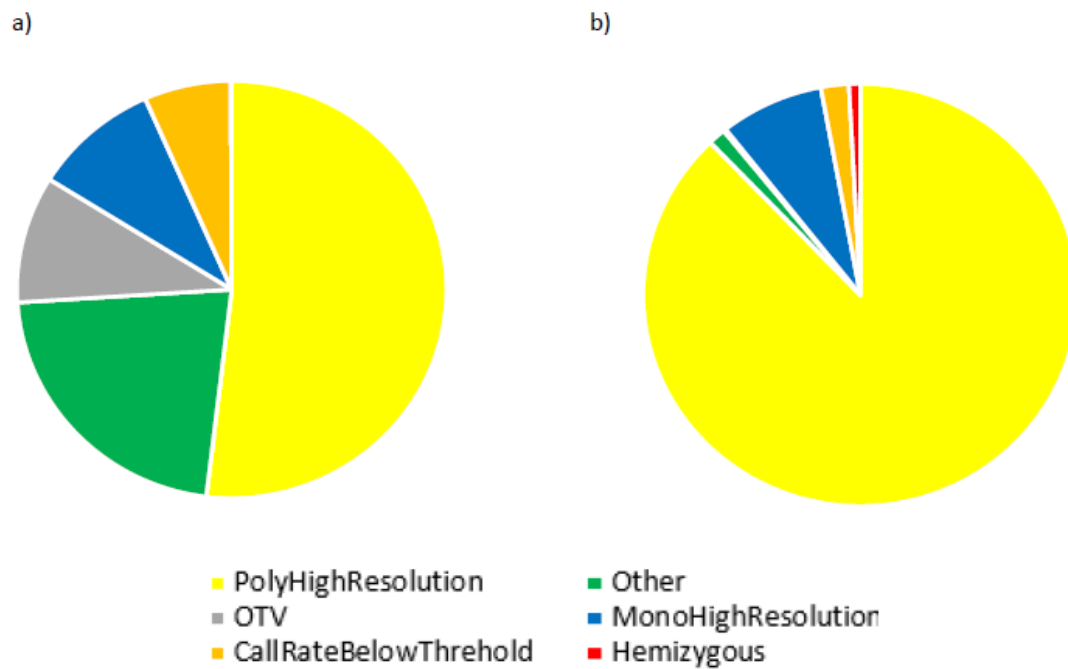


Fig. 3 Genome-wide SNP distribution of Axiom® Citrus HD Genotyping Array, Axiom® Citrus Genotyping Array and heterozygous marker distribution of selected accessions.

SNP distribution of Axiom® Citrus HD Genotyping Array, Axiom® Citrus Genotyping Array and heterozygous markers of selected accessions against Phytozome *C. clementina* v1.0 reference genome represented on a Circos plot. SNP density distribution of Axiom® Citrus HD Genotyping Array (yellow) and Axiom® Citrus Genotyping Array (red) on each of nine chromosomes was binned in non-overlapping 50 kb windows and gene density of *C. clementina* was binned in 50 kb windows (blue). Distribution of heterozygous markers of Chandler pummelo (green), sweet orange (purple) and Pomeroy trifoliate (dark blue) with Axiom® Citrus Genotyping Array data binned in 50 kb windows. There are 4,554 heterozygous markers in Chandler pummelo, 18,159 heterozygous markers in sweet orange, and 1,293 heterozygous markers in Pomeroy trifoliate.



Fig. 4 Comparison of percent heterozygosity estimated with sequence data and Axiom® Citrus HD Genotyping Array data

Percent heterozygosity (proportion of variable sites called heterozygous) was estimated for 38 accessions included in the citrus discovery panel. Highest estimated percent heterozygosity with both sequence data (blue) and array data (orange) was observed with Schaub rough lemon (JAM) and lowest estimated percent heterozygosity was observed with Ethrog citron (ETH). The correlation between sequence data and array data for estimated percent heterozygosity is shown in the plot (upper right corner) with R^2 value of 0.7516.

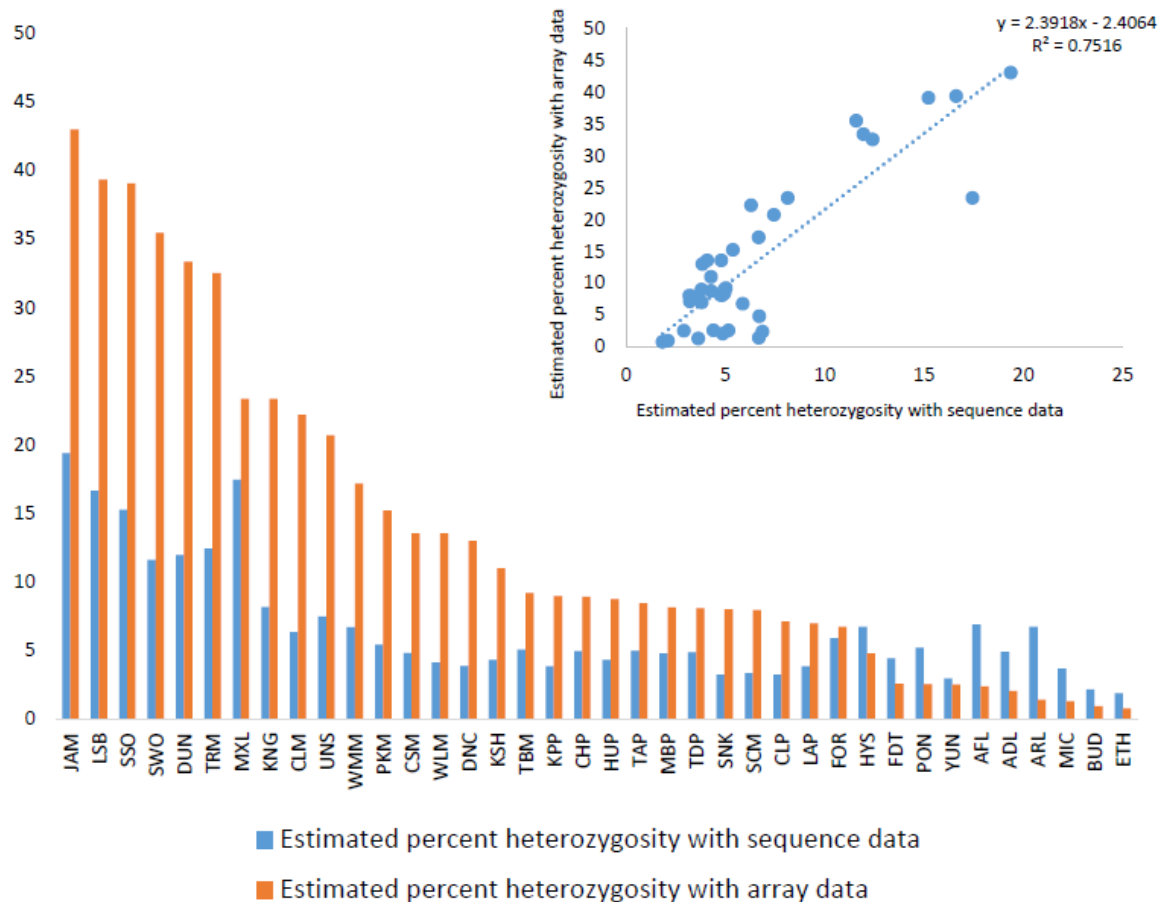


Fig. 5 Genotyping call comparison between sequencing data and Axiom® Citrus HD Genotyping Array data.

Genotyping accuracy of Axiom® Citrus HD Genotyping Array was assessed by comparing genotype calls of 17 accessions from citrus variant discovery panel that were also genotyped with Axiom® Citrus HD Genotyping Array. Genotyping calls at 635,809 PHR loci were considered.

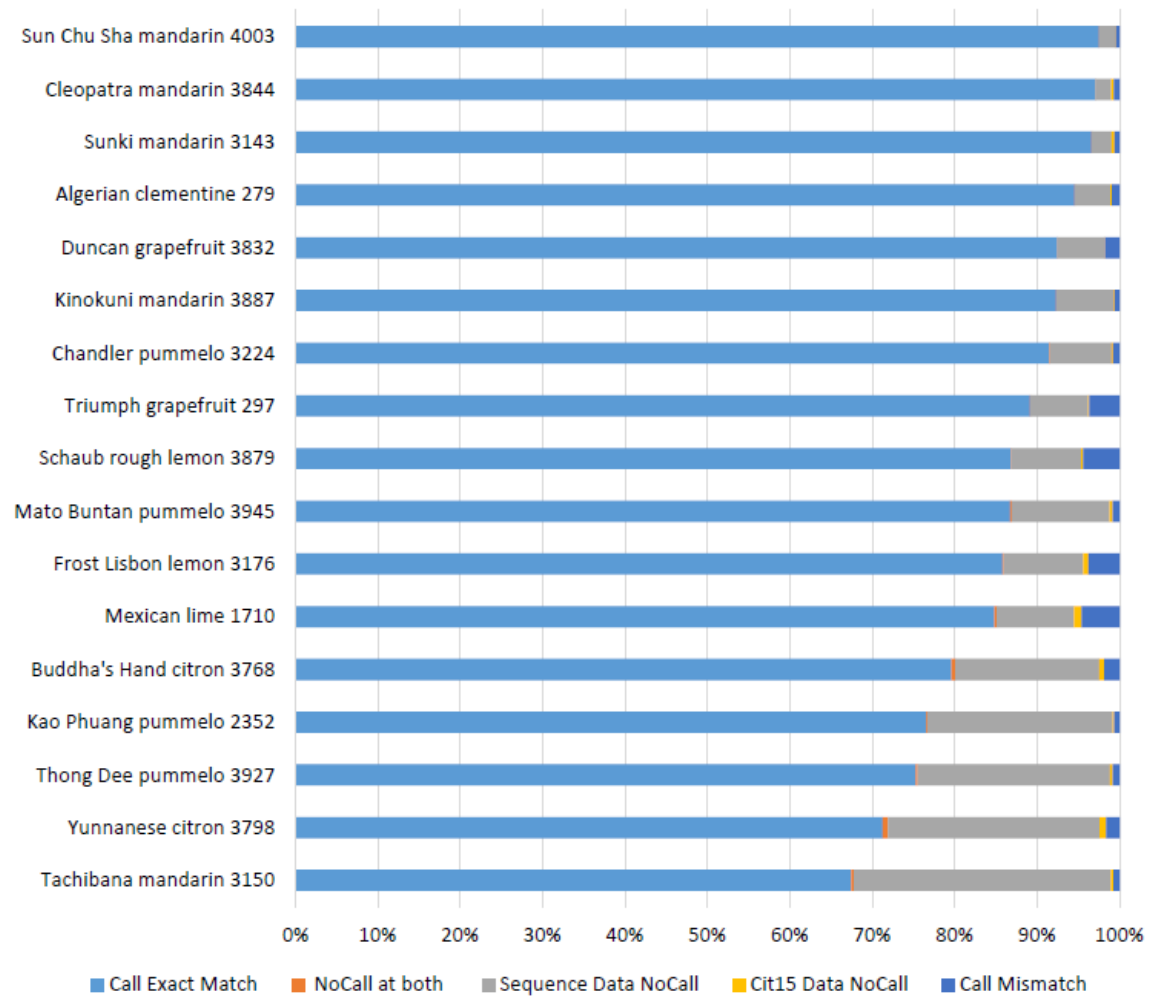


Fig. 6 Reproducibility of genotyping data between Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array was assessed using 28 accessions that were analyzed with both arrays

Genotyping accuracy of Axiom® Citrus Genotyping Array was assessed by comparing genotype calls of 36 accessions from CVC that were also genotyped with Axiom® Citrus HD Genotyping Array. Genotyping calls at 51,296 PHR loci were considered. Figure shows X-axis ranging from 93 to 100%

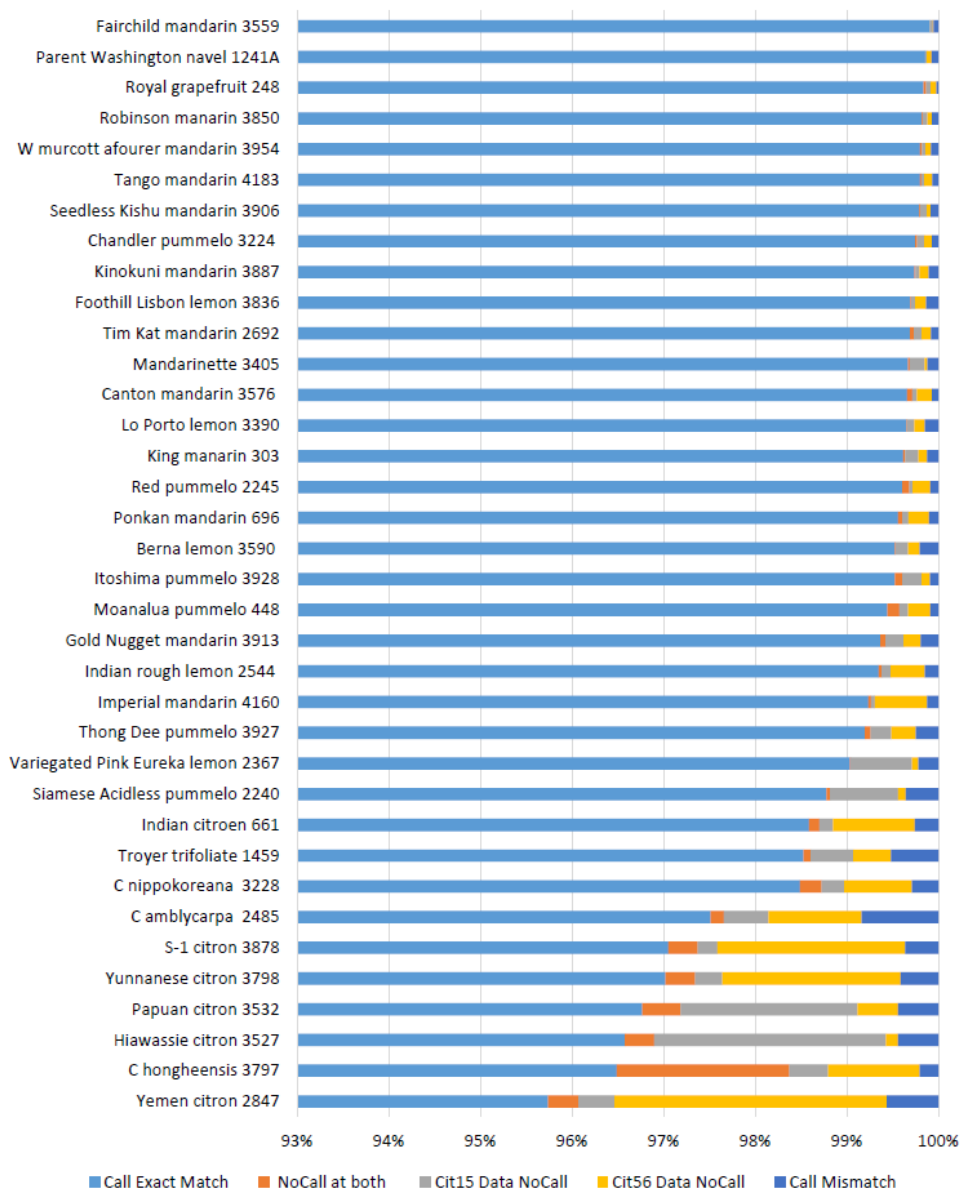


Fig. 7 PCA of citrus variant discovery panel genotyped with Axiom® Citrus HD Genotyping Array. (a) PCA was carried out with 196 genomic DNA samples from CVC (gray) genotyped at PHR loci with Axiom® Citrus HD Genotyping Array. Accessions representing citrus variant discovery panel are highlighted in yellow. PC1(35.6%), PC2(22.5%), and PC3(4.1%) explained 62.2 % of the total variation that exists in the data. The individuals in the citrus variant discovery panel are shown (orange). (b) Cumulative proportion calculated for PCs

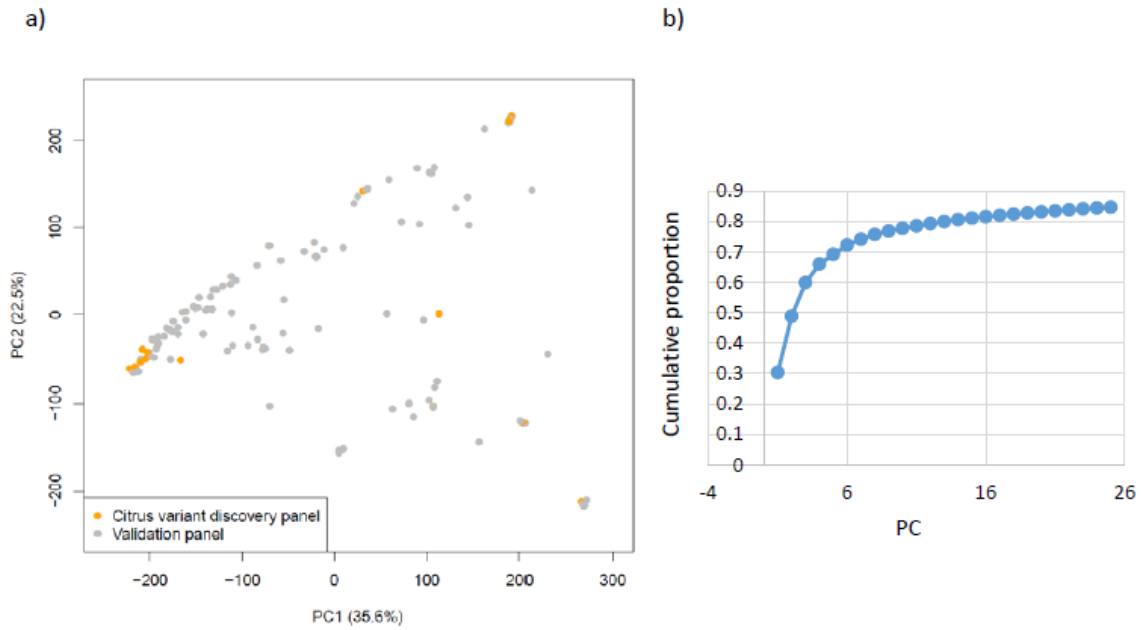


Fig. 8 PCA of 871 diploid accessions in CVC genotyped with Axiom® Citrus Genotyping Array. PCA was carried out with 871 samples genotyped at PHR loci with Axiom® Citrus Genotyping Array. 871 CVC accessions included 20 citron (orange), 94 mandarin (light blue), 58 pummelo (blue), 7 papeda (yellow, 6 kumquat (pink), 11 kumquat hybrid (brown), 33 trifoliate (red), 44 trifoliate hybrid (purple), 4 citrus relative (green) and 594 others (gray). PC1(30.1%), PC2(18.5%), and PC3(11.1%) explained 59.7 % of the total variation that exists in the data. (a) PCA PC1 vs PC2. (b) PCA PC2 vs PC3. (c) PCA PC1 vs PC3. (d) PCA of clonally propagated accessions: navel orange (orange), Valencia orange (dark blue), sweet orange (red), satsuma mandarin (purple), clementine mandarin (green), and Eureka/Lisbon lemon (blue). (e) eigenvalue calculated for PCA

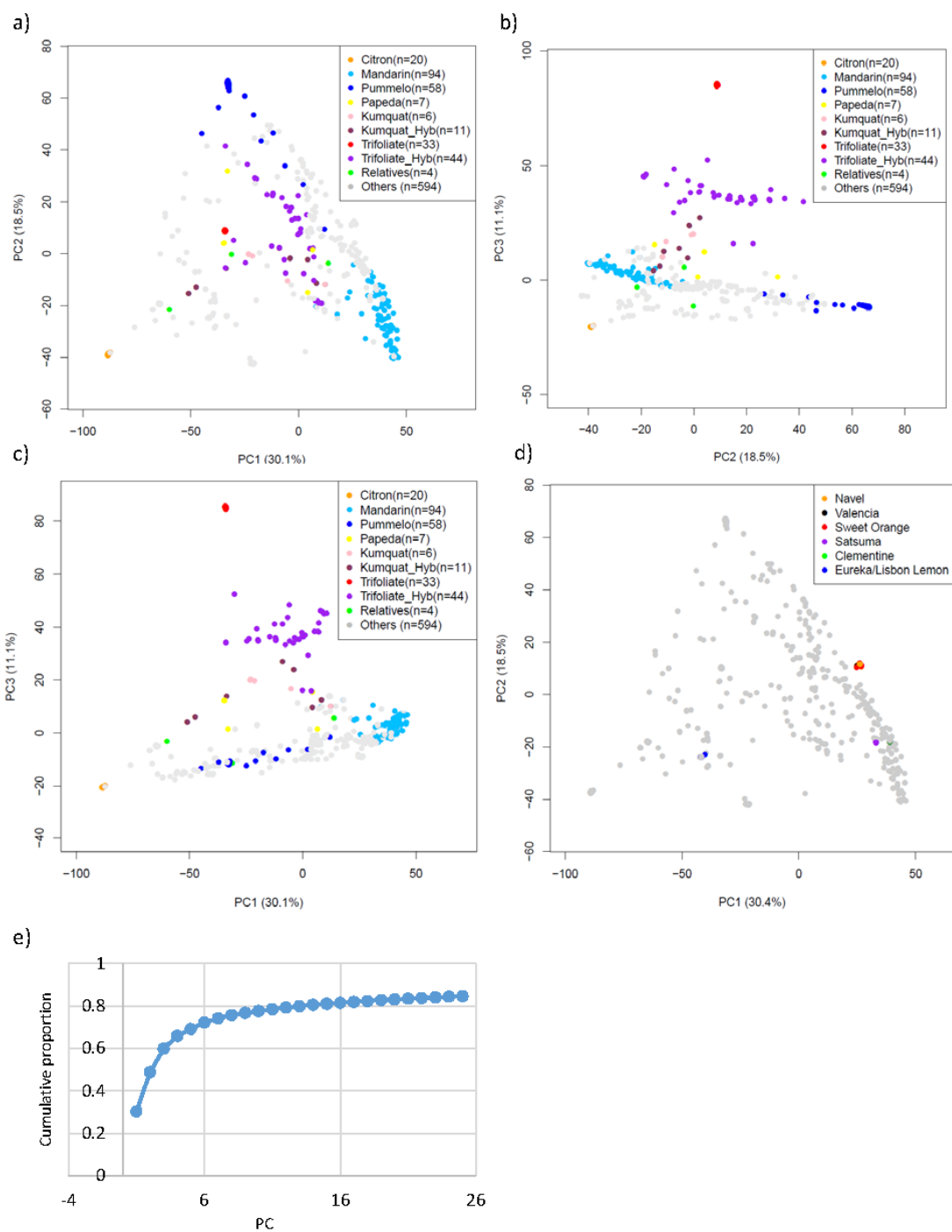


Fig. 9 Maximum likelihood phylogeny trees of 26 accessions in Citrus and close relatives in the citrus variant discovery panel.

Maximum likelihood trees were generated by pipeline SNPhylo and visualized by FigTree v1.4.3. Bootstrap values based on 1000 replications are indicated on the branch nodes. Bootstrap value of above 90% is indicated in bold line. (a) Tree generated by Axiom® Citrus Genotyping Array data based with 41,626 concatenated SNPs. (b) Tree generated by sequence data with 32,754 SNPs also analyzed with Axiom® Citrus Genotyping Array.

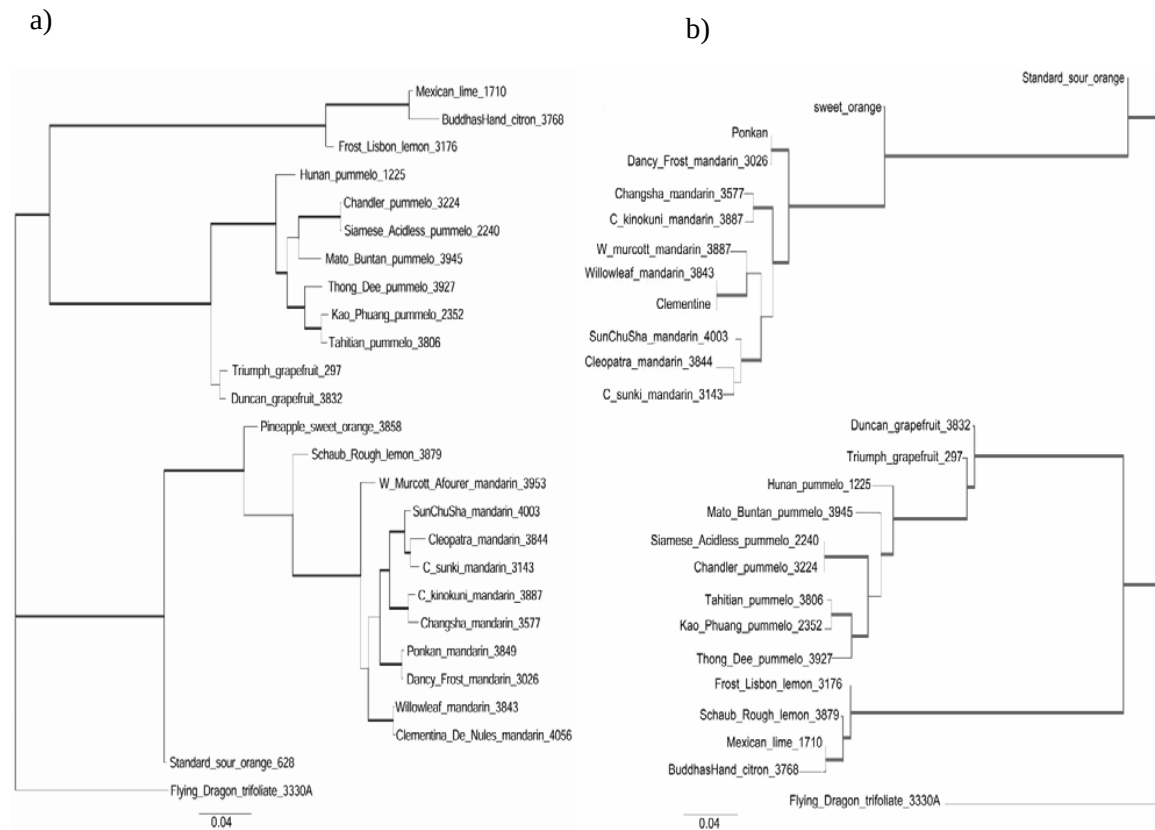


Fig. 10 Maximum likelihood phylogeny tree of 87 accessions in Citrus and close relatives with Axiom® Citrus Genotyping Array data.

Phylogenetic tree was generated with Axiom® Citrus Genotyping Array data based on 41,626 concatenated SNPs generated by SNPhylo. 87 accessions considered for the analysis includes non-hybrid and non-admixed accessions in CVC based on heterozygosity and preliminary admixture results. Maximum likelihood trees were generated by pipeline SNPhylo and visualized by FigTree v1.4.3. Bootstrap values based on 1000 replications are indicated on the branch nodes. Bootstrap value of above 90% is indicated in bold line.

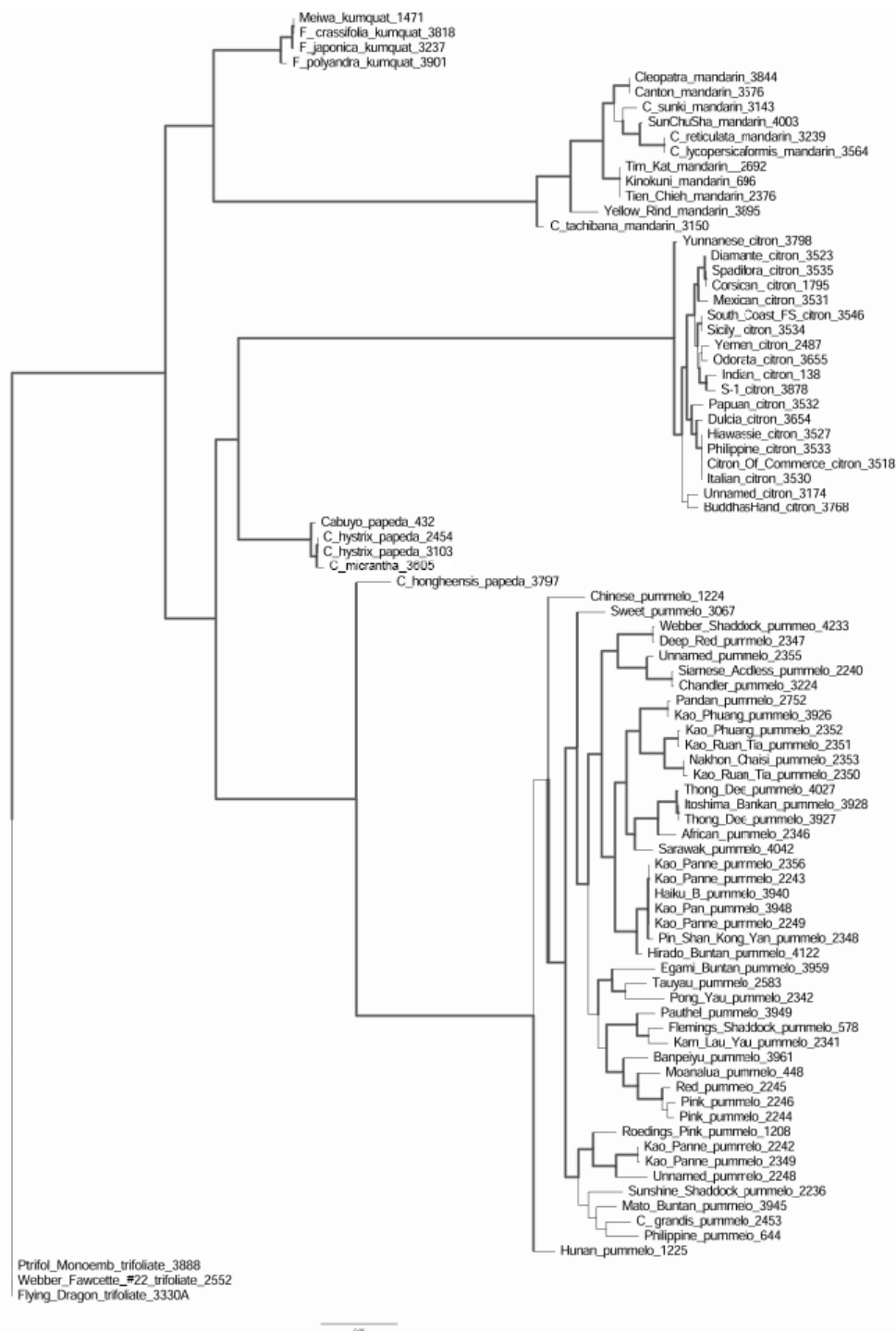


Fig. S1

Venn diagram of 5.3 M (SNPs and INDELs) variants of initial candidate marker loci for designing Axiom® Citrus HD Genotyping Array.

Venn diagram shows the number of polymorphic variants in each group of the variant discovery panel, after variant filtering part 1, submitted to Affymetrix bioinformatics team for *in silico* analysis of the variants to predict conversion performance in the array

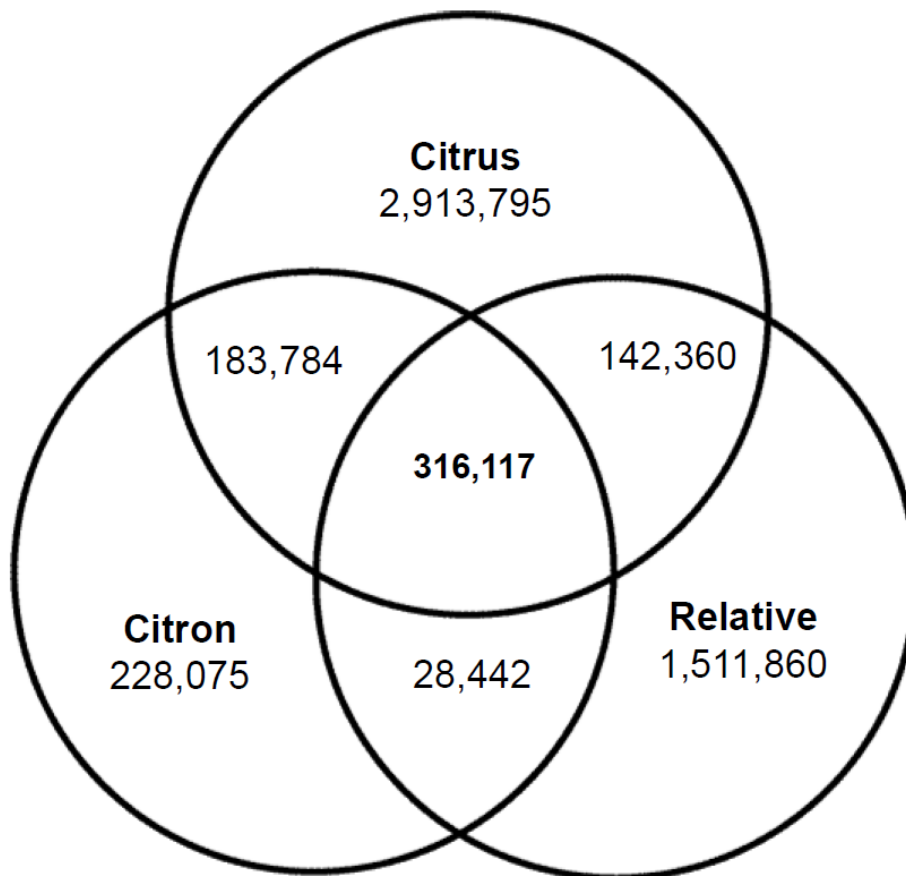


Fig. S2

Venn diagram of variant sites used as QC probes (20,000) for Axiom® Citrus HD Genotyping Array

Venn diagram shows the number of polymorphic variants, after variant filtering part 1, contributed by each group represented in the discovery panel and used as QC probes on Axiom® Citrus HD Genotyping Array

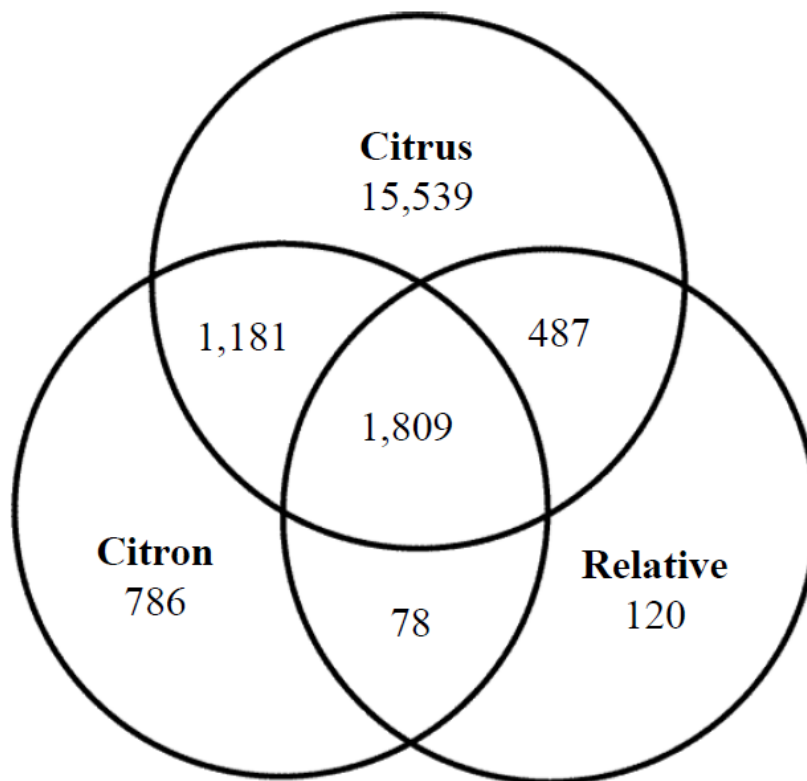


Fig. S3

Venn diagram of 1.4M variants for Axiom® Citrus HD Genotyping Array.

Venn diagram shows the number of polymorphic variants, after variant filtering part 1, contributed by each group represented in the discovery panel and used on Axiom® Citrus HD Genotyping Array

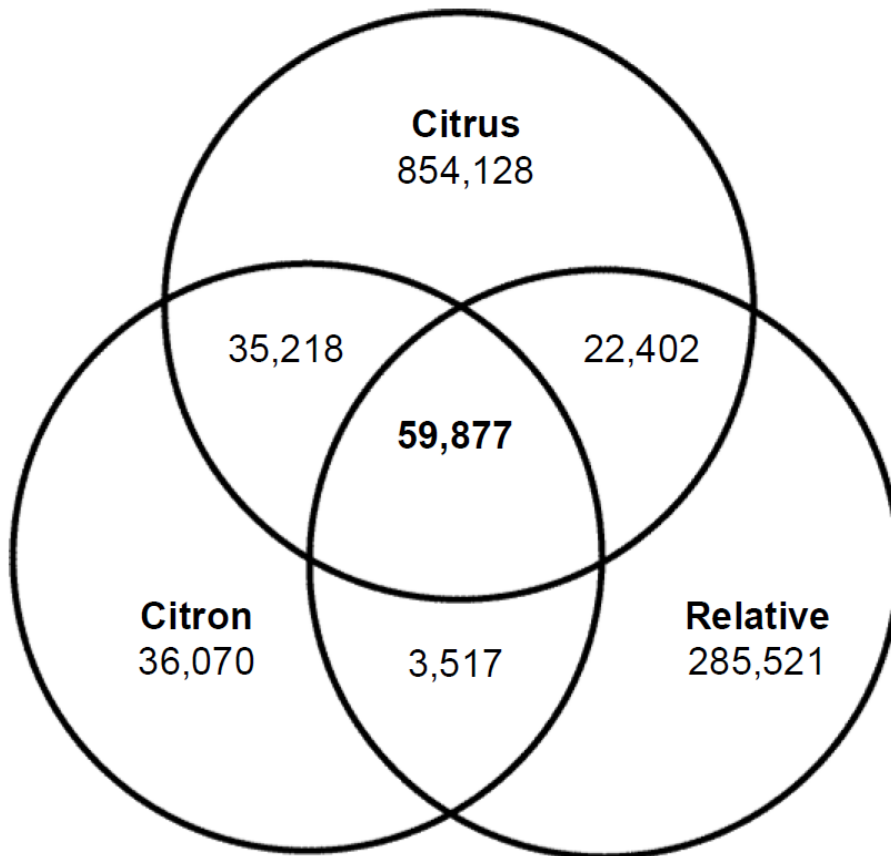
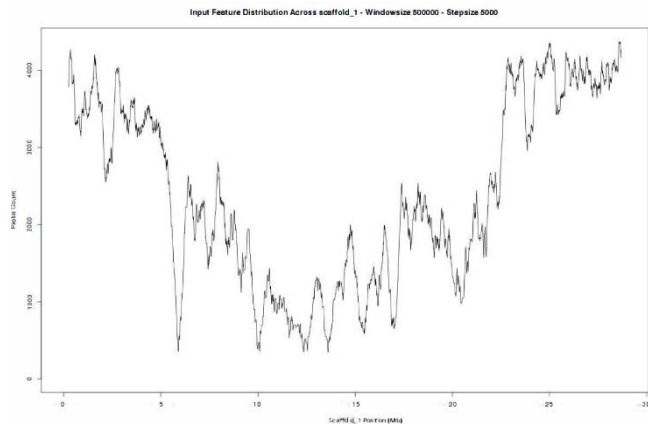


Fig. S4

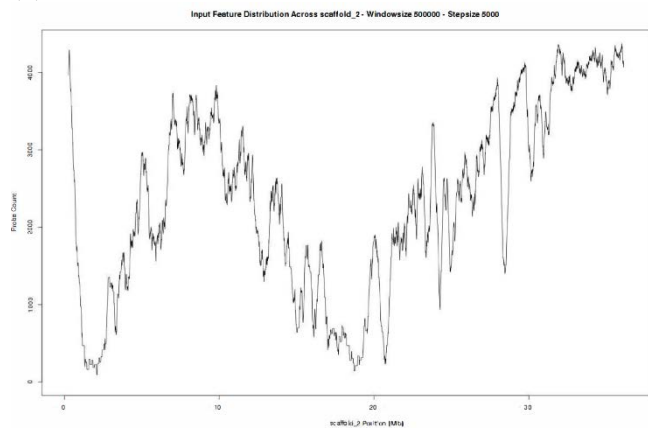
Distribution of SNPs represented on Axiom[®] Citrus HD Genotyping Array on each of nine chromosomes.

SNP distribution is plotted against Phytozome *C. clementina* v1.0 reference genome on chromosome 1-9 (a-i) with 50kb window size and 5kb step size

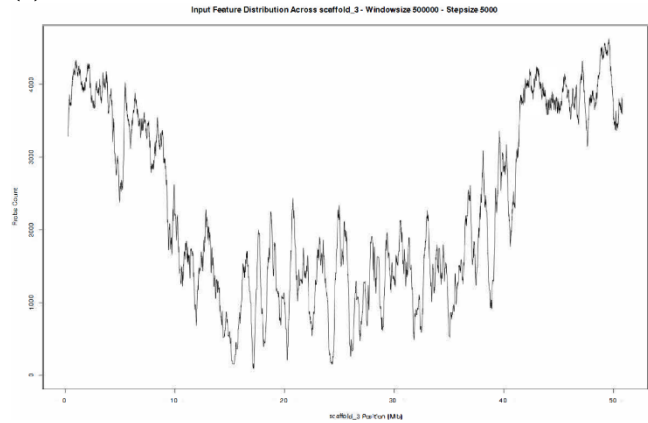
(a)



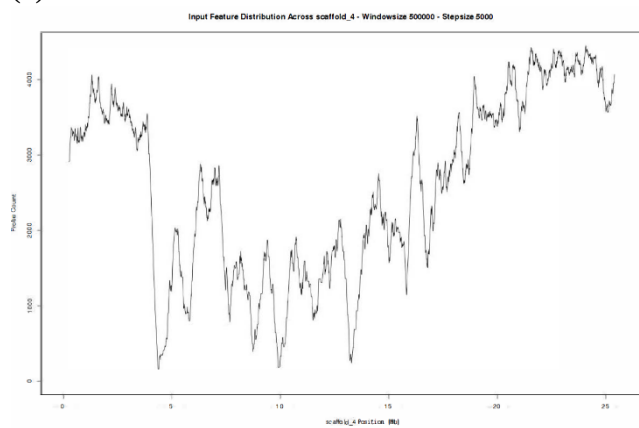
(b)



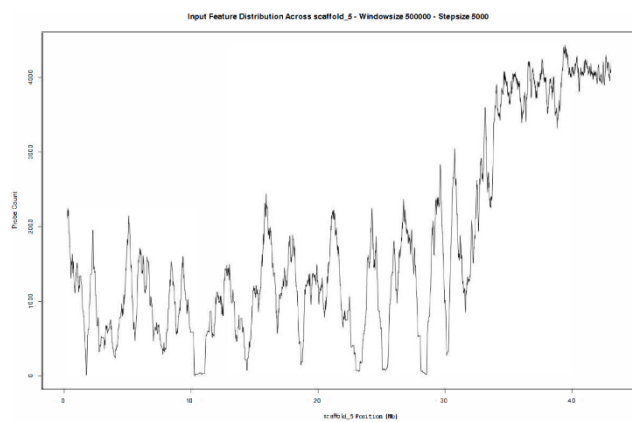
(c)



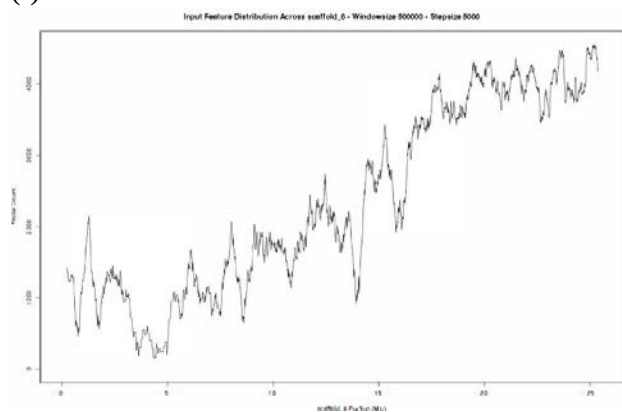
(d)



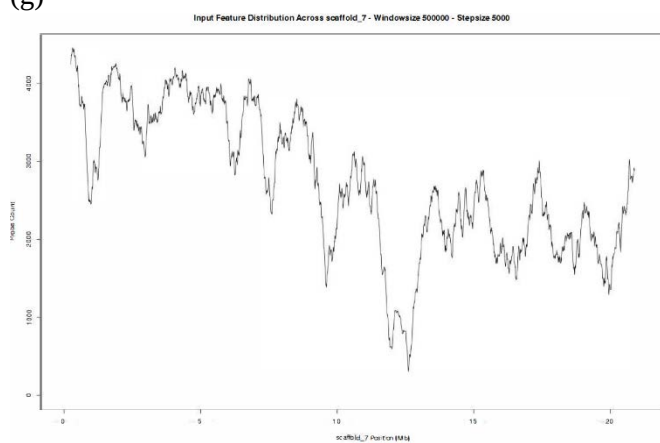
(e)



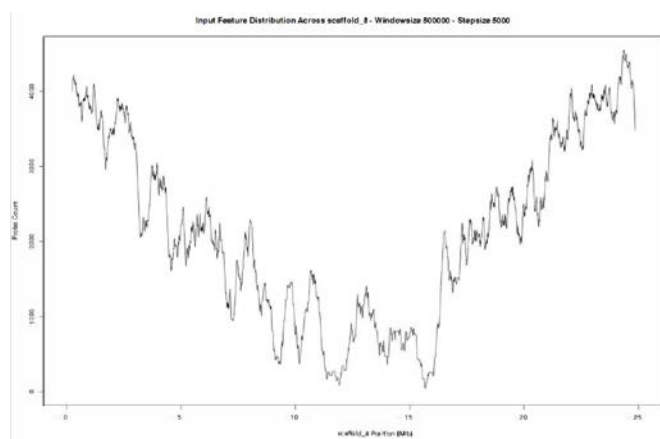
(f)



(g)



(h)



(i)

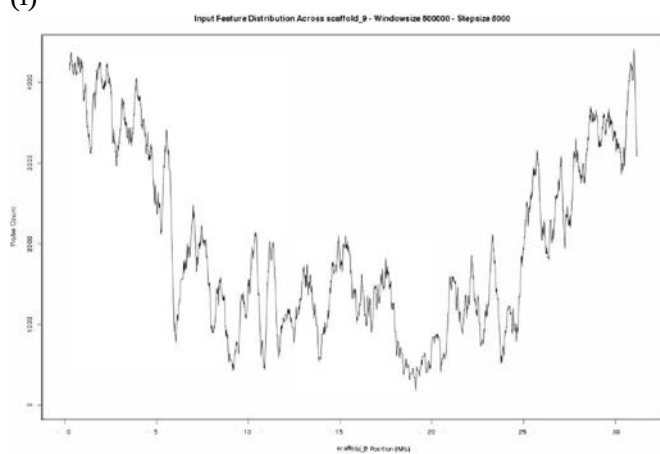


Fig. S5

Venn diagram of 57,933 variants for Axiom® Citrus Genotyping Array.

Venn diagram showing the number of polymorphic variants, after variant filtering part 1, contributed by each group represented in the discovery panel and used on Axiom® Citrus Genotyping Array

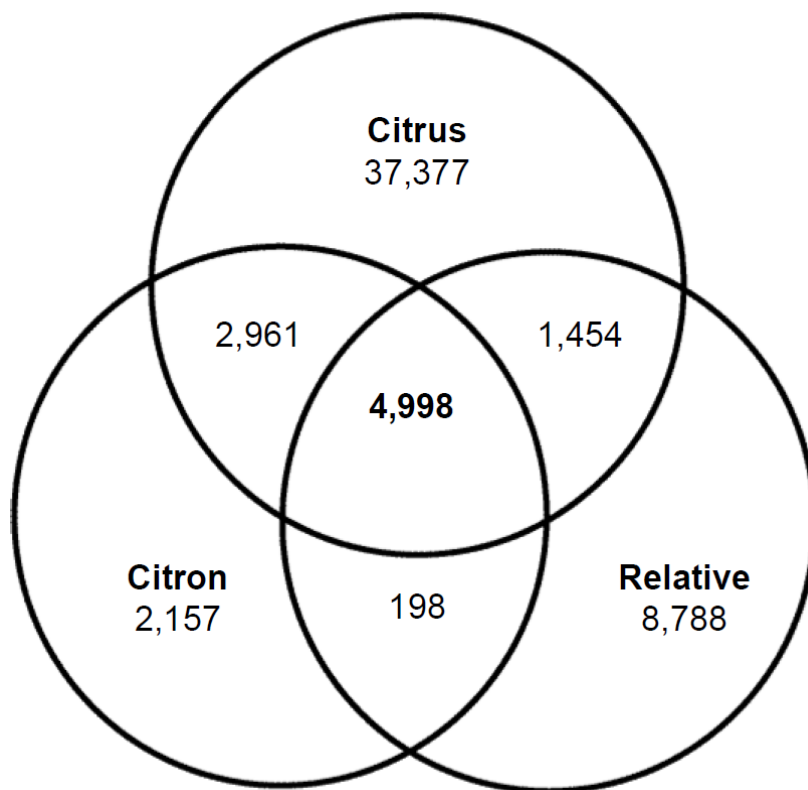
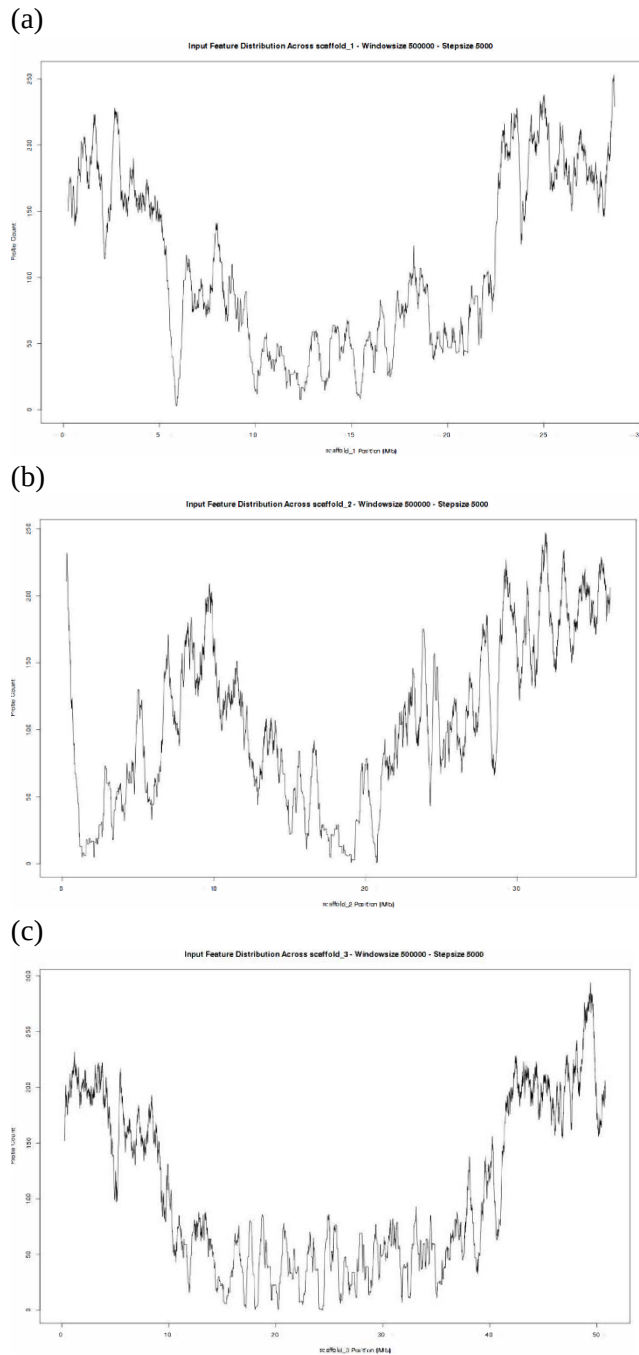
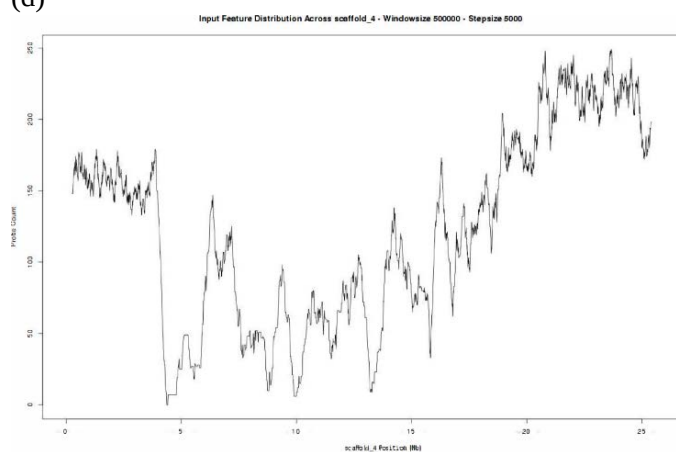


Fig. S6

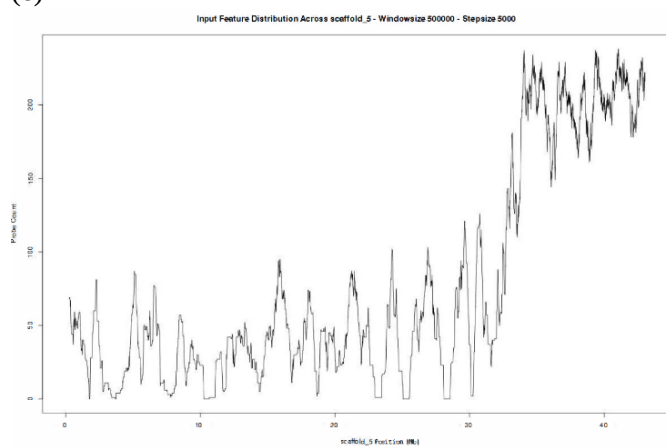
Distribution of SNPs represented on Axiom® Citrus Genotyping Array on each of nine chromosomes. SNP distribution is plotted against Phytozome *C. clementina* v1.0 reference genome on chromosomes 1-9 (a-i) with 50kb window size and 5kb step size



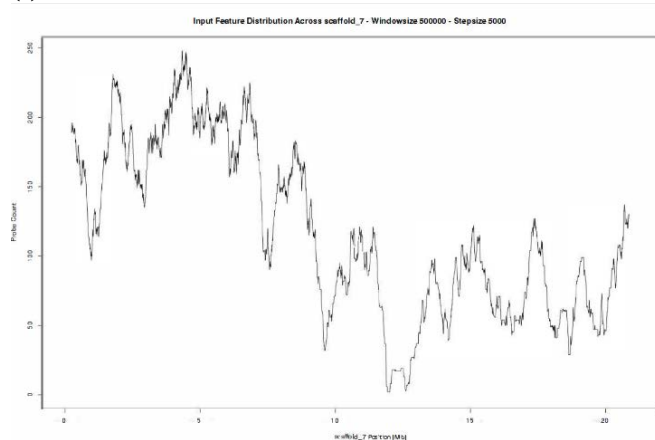
(d)



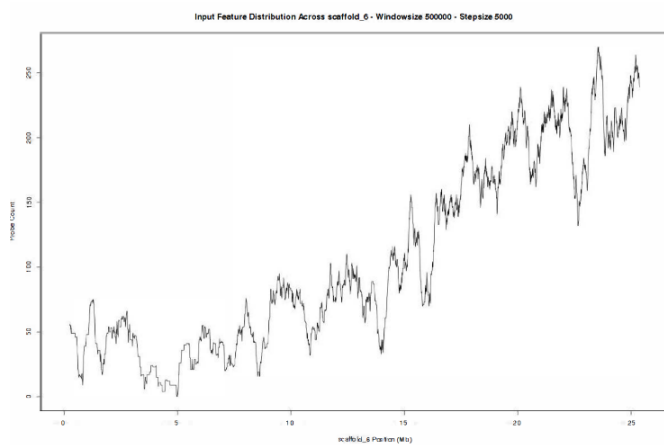
(e)



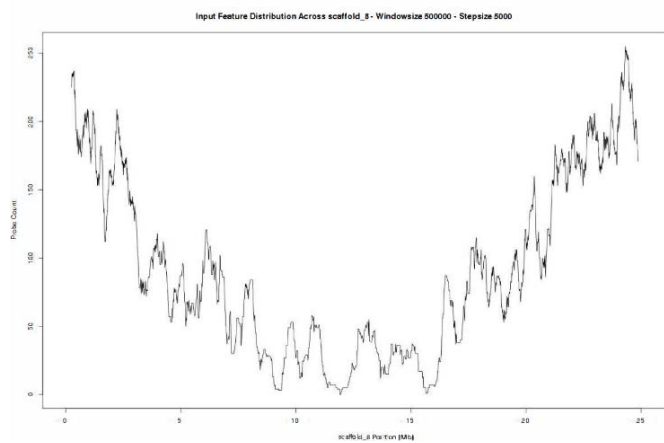
(f)



(g)



(h)



(i)

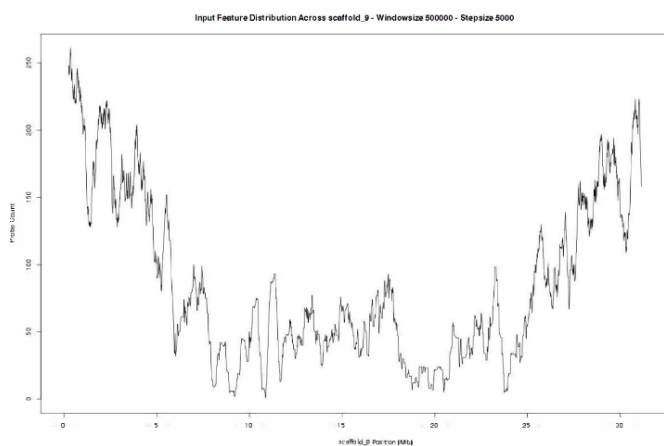
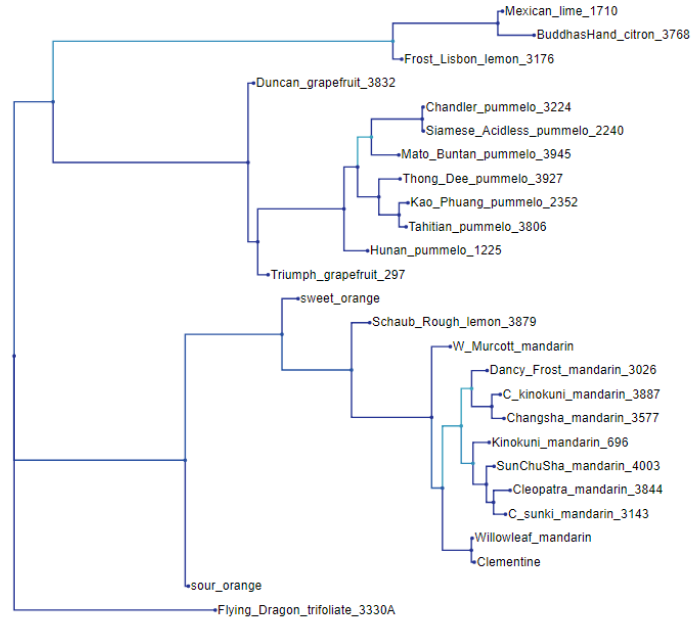


Fig. S7

Topological comparison of two trees generated with data with Axiom[®] Citrus Genotyping Array and sequence data of 26 accessions in citrus variant discovery panel.

Phylogenetic trees with the same 26 accessions and with different topologies were compared side by side with Phylo.io. Branch color scheme of yellow (0) to blue (1) indicates the similarities between the two trees. a) Tree generated by Axiom[®] Citrus Genotyping Array data based on 41,626 concatenated SNPs. (b) Tree generated by sequence data with 32,754 SNPs

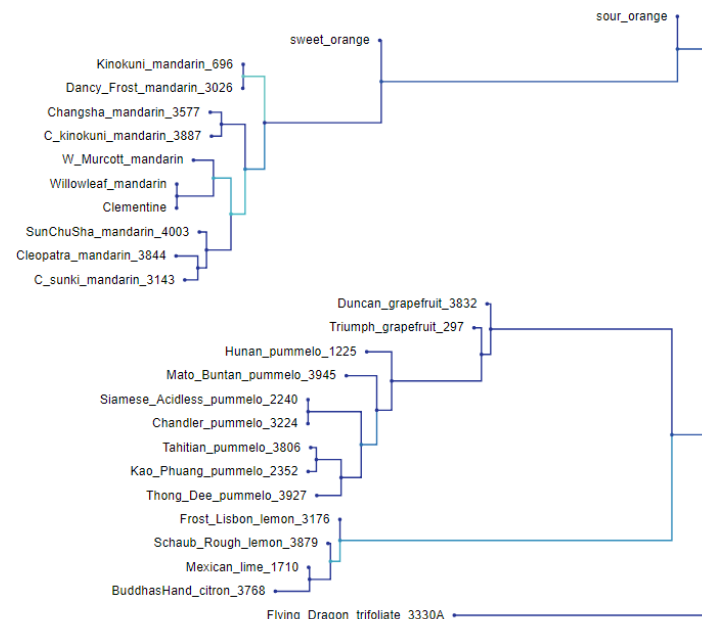
a)



0.12



b)

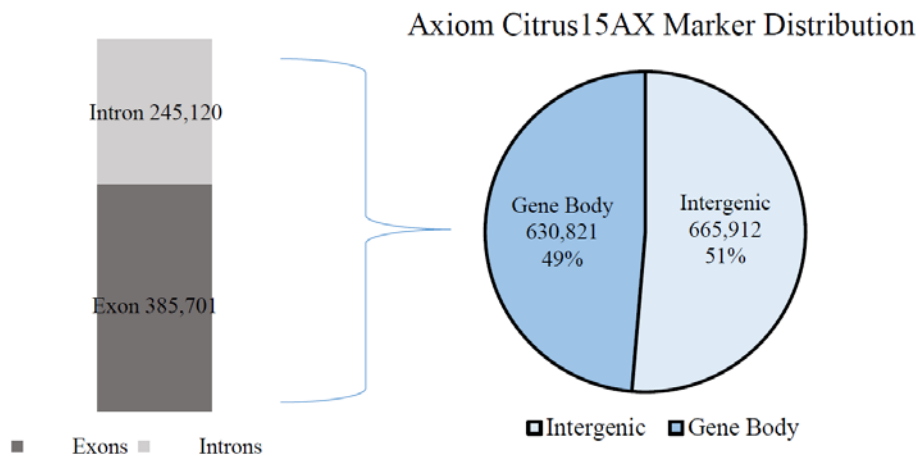


0.12

Fig. S8

Percentage distribution of markers represented on Axiom® Citrus HD Genotyping Array. (a) Percent distribution of Axiom® Citrus HD Genotyping Array markers in gene body versus intergenic regions against Phytozome *C. clementina* v1.0. (b) Percent distribution of Axiom® Citrus HD Genotyping Array markers in genome annotation against Phytozome *C. clementina* v1.0

(a)



(b)

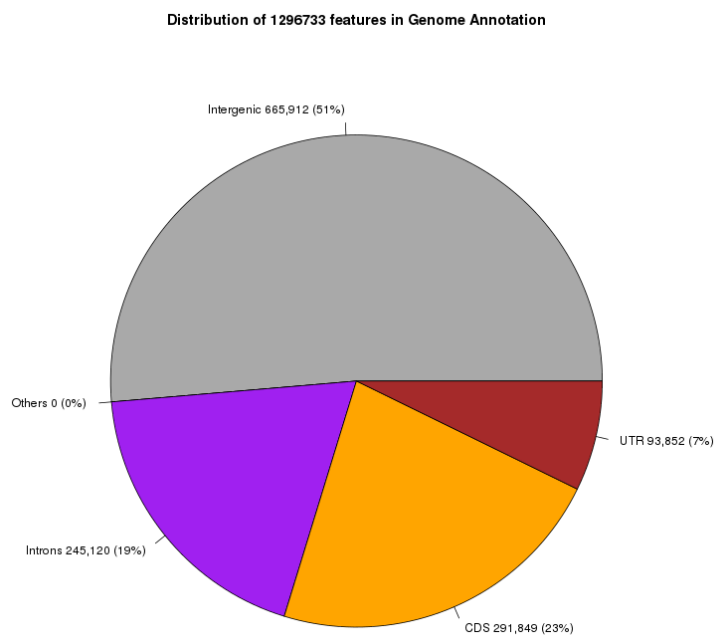
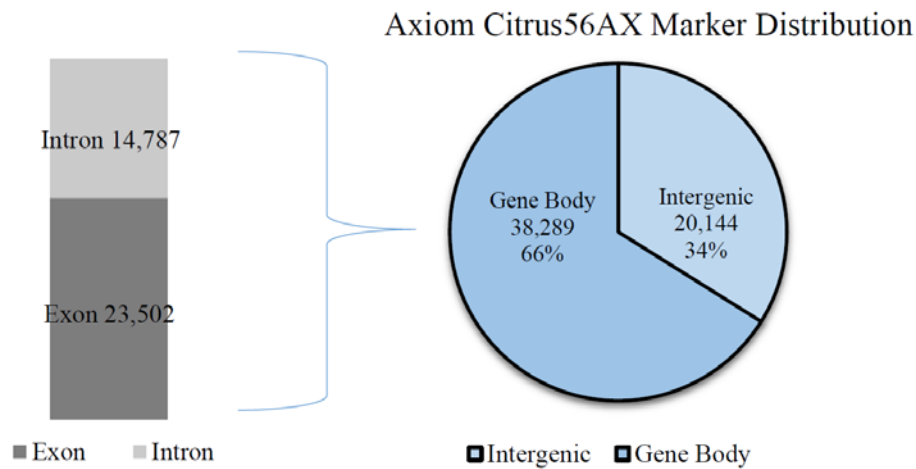


Fig. S9

Percentage distribution of markers represented on Axiom® Citrus Genotyping Array

(a) Percent distribution of Axiom® Citrus Genotyping Array markers in gene body versus intergenic regions against Phytozome *C. clementina* v1.0. (b) Percent distribution of Axiom® Citrus Genotyping Array markers in genome annotation against Phytozome *C. clementina* v1.0.

(a)



(b)

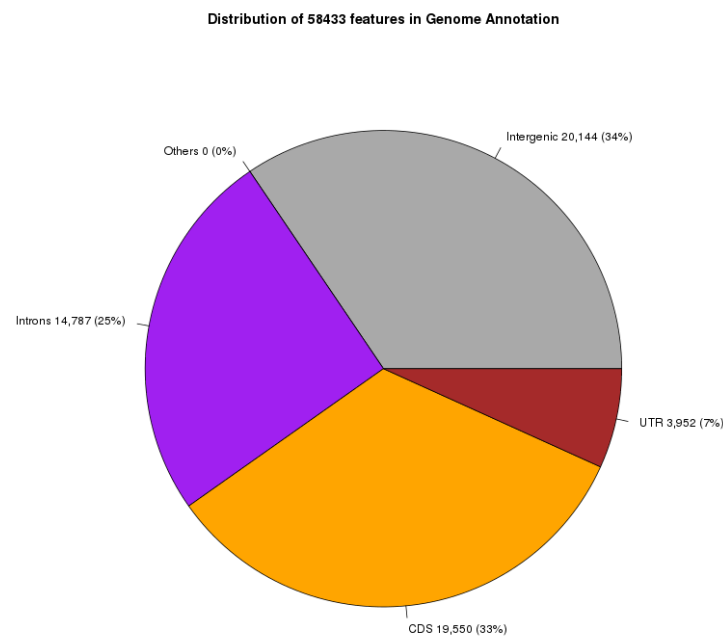


Fig. S10

Genotyping error simulation of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo Citrange CRC 2863.

26 CVC accessions included 1 citron (orange), 9 mandarin (yellow), 6 pummelo (green), 1 trifoliolate (blue), 1 trifoliolate hybrid (dark blue), 1 clementine (gray), 2 grapefruit (pink), 1 lemon (brown), 1 lime (light blue), 1 rough lemon (red), 1 sour orange (purple) and 1 sweet orange (black). PCA was performed with 14,470 loci of Axiom® Citrus Genotyping Array and PC1, PC2, and PC3 in the data was compared. Both PC1 (43.7 %) and PC2 (25 %) showed no differences for 0%, 3% and 5 % error. PC3 was 14.9 %, 14.7 % and 14.8 % for 0%, 3% and 5 % error, respectively. a) PCA with no genotyping call errors. b) PCA with 3 % genotyping error in Flying Dragon CRC 3330A and Carrizo CRC 2863. c) PCA with 5 % genotyping error in Flying Dragon CRC 3330A and Carrizo CRC 2863.

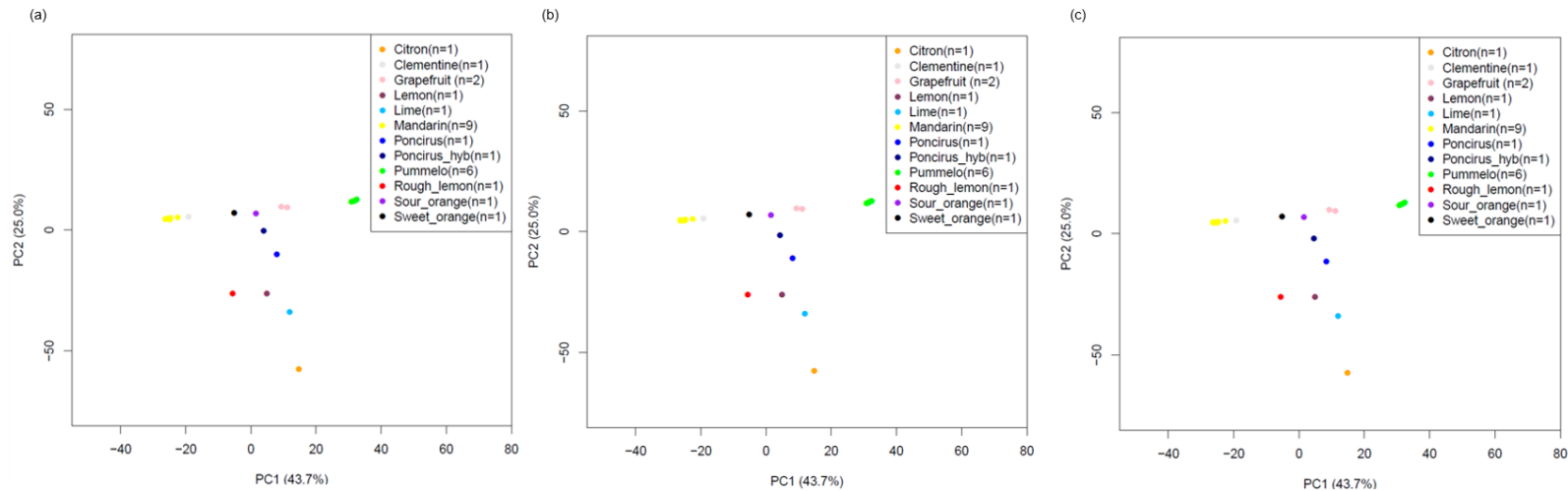


Fig. S11

Genotyping error simulation of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo citrange CRC 2863. Neighbor-joining trees were constructed with 14,470 loci of Axiom® Citrus Genotyping Array and pairwise tree statistical comparison was performed. Carrizo citrange 2863 was not represented in the discovery panel, therefore, haplotype data inferred from whole genome amplified single pollen grains genotyped with Axiom® Citrus Genotyping Array was used as a substitute for the sequence data. Topological differences among trees by PH85 were 0.003595658, 0.002188715, and 0.001411884 for 0 % vs 5 % error, 0 % vs 3% error, and 3 % vs 5% error, respectively and found statistically not significant. a) Neighbor-joining tree with no miss-matched genotyping call against sequence data and haplotype data. b) Neighbor-joining tree with 3 % genotyping error in Flying Dragon CRC 3330A and Carrizo CRC 2863 against sequence data and haplotype data. c) Neighbor-joining tree with 5 % genotyping error in Flying Dragon CRC 3330A and Carrizo CRC 2863 against sequence data and haplotype data.

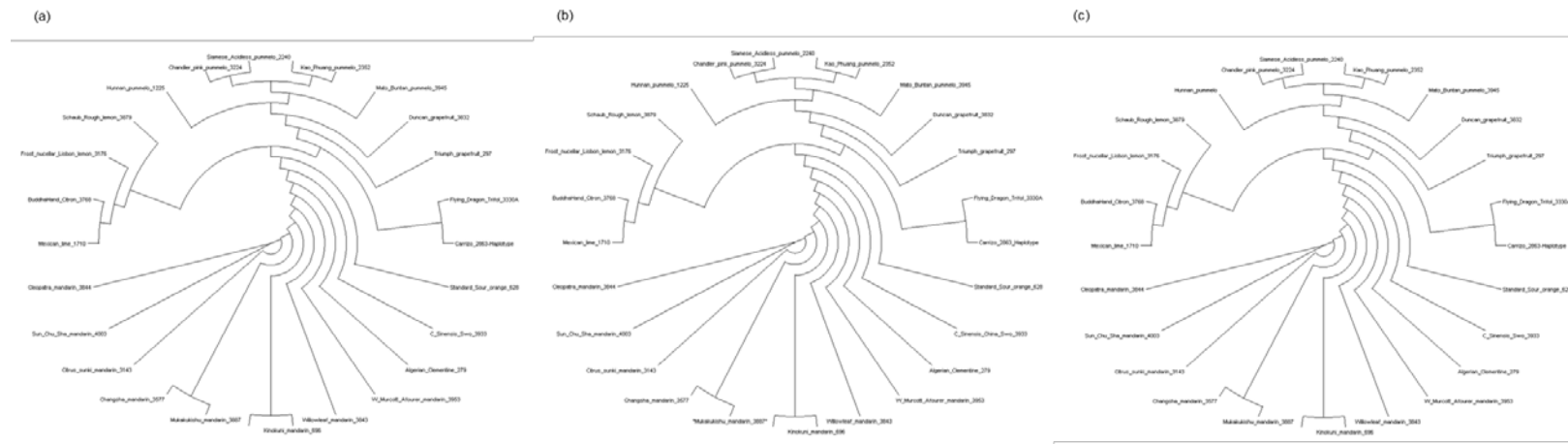


Fig. S12

Genotyping error simulation analysis of 26 CVC accessions including 25 accessions in citrus variant discovery panel and Carrizo Citrange CRC 2863.

ADMIXTURE was performed with 14,470 loci of Axiom® Citrus Genotyping Array. 10-fold cross validation was used for log-likelihood to determine the appropriate number of K in the data. Due to exclusion of some taxa from the analyses, some accessions appear to be admixed with taxa most closely related to the actual donor. There was no difference in the results for 0%, 3% and 5 % error in Flying Dragon and Carrizo at K=4. a) Admixture plot with genotyping call errors at K=4. b) Admixture plot with 3% genotyping call errors at K=4. c) Admixture plot with 5% genotyping call errors at K=4.

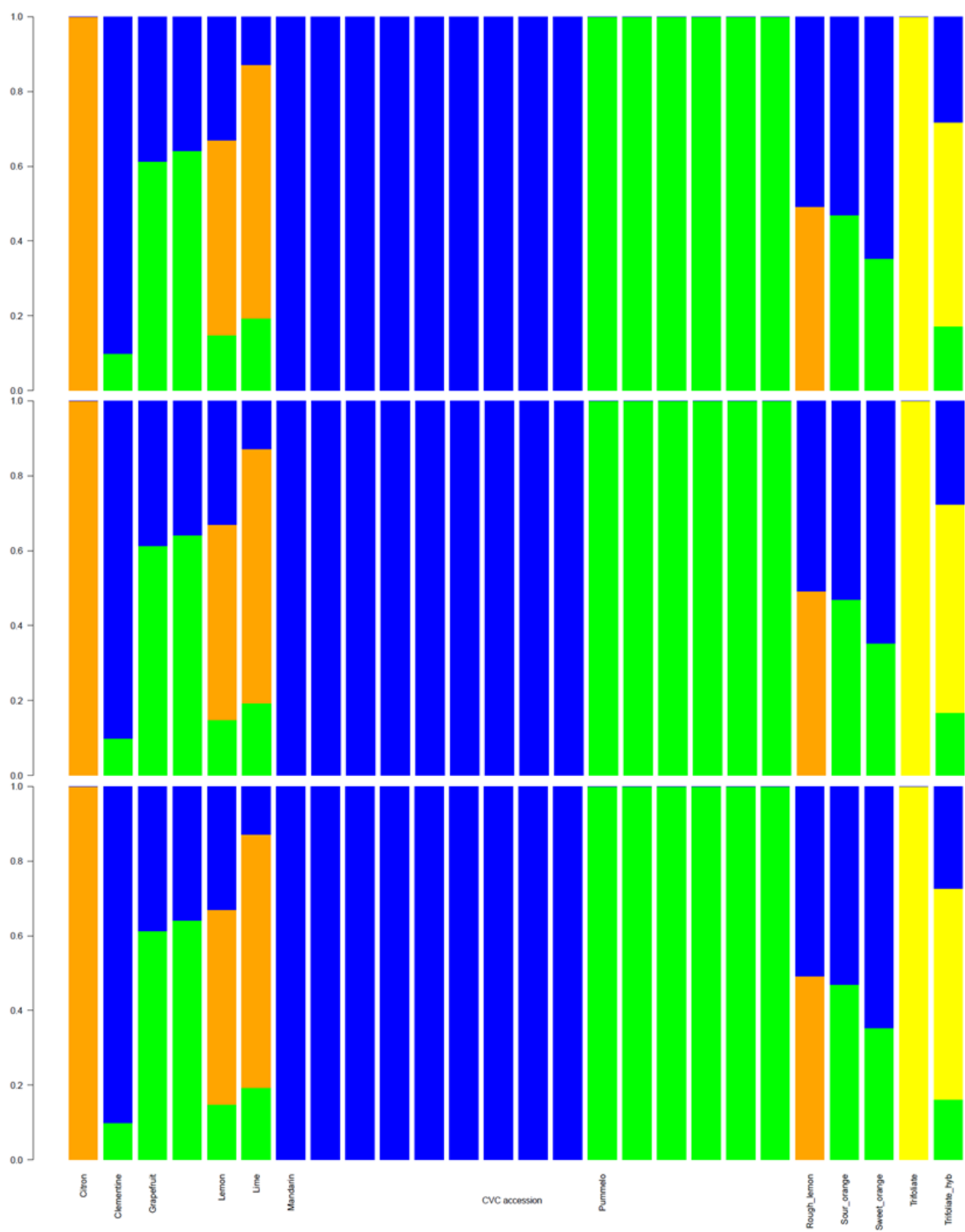


Fig. S13

Neighbor-joining phylogeny tree of 29 accessions in Citrus and close relatives in the citrus variant discovery panel.

Neighbor joining phylogeny tree generated with sequence data based on 2,251,517 concatenated SNPs from sequence data. 29 accessions considered for the analysis included non-hybrid and non-admixed accessions in the variant discovery panel. 2,251,517 SNPs were generated with SNPhylo with filter setting 1 for LD, 0.0 for MAF, and 1 for missing value. Neighbor-joining tree was generated with TASSEL v5.2.31 and visualized with dendroscope 3 (<http://dendroscope.org/>)

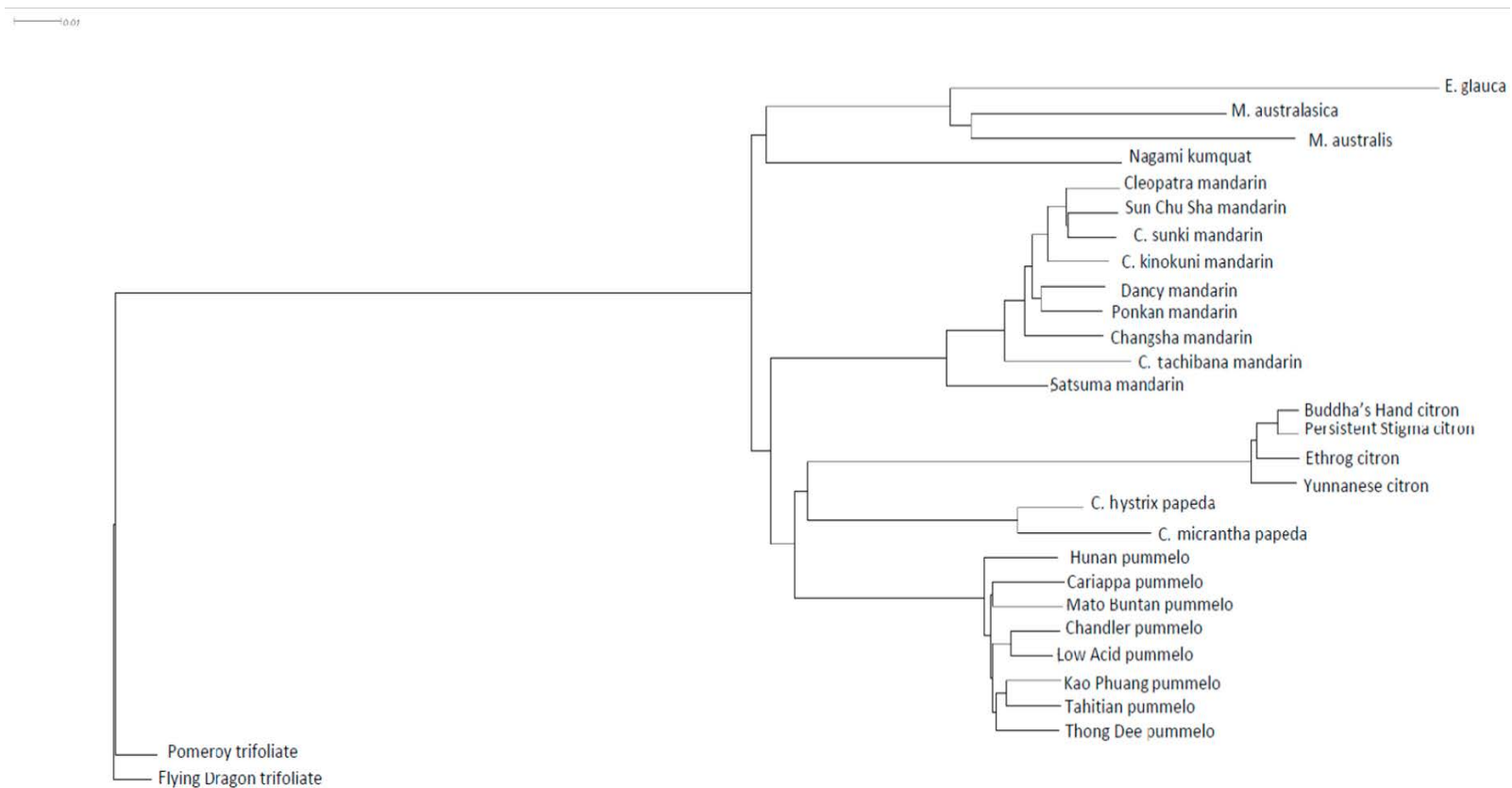


Table S1

Percent heterozygosity at variable bases of accessions in citrus variant discovery panel. Percent heterozygosity was estimated from Variant Call Format (VCF) files of 41 accessions in citrus variant discovery panel and represents the percent of variant positions in the file that were heterozygous in each accession. Group 1 (Citrus, 28 accessions), Group 2 (Citron, 4 accessions), and Group 3 (Citrus relatives, 9 accessions)

Genotype ID	Variety Name	Het calls from Seq	% Heterozygosity	NoCall from Seq	Group
ADL	Australian desert lime	483273	4.87	9981301	3
AFL	Red pulp finger lime	1001776	6.87	5316219	3
ARL	Australian round lime	936095	6.69	5906389	3
BAN	Banpeiyu	703620	4.29	3510948	1
BUD	Buddha's Hand citron (fingered citron)	267659	2.11	7227074	2
CHP	Chandler pummelo	783641	4.94	4053680	1
CLM	Clementine mandarin	1207805	6.30	741438	1
CSM	Changsha mandarin	811384	4.80	3048372	1
CLP	Cleopatra mandarin	615342	3.22	786884	1
CPP	Cariappa CM3 pummelo	749164	5.07	5125608	1
SSO	Sour orange	2714961	15.23	2075827	1
DNC	Dancy mandarin (Frost nucellar)	622367	3.85	3738482	1
DUN	Duncan grapefruit	2148459	11.95	1923954	1
ETH	Ethrog citron	271704	1.84	5115869	2
FDT	Flying Dragon trifoliate	638530	4.4	5405677	3
FOR	Nagami kumquat	947562	5.88	3785107	3
HUP	Hunan pummelo	604913	4.28	5780738	3
HYS	Mauritius papeda	336360	6.71	14893540	3
JAM	Schaub rough lemon	3211273	19.37	3333197	1
KNG	King mandarin	1495489	8.14	1527462	1
KPP	Kao Phuang pummelo	470080	3.8	7523250	1
KSH	Mukakukishu (Kinokuni mandarin)	749336	4.28	2416869	1

Table S1 (cont.)

Genotype ID	Variety Name	Het calls from Seq	% Heterozygosity	NoCall from Seq	Group
LAP	Low acid pummelo	567681	3.82	5043454	1
LSB	Frost nucellar Lisbon lemon	2656909	16.62	3919307	1
MBP	Mato Buntan pummelo	693953	4.76	5324111	1
MIC	Small-fruited papeda	399030	3.64	8942200	1
MXL	Mexican lime	2628024	17.44	4842577	1
PKM	Ponkan mandarin	1024789	5.38	861436	1
PON	Pomeroy trifoliolate	728114	5.16	5791087	1
QIN	Persistent Stigma - OPS citron	373936	2.49	4887646	1
SCM	Sun Chu Sha Kat mandarin	633423	3.32	813007	1
SNK	Sunki mandarin	601625	3.19	1038561	1
SWO	Sweet orange	2116742	11.59	1641714	1
TAP	Tahitian pummelo	785883	4.96	4061901	1
TBM	Tachibana orange	602010	5.03	7929931	1
TDP	Thong Dee pummelo	604569	4.85	7437450	1
TRM	Triumph grapefruit	2145835	12.42	2624569	1
UNS	Satsuma mandarin	1373932	7.45	1459065	1
WLM	Willowleaf mandarin	791813	4.09	539809	1
WMM	W Murcott mandarin	1247065	6.68	1238611	1
YUN	Yunnanese citron	321104	2.92	8921419	1

Table S2

Accessions in Axiom® Citrus HD Genotyping Array validation panel. 232 genomic DNA samples from CVC were selected as a validation panel to be screened with Axiom® Citrus HD Genotyping Array for performance assessment of SNPs and the array

Sample Name	CRC
Aegle marmelos	3140
Algerian clementine	279
Anseikan pummelo	3979
Arajon pummelo	2596
Australian trifoliate	3151
Avana Tardivo di Ciaculli mandarin	4162
Bakrai (mandarin x limetta) hybrid	3776
Balsamocitrus daweei	3514
Banpeiyu pummelo	3961
Bergamot sour orange	2881
Berna lemon	3590
Borneo Rangpur lime	2424
Bouquet de Fleurs sour orange	571
Brazilian sour orange	1689
Buddha's Hand citron	3768
Calamondin	2592
Canton mandarin	3576
Carrizo citrange	2863
Chandler pummelo	3244
Changsha mandarin	3577
Chinese pummelo	1212
Chinotto sour orange hybrid	2375
Christiansen trifoliate	3345
Citrus amblycarpa	2485
Citrus assamensis	3173
Citrus canaliculata	3565
Citrus depressa	2448
Citrus erythrosa	3292
Citrus excelsa	2316
Citrus excelsa	3765
Citrus grandis	2340
Citrus halimii	3780
Citrus hanayu	3469

Table S2 (cont.)

Sample Name	CRC
Citrus hongheensis	3797
Citrus hystrix	3103
Citrus hystrix, Cabuyao	432
Citrus ichangensis	2327
Citrus ichangensis	3931
Citrus keraji	3144
Citrus kharna	3059
Citrus latipes	3052
Citrus leiocarpa	3147
Citrus longispina winged lime	2320
Citrus lycopersicaformis	3564
Citrus macrophylla	3842
Citrus megaloxycarpa	3241
Citrus miaray	3574
Citrus microcarpa	3605
Citrus nippokoreana	3228
Citrus pennivesiculata	2434
Citrus reticulata	3239
Citrus reticulata	3812
Citrus sino-grandis	3148
Citrus succosa	3280
Citrus sudachi	3471
Citrus sunki mandarin	3143
Citrus tachibana	3150
Citrus ujukitsu	3467
Citrus webberii	1455
Citrus webberii var. Montana	767
Citrus yatsushiro	3466
Oroval clementine	4010
Sidi Aissa clementine	3988
Monreal clementine	3773
Cleopatra mandarin	3844
Dancy mandarin	602
Deep Red pummelo	2347
Delta Valencia orange	3963
Diamante citron	3523

Table S2 (cont.)

Sample Name	CRC
Dummett bittersweet orange	2541
Duncan grapefruit	3832
Egami Buntan pummelo	3959
Emerald mandarin	4123
Eremocitrus glauca	3463
Estes rough lemon	3386
Fairchild mandarin	3559
Florida rough lemon	400
Foothill Lisbon lemon	3836
Fortunella crassifolia	1471
Fortunella hindsii	3789
Fortunella hindsii	3790
Fortunella japonica	3237
Fortunella obovata	3475
Fortunella polyandra	3901
Frost nucellar Eureka lemon	3005
Frost nucellar Lisbon lemon	3176
Frost Owari satsuma	3178
Gold Nugget mandarin	3913
Hiawassie citron	3527
Hirado Buntan pummelo	4122
Huang Yen Man Chieh mandarin	3897
Ichang lemon	1215
Imperial mandarin	4160
Improved Meyer lemon	3737
India lime	2450
Indian citron	661
Indian rough lemon	2544
Interdonato lemon	3593
Iraq lemon hybrid	3492
Italian pink fleshed lemon	2899
Itoshima pummelo	3928
Japansche citroen	2875
Kabosu ichandarin	3943
Kao Pan pummelo	2248
Kao Pan pummelo	2355

Table S2 (cont.)

Sample Name	CRC
Kao Phuang pummelo	2352
Kincy mandarin	4144
King tangor	303
Kinnow mandarin	3021
Kinokuni mandarin	696
Kinokuni mandarin (Mukakukishu)	3887
Kiyomi tangor H-12	3952
Lane Late navel orange	3803
Lapithiotiki lemon	4005
Lemonade lemon	4181
Lima sweet orange	950
Limon Real	2317
Little-leaf trifoliolate	4008
Lo Porto lemon	3090
Low thorn Lisbon lemon	NA
Mandarinette	3405
Marsh pink grapefruit	2010
Marsh seedy grapefruit	1565
Marumi (SRA 489)	4125
Mato Buntan pummelo	3945
Mexican lime	1710
Microcitrus warburgina	3782
Microcitrus australasica	3671
Microcitrus australasica	3672
Microcitrus australasica var. Sanguinea	1484
Microcitrus virgata	1485
Miho Wase satsuma	4047
Milam rough lemon	3396
Millsweet limetta	569
Minneola tangelo	3340
Moanalua pummelo	448
Monoembryonic trifoliolate	3888
Moro blood orange	3830
Murraya koenigii	3165
Murraya paniculata	1637
Nagami kumquat	3877

Table S2 (cont.)

Sample Name	CRC
Nagpur orange	3727
Necked mandarin	3738
Nova mandarin	3615
Odorata citron	3655
Okitsu Wase satsuma	3820
Olinda Valencia orange	2750
Olivelands sour orange	2717
Orlando tangelo	3874
Ortanique tangor	3368
Oxanthera neo-caledonica	3793
Page tangelo	3616
Palestine sweet lime	1482
Papuan citron	3532
Parent Washington navel orange	1241A
Parent Washington navel orange	1241B
Philippine pummelo	644
Philippine pummelo hybrid	2343
Pin Shan Kong Yan pummelo	2348
Pineapple sweet orange	3858
Pink pummelo	2246
Pink pummelo	2244
Pomeroy trifoliate	1717
Ponkan mandarin	3848
Pursha lime x Chinotto	3578
Red pummelo	2245
Rex Union grapefruit hybrid	3962
Rhode Red Valencia orange	3995
Rich 16-6 trifoliate	3485
Richard's Special mandarin	3329
Robinson mandarin	3850
Royal grapefruit	248
Ruby orange	3827
Ruvel navel orange	3181
S-1 citron	3878
Sacaton citrumelo	3337
Doblefina blood orange	4000

Table S2 (cont.)

Sample Name	CRC
Santa Barbara Rangpur lime	131
Santa Barbara Rangpur lime	712
Schaub rough lemon	3879
Seedless Kishu	3906
Severinia buxifolia	1492
Siamese acidless pummelo	2240
Siamese pummelo	2421
Siamese seedless pummelo	640
Simmons trifoliolate	3549
Siranui dekopon	4113
Smith's Early Naval orange	574
Som Keowan mandarin	3752
Som Saa mandarin	3853
South Coast Field Station citron	3546
Spadifora citron	3535
SRA 337 mandarin (US Furr)	4238
Standard sour orange	628
Star Ruby grapefruit	3770
Stowe red rough lemon	3185
Sun Chu Sha mandarin	4003
Szinkom mandarin	3085
Tahitian pummelo	3806
Tango mandarin	4183
Temple tangor	2598
Thong Dee pummelo	3927
Thornless Mexican lime	2683
Tibbetts Parent Washington navel orange	4051
Tien Chieh mandarin	2590
Tim Kat mandarin	2692
Tresca pummelo	302
Triumph grapefruit	297
Troyer citrange	1459
Unnamed citron	3174
Unnamed lemon	3885
Unnamed sweet lemon	3093
US 119 hybrid	3998

Table S2 (cont.)

Sample Name	CRC
USDA 6-15-150 mandarin	4165
Valencia orange Late orange	4077
Valentine grapefruit hybrid	4182
Valentine x Kishu #1	NA
Valentine x Kishu #2	NA
Valentine x Kishu #4	NA
Variegated Pink Fleshed Eureka lemon	2367
W. Murcott Afourer mandarin	3953
Washington Sanguine blood orange	3977
Washington seedy navel orange	956
Wild lemon	3300
Wilking mandarin	3020
Willowleaf mandarin	3843
Wincia tangor	4148
Yellow Rind mandarin	3895
Yemen citron	2847
Yunnanese citron	3798
Yuzu	1216
#22 trifoliate	4172

Table S3

Families used in mendelian trio analysis to assess Mendelian inheritance from parents to off-spring for SNP validation

(a) Total of 7 families and 10 trios were used for assessment of mendelian inheritance to validate SNPs represented on the Axiom® Citrus HD Genotyping Array. (b) Total of 10 families and 10 trios were used for assessment of mendelian inheritance to validate SNPs represented on the Axiom® Citrus Genotyping array

(a)

Family	Parent-Male	Parent-Female	Off-spring
F01-a	Orlando_tangelo_3874	Algerian_clementine_279	Fairchild_mandarin_3559
F01-b	Orlando_tangelo_3874	Algerian_clementine_279	Nova_mandarin_3615
F02	Dancy_Weshart_602	Duncan_grapefruit_3832	Minneola_tangelo_3340
F03	Pomeroy trifoliate 1717 Willowleaf mandarin	Parent Washington navel orange 1241A	Troyer citrange 1459
F04	3843	King tangor 303	Kinnow 3021
F05	Ruby orange 3827	Siamese acidless 2240	Raspberry jam 4X Gold Nugget mandarin 3913
F06	Kincy mandarin 3144	Wilking mandarin 3020	
F07-a	Seedless Kishu 3906	Valentine pummelo hybrid 4182	Valentine x Kishu #1
F07-b	Seedless Kishu 3906	Valentine pummelo hybrid 4182	Valentine x Kishu #2
F07-c	Seedless Kishu 3906	Valentine pummelo hybrid 4182	Valentine x Kishu #4

(b)

Family	Parent-Male	Parent-Female	Off-spring
F01	Orlando tangelo 3874	Algerian clementine 279	Fairchild mandarin 3559
F02	Rubidoux Trifol 838	African pummelo 2346	African Shaddock x rubidoux citrumelo 3969
F03	Flying Dragon trifoliate 3330A	Sydney hybrid microcitrus hybrid	Microcitrus x trifoliate trifol hybrid 3566
F04	Frua mandarin 3022	Siamese Acidless pummelo 2240	Cocktail pummelo hybrid 3555
F05	Marumi SRA489 kumquat 4125	West Indian lime 1813	Eustis limequat kumquat Hyb 3603
F06	Washington Songuine 3977	Frost Lisbon 3176	Orban lemon 4140
F07	Siamese Acidless pummelo 2240	Wilking mandarin 3020	13D-5-10
F08	Fairchild mandarin 3559	Fortune mandarin 3560	RS99.201-21
F09	Flying Dragon trifoliate 3330A	Chandler pummelo 3224	RR86-10-99
F10	Pomeroy trifoliate 1717	Chandler pummelo 3224	RR86-10-99

Table S4

Mendelian inheritance error calculated for SNP validation with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array

a) Genotyping calls with Axiom® Citrus HD Genotyping Array were used to assess mendelian inheritance (MI) error from parents to off-spring in 10 trios. MI error was calculated separately for Axiom® Citrus HD Genotyping Array AX1 and AX2 by Affymetrix™ bioinformatics team. (b) Genotyping calls with Axiom® Citrus Genotyping Array were used to assess mendelian error from parents to off-spring in 10 trios involving members of more diverse cross. Number of loci that are inconsistent with mendelian inheritance was calculated with plink v1.9 (https://www.cog-genomics.org/plink/1.9/basic_stats) using 53,206 PHR loci

(a)

Axiom® Citrus HD Genotyping Array	Offspring	Parent-Male	Parent-Female	MI Error
AX1	Fairchild mandarin 3559	Orlando tangelo 3874	Algerian clementine 279	0.283
AX1	Fairchild mandarin 3559	Orlando tangelo 3874	Algerian clementine 279	0.285
AX1	Nova mandarin 3615	Orlando tangelo 3874	Algerian clementine 279	0.322
AX1	Orlando tangelo 3874	Algerian clementine 279	Fairchild mandarin 3559	0.283
AX1	Algerian clementine 279	Orlando tangelo 3874	Nova mandarin 3615	0.322
AX1	Minneola tangelo 3340	Dancy Weshart 602	Duncan grapefruit 3832	0.429
AX1	Dancy Weshart 602	Duncan grapefruit 3832	Minneola tangelo 3340	0.429
AX1	Duncan grapefruit 3832	Dancy Weshart 602	Minneola tangelo 3340	0.429
AX1	Kinnow 3021	Willowleaf mandarin 3843	King tangor 303	0.261
AX1	Willowleaf mandarin 3843	King tangor 303	Kinnow mandarin 3021	0.261
AX1	King tangor 303	Willowleaf mandarin 3843	Kinnow mandarin 3021	0.261
AX1	Gold Nugget mandarin 3913	Kincy mandarin 3144	Wiling mandarin 3020	0.250
AX1	Kincy mandarin 3144	Wiling mandarin 3020	Gold Nugget mandarin 3913	0.250
AX1	Wiling mandarin 3020	Kincy mandarin 3144	Gold Nugget mandarin 3913	0.250
AX1	Valentine x Kishu #2	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.547
AX1	Valentine x Kishu #1	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.538

Table S4 (cont.)

Axiom® Citrus HD Genotyping Array	Offspring	Parent-Male	Parent-Female	MI Error
AX1	Valentine x Kishu #4	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.508
AX1	Seedless Kishu 3906	Valentine pummelo hybrid 4182	Valentine x Kishu #1	0.538
AX1	Valentine pummelo hybrid 4182	Seedless Kishu 3906	Valentine x Kishu #2	0.547
AX2	Fairchild mandarin 3559	Orlando tangelo 3874	Algerian clementine 279	0.288
AX2	Fairchild mandarin 3559	Orlando tangelo 3874	Algerian clementine 279	0.286
AX2	Nova mandarin 3615	Orlando tangelo 3874	Algerian clementine 279	0.303
AX2	Orlando tangelo 3874	Algerian clementine 279	Fairchild mandarin 3559	0.288
AX2	Algerian clementine 279	Orlando tangelo 3874	Fairchild mandarin 3559	0.288
AX2	Minneola tangelo 3340	Dancy Weshart 602	Duncan grapefruit 3832	0.446
AX2	Dancy Weshart 602	Duncan grapefruit 3832	Minneola tangelo 3340	0.446
AX2	Duncan grapefruit 3832	Dancy Weshart 602	Minneola tangelo 3340	0.446
AX2	Kinnow 3021	Willowleaf mandarin 3843	King tangor 303	0.261
AX2	King tangor 303	Willowleaf mandarin 3843	Kinnow 3021	0.261
AX2	Willowleaf mandarin 3843	King tangor 303	Kinnow 3021	0.261
AX2	Gold Nugget mandarin 3913	Kincy mandarin 3144	Wilking mandarin 3020	0.244
AX2	Kincy mandarin 3144	Wilking mandarin 3020	Gold Nugget mandarin 3913	0.244
AX2	Wilking mandarin 3020	Kincy mandarin 3144	Gold Nugget mandarin 3913	0.244
AX2	Valentine x Kishu #2	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.576
AX2	Valentine x Kishu #1	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.551
AX2	Valentine x Kishu #4	Seedless Kishu 3906	Valentine pummelo hybrid 4182	0.517
AX2	Seedless Kishu 3906	Valentine pummelo hybrid 4182	Valentine x Kishu #2	0.576
AX2	Valentine pummelo hybrid 4182	Seedless Kishu 3906	Valentine x Kishu #1	0.551

(b)

Family	Individual	Mendelian Error
F01	Orlando tangelo 3874	113
F01	Algerian clementine 279	56
F01	Fairchild mandarin 3559	168
F02	Rubidoux trifoliolate 838	1020
F02	African pummelo 2346	185
F02	African Shaddock x Rubidoux trifoliolate 3969	1201
F03	Flying Dragon trifoliolate 3330A	478
F03	Sydney hybrid microcitrus hybrid	587
F03	Microcitrus x trifoliolate trifol hybrid 3566	1027
F04	Frua mandarin 3022	186
F04	Siamese Acidless pummelo 2240	337
F04	Cocktail pummelomelo hybrid 3555	519
F05	Marumi SRA489 kumquat 4125	707
F05	West Indian lime 1813	537
F05	Eustis limequat kumquat hybrid 3603	1149
F06	Washington Songvine 3977	247
F06	Frost Lisbon 3176	548
F06	Orban Lemon 4140	782
F07	Siamese Acidless pummelo 2240	292
F07	Wilking mandarin 3020	187
F07	Siamese Acidless pummelo 2240	476
F08	Fairchild mandarin 3559	32
F08	Fortune mandarin 3560	38
F08	RS99.201-21	76
F09	Flying Dragon trifoliolate 3330A	1309
F09	Chandler pummelo 3224	203
F09	RR86-10-99	1497
F10	Pomeroy trifoliolate 1717	1283
F10	Chandler pummelo 3224	192
F10	RR86-10-99	1471

Table S5

Summary table of percent average SNP marker heterozygosity in various groups with Axiom® Citrus HD Genotyping Array

Average percent heterozygous markers was estimated for various citrus groups and relatives using Axiom® Citrus HD Genotyping Array data

Group	Number of Accessions	Avg Heterozygous Markers (%)	Standard Deviation (%)
Oranges	15	33	0.04
Grapefruits	10	31	0.66
Misc Hybrids (lemons, limes, etc.)	38	31	7.84
Clementine	4	21	0.05
Satsuma	3	20	0.01
Mandarin Hybrids	44	18	5.02
Papedas	4	12	8.64
Pummelos	19	11	7.64
Mandarins	11	8.3	1.52
Citrons	11	4.3	0.39

Table S6

Summary table of percent average SNP marker heterozygosity in various groups with Axiom® Citrus Genotyping Array

Average percent heterozygous markers was estimated for various citrus groups and relatives using Axiom® Citrus Genotyping Array data

Group	Number of Accessions	Avg Heterozygous Markers (%)	Standard Deviation (%)
sweet oranges	77	36	1.04
Lemons and limes	81	35	7.96
navel oranges	67	35	0.21
Valencia oranges	25	35	0.08
Blood oranges	17	35	0.05
Grapefruits	34	33	1.98
Trifoliate (Poncirus) Hybrids	44	29	2.52
Clementines	13	24	0.04
Relatives	3	22	1.57
Mandarin Hybrids	132	21	4.71
Satsumas	45	21	0.13
Pummelos	44	8.2	0.88
Mandarins	11	8.1	1.29
Papedas	5	3.6	1.62
Trifoliate (Poncirus)	33	2.6	0.19
Citrons	19	1.4	0.66

Table S7

196 unique accessions from CVC for PCA analysis with data from Axiom® Citrus HD Genotyping Array

Sample Name	CRC
Algerian clementine	279
Anseikan pummelo	3979
Arajon pummelo	2596
Avana Tardivo di Ciaculli mandarin	4162
Bakrai mandarin x limetta hybrid	3776
Banpeiyu pummelo	3961
Bergamot sour orange	2881
Berna lemon	3590
Borneo Rangpur lime	2424
Bouquet de Fleurs sour orange	571
Brazilian sour orange	1689
Buddha's Hand citron (fingered citron)	3768
Calamondin	2592
Canton mandarin	3576
Carrizo citrange	2863
Chandler pummelo	3244
Changsha mandarin	3577
Chinese pummelo	1212
Chinotto orange	2375
Citrus amblycarpa	2485
Citrus assamensis	3173
Citrus canaliculata	3565
Citrus depressa	2448
Citrus erythrosa	3292
Citrus excelsa	2316
Citrus excelsa	3765
Citrus halimii	3780
Citrus hanayu	3469
Citrus hongheensis	3797
Citrus hystrix	3103
Citrus keraji	3144
Citrus kharna	3059
Citrus leiocarpa	3147
Citrus longispina winged lime	2320

Table S7 (cont.)

Sample Name	CRC
Citrus lycopersicaformis	3564
Citrus macrophylla	3842
Citrus megaloxycarpa	3241
Citrus miaray	3574
Citrus nippokoreana	3228
Citrus pennivesiculata	2434
Citrus reticulata	3239
Citrus reticulata	3812
Citrus sino-grandis	3148
Citrus succosa	3280
Citrus sudachi	3471
Citrus sunki mandarin	3143
Citrus tachibana	3150
Citrus ujukitsu	3467
Citrus webberii var. Montana	767
Citrus yatsushiro	3466
Cleopatra mandarin	3844
Dancy mandarin	602
Deep Red pummelo	2347
Delta Valencia orange	3963
Diamante citron	3523
Dummett bittersweet orange	2541
Duncan grapefruit	3832
Egami Buntan pummelo	3959
Emerald mandarin	4123
Estes rough lemon	3386
Fairchild mandarin	3559
Florida rough lemon	400
Foothill Lisbon lemon	3836
Fortunella obovata	3475
Fortunella polyandra	3901
Frost nucellar Eureka lemon	3005
Frost nucellar Lisbon lemon	3176
Frost Owari Satsuma	3178
Gold Nugget mandarin	3913

Table S7 (cont.)

Sample Name	CRC
Hiawassie citron	3527
Hirado Buntan pummelo	4122
Huang Yen Man Chieh mandarin	3897
Ichang lemon	1215
Imperial mandarin	4160
Improved Meyer lemon	3737
India lime	2450
Indian citron	661
Indian rough lemon	2544
Interdonato lemon	3593
Iraq lemon hybrid	3492
Italian pink fleshed lemon	2899
Itoshima pummelo	3928
Japansche citroen	2875
Kabosu ichandarin	3943
Kao Pan pummelo	2248
Kao Pan pummelo	2355
Kao Phuang pummelo	2352
Kincy mandarin	4144
King tangor	303
Kinnow mandarin	3021
Kinokuni mandarin	696
Kinokuni mandarin (Mukakukishu)	3887
Kiyomi tangor H-12	3952
Lane Late navel orange	3803
Lapithiotiki lemon	4005
Lemonade lemon	4181
Lima sweet orange	950
Limon Real	2317
Lo Porto lemon	3090
Low thorn Lisbon lemon	NA
Mandarinette	3405
Marsh pink grapefruit	2010
Marsh seedy grapefruit	1565
Mato Buntan pummelo	3945

Table S7 (cont.)

Sample Name	CRC
Mexican lime	1710
Miho Wase satsuma	4047
Milam rough lemon	3396
Millsweet limetta	569
Minneola tangelo	3340
Moanalua pummelo	448
Monachello lemon	3392
Monreal clementine	3773
Moro blood orange	3830
Nagpur orange	3727
Necked mandarin	3738
Nova mandarin	3615
Odorata citron	3655
Okitsu Wase satsuma	3820
Olinda Valencia orange	2750
Olivelands sour orange	2717
Orlando tangelo	3874
Oroval clementine	4010
Ortanique tangor	3368
Oxanthera neo-caledonica	3793
Page tangelo	3616
Palestine sweet lime	1482
Papuan citron	3532
Parent Washington navel orange	1241-A
Philippine pummelo	644
Philippine pummelo hybrid	2343
Pin Shan Kong Yan pummelo	2348
Pineapple sweet orange	3858
Pink pummelo	2246
Pink pummelo	2244
Ponkan mandarin	3848
Pursha lime x Chinotto	3578
Red pummelo	2245
Rex Union grapefruit hybrid	3962
Rhode red Valencia orange	3995

Table S7 (cont.)

Sample Name	CRC
Richard's Special mandarin	3329
Robinson mandarin	3850
Royal grapefruit	248
Ruby orange	3827
Ruvel navel orange	3181
S-1 citron	3878
Sacaton citrumelo	3337
Doblefina blood orange	4000
Santa Barbara Rangpur lime	131
Santa Barbara Rangpur lime	712
Schaub rough lemon	3879
Seedless Kishu	3906
Seedy navel orange	956
Siamese acidless pummelo	2240
Siamese pummelo	2421
Siamese seedless pummelo	640
Sidi Aissa clementine	3988
Siranui dekopan	4113
Smith's Early navel orange	574
Som Keowan mandarin	3752
Som Saa mandarin	3853
South Coast Field Station citron	3546
Spadifora citron	3535
SRA 337 mandarin (US Furr)	4238
Standard sour orange	628
Star Ruby grapefruit	3770
Stowe Red rough lemon	3185
Sun Chu Sha mandarin	4003
Szinkom mandarin	3085
Tango mandarin	4183
Temple tangor	2598
Thong Dee pummelo	3927
Thornless Mexican lime	2683
Tibbetts Parent Washington navel orange	4051
Tien Chieh mandarin	2590

Table S7 (cont.)

Sample Name	CRC
Tim Kat mandarin	2692
Tresca pummelo	302
Triumph grapefruit	297
Troyer citrange	1459
Unnamed citron	3174
Unnamed lemon	3885
Unnamed sweet lemon	3093
US 119 hybrid	3998
USDA 6-15-150 mandarin	4165
Valencia orange Late orange	4077
Valentine grapefruit hybrid	4182
Variegated Pink Fleshed Eureka lemon	2367
W. Murcott Afourer mandarin	3953
Washington Sanguine blood orange	3977
Wild lemon	3300
Wilking mandarin	3020
Willowleaf mandarin	3843
Wincia tangor	4148
Yellow Rind mandarin	3895
Yemen citron	2847
Yunnanese citron	3798
Yuzu	1216

Table S8

871 unique accessions from CVC for PCA analysis with Axiom® Citrus Genotyping Array data

Sample Name	CRC
Abhayapuri lime	3762
Ac29 lemon	3920
Addanimma lime	3072
African pummelo	2346
African Shaddock x Rubidoux Trifoliate	3969
Aguzdera satsuma	4090
Ain Taoujdate clementine	4212
Akcay Sekeri sweet orange	3802
Algerian clementine	279
Allen Newman#4 Eureka lemon	3496
Allen Variegated Eureka lemon	3007
Allspice tangelo	4143
Altoona tangor	2792
Amber Eureka lemon	2429
Ambersweet sweet orange	4002
Amoa 8 tangor	4237
Anseikan pummelo	3979
Aoshima satsuma	4066
Apopka navel orange	4151
Arajon pummelo	2596
Arancino lemon	3387
Argentina sweet orange	2802
Argentine trifoliate	3206
Armstrong satsuma	4064
Arrufatina clementine	4184
Asahikan pummelo hybrid	3905
Atmore lemon	4092
Atwood navel orange	3791
Australian navel orange	4175
Australian Red Rangpur lime	2319
Australian trifoliate	3151
Autumn Gold navel orange	4037
Avana Apireno mandarin	4163
Avana Tardivo di Ciaculli mandarin	4162

Sample Name	CRC
Aziza sweet orange	3403
Badhri citron hybrid	3723
Bahianinha Araras navel orange	4168
Bahianinha-SRA 513 navel orange	4167
Bahman Persian#1 sweet lime	4174
Bahman Persian#2 sweet lime	4187
Baishaishu Rangpur lime	4129
Bakrai (mandarin x limetta) hybrid	3776
Banokan pummelo	3950
Banpeiyu pummelo	3961
Berna sweet orange	3272
Barnes trifoliolate	2554
Barnfield navel orange	4038
Batangas mandarin	3659
Baygold tangelo	4186
Belady mandarin	3363
Benecke trifoliolate	3338
Bengal citron hybrid	3055
Beni Hassaku pummelo hybrid	3907
Benoit trifoliolate	3547
Benton citrange	3908
Bergamot sour orange	2881
Bergamoto#18 Lisbon lemon	3506
Berna lemon	3590
Bessie sweet orange	245
Bessie sweet orange	1693
Bey navel orange	3406
Bidwells Bar sweet orange	1512
Biondo Comune sweet orange	3866
Biondo Riccio sweet orange	3630
Bishop Red lime	2451
Bitrouni lemon	3265
Bitters C22 citrandarin	4219
Bitters navel orange	4228
Blanca Macetera sweet orange	3898
Blanchard Eureka lemon	3499

Table S8 (cont.)

Sample Name	CRC
Borneo Rangpur lime	2424
Boukhobza blood orange	4156
Bouquet des Fleurs sour orange	571
Bouquetier De Nice sour orange	4157
Bower mandarin	3649
Brazilian sour orange	1689
Bream Tarocco blood orange	4032
Buddha's Hand citron	3768
Burr's Valencia orange	2561
C32 citrange	3911
C35 citrange	3912
Cadena Punchosa sweet orange	3246
Cadenera Fina sweet orange	2856
Cadenera sweet orange	3632
Caffin clementine	3987
Calabrese Ovale orange	3921
Calamondin	2592
Calamondin	3656
Calashu mandarin hybrid	2867
California Rojo navel orange	4201
Campbell Valencia orange	3032
Camulos grapefruit	3139
Canton mandarin	3576
Capuchin sweet orange	3245
Cara Cara navel orange orange	3994
Carpenter C54 citrandarin	4220
Carrizo citrange	2863
Carter navel orange	2008
Cascade Eureka lemon	2703
Cascade#1 Eureka lemon	3498
Castelo lime	4062
Catlin Valencia orange	3860
Cecily grapefruit	2014
Cecily grapefruit	3637
Centennial Variegated kumquat	4082

Table S8 (cont.)

Sample Name	CRC
Ceridwen navel orange	4065
Chandler pummelo	3224
Changsha mandarin	3577
Chapman Valencia orange	4229
China 6-15 satsuma	4180
China 6-18 satsuma	4199
China 6-21 satsuma	4213
China 6-22 satsuma	4214
China S-1 satsuma	4158
China S-11 satsuma	4159
China S-12 satsuma	4154
China S-15 satsuma	4178
China S-17 satsuma	4179
China S-2 satsuma	4060
China S-20 satsuma	4198
China S-5 satsuma	4153
China S-8 satsuma	4177
China S-9 satsuma	4061
China Satsuma#3	4115
China Satsuma#6	4088
China Satsuma#7	4089
Chinese pummelo	1212
Chinese pummelo	1224
Chinotto Broadleaf sour orange	4104
Chinotto sour orange hybrid	2375
Chinotto sour orange hybrid	3728
Chinotto sour orange	3857
Chironja grapefruit hybrid	3909
Chironja grapefruit hybrid	4045
Chislett navel orange	4039
Chulo Key lime	4120
Cipo sweet orange	3894
Citradarin trifoliolate hybrid	2618
Citradia trifoliolate hybrid	1436
Citradia trifoliolate hybrid	1437

Table S8 (cont.)

Sample Name	CRC
Citradia trifoliolate hybrid	1438
Citraldin trifoliolate hybrid	2619
Citrangle#1416 citrange	2866
Citranglequat 19-15-7	3259
Citrangor trifoliolate hybrid	1447
Citremon trifoliolate hybrid	1448
Citremon trifoliolate hybrid	1449
Citron Of Commerce citron	3518
Citrumelo	1452
Citrumelo	3348
Citrus amblycarpa	2485
Citrus assamensis	3173
Citrus benikoji	3149
Citrus canaliculata	3565
Citrus depressa	2448
Citrus erythrosa	3292
Citrus excelsa	2316
Citrus excelsa	2892
Citrus excelsa	3765
Citrus funadoko	3274
Citrus grandis pummelo	2453
Citrus hanayu	3469
Citrus hiroshimana hybrid	3275
Citrus hongheensis	3797
Citrus hystrix	432
Citrus hystrix	2454
Citrus hystrix	3103
Citrus intermedia	3474
Citrus keraji	3144
Citrus kharna	3059
Citrus leiocarpa	3147
Citrus limettoides	919
Citrus limettoides	921
Citrus longispina winged lime	2320

Table S8 (cont.)

Sample Name	CRC
<i>Citrus lycopersicaformis</i>	3564
<i>Citrus macrophylla</i>	2446
<i>Citrus macrophylla</i>	3842
<i>Citrus maderaspatana</i>	3225
<i>Citrus megaloxycarpa</i>	3241
<i>Citrus miaray</i>	3574
<i>Citrus myrtifolia</i>	4164
<i>Citrus natsudaïdai</i>	3235
<i>Citrus nippokoreana</i>	3228
<i>Citrus obovoidae</i>	3465
<i>Citrus otachibana</i>	3470
<i>Citrus pennivesiculata</i>	2434
<i>Citrus reticulata</i>	3239
<i>Citrus reticulata</i>	3812
<i>Citrus reticulata</i>	3813
<i>Citrus rokugatsu</i>	3473
<i>Citrus rugulosa</i>	3556
<i>Citrus shunkokan</i>	3476
Chinese sweet orange	3814
<i>Citrus sinensis</i> China sweet orange	3933
<i>Citrus sinensis</i> China sweet orange	3934
<i>Citrus sino-grandis</i>	3148
<i>Citrus succosa</i>	3280
<i>Citrus sudachi</i>	3471
<i>Citrus sulcata</i>	3257
<i>Citrus sunki</i> mandarin	3143
<i>Citrus tachibana</i> mandarin	3150
<i>Citrus tamurana</i>	3092
<i>Citrus tardiva</i> mandarin	3297
<i>Citrus tengu</i>	3464
<i>Citrus ujukitsu</i>	3467
<i>Citrus volkameriana</i>	3050
<i>Citrus vulgaris</i> sour orange hybrid	760
<i>Citrus webberii</i> var. Montana	767
<i>Citrus yatsushiro</i> mandarin	3466

Table S8 (cont.)

Sample Name	CRC
Citrus yatsushiro mandarin	3880
Citrus yuko	3146
Clausellina satsuma	4012
Clement tangelo	2603
Clementine x Pearl mandarin hybrid	4020
Clementine x Silverhill mandarin hybrid	3731
Cleopatra mandarin	3844
Cluster Sport navel orange	3732
Coarse Valencia orange	1518
Cocktail pummelo hybrid	3555
Cogan navel orange	4050
Cook Eureka lemon	3837
Corona Foothill Eureka lemon	3043
Corpaci lemon	3591
Corrugated Thomson navel orange	1381
Corsican citron	1795
CPB4475 citrumelo	3341
CPB44944A navel orange	3305
CPB44944B navel orange	3306A
CRC343 grapefruit	343
Cuban Shaddock pummelo	1462
Cukurova navel orange	4086
Cunningham citrange	271
Curry sweet orange	3347
Cutter Valencia orange	3030
Dabbe citron hybrid	3190
Daidai sour orange	656
Daisy mandarin	3910
Daisy SL mandarin	4203
Dalandan sweet orange	2373
Dancy Frost mandarin	3026
Dancy Weshart_mandarin	602
Dart North Owari satsuma	4099
Dart South Owari satsuma	4100
Davao papeda hybrid	2427

Table S8 (cont.)

Sample Name	CRC
De Nules clementine	4056
Deep Red pummelo	2347
Delfino blood orange	4114
Delta Valencia orange	3963
Diamante citron	3523
Diller sweet orange	2602
Djerok citron hybrid	2456
Dobashi Beni satsuma	3794
Doblefina blood orange	3152
Doblefina blood orange	4000
Dom Joao Valencia orange	3872
Dr Strong#28 Lisbon lemon	3504
Dream navel orange	3307
Dry navel orange	3354
Dulcia citron	3654
Dummett bittersweet orange	2541
Duncan grapefruit	3832
Dungan satsuma	4101
Dweet tangor	3018
Early tangelo	2560
East Highlands sweet orange	2369
Egami Buntan pummelo	3959
Egyptian lime	2883
Ellendale tangor	3960
Emerald mandarin	4123
Empress mandarin	3613
Encore mandarin	3569
Encore_IR6_mandarin	4204
English Large trifoliate	3548
Entre Fina blood orange	3650
Eremolemon citrus relative hybrid	2439
Espagnole Sans Pepins sweet orange	3249
Estes rough lemon	3386
Eustis limequat	3603
Everglade Lime	2449

Table S8 (cont.)

Sample Name	CRC
Everhard navel orange	4200
Fairchild LS mandarin	4205
Fairchild mandarin	3559
Fallglo mandarin	3990
Faris sweet lemon	2695
Faustrime	2891
Femminello lemon	3500
Femminello Ovale lemon	3388
Femminello Sfusato lemon	3389
Femminello Siracusano 2KR lemon	4166
Feroniella oblata	3510
Fina clementine	4011
Fina Sodea clementine	3999
Finike sweet orange	3870
Fisher navel orange	3135
Fisher navel orange	3645
Flame grapefruit	3993
Fleming's Shaddock pummelo	578
Florida A rough lemon	3385
Florida rough lemon	400
Florida trifoliate	2862
Flying Dragon trifoliate	3330A
Flying Dragon trifoliate	3330B
Foothill Lisbon lemon	3836
Fortune mandarin	3560
Fortunella crassifolia	3818
Fortunella japonica	3237
Fortunella obovata	3475
Fortunella polyandra	3901
Foster Pink grapefruit	1198
Fraser Seville sour orange	4096
Fremont mandarin	3558
Frost Marsh grapefruit	3184
Frost nucellar Eureka lemon	3005
Frost nucellar Lisbon lemon	3176

Table S8 (cont.)

Sample Name	CRC
Frost Nucellar#1 navel orange	3033
Frost Owari satsuma	3178
Frost trifoliolate	3484
Frost Valencia orange	3031
Frua mandarin	3022
Fukumoto navel orange	3903
Fulwood navel orange	4112
Furr C-57 citrandarin	4221
Fuya Menuda sweet orange	3787
Fuzzy Valencia orange	1240
Gabbuchinee sour orange hybrid	3079
Galligan Lisbon lemon	3835
Genetic Dwarf grapefruit	3774
Genoa lemon	565
Giant Key lime	4069
Gillemborg navel orange	4192
Gillette navel orange	3017
Glen Citrangedin	3573
Gold Nugget mandarin	3913
Golden Buckeye navel orange	588
Golden Nugget navel orange	1379
Gomiri rough lemon	2557
Goutoucheng sour orange	4004
Granitos sour orange	2715
Gulgul lemon hybrid	4235
H 56 tangor	3096
Haiku B pummelo	3940
Hall grapefruit	3068
Hall grapefruit x Rubidoux trifoliolate	3821
Hamlin sweet orange	3862
Hangleson Rangpur lime	3932
Hansen mandarin	4152
Hart's Late Valencia orange	314
Hart's Tardiff Valencia orange	570
Hassaku pummelo hybrid	3942

Table S8 (cont.)

Sample Name	CRC
Henderson Ruby grapefruit	3914
Hernandina clementine	4185
Hiawassie citron	3527
Hill mandarin	3847
Hirado Buntan pummelo	4122
Hiroshima Natsuzabon grapefruit	3904
Hiryu trifoliate	3882
Homosassa sweet orange	292
Homosassa sweet orange	1696
Honey mandarin	3177
Huang Yen Man Chieh mandarin	3897
Hudson Foster grapefruit	3638
Hunan pummelo	1225
Hung Kat mandarin	2331
Ichang lemon	1215
Ichang lemon	1219
Imperial grapefruit	596
Imperial mandarin	4160
Imperial Variegated sweet orange	321
Improved Meyer lemon	3737
India lemon	2322
India lemon	2323
India lime	2450
India lime	3070
Indian citron	138
Indian citron	661
Indian River sweet orange	3859
Indian rough lemon	2544
Indio mandarinquat	3759
Interdonato lemon	3593
Iran lemon	3162
Iraq lemon hybrid	3492
Italian citron	3530
Italian pink fleshed lemon	2899
Itoshima pummelo	3928

Table S8 (cont.)

Sample Name	CRC
Iveriya satsuma	4091
Iwasaki satsuma	4218
Iyo San Ponkan mandarin	4118
Iyomikan tangor	3980
Jacobson trifoliolate	3411
Jaffa sweet orange	71
Jamaican Ugli tangelo	4127
Japansche citroen	2875
Jincheng sweet orange	3983
Jochimsen grapefruit	2784
Johnson navel orange	4059
Joppa sweet orange	281
Jullander Khatti rough lemon	4173
Jullundri Khatti rough lemon	3063
Kabosu ichandarin	3943
Kao Pan pummelo	3948
Kao Pan pummelo	2242
Kao Pan pummelo	2243
Kao Pan pummelo	2248
Kao Pan pummelo	2249
Kao Pan pummelo	2349
Kao Pan pummelo	2355
Kao Pan pummelo	2356
Kao Phuang pummelo	2352
Kao Phuang pummelo	3926
Kao Ruan Tia pummelo	2350
Kao Ruan Tia pummelo	2351
Kara mandarin	3019
Karn Lau Yau pummelo	2341
Kawachi Bankan pummelo	3944
Kawano Wase satsuma	1851
Kaweah#1 Lisbon lemon	3010
Keen sour orange	2624
Khalily sweet orange	3251
Khoubs-el-arsa rough lemon	2477

Table S8 (cont.)

Sample Name	CRC
Khoubs-el-arsa rough lemon	2489
Kincy mandarin	4144
King tangor	303
King tangor	3845
Kinkoji Unshiu	3816
Kinnow LS mandarin	4236
Kinnow mandarin	3021
Kinokawa Buntan pummelo	3978
Kinokuni mandarin	696
Kinokuni mandarin (Mukakukishu)	3887
Kirumakki Rangpur lime	4131
Kiyomi tangor H-12	3952
Knorr Rangpur lime	4132
Kobayashi Mikan mandarin	3817
Koethen sweet orange	1106
Konejime sour orange hybrid	3611
Koster tangor	3958
Kryder 28-3 trifoliate	3219
Kryder 5-5 trifoliate	3586
Kryder 55-1 trifoliate	3486
Kryder 60-2 trifoliate	3213
Kryder 8-5 trifoliate	3218
Kryder16-6 trifoliate	3210
Kulobot papeda hybrid	3612
Kulu lemon	3045
Kulu lemon	3487
Kunembo mandarin	3346
Kuno Wase satsuma	4030
Kusaie Rangpur lime	452
Kusner lemon	3194
Lamas lemon	3919
Lane Late navel orange	3803
Lange#3 satsuma	4103
Lapithiotiki lemon	4005
Laranja Cravo Ponkan mandarin	2893

Table S8 (cont.)

Sample Name	CRC
Laranja Selecta sweet orange	1045
Laranja Selecta sweet orange	3754
Lee mandarin	3851
Lemelo pummelo hybrid	1481
Lemelo pummelo hybrid	1775
Lemonime lime hybrid	2459
Leng navel orange	3808
Lima sweet orange	950
Limau Matu Susu citron hybrid	3755
Limoneira 8A IR1 lemon	NA
Limoneira 8A Lisbon lemon	3501
Limonero Fino 95 lemon	4126
Limonero Fino lemon	3970
Limonero Messina lemon	4121
Limonette De Marrakech	3989
Limonreal lime	2317
Limoui Sangui Rangpur lime	3200
Lo Porto lemon	3090
Louisiana Early satsuma	4225
low-thorn Lisbon lemon	NA
Lue Gim Gong Valencia orange	3025
Lumina lemon hybrid	3925
Lunario lemon	3159
Lupe Lisbon lemon	3013
Macetera sweet orange	3871
Madam Vinous sweet orange	3861
Malta blood orange	4085
Maltese Oval navel orange	274
Mandarinette	3405
Mandarin Sanguine mandarin	3367
Mapo tangelo	4222
Marisol clementine	4055
Marks trifoliate	3588
Marlow tangelo	4145
Marr's Early sweet orange	3865

Table S8 (cont.)

Sample Name	CRC
Marsh pink grapefruit	2010
Marsh seedy grapefruit	1565
Mary Ellen sweet lime	4053
Mato Buntan pummelo	3945
Mazoe rough lemon	1222
McEwen satsuma	4098
Mcfadden navel orange	609
Mediterranean sweet orange	4176
Meiwa kumquat	1471
Melogold grapefruit	3764
Mency tangor	3183
Mesero lemon	3892
Messina lemon	3009
Mexican citron	3531
Mexican lime	1710
Michal mandarin	4049
Microcitrus x P trifoliata	3566
Midknight Valencia orange	3955
Miho Wase satsuma	4047
Milam lemon	3396
Millsweet limetta	569
Minneola tangelo	3340
Mitha Tulia limetta	3051
Miyagawa satsuma	4074
Miyamoto satsuma	4216
Miyauchi Iyo tangor	3946
Moanalua pummelo	448
Moli Kurikuri pummelo hybrid	579
Monachello lemon	3392
Monoembryonic trifoliata	3888
Monreal clementine	3773
Monroe Lisbon lemon	3839
Moro blood orange	3830
Morton citrange	1463
Morton citrange	2748

Table S8 (cont.)

Sample Name	CRC
Murcott mandarin	3846
Myrtifolia sour orange	3290
Nagpur mandarin	3727
Nakhon Chaisi pummelo	2353
Nansho Daidai sour orange	2588
Navelate navel orange	4022
Navelencia navel orange	287
Navelina navel orange	4015
Necked mandarin	3738
Nepolitana satsuma	3848
New Zealand Lemonade	4181
Newhall navel orange	3014
New Zealand grapefruit	2381
New Zealand grapefruit	3769
Nicaraguan lemon	3841
Nichinan Ichigo satsuma	4217
Nicholson navel grapefruit	3398
Nippon orangequat	3360
Nissv E mandarin hybrid	3965
Nour clementine	4057
Nova IR10 mandarin	4206
Nova mandarin	3615
Novelty x Ellendale mandarin hybrid	3956
Odorata citron	3655
Okitsu Wase satsuma	3820
Old Indian lemon	4211
Old Line Eureka lemon	4094
Olinda McEwen Valencia orange	4230
Olinda Mulholland Valencia orange	4231
Olinda Valencia orange	2750
Olivelands sour orange	2717
Olivelands sweet orange	3863
Olympic Gold navel orange	4209
Orange A Feuilles Lachiniees sweet orange	4202
Orange De Nice sweet orange	3624

Table S8 (cont.)

Sample Name	CRC
Orange Of Heaven acidless sweet orange	371
Orban lemon hybrid	4140
Orlando bittersweet sour orange	1588
Orlando tangelo	3874
Oroblanco grapefruit hybrid	3602
Orogold sour orange hybrid	693
Oroval clementine	4010
Ortanique tangor	3368
Otaheite mandarine hybrid	2709
Ovale sweet orange	3625
Oxanthera neo-caledonica	3793
P trifoliata	4006
P trifoliata	4007
P trifoliata	4009
Paak Ling Mung lemon hybrid	3193
Page tangelo	3616
Palestine sweet lime	1482
Palmer navel orange	4013
Palmer navel orange	3964
Pan Dan pummelo	2752
Papuan citron	3532
Paradise navel orange	2853
Paraguay bittersweet sour orange	660
Parent Washington navel orange	1241-A
Parent Washington navel orange	1241-B
Parson Brown sweet orange	3873
Parson's Special mandarin	300
Pauthel pummelo	3949
Pearl tangelo	2849
Pehrson#3 Valencia orange	4196
Pehrson#4 Valencia orange	4197
Pera sweet orange	3864
Perao Valencia orange	3599
Peretta lemon hybrid	3924
Perlis#1 grapefruit	3886

Table S8 (cont.)

Sample Name	CRC
Perlis#2 grapefruit	3883
Perrine lemon hybrid	2458
Philippine Calamondin	3073
Philippine citron	3533
Philippine pummelo	644
Philippine pummelo hybrid	2343
Philippine Rangpur lime	2318
Pin Shan Kong Yan pummelo	2348
Pineapple sweet orange	3858
Pink pummelo	2244
Pink pummelo	2246
Pixie mandarin	3568
Pollock T4 navel orange	4193
Pomelit pummelo hybrid	4026
Pomeroy trifoliolate	1717
Pomona sweet lemon	4068
Ponderosa lemon hybrid	294
Pong Yau pummelo	2342
Ponkan mandarin	3849
Poona Rangpur lime	4135
Poona Srirampur Rangpur lime	4136
Powell navel orange	4040
Precoce De Valence sweet orange	3373
Primofiore lemon	3491
Prior Lisbon lemon	3505
Procimequat kumquat hybrid	3295
Procimequat X Clymenia	4109
Puma pummelo hybrid	3941
Pursha lime x Chinotto	3578
Rangpur lime x Troyer citrange	3997
Ray Ruby grapefruit	3915
Red Aranyan pummelo	2608
Red pummelo	2245
Red pummelo	2338
Redblush grapefruit	2850

Table S8 (cont.)

Sample Name	CRC
Reed Marsh grapefruit	3128
Reinking pummelo	3805
Rex Union grapefruit hybrid	3962
Rhode Red Valencia orange	3995
Ricalate navel orange	4058
Rich 16-6 trifoliate	3485
Richard's Special mandarin	3329
Richard's Special mandarin	4117
Ricote lemon	3893
Rio Farms Vaniglia acidless sweet orange	4128
Rio Grande lemonquat	4097
Rio Grande navel orange	3315
Rio Red grapefruit	3916
Road 164 satsuma	4102
Robertson navel orange	4023
Robinson mandarin	3850
Robyn navel orange	4063
Rocky Hill navel orange	4232
Roeding's Pink pummelo	1208
Rohde navel orange	4084
Rojo Blanco pummelo hybrid	4029
Rosenberger Lisbon lemon	3840
Ross Eureka lemon	3838
Rotuma Island sweet orange	3867
Royal grapefruit	248
Rubidoux mandarin	4031
Rubidoux pummelo	4028
Rubidoux sour orange	3855
Rufert A sweet orange	4141
Unnamed rough lemon	3060
Rusk citrange	301
Rusk citrange	1441
Ruvel navel orange	3181
Ruvel-nucellar navel orange	4142
S-1 citron	3878

Table S8 (cont.)

Sample Name	CRC
S281 citrangelo	3552
S302 citranguma	3415
Sacaton citrumelo	3337
Sacaton citrumelo	3414
Sacaton citrumelo	3417
Sacaton tangelo	2609
Sacaton tangelo	3331
Salustiana sweet orange	3868
Sampson tangelo	2418
San Jacinto tangelo	2011
San Marino Valencia orange	2689
Sanford citrange	276
Sanguine Gross Ronde blood orange	3401
Sanguinelli blood orange	3828
Sanguinello a Pignu blood orange	3829
Santa Barbara Rangpur lime	131
Santa Barbara Rangpur lime	712
Santa Barbara tangelo	4073
Santa Teresa lemon	3894
Sarah sweet orange	3584
Sarawak pummelo	4042
Sathgudi sweet orange	3083
Savage citrange	275
Scarlet Emperor Ponkan mandarin	3226
Schaub rough lemon	3879
Secret Garden navel orange	4194
Seedless Kishu mandarin	3906
Seedless lemon	3001
Seedless Valencia orange	2776
Seedy navel orange	65
Selma satsuma	4119
Seminole tangelo	2604
Seville sour orange	3856
Sexton tangelo	2543
Shaddock x St Michael orange pummelo hybrid	42

Table S8 (cont.)

Sample Name	CRC
Shambar grapefruit	3831
Shamouti sweet orange	3746
Shirokolistvennyi satsuma	4091
Siamelo tangelo	2586
Siamese Acidless pummelo	2240
Siamese pummelo	2421
Siamese seedless pummelo	640
Sicily citron	3534
Sidi Aissa clementine	3988
Silverhill satsuma	4072
Sinton Citrangequat	3642
Siranui dekopan	4113
Skaggs Bonanza navel orange	3971
Small Leaf trifoliolate	4017
Smith Red blood orange	4093
Smith's Early navel orange	574
Smooth Flat Seville sour orange	4079
Sniff acidless Rangpur lime	2484
Soghi lime	4130
Soh Himtra mandarin	4095
Soh Long lemon hybrid	3199
Soh Niamtra mandarin	3260
Soh Niangrang papeda hybrid	3203
Soh Synteng lemon hybrid	3216
Solid Scarlet mandarin	3328
Som Kaeo mandarin	3852
Som Keowan mandarin	3752
Som Saa mandarin	3853
sour pummelo	3066
South African rough lemon	2325
South Coast Field Station citron	3546
Spadifora citron	3535
Spanish sweet orange x P trifoliata	3336
Spring navel orange	4024
SRA337 mandarin (US Furr)	4238

Table S8 (cont.)

Sample Name	CRC
SRA63 clementine	4076
SRA92 clementine	4052
Srirampur Rangpur lime	4137
St.Michael Paperrind sweet orange	591
Standard sour orange	628
Star Ruby grapefruit	3770
Star sweet orange	366
Stow#20 sour orange	2192
Stow#7 Key lime	2188
Stowe red rough lemon	3185
Sue Linda Temple tangor	3810
Suisho Buntan pummelo	3947
Sukega grapefruit hybrid	4149
Summer Gold navel orange	4041
Sun Chu Sha mandarin	4003
Sunburst mandarin	3809
Sunrise tangelo	2606
Sunshine pummelo	2236
Sunshine tangelo	2788
Unnamed sweet lemon	3093
sweet pummelo	3067
Swingle citrumelo	3767
Swingle citrumelo	3771
Szinkom mandarin	3085
Tahiti sweet orange	3869
Tahitian Pummelo x Star Ruby grapefruit hybrid	3781
Tango mandarin	4183
Tankan tangor	2224
Tankan tangor	3875
Tarocco#7 blood orange	3596
Tau Yau pummelo	2583
Tavares limequat	4048
Taylor Eureka lemon	4014
Taylor trifoliate	3571
Temple tangor	2598

Table S8 (cont.)

Sample Name	CRC
Temple x Dancy mandarin hybrid	4021
Texas trifoliolate	2861
Thermal Tarocco blood orange	4018
Thomasville citrangequat	1440
Thomson navel orange	969
Thomson navel orange	983
Thong Dee pummelo	3927
Thong Dee pummelo	4027
Thornless Mexican lime	2683
Thornton tangelo	2013
Tibbets Washington navel orange	4051
Tien Chieh mandarin	2376
Tien Chieh mandarin	2590
Tim Kat mandarin	2692
Tosu sour orange hybrid	3607
Towne F trifoliolate	3572
Tresca pummelo	302
Trifeola trifoliolate hybrid	3954
Triumph grapefruit	297
Trovita sweet orange	3182
Troyer citrange	1459
Tuningmeng Rangpur lime	4139
Tunisian sour orange	2438
UCR Early navel orange	4087
Ueno satsuma	4223
Ugli tangelo	2780
Ugli tangelo	4171
Unnamed lemon hybrid	3172
Unnamed citron	3174
Unnamed India sour orange hybrid	3175
Unnamed kumquat hybrid	1044
Unnamed lemon	3885
Unnamed lemon hybrid	3154
Unnamed lemon hybrid	3748
Unnamed sweet lemon	3093

Table S8 (cont.)

Sample Name	CRC
Unnamed sweet lime	363
Unnamed sweet orange hybrid	648
Unnnamed lemon hybrid	3155
US 119 hybrid	3998
USDA 1-77-19 tangelo (US Seedless Surprise)	4208
USDA 6-15-150 mandarin	4165
USDA 88-2 mandarin hybrid	3991
USDA 88-3 mandarin hybrid	3992
USDA trifoliate	1498
Uvalde citrange	2865
Vaccaro blood orange	3242
Vaccaro blood orange	3627
Vainiglia Sanguigno acidless sweet orange	3801
Valencia Late sweet orange	4077
Valentine pummelo hybrid	4182
Vangasay 807 rough lemon	3996
Variegated Calamondin	3087
Variegated Cara Cara navel orange	4207
Variegated Eureka lemon	599
Variegated navel orange	1203
Variegated Pink Fleshed Eureka lemon	2367
Variegated sour orange	622
Variegated Valencia orange	4001
Verna sweet orange	3807
Villafranca lemon	280
Villafranca lemon	390
W. Murcott Afourer mandarin	3953
Warren Limequat	3069
Warren Marsh grapefruit	4150
Washington Sanguine blood orange	3977
Washington seedy navel orange	956
Webber tangelo	2746
Webber Fawcett#22 trifoliate	2552
Webber Shaddock pummelo	4233
Weirick acidless Rangpur lime	1684

Table S8 (cont.)

Sample Name	CRC
Wekiwa tangelo	2012
Werley Valencia orange	3339
West Indian lime	1813
Wetumpka sweet orange	3550
Wheeny grapefruit	2885
Whitley navel orange	4195
Whitney Marsh grapefruit	3804
Wiffen navel orange	4083
Wild lemon	3300
Wilking mandarin	3020
Willial tangelo	4147
Williams tangelo	2787
Willowleaf mandarin	3843
Wincia tangor	4148
Workman navel orange	3320
X639 trifoliate hybrid	3957
Xianfengcheng sweet orange	3982
Xie Shan satsuma	4081
Xinhuicheng sweet orange	3985
Xuegan sweet orange	3984
Yalaha tangelo	2559
Yamaguchi trifoliate	3412
Yellow Rind mandarin	3895
Yemen citron	2847
Yen Ben Lisbon lemon	4067
Yuma citrange	3205
Yuma Ponderosa lemon	3488
Yunnanese citron	3798
Yuzu	1216
Zhuluan sour orange	3930
Zhuluan sour orange	3981
#22 trifoliate	4172

Table S9

New classification assignment according to the marker analysis performed during this study

Sample Name	CRC	New Classification
New Zealand Lemonade Lemon	4181	Lemon Hybrid
Orban Lemon	4140	Lemon Hybrid
Naranja Selecta sweet orange	1045	Valencia orange
Burris Valencia orange	2561	Blood orange
Precoce De Valence Valencia orange	3373	sweet orange

CHAPTER 2

Citrus Ancestry: Diversity Study and Admixture Inference in Diverse Accessions of Citrus with High Density SNP Genotyping Array Data

Abstract

Citrus is one of the most economically important and cultivated fruit tree crops in the world. More than a thousand years of domestication and dispersal resulted in the establishment of diverse citrus genotypes that show a wide range of phenotypic and genetic diversity. Recent studies have revealed the ancestry and history of the most commercially important citrus taxa. However, such information is not available for most of the accessions in diverse citrus germplasm collections. We genotyped 924 accessions of *Citrus* and related genera with high-density SNP genotyping arrays to study genetic diversity, infer phylogeny, and determine patterns of introgression to better understand relationships among diverse citrus accessions, in terms of their inferred ancestry and history of admixture. The multiple complementary analyses identified non-admixed individuals belonging to citrus ancestral species, including three candidate type 1 mandarins, in addition to identifying admixed individuals that have never been reported previously in the University of California Riverside Citrus Variety Collection (CVC). In addition, our observations suggested that many of Tanaka's species are admixed and papeda is a polyphyletic group, contrary to Swingle's classification. We also identified that *Citrus hongheensis*, currently classified as papeda, is a possible progenitor of pummelo.

Introduction

Citrus, a member of the family Rutaceae, is one of the most widely cultivated and economically valued fruit tree crops in the world. It is cultivated in more than 100 countries [1] and there are more than 10,000 presumably differently named citrus varieties worldwide resulting from thousands of years of domestication, selection and environmental adaptation. The origin of citrus is now thought to be in Southeast Asia prior to dispersal to the rest of the world [2]. The genomes of many commercially important *Citrus* species are highly heterozygous via natural or artificial interspecific hybridization among ancestral *Citrus* species [3]. Recent citrus genomic studies suggest that there are nine progenitor citrus species, including three Australian species (*Microcitrus australasica* (F. Muell.) Swingle, *M. australis* (Planch.) Swingle and *Eremocitrus glauca* Lindl.) Swingle) and five Asian species (*Citrus reticulata*, *Citrus Maxima* (Burm. F.) Merr., *Citrus Medica* L, *Citrus micrantha* and species of *Fortunella*) and species of *Poncirus* [4]. These five Asian ancestral species are considered to contribute most of the genetic diversity that exists in major commercial citrus species [5]. Due to their highly heterozygous genomes and long juvenility, these hybrid citrus crops, often referred to as secondary species such as oranges (*C. sinensis*), grapefruit (*C. paradisi*), lemons (*C. limon*), limes (*C. aurantifolia*) and many mandarins (*C. reticulata*), are propagated clonally, allowing true to type propagation of the desired genotype. Much of the natural genetic variability existing within each of these groups originated from sport mutation or selections of nucellar variants [5–7], but in some cases hybridization also contributes [5]. There are several factors which complicate genetic studies in *Citrus*, such as the presence of nucellar embryony, long juvenility and frequent interspecific hybridization [6].

Repeated hybridization events give rise to reticulated lineages which contain genetic materials derived from two or more ancestral progenitor species. Such events complicate the process of phylogeny inference and reconstruction of hybrid speciation events. The role of

reticulate evolution in plant diversification has long been recognized [8–15], and hybridization and subsequent introgression has been reported in an increasing number of species in recent years [5, 13, 16–21]. Such reticulate evolution is thought to contribute to increased genome diversity, and broader environmental adaptation [12, 22–25]. Many citrus genotypes are believed to have evolved through reticulated evolution, resulting in their genomes showing complex patterns of admixture [3, 5, 26, 27].

Citrus taxonomy and phylogeny have been historically controversial and confounded due to a long history of cultivation, its complex genetics and wide diversity, both in terms of genetics and morphology. The classical *Citrus* taxonomists, Tanaka and Swingle, grouped citrus into 16 species and 172 species [28], respectively. The major discrepancy between these two taxonomic systems stemmed from whether or not hybrid citrus should be considered a species. Recently, efforts have been made to resolve citrus ancestry and phylogeny [3, 4, 29–33]. It should be noted that citrus taxonomy is currently under revision and we use the nomenclature currently used by the CVC.

Identification of genetic variants relevant to environmental adaptation, disease resistance and tolerance are important for breeding, and their discovery can be aided by studying the ancestry and admixture events involving *Citrus* ancestral progenitor species. However, the genetic diversity and genetic signatures which have arisen through environmental adaptation and cultivation can easily be eroded by current breeding practices where new cultivars are developed by often successive introgression with existing elite cultivars [34, 35]. Such knowledge is important and can lead to the development of new superior cultivars that can withstand a range of environmental conditions and diseases. In order to further understand the history of admixture, ancestry, and genetic variants, the genetic diversity that exists in citrus germplasm should be assessed in terms of population structure, phylogeny and admixture events.

The most commonly used statistical approaches to infer genetic diversity and ancestry are principal component analysis (PCA) and model based clustering approaches such as STRUCTURE or ADMIXTURE. The Bayesian algorithm based STRUCTURE was first developed [36], then ADMIXTURE was introduced in order to analyze data produced with high-density marker systems [37] such as high density SNP array data. Phylogeny studies are often conducted with bifurcating tree analyses to infer evolution and taxonomy, in addition to understanding evolutionary diversity. However, such an approach is not suitable to describe the evolutionary relationships among diverse accessions of *Citrus* which include admixed and hybrid accessions [38]. An assumption of the conventional bifurcating phylogenetic tree approach is that the evolutionary process of one lineage is independent of that of another [39–41]. A network approach, on the contrary, can model evolutionary processes where the assumption of independence of lineages is violated, allowing us to study relationships of lineages that arise through complete or partial merging of two or more ancestors [42]. Phylogenetic network analyses have been applied to study reticulation events [43–48] such as horizontal gene transfer and hybridization [49–53]. Therefore, a network approach is better suited for analyzing citrus data with many admixed and hybrid individuals.

Previous studies have reported diversity in *Citrus* [54–57], however, these studies have focused on a rather small number of selected accessions. The diversity, admixture events, and ancestry information has not been investigated in depth for the majority of the diverse germplasm collection available from the University of California Riverside Citrus Variety Collection (CVC). The aim of this study was to conduct detailed genetic characterization of accessions in CVC with high-density SNP genotype data generated by a recently developed SNP genotyping array for Citrus, Axiom™ Citrus Genotyping Array (Affymetrix, Inc.) (58K autosomal and 500 Chloroplast SNPs), combined with whole genome sequence data of the variant discovery panel used to design the SNP genotyping arrays (Hiraoka et al. 2020). In order to reveal the complex history of

admixture and ancestry of citrus, multiple complementary methods of population genetics, such as PCA, admixture, phylogenetic network, and graphical genotyping analyses, were applied to array datasets, and the ancestry of many selected citrus accessions is proposed from integration of the results.

Methods

Sample preparation for array hybridization

Leaf samples (young flush) of the accessions of *Citrus* and its relatives were collected on ice in the field from the CVC (Riverside, CA) from 2014-2018. Each tree from which a leaf sample was collected was tagged with a unique identifier upon collection. Collected leaf samples were then returned to the laboratory on ice and the leaf surface was quickly and gently wiped with chloroform prior to drying with silica packs in a sealed, labeled sample bag. Silica dried leaf samples were homogenized with FastPrep-24™ (MP Biomedicals, Santa Ana, CA) with ¼” ceramic spheres or ¼” stainless steel beads. High molecular weight genomic DNA was then isolated from homogenized dried leaf samples with DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). The quality of DNA was assessed by electrophoresis, 260/280 and 260/230 ratios with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) and genomic DNA was quantified with a Qubit fluorometric system (Thermo Fisher Scientific, Wilmington, USA) according to manufacturer’s instructions. Genomic DNA of concentration of 25 ng/µl in 50 µl and 30 ng/µl in 25 µl for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array, respectively were prepared in Affymetrix™ bar-coded 96 well plates.

Genotyping and SNP calling

Hybridization data of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array for 924 CVC accessions (Table 1) was analyzed with Affymetrix Axiom® analysis suite software v.4.0.3.3 by Affymetrix™ at Axiom® BestPractice workflow default setting (default DQC, QC, and % passing) to identify “*Poly High Resolution*” (PHR) probes. PHR probes are the most informative and accurate probe performance class on the Axiom® platform based on the genotype call cluster separation, cluster variance and cluster position. In addition to PHR probes, “NoMinorHomo” probes (NMH) were included because alleles of these loci were useful for more distant taxa such as *Poncirus*. The most stringent threshold QC97 was applied first, followed by the lower threshold of QC95 to allow inclusion of accessions belonging to distant taxa such as *Poncirus*. All 126 CVC accessions that were analyzed with Axiom® Citrus HD Genotyping Array passed with QC97. A total of 799 CVC accessions were analyzed with Axiom® Citrus Genotyping Array, of which 754 passed with QC97, while 45 passed with QC95. Genotyping data obtained from Axiom® Citrus Genotyping Array was further filtered with R package SNPfisher (SNPfisher, v2.0.3, Affymetrix™) using a lowered confidence threshold of 0.01. A combined dataset with 56,367 PHR and NMH loci for 924 CVC accessions was then generated by adding Axiom® Citrus HD Genotyping Array genotyping calls to those of Axiom® Citrus Genotyping Array.

Percent identity among accessions in groups composed of clonal accessions and other taxonomic groups

The pairwise concordance of a genotype between the two data sets was defined as the fraction of the SNP genotyping calls for a given SNP that are in agreement with each other. “NoCall” was

excluded from the calculation. A total of 924 accessions analyzed with either the Axiom® Citrus Genotyping Array (799 accessions) or only the Axiom® Citrus HD Genotyping Array (126 accessions) were included in the analysis after removal of clonal and technical duplicate samples. Percent identity of accessions in groups composed of clonal accessions was assessed by comparison of genotype calls at 53,206 PHR loci within each group. Groups assessed included Navel orange (*C. sinensis*, 67 accessions), Valencia orange (*C. sinensis*, 25 accessions), Satsuma mandarin (*C. unshiu*, 45 accessions), Eureka and Lisbon lemon (25 accessions) and Clementine (*C. clementina*, 17 accessions), *P. trifoliata* (48 accessions). In each major clonal group, only those accessions not exceeding a percent identity threshold were included in subsequent diversity analyses. In addition to reducing the accessions in the major clonal groups, additional sets of identical to near-identical accessions were also identified from the array data and only one representative of nearly all such groups was included in the diversity analyses, resulting in the 382 UCR diversity set (Table 2). The pairwise concordance was also calculated using 500 chloroplast SNPs for 924 accessions.

Genetic parameter estimation

Genetic diversity parameters were estimated including observed heterozygosity (H_o), expected heterozygosity (H_e), polymorphic information content (PIC) and percent polymorphic loci with genotyping calls at 56,367 PHR and NMH loci in the combined dataset for 924 and the 382 UCR accessions. The level of genetic divergence between the populations using the 382 UCR accessions was estimated by F_{st} [58] using Plink [59]. Accessions analyzed include non-hybrid/non-admixed individuals representing all ancestral taxa and selected from the 382 UCR diversity set based on results obtained from admixture and introgressions analyses.

PCA

Principal component analysis (PCA) with the genotyping calls of 56,367 PHR and NMH autosomal loci was carried out with 924 CVC accessions and the 382 diversity set and with 500 chloroplast markers with the 382 diversity set. Both autosomal and chloroplast PCA analyses were carried out with software TASSEL v5.2.31 (<http://www.maizegenetics.net/tassel>).

Admixture

Admixture analysis was performed to infer the genetic structure of the ancestral population using unsupervised hierarchical clustering model based clustering software ADMIXTURE [37]. Admixture analysis was carried out to infer the number of ancestral populations (K) in the 382 UCR diversity set using both autosomal and chloroplast datasets. Genotyping data at 56,367 autosomal PHR and NMH loci was converted to bedfile using Plink prior to admixture analysis. The optimal K value was chosen based on Cross-validation (CV) filter of 10 for K1 to 10, in addition to biological inference from previous studies.

Phylogeny Inference and Phylogeny network analysis

For phylogenetic analysis of non-hybrid and non-admixed CVC accessions, 85 accessions, selected from the UCR diversity set, representing major citrus taxa and close relatives belonging to citrus ancestral species were selected. The SNP genotyping dataset was filtered by the method described above. All 85 accessions (Table 3) were selected based on the results obtained from admixture and introgression analyses (see below). The pipeline SNPhylo was used to perform phylogenetic analysis with *Poncirus* as an outgroup genus. SNPhylo first concatenated and aligned the SNPs,

then performed a maximum likelihood distance estimation without molecular clock and bootstrapping analysis. For SNPhylo the parameter settings were the minor allele frequency filter was set to 0, the maximum missing data filter was set to 1, and the LD filter was set to 1 to remove redundant SNPs by SNPrelate, resulting in 41,147 SNPs to be used in the analysis. The maximum likelihood tree by DNAML with bootstrap analysis with 1000 replications by Phangorn was constructed. The phylogenetic tree was then generated and visualized using FigTree v1.4.3. (<http://tree.bio.ed.ac.uk/software/figtree/>).

Phylogenetic network analysis was carried out with the 382 sample UCR diversity set and 36 accessions that were representative of the citrus variant discovery panel. Genotyping calls of 12,183 and 43,614 PHR loci were used for the analysis after SNPhylo filtering for 382 UCR diversity set and 36 citrus variant discovery accessions, respectively. For network analysis with the 382 UCR diversity set, the multiple alignment file was generated using SNPhylo. SNPhylo parameter settings were as follows; the minor allele frequency filter was set to 0, the maximum missing data filter was set to 0.02, and the LD filter was set to 0.7 to remove redundant SNPs by SNPrelate. For network analysis with 38 accessions, the parameter settings were the minor allele frequency filter was set to 0, the maximum missing data filter was set to 1, and the LD filter was set to 1 to remove redundant SNPs by SNPrelate. Changes in the SNPhylo parameter settings were primarily done to optimize for efficient run time for the 382 UCR diversity set. The maximum likelihood tree by DNAML with bootstrap analysis with 1000 replications by Phangorn was constructed for both datasets. Reticulation indicates phylogenetic incongruence, therefore, it is indicative of the presence of hybridization between the genomes of accessions being analyzed in the data. Neighbor-net analysis was carried out with default setting for both the 382 UCR diversity set and the 36 accessions using the filtered data sets described above using SplitsTree v4.14.6 to investigate a reticulated structure in citrus. Phylogeny analysis with a neighbor-joining tree was

also carried out for the 382 UCR diversity set using 500 chloroplast markers without any filtering and visualized using FigTree v1.4.3.

Graphical genotyping and identification of chromosomal level introgression.

“Species-specific” alleles were identified as those for which the difference in allele frequency between the putative ancestral taxa was > 0.9 . In admixed accessions, each 0.5 Mb bin was classified as homozygous for *C. reticulata*/*C. reticulata* (candidate type 1) [4] alleles (MM), *C. maxima*/*C. maxima* alleles (PP), heterozygous *C. reticulata*/*C. maxima* (MP), or undetermined (ND) when the frequency of one genotype was greater than 5 and that of alternate genotypes was less than five. The same procedure was followed for identification of chromosomal level introgression between papada (*C. micrantha* type) and *C. medica*, *C. medica* and *C. maxima*, *C. reticulata* and *C. medica*, *C. reticulata* and papada (*C. micrantha* type) and *C. reticulata* and *P. trifoliata* (Supplemental Data).

Results

Genetic parameters estimation

The genetic diversity was estimated using allele frequency, observed heterozygosity (H_o) and expected heterozygosity (H_e), polymorphic information content and percent polymorphic markers (Table 4). Genotyping data of 924 CVC accessions and the 382 UCR diversity set at 56,367 PHR and NMH loci were used. Average minor allele frequency, average H_o , average H_e , and mean PIC were 0.189, 0.241, 0.262, and 0.223 for 924 CVC accessions, and 0.19, 0.198, 0.268, and 0.251 for the 382 UCR diversity set, respectively. The percent polymorphic of markers with $q > 0.05$, $q > 0.01$,

and $q > 0.0$ were 0.802, 0.963, and 1 for 924 CVC accessions and 0.831, 0.979 and 1 for 382 UCR diversity set, respectively.

Population divergence parameter

Fst values indicate the allele frequency differentiation between pairs of populations as the difference between expected heterozygosity (under Hardy-Weinberg) in separate populations from that of a combined population [60]. The level of divergence between different ancestral taxa was estimated using Fst statistics. Both mean and weighted Fst were calculated given that there are different numbers of accessions representing each taxonomic group (Supplemental Table 1). Accessions were selected for analysis based on the results obtained from admixture and introgressions analyses. The smallest and largest mean Fst values were 0.515 and 0.845 and were observed between kumquat and microcitrus and between citron and trifoliolate, respectively. The smallest and largest weighted Fsts were 0.685 and 0.941 and observed between papeda and pummelo, and between citron and trifoliolate, respectively.

PCA

PCA was carried out with autosomal markers for 924 CVC accessions and the 382 UCR diversity set and with chloroplast markers for the 382 UCR diversity set. PCA with 924 CVC accessions show that the first three principal components (PCs), PC1, PC2 and PC3 explained 57.5% (27.9%, 17.1%, and 12.5% respectively) of the total variation in the dataset (Supplemental Table 2). The representatives in the citrus variant discovery panel appear at three of the corners of the 924 unique accessions plotted. PCA for the 382 UCR diversity set shows that the first three PCs explained

56.4% (29.3%, 20.1% and 7.1 % respectively) of the total variation in the dataset (Figure 1). PC4 explained 5.5% and 2.9% of the variation in the dataset for 924 CVC accessions and 382 UCR diversity set, respectively. PCA showed that accessions in groups constituted of clonal accessions, *C. sinensis*, *C. unshiu*, *C. clementina*, *Citrus limon* (L.) Burm. f. clustered tightly together, showing almost no variation within each group. All three PCA analyses showed that three classical *Citrus* ancestral taxa, mandarins (*C. reticulata*), pummelos (*C. maxima*), and citrons (*C. medica*) form discrete, well-separated clusters with non-admixed representatives clustering at the very tip of each cluster. In addition, PC3 in all three PCA analyses is mostly explained by inclusion of accessions of *Poncirus* (dark pink), its hybrids with *Citrus* (light pink), kumquat (*Fortunella*) (coral) and papeda (green, dark green, aquamarine) in the data analyzed. PCA carried out with chloroplast markers for the 382 UCR diversity set show that the first three principal components (PCs), PC1, PC2 and PC3 explained 79.1% (50.8 %, 18.2%, and 10.1% respectively) of the total variation in the dataset.

Admixture analysis

ADMIXTURE was carried out separately for autosomal and chloroplast markers for the 382 UCR diversity set. For autosomal admixture analysis, the appropriate number of K was estimated to be K4 to K6 based on the admixture clustering results. CV error decreased rapidly until K=4 (Δ CV error 0.31772) and changed little from K=4 to K=6 (Δ CV error 0.02499). The lowest CV error was 0.19632, corresponding to K=10 (Figure 2). The optimal K for autosomal admixture clustering was estimated to be at K=6 (Figure 3a, Table 5). For K=5, there was clustering of citron, mandarin, pummelo, trifoliolate but differences between papeda, kumquat and microcitrus were not resolved. At K=6 a cluster of papeda accessions was found. At K=8, non-admixed type 1 mandarins and a

few type 2 mandarins, such as *C. kinokuni* CRC 3887, Cleopatra CRC 3844, and *C. sunki* CRC 3143 form a separate cluster (Table 6).

Phylogeny Inference

Phylogeny was inferred for 85 accessions of non-hybrid and non-admixed CVC accessions representing all nine *Citrus* progenitor species, *C. medica*, *C. maxima*, *C. reticulata*, and *Citrus* relatives, *Fortunella*, papeda, and outgroup genus *P. trifoliata* (Figure 4). 41,147 LD filtered SNPs were included for phylogenetic tree construction. The bootstrap values of above 90% are shown as bold lines. 54.9% of total bootstraps were actually 100% in the tree generated. The citrus cluster is well separated from outgroup *Poncirus* with 100 percent bootstrap support and the five *Citrus* and *Fortunella* taxa form discrete clusters with complete monophyly and 100% bootstrap support. The accessions belonging to mandarin, pummelo and citron are confined to three of the major clusters, and those belonging to *C. reticulata* appear to be more divergent within the cluster based on their branch length in comparison to other clusters in the analysis. Accessions belonging to papeda, *Fortunella* and *P. trifoliata* are also in separate clusters. Interestingly, accessions currently classified as papeda were assigned to three separate clusters. Two *C. ichangensis* accessions *C. ichangensis* CRC 2327 and CRC 3931 cluster together and nested with the mandarin group, *C. hongheensis* CRC 3797, nested with the pummelo cluster, and Cabuyao CRC 432, Kalpi CRC 1456, *C. micrantha* CRC 3605, and *C. hystrix* CRC 3103 cluster together separately from other papeda groups. To investigate the level of genetic similarities within 85 accessions, a phylogeny tree was also constructed with the same data without placing *Poncirus* as an outgroup. The bootstrap values of above 90% are shown as bold lines, and 82.5 % of the above 90% bootstrap values were actually 100% in this tree. Interestingly, the citron cluster separates from the citrus

cluster and *Poncirus*. *Poncirus* clusters with *Citrus*, and is nested with two *C. ichangensis* CRC 2327 and CRC 3931 and mandarin cluster.

Phylogeny network

Phylogenetic relationships in selected accessions of citrus and close relatives were inferred using phylogeny network analysis. Neighbor-net analysis implemented in software SplitsTree v4.14.6 was used to generate a phylogenetic network for 36 citrus variant discovery panel accessions (Supplemental Table 2, Supplemental Figure 2) and the 382 UCR diversity set (Table 7, Figure 5) to infer the level of reticulation events in a diverse set of citrus accessions. Neighbor-net analysis with 36 citrus variant discovery panel accessions estimated average delta score [61, 62], and average Q-residual score [63] to be 0.227 and 0.02366, respectively (Supplementary Table 3). Neighbor-net analysis with the 382 UCR diversity set estimated average delta score, and average Q-residual score to be 0.309 and 0.0128, respectively (Supplementary Table 4), therefore, making this dataset non-tree like [63]. The smallest delta score observed was 0.278 for Clementine de Nules CRC 4056 and 0.262 for Thong Dee pummelo CRC 4027 and the highest delta score observed was 0.377 for Etrog citron CRC 3891 and 0.375 for Orban lemon CRC 4140 with 36 citrus variant discovery panel accessions and the 382 UCR diversity set, respectively. The smallest Q-residual score observed was 0.0143 for Nagami kumquat CRC 3877 and 0.00860 for Soh Niangrang papeda CRC 3203, and the highest Q-residual score observed was to be 0.0611 for Schaub rough lemon CRC 3879 and 0.0236 for Ponderosa lemon hybrid CRC 294 with 36 citrus variant discovery panel accessions and the 382 UCR diversity set, respectively. Both analyses (with and without *Poncirus* as an outgroup) suggest papeda's polyphyly, in agreement with network approach and chloroplast phylogeny.

Discussion

High density SNP genotyping arrays provide tools that are efficient and cost-effective for genome-wide genotyping to rapidly characterize elite germplasm and different breeding populations, allowing breeders to develop informed strategies in order to create a new superior cultivar. In addition, for germplasm represented in the CVC, there is little effect of ascertainment bias with this citrus array genotyping method, due to the array design panel being adequately diverse for the interpretation of analyses such as PCA, admixture and phylogeny analyses. Using SNP array data, we conducted a systematic diversity study in the diverse set of accessions of *Citrus*, close relatives and more distant taxa in CVC. The advent of cost-efficient sequencing technologies has made it possible to reveal genetic profiles and admixture history of *Citrus* [3–5, 26], but these efforts have studied a much more limited set of selected accessions than investigated here. However, there is growing interest in the genetics and the relationships among diverse accessions of citrus including related genera because of the wide phenotypic diversity and unique adaptation conferred by their genomes. Understanding genetic diversity and population structure is useful to plan for genetic improvement of citrus. Therefore, characterizing the CVC to gain such information is important for its potential contribution to problems associated with environmental changes and diseases, especially under the imminent threat of citrus greening (Huanglongbing). Additionally, the results obtained by model based clustering analysis such as ADMIXTURE are difficult to interpret. The model is built on the assumption that all the populations have risen from a single ancestral species and ignores historical models. Therefore, our study included PCA, ADMIXTURE, phylogeny inference and inferred introgression/admixture using both nuclear genome and chloroplast SNPs.

Genetic diversity in major citrus taxa in CVC

The level of genetic diversity was estimated using genetic parameters such MAF, Ho, He, PIC and percent polymorphic markers in 924 CVC accessions and UCR diversity set. The diversity assessment shows that the UCR diversity set represents the diversity that exists in the CVC, except for more distant taxa, excluded from 924 CVC accessions data. Population differentiation estimated as F_{st} varied from 0.685 to 0.941, between papeda and pummelo and between citron and trifoliata, respectively. Admixture analysis suggested papeda to be more closely related to kumquat and microcitrus than to citrus as expected based on the population differentiation estimation and phylogeny analysis.

PCA

PCA takes in the whole variation data such as high-density SNP data. PCA conducted with autosomal SNPs for 924 and 382 accessions successfully separated ancestral *Citrus* [64] and more distant taxa into separate clusters. *C. reticulata*, *C. maxima*, *C. medica* and *P. trifoliata*, a *Citrus* relative often used in rootstock breeding, clustered separately explaining the most variation within the dataset (57.5% and 56.4% for 924 and 382 accessions, respectively). PCA with 382 accessions show that non-admixed accessions belonging to each taxon, as identified in admixture and graphical genotyping analyses, appear at the tip of each represented taxon cluster. Type-1 mandarins *C. lycopersicaformis* CRC 3564, and Sun Chu Sha CRC 4003 appear at the tip of mandarin cluster, and divergent type 1 mandarin *C. tachibana* CRC 3150 appear away from the mandarin cluster, along with Yellow Rind CRC 3895 (Figure 1). PCA conducted with chloroplast SNPs and 382 accessions shows less overall diversity within the dataset (76.1 % explained by PC1, PC2 and PC3), in comparison to autosomal PCA analysis, but the clustering pattern was similar.

Population structure and inferred admixture

Estimating population structure and identifying genetically homogenous clusters is important in many fields of biology. In addition, this information is important in order to assess the appropriate marker density so that association studies are not affected by spurious associations. Population structure of the CVC was investigated using a reduced dataset comprising 382 accessions that represent the most of the diversity existing in CVC. Filtering of clonal and most of the near-identical accessions was performed in order to avoid spurious clustering results. The software ADMIXTURE infers population structure from high-density SNP data by a maximum-likelihood method to estimate the proportion of individual admixture and it runs based on the model that is built on the assumption that all the population have risen from a single ancestral species and ignores a historical model. The analysis revealed some signatures of admixture and genetic relationships among accessions under study. For the autosomal genome, at $K=6$, citron, pummelo, mandarin, trifoliate, papeda and kumquat/microcitrus form clusters and represent each K .

Washington navel is one of the most important sweet orange varieties cultivated worldwide. Navel orange arose from natural somatic mutation from sweet orange which is a natural interspecific hybrid of pummelo and mandarin which was further successively hybridized by mandarin pollen donor [3, 4]. At $K=6$, the genomic composition of Parent Washington navel CRC 1241A is pummelo (33%) and mandarin (67%). For Standard sour orange CRC 628, which is a supposed simple F1 hybrid between pummelo and mandarin [3, 4], admixture composition was pummelo (44%) and mandarin (56%). The accessions belonging to F1 hybrid groups such as citrandarin, citrange, citrumelo, grapefruit, lemons, limes, Rangpur, rough lemon, and tangor, appeared to have the genomic compositions expected of F1 hybrids. Admixture analysis also revealed that there are accessions currently classified as non-hybrid/non-admixed that are actually hybrid/admixed, in addition to differences in the proportions of the genomic components in the

accessions belonging to some hybrid groups. For example, Sexton CRC 2543 is described as tangelo of unknown parentage and USDA 1-77-19 CRC 4208 (US Seedless Surprise) is an irradiated USDA 75-8 which is a tangelo x grapefruit hybrid. Sexton CRC 2543 and USDA 1-77-19 CRC 4208 have higher pummelo genomic composition (56% to 58%) and lower mandarin genomic composition compared to other accessions of tangelo and tangor analyzed in UCR diversity set. Admixture analysis shows that the pummelo genomic composition of tangelo and tangor ranges from 6.3% with Koster CRC 3958 to 48% with USDA 177-19 CRC 4108, with average and median pummelo content of 26% and 24%, respectively. It should be noted that Koster CRC 3958 is a supposed mutant of Ellendale CRC 3960, however, parentage of Ellendale CRC 3960 is unknown (CVC archives). The percent identity between two accessions is above 99.8% and the plots showing chromosomal level introgression revealed remarkable similarities between the two accessions except for a small region of *C. reticulata*/*C. reticulata* on chromosome 1 (24.4-25Mb) of Koster CRC 3958, supporting its supposed origin (Figure 6, 7). Phylogeny network and chloroplast phylogeny places Koster CRC 3958 near admixed mandarins. Mency CRC 3183, recorded as a hybrid of Mediterranean sweet orange and Dancy mandarin, also has low genomic contribution by pummelo (7.0%) based on admixture analysis. However, in contrast to Koster CRC 3958, phylogeny network analysis places Mency CRC 3183 with other tangors such as Ruddy CRC 4146, and Dweet CRC 3183. These observations suggest the possibility that Koster CRC 3958, and therefore Ellendale, are not tangors.

Mandarin

Mandarin, *C. reticulata*, classified under genus *Citrus* [2], is one of the five Asian *Citrus* progenitors, and a remarkably diverse taxonomic group in terms of phenotype and genotype. It is

also a very economically important citrus group ranking as having the second highest citrus production after oranges. Many studies indicate that many cultivated mandarins are introgressed with pummelo [3, 5, 26, 27], contrary to widespread beliefs before the citrus genomic era. Swingle classified accessions of mandarins into four different groups, Satsuma, king, mandarin, and tangelo [2], while Tanaka classified them into five groups. In addition, Tanaka's system included 36 species in the five mandarin groups based on morphological characteristics [28].

C. amblycarpa CRC 2485, described as a small sour mandarin [2], showed mandarin (52%) and papeda (48%) ancestry, which is in agreement with previously published studies reporting it is a mandarin and papeda hybrid based on 24 SSR markers [54], 103 species diagnostic markers [65], and 1,457 SNP markers [66]. Phylogeny network analysis places *C. amblycarpa* CRC 2485 with papeda x mandarin hybrid accessions such as Kulobot CRC 3612 and *C. webberri* CRC 767, which are both papeda hybrids with mandarin. Chloroplast phylogeny clusters *C. amblycarpa* CRC 2485 with *C. micrantha* CRC 3605 and *C. hystrix* CRC 3103, also supporting the possibility that papeda, *C. micrantha*, is the maternal ancestral genotype of this accession [67]. Graphical genotyping showing the pattern of chromosomal level introgression of *C. amblycarpa* CRC 2485 (Figure 8) demonstrates that its chromosomes are indeed all admixed confirming it is a mandarin and papeda hybrid.

C indica was reported to be the most primitive species and possible progenitor of cultivated Citrus [68]. Previous study of CRC 3163 with SSR markers reported admixture between citron, papeda and kumquat groups questions the description that it is a mandarin hybrid [54]. The more detailed admixture analysis reported here suggested *C indica* CRC 3163 having citron (60.7%), papeda (28%), and kumquat/microcitrus (12%) genomic ancestry. *C indica* CRC 3163 shared no admixture signature with mandarin, despite sometimes being classified as a mandarin hybrid. Phylogeny network analysis places *C indica* CRC 3163 with papeda hybrids in between the citron

cluster and *C. hongheensis* CRC 3797, and it is nested with citron cluster according to chloroplast phylogeny. These results suggest a complex history of admixture of *C. indica* CRC 3163.

C. tachibana CRC 3150 is reported to be a genetically divergent type 1 mandarin with no admixture with pummelo [4]. Admixture analysis of *C. tachibana* CRC 3150 showed mandarin (93%) and kumquat/microcitrus (7%) genomic ancestry at K=8, where non-admixed type 1 mandarins and a few type 2 mandarins, such as *C. kinokuni* CRC 3887, Cleopatra CRC 3844, and *C. sunki* CRC 3143 form a separate cluster. Phylogeny network analysis places *C. tachibana* CRC 3150 with *C. depressa* CRC 2448 and near *C. nippokoreana* CRC 3228, *C. hanayu* CRC 3469 and Yuzu CRC 1216. *C. hanayu* CRC 3469, described as a Yuzu type (CVC archives), is reported to have Yuzu as a seed parent and *C. tachibana* as a pollen parent [27] which is in agreement with our observation. Interestingly, admixture analysis of *C. hanayu* CRC 3469 revealed its genomic composition to be mandarin (55%) and kumquat/micrantha (45%). The genomic contribution by kumquat/micrantha to *C. hanayu* CRC 3469 is explained later in the discussion in regards to the ancestry of Yuzu CRC 1216. Chloroplast marker analysis places *C. tachibana* CRC 3150 between a cluster composed of two *C. ichangensis* CRC 2327 and CRC 3931, *C. nippokoreana* CRC 3228, *C. hanayu* CRC 3469 and Yuzu CRC 1216 and the mandarin cluster. It should be noted that the Japanese indigenous mandarin species, *C. tachibana* CRC 3150, *C. depressa* CRC 2448, and *C. nippokoreana* CRC 3228 all showed a similar genomic contribution by kumquat/microcitrus.

C. depressa CRC 2448, also known as Shekwasha, is described as a small sour mandarin [2] and a possible hybrid between *C. tachibana* and some form of mandarin [69]. Admixture analysis shows its genomic composition to be mandarin (96%) and kumquat/microcitrus (4%) at K=8. Phylogeny network analysis revealed its close relationship to *C. tachibana* CRC 3150, in addition to its placement in the mandarin cluster by chloroplast phylogeny, supporting that the

observation made previously by Hayata that *C. depressa* CRC 2448 is likely to be a *C. tachibana* hybrid [69]

C nippokoreana CRC 3228, also known as Korai tachibana, shows genetic admixture pattern similar to that of *C tachibana* CRC 3150 and *C. depressa* CRC 2448, with a higher proportion of and kumquat/microcitrus (33%) at K=8. In contrary to what the name implies, morphological traits and tree growth habit of *C. nippokoreana* differs from those of *C. tachibana* [70]. *C. nippokoreana*, originally classified as *C. tachibana*, was reclassified as *C nippokoreana* by Tanaka [28]. Phylogeny network analysis clusters *C nippokoreana* CRC 3228 with *C. hanayu* CRC 3469 and Yuzu CRC 1216. Chloroplast phylogeny also places *C nippokoreana* CRC 3228 with *C. hanayu* CRC 3469 and *C. ichangensis* CRC 2327 and CRC 3991. These results suggest the possibility that *C nippokoreana* CRC 3228 is an *ichangensis* hybrid with tachibana or some other tachibana type mandarin. In addition, based on 246 SSR/indel markers it is reported that Yuzu and *C. tachibana* type are the possible maternal and pollen donors of *C nippokoreana* CRC 3228 [27], aligning with the observation from our study.

The genomic composition of both *C leiocarpa* CRC 3147 and *C. keraji* CRC 3144 also includes kumquat/microcitrus as observed with the three Japanese indigenous species of mandarins. However, graphical genotyping clearly revealed a very different pattern from *C. tachibana* CRC 3150 suggesting a much higher level of introgression of kumquat/microcitrus into the genome (Figure 9, 10, 11).

C. leiocarpa CRC 3147, also known as Koji orange, is described as seedy small mandarin. Its close relationship with *C. tachibana* was suggested first by Tanaka [71]. However, Swingle described *C. leiocarpa* to be a *ichangensis* hybrid [72]. Admixture analysis shows its genomic composition includes pummelo (6%) and kumquat/microcitrus (11%) at K=6. Phylogeny network

analysis places *C. leiocarpa* CRC 3147 between Kabosu ichandarin CRC 3943 and *C. benikoji* tangor CRC 3149. Kabosu CRC 3943 is believed to be papeda and sour orange hybrid (CVC archives), however, its parentage was recently reported to be Kunembo and Yuzu [27]. Chloroplast phylogeny suggests *C. leiocarpa* CRC 3147 has close relationship to *C. benikoji* CRC 3149 and other lemon hybrids. Graphical genotyping analysis shows remarkable similarities between *C. leiocarpa* CRC 3147 and *C. benikoji* tangor CRC 3149 for chromosomes 2, 5, 6, and 7 (Figure 12). These observations might suggest that *C. leiocarpa* CRC 3147 is hybrid between *C. tachibana* and *C. benikoji*, aligned with the recent reported parentage *C. leiocarpa* being koji type and *C. tachibana* [27].

C. keraji CRC 3144 is a small fruit mandarin that has the taste of tangerine (CVC archives). Admixture analysis revealed its genomic composition to include mandarin, plus pummelo (12%) and kumquat/microcitrus (12%). Phylogeny network identifies its close relationship with *C. depressa* CRC 2448, *C. tachibana* CRC 3150, and Pursha x Chinotto hybrid CRC 3578, which all shows an evidence of kumquat/microcitrus ancestry. Chloroplast phylogeny clusters *C. keraji* CRC 3144 with Kunembo CRC 3346 in the sweet orange cluster, while *C. depressa* CRC 2448, Chinotto CRC 3728, and Pursha x Chinotto hybrid CRC 3578 cluster with mandarins. The pattern of chromosomal level introgression showed remarkable similarities between some of the chromosomes between *C. keraji* CRC 3144 (Figure 10) and Kunembo CRC 3346 (Figure 13), also suggesting the possibility that they are related. It has been reported that *C. keraji* is a offspring of Kunembo based on 246 SSR/indel and inter simple sequence repeat (ISSR) markers [27, 73], in agreement with our observation that maternal genotype of *C. keraji* CRC 3144 is likely to be Kunembo CRC 3346. Interestingly, our observation suggests that there was some genetic basis for Tanaka's system to separate *C. tachibana*, *C. depressa*, *C. keraji*, *C. nippokoreana*, and *C. leiocarpa* [74].

In addition, these results support the published report of *C. tachibana* being divergent type 1 mandarin, rather than it being admixed with kumquat/microcitrus is an artifact of the software algorithm [75]. Graphical genotyping analysis indicates no evidence of admixture along the genome of *C. tachibana* CRC 3150 (Figure 11), while *C. depressa* CRC 2448 (Figure 14) shows a pattern of introgression similar to type 2 mandarins such as Cleopatra CRC 3844 (Figure 15), and *C. sunki* CRC 3143 (Figure 16) with minimal introgression from pummelo, [4].

Our study identified candidate type 1 mandarin accessions in CVC that have never been reported previously. The candidate type 1 mandarin accessions include Yellow Rind mandarin CRC 3895, *C. lycopersicaformis* CRC 3564 and *C. reticulata* CRC 3239 (Figure 17, Supplemental Figure 3), along with reported type 1 mandarin accessions *C. tachibana* CRC 3150 and Sun Chu Sha CRC 3150. *C. reticulata* CRC 3239 is not included in the UCR diversity set due to its high concordance with *C. lycopersicaformis* CRC 3564 (99.9%). In addition, these candidate type 1 mandarins show low percent genome-side heterozygosity in comparison to other hybrid mandarin types. Average percent heterozygous markers for candidate type 1 mandarins were 7.1% and the median was 7.4%. Phylogeny network analysis shows that Sun Chu Sha mandarin CRC 4003, *C. lycopersicaformis* CRC 3564, and Yellow Rind mandarin CRC 3895 cluster near each other, as a sister group to *C. tachibana* CRC 3150 cluster. Our study also identified candidate type 2, Cleopatra/*C. sunki* like accessions, Canton CRC 3576, and Mihrio CRC 4248 that have not been reported previously (Figure 18, Supplemental figure 4). Average and median percent heterozygous markers was 7.8% and 7.1%, while it was 16.0% and 14.3%, respectively for admixed mandarins, including type 2 mandarins.

Citron

Citron, *C. medica*, is one of the five Asian *Citrus* progenitors giving rise to Rangpur, lime, lemon, and rough lemon. Citron is self-compatible and reported to be largely monophyletic [76]. Indian CRC 661 was reported to be a hybrid with non-citron genotype (40%) [54] and it was since suspected to have pummelo ancestry (CVC archives). Admixture analysis revealed its genomic composition to be citron (54%) and papeda (46%). Phylogeny network analysis clusters Indian citron hybrid CRC 661 with *C. pennivesiculata* CRC 2434. Admixture analysis of *C. pennivesiculata* 2434 also shows its genomic composition to be citron (55%) and papeda (49%). Chloroplast phylogeny places Indian CRC 661 in the cluster with Soh Long CRC 3199, a supposed lemon hybrid (CVC archives), which is a papeda x citron hybrid according to the admixture results. These observations imply that Indian CRC 661 is a citron hybrid with papeda, rather than with pummelo.

Both Badhri CRC 3723 and Dabbe CRC 3190 are described as citron hybrids of unknown parentage. Admixture analysis suggests citron, pummelo and mandarin ancestry. Chloroplast phylogeny places Badhri CRC 3723 with the pummelo group along with Djerok CRC 2456 and Bengal CRC 3055. Djerok CRC 2456 is a supposed citron lemon hybrid (CVC archives), though it is reportedly a hybrid of citron (55%) and pummelo (45%) based on SSR markers [54].

Bengal CRC 3055 is described as a citron hybrid of unknown parentage, which was reported to be a citron and pummelo hybrid [54]. Admixture analysis suggests both Djerok CRC 2456 and Bengal CRC 3055 are pummelo citron hybrids with genomic composition of pummelo (50.4% and 67.1% for Djerok CRC 2456 and Bengal CRC 3055, respectively and citron (49.5% and 32.9% for Djerok CRC 2456 and Bengal CRC 3055, respectively). Our study aligns with the report that Djerok CRC 2456 is a pummelo citron hybrid [54] since there is no evidence of mandarin

ancestry and in particular, chromosomal level introgression showed the pattern observed with other F1 hybrids (Figure 19). In addition, it was reported that the majority of alleles of Bengal CRC 3055 were donated by *C. maxima* [54], which agrees with our observation as there is estimated 30 % *C. maxima/C. maxima* region at the chromosomal level (Figure 20). Dabbe CRC 3190 clusters with lemon and other sour and bittersweetorange types on the basis of chloroplast phylogeny. These results suggest Dabbe CRC 3190 could be a citron and lemon hybrid while ancestry of Badhri CRC 3723 needs to be further investigated. Our study identified 26 non-admixed citrons in CVC. Spadifora CRC 3535, Philippine CRC 3533 and S-1 CRC 3878 were not included in the 382 CVC diversity due to their high concordance with Corsican CRC 1795, Dulcia CRC 3654 and Etrog CRC 3891, respectively.

Lemon

Lemon is the third largest citrus group by production [77]. Lemon is classified as a natural hybrid of sour orange and citron [78, 79], which was confirmed by sequencing study [4, 57, 80]. However, the ancestry of most accessions represented in the lemon taxonomic group still largely remains controversial. The lemon group is highly polymorphic in terms of phenotype and genotype [80].

Interdonato is described as a lemon citron hybrid [69, 81]. Admixture analysis showed that there is a higher percentage of genomic contribution by citron (72%) in Interdonato CRC 3593 in comparison to the other lemon types, which have an average citron genomic contribution of 50%, as expected based on reported ancestry of most. The average genomic contribution by citron, mandarin and pummelo in the lemon group is 50% (stdev 16%), 26% (stdev 8%), and 15% (stdev 15%), respectively.

Bitrouni CRC 3265, Lumina CRC 3925, Soh Synteng lemon CRC 3261, Iran CRC 3162, Unnamed lemon hybrid CRC 3154, and Unnamed Morocco CRC 3155, all described as lemon or lemon hybrids (CVC archives), have significantly higher proportion of genomic contribution by pummelo and less by mandarin in comparison to other lemons. It should be noted that mandarin genomic contribution (2%) in Bitrouni CRC 3265 genome could result from an artifact of the program. However, when the pattern of chromosomal level introgression was compared to those of pummelo citron hybrids such as Bengal CRC 3055, Djerok CRC 2456, Ponderosa CRC 294, and Siamese CRC 2421, the pattern observed in Bitrouni CRC 3265 suggests that it could be possible that Bitrouni CRC 3265 has mandarin ancestry because of the larger proportion of undetermined (ND) region often observed when the accession is introgressed with more than two ancestors (Supplemental Data).

Orban CRC 4140 and Improved Meyer CRC 3737, described as lemon hybrids with orange types (CVC archives), and New Zealand lemonade CRC 4181, Limonette de Marrakech CRC 3989, and Iraq lemon CRC 3492, described as lemon hybrids of unknown parentage (CVC archives), have significantly more mandarin genomic contribution than the main lemon group. Interestingly, Unnamed lemon CRC 3885 has evidence of papeda ancestry (6%). Phylogeny network and chloroplast phylogeny place Unnamed lemon CRC 3885 with other lemon types, suggesting a complex ancestry with introgression from papeda.

C. volkameriana CRC 3040, reported to be a hybrid between sour mandarin and citron [82], shows mandarin (52%) and citron (48%) ancestry as expected.

India CRC 2323, Paak Link Mung CRC 3193, and Unnamed lemon hybrid CRC 3172, described as lemon hybrids of unknown parentage (CVC archives), show no evidence of pummelo ancestry and have genomic contribution from mandarin and citron, suggesting that they are rough

lemon types. In addition, all chromosomes seem to be heterozygous for mandarin and citron alleles as expected for F1 hybrid types. However, while the phenotype of India CRC 2323 resembles that of rough lemon types, Paak Link Mung CRC 3193, and Unnamed lemon hybrid CRC 3172 resemble each other phenotypically more so than they do rough lemon types. Interestingly, chloroplast phylogeny places India CRC 2323 with other rough lemons such as South African CRC 2325, Schaub CRC 3879, and Vangasay CRC 3996, while Paak Link Mung CRC 3193, and Unnamed lemon hybrid CRC 3172 formed a different cluster, separate from the rough lemon cluster, suggesting a different maternal ancestry from that of India CRC 2323.

Our study also found that there are accessions currently classified as lemon or lemon hybrid (CVC archives) with indication of papeda ancestry, while showing no evidence of pummelo ancestry. *C. excelsa* CRC 2316, Nicaraguan CRC 3841, and Soh Long CRC 3199 all have genomic contributions from papeda ranging from 27% with Nicaraguan CRC 3841 to 48% with *C. excelsa* CRC 2316. Phylogeny network places *C. excelsa* CRC 2316, Nicaraguan CRC 3841, and Soh Long CRC 3199 with Davao CRC 2427, and Mexican lime CRC 1710, suggesting that these accessions are limes types, rather than lemon types. Interestingly, all the chromosomes of *C. excelsa* CRC 2316 and Soh Long CRC 3199 seem to be heterozygous for papeda and citron alleles, as expected for F1 hybrid type, while Nicaraguan CRC has citron specific regions spanning a very large portion of its genome. Our observations suggest that *C. excelsa* CRC 2316 and Soh Long CRC 3199 have similar ancestry to that of limes., while Nicaraguan CRC 3841 is likely to be a lime hybrid with citron.

Ponderosa CRC 294, described as a possible hybrid between citron and lemon (CVC archives), shows evidence of pummelo (51%) and citron (49%) ancestry. The admixture pattern of chromosomes also indicated that Ponderosa CRC 294 is likely to be a F1 hybrid type, rather than lemon hybrid with citron. Phylogeny network analysis shows its close relationship to other

pummelo x citron hybrids, while clustering with a pummelo group based on chloroplast phylogeny, indicating the Ponderosa CRC 294 is likely a pummelo x citron hybrid.

Mitha Tulia CRC 3051, a supposed limetta/sweet lemon (CVC archives), shows citron (48%), mandarin (52%) ancestry but no evidence of a genomic contribution by pummelo. In addition, both phylogeny network and chloroplast phylogeny place Mitha Tulia CRC 3051 with other rough lemons. Based upon these observations and its highly admixed chromosomal patterns support that Mitha Tulia 3051 is likely to have a similar ancestry to that of rough lemons.

Lime

Lime is the third largest citrus group along with lemon [77]. Lime is believed to be a natural hybrid of citron and *C. micrantha* or close papada species [4, 65, 78]. However, the ancestry of the lime taxonomic group also remains controversial as in the case of lemons.

Excalibur lime CRC 4239, a presumed red lime hybrid derived from a cross between kumquat and Rangpur (CVC archives, Richard Wilson of Excalibur Nursery), showed evidence of Kumquat/microcitrus (53%), mandarin (33%) and pummelo (15 %) ancestry at K=6. At K=8, there is an admixture signature of kumquat/microcitrus (52%) and orange type (48%), suggesting the possibility that Excalibur lime CRC 4239 is a hybrid between kumquat and some orange type, rather than with Rangpur (which is mandarin x citron). In addition, chloroplast marker phylogeny clusters Excalibur lime CRC 4239 with other kumquat hybrids, confirming its maternal ancestry. Phylogeny network analysis placed Excalibur lime CRC 4239 near kumquat mandarin hybrids such as Indio CRC 3759, Citrangequat-19-15-7 CRC 3259, and Rio Grande Valley CRC 4097. These accessions show similar pummelo and kumquat genomic composition with pummelo content ranging from 4% with Citrangequat 19-15-7 CRC 3259 to 15% with Excalibur lime CRC 4239,

and with kumquat content ranging from 52% with Rio Grande CRC 4097 to 54% with Citrangequat 19-15-7 CRC 3259.

True lime type accessions such as Mexican lime CRC 1710, and Davao CRC 2427 have genomes consistent with an origin as F1 hybrids of papeda and citron, as suggested by other authors [4, 65]. Our study revealed that the same pattern was also observed for the chromosomes of *C. pannivesiculata* CRC 2434, *C. excelsa* CRC 2316, Soh Long CRC 3199, *C. vulgaris* CRC 760, India CRC 661, and possibly *C. macrophylla* CRC 2446. Similar patterns of genomic composition and the introgression patterns along the genome (Figure 21), suggest that these accessions are all likely to be lime. However, it should be noted that *C. macrophylla* CRC 2446 needs to be further investigated whether it is truly an F1 hybrid lime type. Chromosome five of *C. macrophylla* CRC 2446 (Figure 21f) has a small citron specific region. However, this could be an error because this region has relatively few markers.

Papeda

Papeda is one of the five Asian *Citrus* progenitor species. Subgenus *Papeda* [2] is the least known group in *Citrus*. Nonetheless, it is one of the important ancestral progenitor species of *Citrus*. Papeda phylogeny has also been controversial for many years. Recent studies reported papeda to be polyphyletic, questioning Swingle's classification of subgenus papeda [4, 54, 80, 83–85]. *C. ichangensis* CRC 2327 and CRC 3931, commonly known as Ichang papeda, belong to subgenus *Papeda* by Swingle's classification system and are described as the hardiest evergreen species of *Citrus* known [2]. Our study identified six non-admixed papeda accessions: Cabuyo CRC 432, *C. micrantha* CRC 3605, Kalpi CRC 1456, *C. hystrix* CRC 3103, *C. hystrix* cabuyo RRUT 361, and

C. hystrix 2454, *C. hystrix* cabuyo RRUT 361, and *C. hystrix* 2454 were not included in the UCR diversity set due to its high concordance with *C. hystrix* 3103 (99.8%).

Although it remains controversial, *C. ichangensis* is thought to be a more genetically divergent type of papeda [4]. Interestingly, admixture analysis reveals that two *C. ichangensis* accessions, CRC 2327 and CRC 3931 cluster with the kumquat/microcitrus group rather than with the papeda group with genomic composition of 91% and 9%, respectively. This observation also indicates the likelihood *C. ichangensis* CRC 2327 and CRC 3931 are divergent [4] from a non-admixed papeda group and is in agreement with our phylogeny analysis (Figure 4). The clustering of *C. ichangensis* CRC 2327 and CRC 3931 with kumquat/microcitrus at K=6 in admixture analysis is likely the result of an artifact and ascertainment bias introduced by array design [75], as phylogeny analysis on 85 accessions shows that the two *C. ichangensis* accessions cluster away from the non-admixed papeda, kumquat, and microcitrus clusters (Figure 4). Chloroplast phylogeny analysis revealed these two *C. ichangensis* accessions nesting closely with *C. hanayu* CRC 3469 and *C. nippokoreana* CRC 3228, and away from other papeda accessions. The relationships between *C. hanayu* CRC 3469 and *C. nippokoreana* CRC 3228 with *C. ichangensis* were already discussed. These observations strongly support the polyphyletic nature of the papeda group.

Interestingly, admixture analysis suggests that Yuzu CRC 1216, Kabosu CRC 3943, *C. yuko* CRC 3146 and Ichang lemon CRC 1215, all classified as papeda hybrids (CVC archives), show no evidence of papeda ancestry. Yuzu CRC 1216, a supposed *C. ichangensis* and satsuma hybrid (CVC archives), was described by Swingle to be a chance seedling of *C. ichangensis* [72]. Admixture analysis of Yuzu CRC 1216 shows genomic contributions by kumquat/microcitrus (55%) and mandarin (45%), similarly to that of *C. hanayu* CRC 3469. Given that admixture analysis cluster *C. ichangensis* CRC 2327 and CRC 3931 with kumquat/microcitrus due to possible

program artifact and ascertainment bias, results obtained from network analysis and chloroplast phylogeny for Yuzu CRC 1216 is in agreement with the previous report, confirming *C. ichangensis* as the seed parent. However, admixture analysis shows no evidence of pummelo ancestry making it unlikely that satsuma is the pollen parent since satsuma has regions heterozygous for mandarin-pummelo alleles on each chromosome. In addition, the close relationship between Yuzu CRC 1216 and *C. tachibana* CRC 3150 was already discussed above, where it is suggested that *C. tachibana* CRC 3150 may be the pollen parent of Yuzu CRC 1216.

Kabosu CRC 3943 is a presumed hybrid between papeda and sour orange (CVC archives) which was later reported to be Kunembo and Yuzu [27]. Admixture analysis showed Kabosu CRC 3943 having genomic contributions by mandarin (55.1%), kumquat/microcitrus (32.7%) and pummelo (12.2%) and no evidence of papeda ancestry, as expected if *C. ichangensis* ancestry is suspected. Network analysis places Kabosu CRC 3943 between *C. nippokoreana* CRC 2448, *C. hanayu* CRC 3469, and tangors. Chloroplast phylogeny clusters Kabosu CRC 3943 and with Kunembo CRC 3346, admixed mandarin and orange types, while Ichang lemon CRC 1215 clusters with admixed pummelo and sour orange types. These observations, the possible pollen parent of Yuzu CRC 1216 being *C. tachibana* CRC 3150 and the relationship between *C. tachibana* CRC 3150, *C. nippokoreana* CRC 2448, and *C. hanayu* CRC 3469 as discussed above, suggest that the maternal ancestor Kabosu CRC 3943 could be Kunembo and the pollen donor could be Yuzu CRC 1216 as reported [27], though this should be further investigated.

C. latipes CRC 3052, known as Khasi papeda, was described as a small papeda similar to *C. ichangensis* based on morphology [69]. *C. latipes* CRC 3052, *C. hongheensis* CRC 3797 and Unnamed CRC 3056 show evidence of pummelo (average 51%), papeda (average 26%) and kumquat/microcitrus (average 22%) ancestry based on admixture analysis. Phylogeny network analysis clusters *C. latipes* CRC 3052, *C. hongheensis* CRC 3797, and Unnamed CRC 3056

together and away from other papeda and ichangensis clusters, while chloroplast phylogeny also places these accessions together and separate from papeda and ichangensis clusters, except for *C. hongheensis* CRC 3797 which is due to their kumquat/microcitrus genomic contribution, we could hypothesize that they are all derivatives of *C. ichangensis*. In addition, while it remains possible, the admixture clustering pattern and low observed percent heterozygous markers in these accessions (Supplemental Data) (2.7% 2.2% and 2.5% for *C. latipes* CRC 3052, *C. hongheensis* CRC 3797 and Unmanned CRC 3056, respectively) makes it unlikely that they are simple hybrids between *C. ichangensis* and pummelo. In addition, the predicted percent heterozygous markers was 12.8 % using two ichangensis accessions and admixed and non-admixed pummelo accessions Hunan CRC 1225 and Siamese acidless CRC 2240, thereby implying the ascertainment bias alone cannot explain such low observed percent heterozygous markers. Further investigation to elucidate the relationships between these accessions in relation to *C. ichangensis* and the pummelo group is necessary.

Soh niangrang CRC 3203, classified as an unknown papeda hybrid (CVC archives), shows evidence of mandarin (55%), pummelo (25%), papeda (12%) and kumquat/microcitrus ancestry (9%) according to admixture analysis. Phylogeny network analysis places *Soh niangrang* CRC 3203 among mandarin pummelo hybrids. Chloroplast phylogeny shows *C. latipes* CRC 3052, *C. hongheensis* CRC 3797, Unnamed CRC 3056, and *Soh niangrang* CRC 3203 cluster together and away from the papeda and ichangensis cluster. In addition, the observed percent heterozygous markers with *Soh niangrang* CRC 3203 (27%), suggests it is likely a hybrid and the chloroplast SNPs indicate it has the same maternal ancestral genotype as *C. latipes* CRC 3052, *C. hongheensis* CRC 3797 and Unnamed CRC 3056.

Kulobot CRC 3612, a presumed papeda hybrid (CVC archives), shows the evidence of papeda (53%) and mandarin (47%) ancestry according to admixture analysis. Phylogeny network

analysis places Kulobot CRC 3612 with other papeda mandarin hybrid accessions, *C. amblycarpa* CRC 2485 and *C. webberii* CRC 767, while chloroplast phylogeny clusters Kulobot CRC 3612 with other non-admixed papeda accessions, suggesting the possibility of a non-admixed papeda (*C. micrantha* type) as its seed parent and mandarin to be the pollen parent as observed with *C. amblycarpa* CRC 2485 and *C. webberii* CRC 767. *C. macrophylla* CRC 2446, described as a papeda hybrid (CVC archives), has evidence of citron (51%) and papeda (49%) ancestry according to admixture analysis. Phylogeny network places *C. macrophylla* CRC 2446 with *C. excelsa* CRC 2316, Mexican lime CRC 1710 and lemon hybrids, while chloroplast phylogeny suggests a non-admixed papeda (*C. micrantha* type) as its seed parent, thereby confirming that Kulobot CRC 3612 is a papeda hybrid with mandarin. Our study finds there are at least two distinct papeda groups, one a non-admixed papeda group included Cabuyo CRC 432, *C. micrantha* CRC 3605, Kalpi CRC 1456 and *C. hystrix* CRC 3103, *C. hystrix* cabuyo RRUT 361, and *C. hystrix* 2454 and the other group consisting of two ichangensis accessions, CRC 2327 and CRC 3931 (Figure 4).

Kumquat

Genus *Fortunella* [2], commonly known as kumquat, is one of the five Asian *Citrus* progenitors and means “golden orange” in Chinese and Japanese [86]. Morphology of *Fortunella* is quite different from that of other *Citrus*, including having small flower buds, small fruits with acidic pulp and a sweet edible pulpy skin, nearly veinless leaves with small glandular dots, cavernous stigma containing large lysigenous oil glands, two colateral ovules in each locule and isomerous or hypomerous ovary. [2]. Our study found *F. polyandra* CRC 3901, Marumi SRA489 CRC 4125, Nordmann CRC 4070, *F. crassifolia* CRC 3818, and Meiwa CRC 3833 to be non-admixed kumquat

accessions in CVC. *F. crassifolia* CRC 3818, and Meiwa CRC 3833 were not included in CVC diversity set due to their high concordance to Marumi SRA489 CRC 4125.

F. obovata CRC 3475, also called Fukushu kumquat in Japan (CVC archives), was considered to be a chance hybrid between two *Fortunella* species by Swingle [69]. Interestingly, admixture analysis revealed that there is a genomic contribution by mandarin (23%). Phylogeny network analysis also places *F. obovata* CRC 3475 between Thomasville citrangequat CRC 1440 and the non-admixed kumquat cluster, while chloroplast phylogeny suggests a kumquat to be a seed parent, suggesting *F. obovata* CRC 3475 is a kumquat mandarin hybrid.

Citranequat 19-15-7 CRC 3259 is documented as a *Fortunella* x (trifoliolate *Citrus* hybrid) hybrid (CVC archives), with the *Citrus* component of the male parent not identified. However, admixture analysis shows its genetic composition to be kumquat/microcitrus (54%), and mandarin (36%). This observation differs from that for the other citrangequats, Thomasville CRC 1440 and Sinton CRC 3642 which have the expected proportion of trifoliolate ancestry. Phylogeny network analysis places Citrangequat 19-15-7 CRC 3259 between Rio Grande CRC 4097 and Excalibur lime CRC 4239, and chloroplast phylogeny clusters Citrangequat 19-15-7 CRC 3259 in the mandarin cluster, suggesting that Citrangequat 19-15-7 CRC 3259 could be a mandarin kumquat hybrid.

Pummelo

Pummelo, *C. maxima*, classified as *C. grandis* (L) by Swingle [72], is one of the five Asian *Citrus* progenitor species [3, 4]. Recent studies have been showing the importance of pummelo introgression for many cultivated mandarins [3–5, 26]. Our study identified 39 accessions in CVC that are considered not admixed. Haiku B CRC 3940 and Sarawak CRC 4042 are not included

in the UCR diversity set due to their high concordance to *C. grandis* CRC 3282 and Tahitian CRC 3806, respectively.

C. pennivesticulata CRC 2434, Sour CRC 3066, Lemelo CRC 1775, Lemelo CRC 148 and Moli kurikuri CRC 579 all show no evidence of pummelo ancestry despite them being classified as pummelo and pummelo hybrids (CVC archives). *C. pennivesticulata* CRC 2434, described as pummelo like (CVC archives), has genomic composition of citron (55%) and papeda (45%) as discussed above, similar ancestry to that of limes.

Sour CRC 3066 is currently classified as pummelo (CVC archives), however, admixture analysis reveals pummelo (90%), citron (6%), papeda (3%) ancestry. Phylogeny network analysis places Sour CRC 3066 between more divergent non-admixed pummelo accessions and `pummelo CRC 2421, which also has evidence of citron ancestry (18%) with a majority of alleles being donated by *C. maxima* (Figure 22). Chloroplast phylogeny places Sour CRC 3066 with lime and limequat accessions, confirming the involvement of papeda, and suggesting a complex, multigeneration ancestry.

Two Lemelo pummelo hybrid accessions, CRC 1775, classified as a presumed lemon and pummelo hybrid and CRC 1481 of unknown parentage, are described as not alike to each other (CVC archives). Admixture analysis shows Lemelo CRC 1481 has a genomic composition of pummelo (55.6%), citron (29.3%) and mandarin (15.1%) as opposed to Lemelo CRC 1775 having a genomic composition of pummelo (77.8%), citron (19.5%) and mandarin (2.7%). Lemelo CRC 1481 is reported to be almost edible while Lemelo CRC 1775 is very sour (CVC archives). This could be explained by a higher proportion of genetic contribution from mandarin in Lemelo CRC 1481 (15.1%), compared to Lemelo CRC 1775 (2.7%).

Moli kurikuri CRC 579 is described as a Fiji name for *C. macroptera* [69]. *C. macroptera* is described as papeda like [69], however, CRC 579 shows no evidence of papeda ancestry according to admixture analysis which indicates a genomic composition of pummelo (57%), citron (30%) and mandarin (12%). In addition, phylogeny network places Moli kurikuri CRC 579 with Lumina lemon CRC 3925 and Bengal citron CRC 3055 and chloroplast phylogeny suggests pummelo as seed parent, providing no evidence of its supposed *C. macroptera* ancestry.

Cuban Shaddock CRC 1462, described as resembling citron or lemon in most of its characters [69] (citrus industry), shows the highest proportion of genomic contribution by citron (49.5%) among pummelo hybrid accessions analysed from the 382 UCR diversity set. Phylogeny network places Cuban Shaddock CRC 1462 with Ponderosa CRC 294, a pummelo citron hybrid (CVC archives), suggesting Cuban Shaddock CRC 1462 is a pummelo citron hybrid.

Suisho buntan CRC 3947, *C. hirosimana* CRC 3275, also called Hiroshima-natsuzabon in Japan (CVC archives), *C. sino-grandis* CRC 3148, Chinese pummelo CRC 1212, Kawachi Bankan CRC 3944, a natural hybrid of pummelo found in Japan, Arajon CRC 2596, Banokan CRC 3950, and Kinokawa Buntan CRC 3978 are all classified as pummelo (CVC archives). However, our study identified them as admixed with mandarin with the mandarin genomic composition ranging from 14% (Suisho buntan CRC 3947) to 59% (Kinokawa Buntan CRC 3978).

Pummelo hybrid accessions analyzed in the 382 UCR diversity sets which identified to be admixed with mandarin shows mandarin genomic composition ranging from 13% to 62%, with average 37% and median 37% (Stdev 15%). Additionally, graphical genotyping and genome-wide heterozygous marker distribution of pummelo accessions in UCR diversity revealed that Banpeiyu CRC 3961, Red pummelo CRC 2245, Pong Yau CRC 2342, Moanalua CRC 448, and Pink

pummelo CRC 2244 are Chandler CRC 3224 type with a *C. maxima*/*C. reticulata* region on chromosome 2 (24 Mb to 25.5 Mb) (Figure 23) [4]

Orange, Sour Orange, and Bittersweet orange

Parent Washington CRC 1241A, is one of the most important cultivated citrus in the world. Navel oranges arose from somatic mutation from sweet orange (*C. sinensis*), which was revealed to have a complex history of admixture between *C. reticulata* and *C. maxima* [3, 4]. The genomic composition, based on admixture, was estimated to be mandarin (67%) and pummelo (33%), similar to values previously published [3]. The pattern of introgression between mandarin and pummelo at the chromosome level (Figure 24) was also assessed and found to be similar to previously published report [3, 4], except for the *C. maxima*/*C. maxima* region at 0 to 6 Mb on chromosome 8 which our analysis shows undetermined. These observations could have resulted from our analysis including 37 pummelo accessions, a significantly larger number than one used in the reported study [3]. In addition, chloroplast phylogeny and chloroplast percent identity identified only Hunan CRC 1225 (no other pummelo accessions with high concordance identified in the CVC) clusters with sweet oranges among pummelo and admixed-pummelo accessions.

Ambersweet CRC 4002, classified a sweet orange hybrid, and its reported parentage is (*C. clementina* x Orlando tangelo) x *C. sinensis* (CVC archives) [87]. It's overall genomic composition is very similar to that of Parent Washington CRC 1241A, however, graphical genotyping reveals significant differences in the pattern of chromosomal level introgression with much higher proportion of the chromosomes assigned as *C. reticulata*/*C. reticulata* and *C. maxima*/*C. maxima*, suggesting much larger contribution by *C. maxima* alleles as expected from such cross (Figure 25)

Sour orange is a supposed F1 hybrid resulting from admixture of *C. reticulata* and *C. maxima* [4] and is an important group used mainly as rootstock and for flavoring and perfumery [88, 89].

Chinotto CRC 3728, Bergamot CRC 2881, *C. canaliculata* CRC 3565 and *C. vulgaris* CRC 760, classified as sour orange hybrids (CVC archives), all show evidence of citron ancestry. Possible citron ancestry of Chinotto CRC 3728 has been previously reported [54]. Admixture analysis shows the genomic composition of Chinotto CRC 3728 to be mandarin (57%), pummelo (22%), and citron (18%). Phylogeny network analysis places Chinotto CRC 3728 with *C. rokugatsu* CRC 3473 and Pursha lime x Chinotto CRC 3578.

C. rokugatsu CRC 3473, also known as Zadaidai sour orange [69], has genomic composition of mandarin (66%), pummelo (21%) and kumquat/microcitrus (13%). Phylogeny network shows its close relationship to Pursha lime x Chinotto CRC 3578 as described above and places it near *C. depressa* CRC 2448 and *C. tachibana* CRC 3150, while clustering with sour orange and lemon types based on chloroplast. *C. rokugatsu* CRC 3473 has been suggested to be a hybrid with mandarin based on horticultural characteristics (CVC archives). Our observations suggest that it is possible that *C. rokugatsu* CRC 3473 is a sour orange hybrid with a tachibana type mandarin.

The genomic composition of Pursha lime x Chinotto CRC 3578 was revealed to be mandarin (69%), citron (26%), and kumquat/microcitrus (5%). Pursha lime is a supposed Rangpur and *C. depressa* hybrid (CVC archives) and the presence of kumquat/microcitrus (5%) could confirm *C. depressa* as an ancestor of Pursha lime as discussed above that *C. depressa* CRC 2448 has genomic contribution by kumquat/microcitrus based on admixture analysis. In addition, chloroplast phylogeny places Pursha lime x Chinotto CRC 3578 with Rangpur accessions and

Chinotto CRC 3728, suggesting that either Rangpur x *C. depressa* or Chinotto CRC 3728 could be the seed parent of Pursha lime x Chinotto CRC 3578.

Bergamot CRC 2881, presumed hybrid of sour orange and citron or lemon (CVC archives), is one of the most popular citrus used for perfumery [90]. Network analysis revealed its close relationship to lemon and rough lemon accessions, while chloroplast phylogeny places Bergamot CRC 2881 with Standard sour orange CRC 628 and other lemons. Admixture analysis shows its genomic composition to be pummelo (36%), citron (32%), and mandarin (31%). Chloroplast marker concordance (500 chloroplast SNPs) among Bergamot CRC 2881 and Standard sour orange CRC 628 and some lemons such as Frost Lisbon CRC 3176 is 99.8% as expected based on the phylogeny analysis. Further investigation is needed to elucidate the exact ancestry of Bergamot CRC 2881.

C. canaliculata CRC 3565, also known as Kikudaidai, is a presumed hybrid between sour orange and grapefruit (CVC archives). Its genomic composition shows the smallest proportion of citron (9%), and the highest pummelo genomic contribution (68%) among sour orange and sour orange hybrid accessions that are admixed with citron in our study. It has 14.8 % mandarin genome according to Admixture. The citron ancestry apparently rules out the existing parentage hypothesis.

C. vulgaris CRC 760 is described as a possible sour orange and citron hybrid (CVC archives), however, admixture analysis suggests citron (53%) and papeda (47%) ancestry. Phylogeny network analysis clusters *C. vulgaris* CRC 760 with other papeda citron hybrid accessions and chloroplast phylogeny suggests papeda as seed parent. These observations suggest that *C. vulgaris* CRC 760 is a papeda citron hybrid rather than a sour orange citron hybrid.

Indian CRC 3175 and *C. maderaspatana* CRC 3225, both classified as sour orange hybrids (CVC archives), have genomic compositions of mandarin (63% for Indian sour orange hybrid CRC

3175, 55% for *C. maderaspatana* CRC 3225) and papeda (37% with Indian sour orange hybrid CRC 3175, 45% with *C. maderaspatana* sour orange hybrid CRC 3225). Phylogeny network places them between the ichangensis cluster and a cluster with mandarin papeda hybrid accessions. Interestingly, chloroplast phylogeny places both Indian CRC 3175 and *C. maderaspatana* CRC 3225 in the mandarin cluster that includes Sun Chu Sha CRC 4003, and *C. lycopersicaformis* CRC 3564, candidate type 1 mandarins. Previous study suggested Indian CRC 3175 and *C. maderaspatana* CRC 3225 were closer to mandarins than pummelos [54]. These observations could suggest that Indian CRC 3175 and *C. maderaspatana* CRC 3225 are mandarin and ichangensis hybrids rather than sour orange hybrids. Both have mosaic genome structures, with substantial portions of many chromosomes homozygous for mandarin, pummelo and heterozygous for alleles from both ancestral taxa (Figure 26), a pattern which indicates both parents of these accessions were themselves interspecific hybrids.

Zhuluan CRC 3930, Myrtifolia CRC 3290, Nansho daidai CRC 2588, Standard sour orange CRC 628, Goutoucheng CRC 4004, Smooth Flat Seville CRC 4079, *C. natsudai* CRC 3235, *C. intermedia* CRC 3474, *C. aurantium* var. *salicifolia* CRC 3289, and Orogold CRC 693 are all classified as sour orange or sour orange hybrid (CVC archives). Standard sour orange CRC 628 represents a group of 9 nearly identical sour orange accessions based on the concordance analysis. Admixture analysis is in agreement with such description of these accessions showing mandarin (44% with Orogold CRC 693 and 60% with Zhuluan CRC 3930) and pummelo (40% with Zhuluan CRC 3930 and 56 % with Orogold CRC 693) ancestry. Phylogeny network clusters Zhuluan CRC 3930, Myrtifolia CRC 3290, Nansho daidai CRC 2588, Standard sour orange CRC 628, Goutoucheng CRC 4004, Smooth Flat Seville CRC 4079, *C. natsudai* CRC 3235 and *C. aurantium* var. *salicifolia* CRC 3289 together, except for *C. intermedia* CRC 3474 and Orogold CRC 693 that cluster together and away from the rest. In addition, *C. intermedia* CRC 3474 clusters

with lemon and orange types and Orogold CRC 693 clusters with pummelo type while all the other sour oranges cluster together based on chloroplast phylogeny. Interestingly, the comparison of chromosomal level introgression patterns among *C. intermedia* CRC 3474 and Orogold CRC 693 (Figure 27) revealed remarkable similarities among some of the chromosomes (chromosome 2, 8, 9) suggesting the possibility of shared of ancestry between these two accessions.

Orlando CRC 1588, classified as bittersweet orange (CVC archives), shows similar genomic composition with sour orange accessions with mandarin (56%) and pummelo (44%). Phylogeny network and chloroplast phylogeny also suggest its close relationship with sour orange type. Graphical genotyping reveals that the pattern of chromosomal level introgression of Orlando CRC 1588 is the same as Smooth Flat Seville CRC 4079, but these accessions differ in the pattern of heterozygosity of their chromosomes and in their chloroplast genome with Orlando CRC 1588 identical to the main group of sour oranges and Smooth Flat Seville CRC 3859 identical to Duncan grapefruit CRC 3921. Graphical genotyping of these accessions is consistent with their mandarin parent being similar to the Cleopatra type-2 mandarin which is heterozygous for a pummelo introgression on chromosome 3 at the same position as these accessions are homozygous for a pummelo segment (Figure 28).

Conclusion

The systematic diversity study using accessions in CVC was conducted with high density SNP data from Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. Our study integrated multiple complementary analyses that elucidate many complex relationships among different accessions in CVC. Our study included many accessions of *Citrus* and related genera that have never been analyzed with a high density marker system, allowing us to gain

information regarding their genetic profiles. Extensive inventory of the genetic profile of the accessions in the CVC can now provide us with the information necessary to identify the scale of diversity, and assist in efficient germplasm conservation and utilization. The diversity study conducted with CVC offered an important insight into the history of admixture and ancestry of citrus, and could be used to update the sometimes questionable passport records of many accessions added to the CVC during its more than 100 years of history.

References cites

1. Fao. FAO Statistical Yearbook: World Food and Agriculture. FAO; 2013.
2. Swingle, T W. The botany of Citrus and its wild relatives. The citrus industry. 1967;;190–430. <https://ci.nii.ac.jp/naid/10006100992/>. Accessed 14 May 2018.
3. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat Biotechnol.* 2014;32:656–62. doi:10.1038/nbt.2906.
4. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of Citrus. *Nature.* 2018;32:656. doi:10.1038/nature25447.
5. Ahmed D, Comte A, Curk F, Costantino G, Luro F, Dereeper A, et al. Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: a case study in diploid and polyploid citrus. *Ann Bot.* 2019;123:1231–51. doi:10.1093/aob/mcz029.
6. Frost HB. Nuclear embryony and juvenile characters in clonal varieties of citrus. *J Hered.* 1938;29:423–92. doi:10.1093/oxfordjournals.jhered.a104431.
7. Soost RL, Roose ML. Citrus. Fruit Breeding, Vol. 1: Tree and Tropical Fruits. In Janick J. and J.N.Moore (eds.), John Wiley&Sons, Inc. 1996;;257-323
8. Marhold K, Lihová J. Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. *Plant Syst Evol.* 2006;259:143–74. doi:10.1007/s00606-006-0417-x.
9. Arnold ML. Natural hybridization and evolution. Oxford University Press, USA; 1997. <https://play.google.com/store/books/details?id=RXU8DwAAQBAJ>.
10. Wissemann V. Evolution by hybridisation. The influence of reticulate evolution on biosymmetrical patterns and processes in plants. *Theory Biosci.* 2005;123:223–33. doi:10.1016/j.thbio.2004.09.003.
11. Arnold ML, Kentner EK, Johnston JA, Cornman S, Bouck AC. Natural hybridisation and fitness. *Taxon.* 2001;50:93–104. doi:10.2307/1224513.
12. Grant V. Plant Speciation. 1981. doi:10.7312/gran92318.
13. Doyle JJ, Doyle JL, Rauscher JT, Brown AHD. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine subgenus Glycine*). *New Phytol.* 2003;;121–32. doi:10.1046/j.1469-8137.2003.00949.x@10.1002/(ISSN)1469-8137(CAT)SpecialIssues(VI)Plantspeciation.
14. Cronn R, Wendel JF. Cryptic trysts, genomic mergers, and plant speciation. *New Phytologist.* 2003;161:133–42. doi:10.1111/j.1469-8137.2004.00947.x.
15. Linder CR, Rieseberg LH. Reconstructing patterns of reticulate evolution in plants. *Am J Bot.*

2004;91:1700–8. <https://www.ncbi.nlm.nih.gov/pubmed/18677414>.

16. Liu Y, Liu Y, Huang H. Genetic variation and natural hybridization among sympatric *Actinidia* species and the implications for introgression breeding of kiwifruit. *Tree Genet Genomes*. 2010;6:801–13. doi:10.1007/s11295-010-0293-y.

17. Lexer C, Fay MF, Joseph JA, Nica M-S, Heinze B. Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Mol Ecol*. 2005;14:1045–57. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2005.02469.x>.

18. Chat J, Jáuregui B, Petit RJ, Nadot S. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. *Am J Bot*. 2004;91:736–47. doi:10.3732/ajb.91.5.736.

19. Willyard A, Cronn R, Liston A. Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Mol Phylogenet Evol*. 2009;52:498–511. doi:10.1016/j.ympev.2009.02.011.

20. Martis MM, Zhou R, Haseneyer G, Schmutzer T, Vrána J, Kubaláková M, et al. Reticulate evolution of the rye genome. *Plant Cell*. 2013;25:3685–98. doi:10.1105/tpc.113.114553.

21. Goh WL, Chandran S, Franklin DC, Isagi Y, Koshy KC, Sungkaew S, et al. Multi-gene region phylogenetic analyses suggest reticulate evolution and a clade of Australian origin among paleotropical woody bamboos (Poaceae: Bambusoideae: Bambuseae). *Plant Syst Evol*. 2013;299:239–57. doi:10.1007/s00606-012-0718-1.

22. Gallardo MH. Phylogenetics, Reticulation and Evolution. *Phylogenetics*. 2017. doi:10.5772/intechopen.68564.

23. Stebbins GL. The role of hybridization in evolution. *Proc Am Philos Soc*. 1959;103:231–51. <http://www.jstor.org/stable/985151>.

24. Lewontin RC, Birch LC. Hybridization as a source of variation for adaptation to new environments. *Evolution*. 1966;20:315–36. doi:10.1111/j.1558-5646.1966.tb03369.x.

25. Arnold ML. Evolution through genetic exchange. OUP Oxford; 2006. <https://play.google.com/store/books/details?id=R5QUDAAAQBAJ>.

26. Minamikawa MF, Nonaka K, Kaminuma E, Kajiya-Kanegae H, Onogi A, Goto S, et al. Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. *Sci Rep*. 2017;7:4721. doi:10.1038/s41598-017-05100-x.

27. Shimizu T, Kitajima A, Nonaka K, Yoshioka T, Ohta S, Goto S, et al. Hybrid origins of citrus varieties inferred from DNA marker analysis of nuclear and organelle genomes. *PLoS One*. 2016;11:e0166969. doi:10.1371/journal.pone.0166969.

28. Tanaka T, Others. Species problem in *Citrus*. 1954. <http://agris.fao.org/agris-search/search.do?recordID=US201300690891>.

29. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al. The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet.* 2013;45:59–66. doi:10.1038/ng.2472.
30. Bayer RJ, Mabberley DJ, Morton C, Miller CH, Sharma IK, Pfeil BE, et al. A molecular phylogeny of the orange subfamily (Rutaceae: Aurantioideae) using nine cpDNA sequences. *American Journal of Botany.* 2009;96:668–85. doi:10.3732/ajb.0800341.
31. Mabberley DJ. The species of *Citrus* (Rutaceae) with pinnate leaves. *Blumea - Biodiversity, Evolution and Biogeography of Plants.* 2010;55:73–4. doi:10.3767/000651910X499222.
32. Zhang DX, Mabberley DJ. *Citrus*. *Flora of China.* 2008;11:90–6.
33. Mabberley DJ, Kodala PG. Lectotypification of *Citrus cavaleriei* H.Lév. ex Cavalerie (Rutaceae: Aurantioideae). *Telopea*. 2015;18:115–9. doi:10.7751/elopea8510.
34. Van de Wouw M, Kik C, van Hintum T, van Treuren R, Visser B. Genetic erosion in crops: concept, research results and challenges. *Plant Genet Resour.* 2010;8:1–15. <https://www.cambridge.org/core/journals/plant-genetic-resources/article/genetic-erosion-in-crops-concept-research-results-and-challenges/21260D7D924DEA9F27327C6C0B22E84C>.
35. Donini P, Law JR, Koeber RMD, Reeves JC, Cooke RJ. The impact of breeding on genetic diversity and erosion in bread wheat. *Plant Genet Resour.* 2005;3:391–9. doi:10.1079/PGR200593.
36. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155:945–59. <https://www.ncbi.nlm.nih.gov/pubmed/10835412>.
37. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;19:1655–64. doi:10.1101/gr.094052.109.
38. Baptiste E, van Iersel L, Janke A, Kelchner S, Kelk S, McInerney JO, et al. Networks: expanding evolutionary thinking. *Trends Genet.* 2013;29:439–41. doi:10.1016/j.tig.2013.05.007.
39. Gregory TR. Understanding evolutionary trees. *Evolution: Education and Outreach.* 2008;1:121. <https://www.biomedcentral.com/openurl?doi=10.1007/s12052-008-0035-x>.
40. Brinkman FS, Leipe DD. Phylogenetic analysis. *Methods Biochem Anal.* 2001;43:323–58. doi:10.1002/0471223921.ch14.
41. Baptiste E, Boucher Y, Leigh J, Doolittle WF. Phylogenetic reconstruction and lateral gene transfer. *Trends Microbiol.* 2004;12:406–11. doi:10.1016/j.tim.2004.07.002.
42. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* 2006;23:254–67. doi:10.1093/molbev/msj030.
43. Wen D, Yu Y, Hahn MW, Nakhleh L. Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. *Mol Ecol.* 2016;25:2361–72. doi:10.1111/mec.13544.
44. Nakhleh L. Evolutionary phylogenetic networks: models and issues. In: Heath LS,

- Ramakrishnan N, editors. Problem Solving Handbook in Computational Biology and Bioinformatics. Boston, MA: Springer US; 2011. p. 125–58. doi:10.1007/978-0-387-09760-2_7.
45. Karimi N, Grover CE, Gallagher JP, Wendel JF, Ané C, Baum DA. Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). *Syst Biol*. 2019. doi:10.1093/sysbio/syz073.
 46. Daco L, Maurice T, Muller S, Rossa J, Colling G. Genetic status of the endangered plant species *Gladiolus palustris* in the western part of its distribution area. *Conserv Genet*. 2019;20:1339–54. doi:10.1007/s10592-019-01213-0.
 47. McBreen K, Lockhart PJ. Reconstructing reticulate evolutionary histories of plants. *Trends Plant Sci*. 2006;11:398–404. doi:10.1016/j.tplants.2006.06.004.
 48. Huson DH, Rupp R, Scornavacca C. Phylogenetic networks: concepts, algorithms and applications. Cambridge University Press; 2010. <https://play.google.com/store/books/details?id=0rB5I5GxveAC>.
 49. Martin W, Brinkmann H, Savonna C, Cerff R. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc Natl Acad Sci U S A*. 1993;90:8692–6. doi:10.1073/pnas.90.18.8692.
 50. Sato M, Miyazaki K. Phylogenetic network analysis revealed the occurrence of horizontal gene transfer of 16S rRNA in the genus *Enterobacter*. *Front Microbiol*. 2017;8:2225. doi:10.3389/fmicb.2017.02225.
 51. Kunin V, Goldovsky L, Darzentas N, Ouzounis CA. The net of life: reconstructing the microbial phylogenetic network. *Genome Res*. 2005;15:954–9. doi:10.1101/gr.3666505.
 52. Roberts WR, Roalson EH. Phylogenomic analyses reveal extensive gene flow within the magic flowers (*Achimenes*). *Am J Bot*. 2018;105:726–40. doi:10.1002/ajb2.1058.
 53. Zhang Q-P, Hu W-F, Zhou T-T, Kong S-S, Liu Z-F, Zheng R-Q. Interspecies introgressive hybridization in spiny frogs *Quasipaa* (Family Dicroglossidae) revealed by analyses on multiple mitochondrial and nuclear genes. *Ecol Evol*. 2018;8:1260–70. doi:10.1002/ece3.3728.
 54. Barkley NA, Roose ML, Krueger RR, Federici CT. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor Appl Genet*. 2006;112:1519–31. doi:10.1007/s00122-006-0255-9.
 55. Herrero R, Asíns MJ, Carbonell EA, Navarro L. Genetic diversity in the orange subfamily Aurantioideae. I. Intraspecific and intragenus genetic variability. *Theor Appl Genet*. 1996;92:599–609. doi:10.1007/BF00224564.
 56. Luro F, Gatto J, Costantino G, Pailly O. Analysis of genetic diversity in *Citrus*. *Plant Genet Resour*. 2011;9:218–21. doi:10.1017/S1479262111000189.
 57. García-Lor A, Luro F, Navarro L, Ollitrault P. Comparative use of InDel and SSR markers in deciphering the interspecific structure of cultivated citrus genetic diversity: a perspective for

genetic association studies. *Mol Genet Genomics*. 2012;287:77–94. doi:10.1007/s00438-011-0658-4.

58. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*. 1965;19:395–420. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1965.tb01731.x>.

59. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–75. doi:10.1086/519795.

60. Nei M. F-statistics and analysis of gene diversity in subdivided populations. *Ann Hum Genet*. 1977;41:225–33. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1977.tb01918.x>.

61. Holland BR, Jermiin LS. Improved consensus network techniques for genome-scale phylogeny. *Mol Biol*. 2005. <https://academic.oup.com/mbe/article-abstract/23/5/848/1058339>.

62. Holland BR, Huber KT, Dress A, Moulton V. δ Plots: A Tool for Analyzing Phylogenetic Distance Data. *Mol Biol Evol*. 2002;19:2051–9. doi:10.1093/oxfordjournals.molbev.a004030.

63. Gray RD, Bryant D, Greenhill SJ. On the shape and fabric of human history. *Philos Trans R Soc Lond B Biol Sci*. 2010;365:3923–33. doi:10.1098/rstb.2010.0162.

64. Barrett HC, Rhodes AM. A numerical taxonomic study of affinity relationships in cultivated citrus and its close relatives. *Syst Bot*. 1976;1:105–36. doi:10.2307/2418763.

65. Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet J-P, Garcia-Lor A, et al. Nuclear species-diagnostic SNP markers mined from 454 amplicon sequencing reveal admixture genomic structure of modern citrus varieties. *PLoS One*. 2015;10:e0125628. doi:10.1371/journal.pone.0125628.

66. Ollitrault P, Terol J, Garcia-Lor A, Bérard A, Chauveau A, Froelicher Y, et al. SNP mining in *C. clementina* BAC end sequences; transferability in the Citrus genus (Rutaceae), phylogenetic inferences and perspectives for genetic mapping. *BMC Genomics*. 2012;13:13. doi:10.1186/1471-2164-13-13.

67. Froelicher Y, Mouhaya W, Bassene J-B, Costantino G, Kamiri M, Luro F, et al. New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. *Tree Genet Genomes*. 2011;7:49–61. doi:10.1007/s11295-010-0314-x.

68. Tanaka T. Citrus fruits of India. Allahabad Farmer. 1937;2:53–7.

69. Reuther W, Webber HJ, Batchelor LD. The citrus industry volume I. University of California. 1967.

70. Asadi Abkenar A, Matsumoto R, Yamamoto M. Relatedness among mandarins (*Citrus* spp.) and their close relatives revealed by cpDNA polymorphism. Laboratory of Fruit Science, Faculty of Agriculture, Saga University. 2019;104:1–11. http://portal.dl.saga-u.ac.jp/bitstream/123456789/123966/1/asadiabkenar_201901.pdf.

71. Tanaka T. Citrologia. Semi-centennial commemoration papers on citrus studies Citrologia Supporting Foundation, Osaka. 1961.
72. Swingle, T W. The botany of Citrus and its wild relatives in the orange subfamily. The citrus industry. 1943;1:128–474. <https://ci.nii.ac.jp/naid/10018851984/>. Accessed 13 Apr 2020.
73. Yamamoto M, Fukuda M, Koga T, Kubo T, Tominaga S, Others. Examination of the origin of Keraji (*Citrus keraji*), local citrus of Kikaijima island, Kagoshima prefecture. Horticultural Research (Japan). 2010;9:7–12. <https://www.cabdirect.org/cabdirect/abstract/20103064947>.
74. Tanaka, T. Fundamental discussion of Citrus classification. Studia Citrologica. 1977;14:1–6. <https://ci.nii.ac.jp/naid/10006285025/>. Accessed 18 Apr 2020.
75. Lawson DJ, van Dorp L, Falush D. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nature Communications. 2018;9. doi:10.1038/s41467-018-05257-7.
76. Ramadugu C, Keremane ML, Hu X, Karp D, Federici CT, Kahn T, et al. Genetic analysis of citron (*Citrus medica* L.) using simple sequence repeats and single nucleotide polymorphisms. Sci Hortic . 2015;195:124–37. doi:10.1016/j.scienta.2015.09.004.
77. Fao U, Others. FAOstat. Retrieved Feb. 2014;2014.
78. Scora RW. On the history and origin of Citrus. Bull Torrey Bot Club. 1975;;369–75. <https://www.jstor.org/stable/2484763>.
79. Barrett HC, Rhodes AM. A numerical taxonomic study of affinity relationships in cultivated citrus and its close relatives. Systematic Botany. 1976;1:105. doi:10.2307/2418763.
80. Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P. Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. Ann Bot. 2016;117:565–83. doi:10.1093/aob/mcw005.
81. Gulsen O, Roose ML. The origin of Interdonato lemon inferred from cpRFLP, SSR, isozyme and ISSR markers. In: Proc. Int. Soc. Citricult. IX Congr. 2000. p. 158–9.
82. Rouiss H, Bakry F, Froelicher Y, Navarro L. Origin of *C. latifolia* and *C. aurantiifolia* triploid limes: the preferential disomic inheritance of doubled-diploid “Mexican” lime is consistent with an interploid hybridization hypothesis. Ann Bot. 2018;121:571–585 <https://academic.oup.com/aob/article-abstract/121/3/571/4775246>.
83. Butelli E, Garcia-Lor A, Licciardello C, Las Casas G, Hill L, Recupero GR, et al. Changes in anthocyanin production during domestication of *Citrus*. Plant Physiol. 2017;173:2225–42. doi:10.1104/pp.16.01701.
84. Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. A Phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. Mol Biol Evol. 2015;32:2015–35. doi:10.1093/molbev/msv082.
85. Yang X, Li H, Yu H, Chai L, Xu Q, Deng X. Molecular phylogeography and population

evolution analysis of *Citrus ichangensis* (Rutaceae). *Tree Genet Genomes*. 2017;13:29. doi:10.1007/s11295-017-1113-4.

86. Palma A, D'Aquino S. Kumquat—*Fortunella japonica*. In: Rodrigues S, de Oliveira Silva E, de Brito ES, editors. *Exotic Fruits*. Academic Press; 2018. p. 271–8. doi:10.1016/B978-0-12-803138-4.00035-6.

87. Hearn CJ. Yield and fruit quality of“ Ambersweet” orange hybrid on different rootstocks. In: *Proc. Fla, State Hort. Soc.* 1989. p. 75–8. <https://pdfs.semanticscholar.org/1584/3626fa483c5e47af63bcafb336a491854488.pdf>.

88. Putzbach K, Rimmer CA, Sharpless KE, Sander LC. Determination of bitter orange alkaloids in dietary supplements standard reference materials by liquid chromatography with ultraviolet absorbance and fluorescence detection. *J Chromatogr A*. 2007;1156:304–11. doi:10.1016/j.chroma.2006.12.027.

89. Arctander S. *Perfume and Flavor Materials of Natural Origin*. Lulu.com; 2017. <https://play.google.com/store/books/details?id=8Os5DwAAQBAJ>.

90. Draelos ZD. *Cosmetic Dermatology: Products and Procedures*. John Wiley & Sons; 2015. <https://play.google.com/store/books/details?id=D2a-CgAAQBAJ>.

Table. 1 Accessions genotyped with Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array
924 accessions used in the study that are genotyped at 56,367 PHR and NMH loci for the study

Sample Name (CVC) and CRC	Type
Orlando_bittersweet_orange_1588	Bittersweet_Orange
Paraguay_bittersweet_orange_660	Bittersweet_Orange
Dummett_bittersweet_orange_2541	Bittersweet_Orange
Amoa_8_Blood_mandarin_4237	Blood_Mand
Bream_Tarocco_Blood_orange_4032	Blood_Orange
Burris_Blood_Valencia_orange_2561	Blood_Orange
Delfino_blood_orange_4114	Blood_Orange
Dobla_fina_blood_orange_3152	Blood_Orange
Entre_fina_bood_orange_3650	Blood_Orange
Malta_blood_orange_4085	Blood_Orange
Ruby_4N_RRUT263	Blood_Orange
Ruby_blood_orange_3827	Blood_Orange
Grosse_Ronde_blood_orange_3401	Blood_Orange
Sanguinelli_blood_orange_3828	Blood_Orange
Sanguinello_a_Pignu_blood_orange_3829	Blood_Orange
Smith_Red_blood_orange_4093	Blood_Orange
Tarocco_#7_blood_orange_3596	Blood_Orange
Thermal_Tarocco_blood_orange_4018	Blood_Orange
Vaccaro_blood_orange_3242	Blood_Orange
Vaccaro_blood_orange_3627	Blood_Orange
Washington_Sanguine_blood_orange_3977	Blood_Orange
Doblafina_blood_orange_4000	Blood_Orange
Ruby_blood_orange_3827	Blood_Orange
Moro_blood_orange_3830	Blood_Orange
Calamondin_Sport_3656	Calamondin
Variegated_calamondin_3087	Calamondin
Calamondin_2592	Calamondin
Carpenter_C54_4220	Citrandarin
US_852_4245	Citrandarin
Benton_citrangle_3908	Citrangle
Bitters_C22_4219	Citrangle
C32_citrangle_3911	Citrangle
C35_citrangle_3912	Citrangle
Citrangle_#1416_citrangle_2866	Citrangle
Cunningham_citrangle_271	Citrangle
Furr_C57_citrangle_4221	Citrangle

Morton_citrange_1463	Citrange
Morton_citrange_2748	Citrange
Rusk_citrange_1441	Citrange
Rusk_citrange_301	Citrange
Sanford_citrange_276	Citrange
Savage_citrange_275	Citrange
Spanish_sweet_orange_X_Ptrifoliate_3336	Citrange
Troyer_citrange_1459	Citrange
Uvalde_citrange_2865	Citrange
Yuma_citrange_3205	Citrange
Carrizo_citrange_2863	Citrange
Assads_citron_RSD2008001	Citron
Badhri_citron_hybrid_3723	Citron
Bengal_citron_3055	Citron
Buddha's_Hand_citron_3768	Citron
Citron_of_Commerce_3518	Citron
Corsican_citron_1795	Citron
Dabbe_citron_hybrid_3190	Citron
Djerok_citron_hybrid_2456	Citron
Dulcia_citron_3654	Citron
Etrog_citron_3891	Citron
GouTouXiangYuan_RRUT395	Citron
GouTouXiangYuan_RRUT396	Citron
GouTouXiangYuan_RRUT398	Citron
Hiawassie_citron_3527	Citron
Indian_citron_138	Citron
Indian_citron_661	Citron
Italian_citron_3530	Citron
Limau_Mata_Susu_citron_hybrid_3755	Citron
Mexican_citron_3531	Citron
Morningsong_Temoni_citron-Karp	Citron
Papuan_citron_3532	Citron
Philippine_citron_3533	Citron
S-1_citron_3878	Citron
Sicily_citron_3534	Citron
Yemen_citron_2847	Citron
C_Halimi_citron_hybrid_3900	Citron_Hyb
Odorata_citron_3655	Citron
Spadifora_citron_3535	Citron
Diamante_citron_3523	Citron

Yunnanese_citron_3798	Citron
Unnamed_citron_hybrid_3174	Citron
South_Coast_Field_Station_citron_3546	Citron
African_Shaddock_X_Rubidoux_trifoliolate_3969	Citrumelo
Citrumelo_1452	Citrumelo
Citrumelo_3348	Citrumelo
Citrumelo_3341	Citrumelo
Hall_grapefruit_X_Rubiboux_trifoliolate_3889	Citrumelo
Sacaton_citrumelo_3414	Citrumelo
Sacaton_citrumelo_3417	Citrumelo
Swingle_citrumelo_3767	Citrumelo
Swingle_citrumelo_3771	Citrumelo
Sacaton_citrumelo_3337	Citrumelo
Ain-Taoujdate_clementine_4212	Clem
Algerian_clementine_279	Clem
Arrufatina_clementine_4184	Clem
Clementina_de_Nules_4056	Clem
Clementina_Fina_Sodea_3999	Clem
Clementina_Marisol_4055	Clem
Clementina_3_4N_RRUT270	Clem
Caffin_clementine_3987	Clem
Clementina_Fina_4011	Clem
Clementine_X_Pearl_mandarin_hybrid_4020	Clem
Clementine_X_Silverhill_mandarin_hybrid_3731	Clem
Hernandina_clementine_4185	Clem
SRA63_clementine_4076	Clem
SRA92_clementine_4052	Clem
Monreal_clementine_3773	Clem
Clementine_Sidi_Aissa_3988	Clem
Clementina_Oroval_4010	Clem
Allen_variegated_Eureka_lemon_3007	Eureka_Lemon
Amber_Eureka_lemon_2429	Eureka_Lemon
Blanchard_Eureka_lemon_3499	Eureka_Lemon
Cascade_Eureka_lemon_3498	Eureka_Lemon
Cascade_Eureka_lemon_2703	Eureka_Lemon
Cook_Eureka_lemon_3837	Eureka_Lemon
Corona_Foothill_Eureka_lemon_3043	Eureka_Lemon
Frost_Eureka_lemon_3005	Eureka_Lemon
Old_line_Eureka_lemon_4094	Eureka_Lemon
Ross_Eureka_lemon_3838	Eureka_Lemon

Taylor_Eureka_lemon_4014	Eureka_Lemon
Variegated_Pink_Eureka_lemon_2367	Eureka_Lemon
Variegated_Eureka_lemon_599	Eureka_Lemon
Camulos_grapefruit_3139	Gft
Cecily_grapefruit_2014	Gft
Cecily_grapefruit_3637	Gft
CRC343_343	Gft
Duncan_grapefruit_3832	Gft
Flame_grapefruit_3993	Gft
Foster_Pink_grapefruit_1198	Gft
Frost_Marsh_grapefruit_3184	Gft
Genetic_Dwart_grapefruit_3774	Gft
Hall_grapefruit_3068	Gft
Henderson_Ruby_grapefruit_3914	Gft
Citrus_hiroshimana_Natsuzabon_3904	Gft
Hudson_foster_grapefruit_3638	Gft
Hudson4N_RRUT385	Gft
Imperial_grapefruit_596	Gft
Jochimsen_grapefruit_2784	Gft
New_Zealand_grapefruit_2381	Gft
Nicholson_navel_grapefruit_3398	Gft
Perlis_#1_grapefruit_3886	Gft
Perlis_#2_grapefruit_3883	Gft
Ray_Ruby_grapefruit_3915	Gft
Redblush_grapefruit_2850	Gft
Reed_Marsh_grapefruit_3128	Gft
Rio_Red_grapefruit_3916	Gft
Shambar_grapefruit_3831	Gft
Star_Ruby_grapefruit_3770	Gft
Sukega_grapefruit_hybrid_4149	Gft
Tetraploid_grapefruit_3968	Gft
Warren_Marsh_grapefruit_4150	Gft
Whitney_Marsh_grapefruit_3804	Gft
Marsh_pink_grapefruit_2010	Gft
Marsh_seedy_grapefruit_1565	Gft
Triumph_grapefruit_297	Gft
Royal_grapefruit_248	Gft
Citrus_otachibana_grapefruit_hybrid_3470	Gft_Hyb
Citrus_rugulosa_pummelo_hybrid_3556	Gft_Hyb
Chironja_grapefruit_hybrid_4045	Gft_Hyb

Chironja_grapefruit_hybrid_3909	Gft_Hyb
New_zealand_grapefruit_hybrid_3769	Gft_Hyb
Wheeny_grapefruit_hybrid_2885	Gft_Hyb
Rex_Union_grapefruit_hybrid_3962	Gft_Hyb
Valentine_pummelo_hybrid_4182	Gft_Hyb
Rojo_Blanco_pummelo_hybrid_4029	Gft_Pum_Hyb
Centennial_variegated_kumquat_4082	Kumquat
Fortunella_crassifolia_3818	Kumquat
Fortunella_polyandra_3901	Kumquat
Marumi_SRA489_kumquat_4125	Kumquat
Meiwa_kumquat_3833	Kumquat
Nordmann_Seedless_kumquat_4070	Kumquat
Fortunella_obovata_3475	Kumquat
Citrangquat_19_15_7_3259	Kumquat_Hyb
Eustis_limequat_3603	Kumquat_Hyb
Indio_mandarinquat_3759	Kumquat_Hyb
Nippon_orangequat_3360	Kumquat_Hyb
Procimequat_kumquat_hybrid_3295	Kumquat_Hyb
Rio_Grande_Valley_lemonquat_4097	Kumquat_Hyb
Sinton_citrangquat_3642	Kumquat_Hyb
Tavares_kumquat_hybrid_4048	Kumquat_Hyb
Thomasville_citrangquat_1440	Kumquat_Hyb
Unnamed_kumquat_hybrid_1044	Kumquat_Hyb
Warren_limequat_3069	Kumquat_Hyb
Allen_Newman_Eureka_lemon_3496	Lemon
Arancino_lemon_3387	Lemon
Atmore_lemon_4092	Lemon
Bergamotto_#18_Lisbon_lemon_3506	Lemon
Berna_lemon_3590	Lemon
Bitrouni_lemon_3265	Lemon
Citrus_limon_AC29_lemon_3920	Lemon
Lamas_lemon_3919	Lemon
Carpaci_lemon_3591	Lemon
Chinese_lemon_710	Lemon
Femminello_Siracusano_white_flower_lemon_4226	Lemon
femminello_Ovale_lemon_3388	Lemon
Femminello_Sfusato_lemon_3389	Lemon
Femminello_Siracusano_2Kr_lemon_4166	Lemon
Genoa_lemon_565	Lemon
India_lemon_2322	Lemon

India_lemon_2323	Lemon
Indian_rough_lemon_2544	Lemon
Interdonato_lemon_hybrid_3593	Lemon
Iran_lemon_3162	Lemon
Italian_Pink_fleshed_lemon_2899	Lemon
Khoubs_el_arsa_lemon_2477	Lemon
Khoubs_el_ersa_lemon_2489	Lemon
Kulu_lemon_3045	Lemon
Kusner_lemon_3194	Lemon
Lapithiokii_lemon_4005	Lemon
Limonero_Fino_95_lemon_4126	Lemon
Limonero_Fino_lemon_3970	Lemon
Limonero_Messina_lemon_4121	Lemon
Lo_Porto_lemon_3090	Lemon
Lumia_lemon_hybrid_3925	Lemon
Lunario_lemon_3159	Lemon
Mesero_lemon_3892	Lemon
Messina_lemon_3009	Lemon
New_Zealand_lemonade_lemon_4181	Lemon
Old_Indian_lemon_4211	Lemon
Orban_lemon_hybrid_4140	Lemon
Primofiore_lemon_3491	Lemon
Ricote_lemon_3893	Lemon
Santa_Teresa_lemon_3894	Lemon
Seedless_lemon_3001	Lemon
Soh_Synteng_lemon_hybrid_3216	Lemon
Unnnamed_lemon_hybrid_3155	Lemon
Villafranca_lemon_280	Lemon
Villafranca_lemon_390	Lemon
Yen_Ben_Lisbon_lemon_4067	Lemon
Estes_rough_lemon_3386	Lemon
Citrus_excelsa_2316	Lemon
Wild_lemon_hybrid_3300	Lemon
Citrus_excelsa_3765	Lemon
Monachello_lemon_3392	Lemon
Citrus_kharna_3059	Lemon
Improved_Meyer_lemon_3737	Lemon
Unnamed_lemon_3885	Lemon
Citrus_volkameriana_3050	Lemon_Hyb
Gul-gul_lemon_hybrid_4235	Lemon_Hyb

Lemonime_lemon_hybrid_2459	Lemon_Hyb
Limonette_De_Marrakech_lemon_hybrid_3989	Lemon_Hyb
Nicaraguan_lemon_3841	Lemon_Hyb
Paak_Ling_Mung_lemon_hybrid_3193	Lemon_Hyb
Peretta_lemon_hybrid_3924	Lemon_Hyb
Perrine_lime_hybrid_2458	Lemon_Hyb
Ponderosa_lemon_hybrid_294	Lemon_Hyb
Soh_Long_lemon_hybrid_3199	Lemon_Hyb
Unnamed_lemon_hybrid_3172	Lemon_Hyb
Unnamed_lemon_hybrid_3154	Lemon_Hyb
Unnamed_lemon_hybrid_3748	Lemon_Hyb
Iraq_lemon_limetta_3492	Lemon_Hyb
Citrus_hanayu_3469	Lemon_Hyb
Citrus_succosa_3280	Lemon_Hyb
Abhayapuri_lime_3762	Lime
Castelo_lime_4062	Lime
Chulo_Key_lime_4120	Lime
Davao_papeda_hybrid_2427	Lime
Egyptian_lime_2883	Lime
Everglade_lime_2449	Lime
Excalibur_red_lime_hybrid_4239	Lime
Excalibur_lime_RRUT332	Lime
Giant_Key_lime_4069	Lime
India_lime_3070	Lime
Limon_real_2317	Lime
Soghi_lime_4130	Lime
Stow_#7_Key_lime_2188	Lime
Unnamed_Philippine_lime_2892	Lime
Variegated_Tahiti_lime_4246	Lime
West_Indian_lime_1813	Lime
Bakrai_X_limetta_3776	Lime
Mexican_lime_1710	Lime
Millsweet_limetta_569	Lime
India_lime_2450	Lime
Citrus_Logispina_2320	Lime
Thornless_Mexican_lime_2683	Lime
Addanimma_lime_hybrid_3072	Lime_Hyb
Limoui_Sangui_Rangpur_lime_3200	Lime_Hyb
Rangpur_X_Troyer_hybrid_3997	Lime_Hyb
Dr_Strong_Lisbon_lemon_3504	Lisbon_Lemon

Femminello_lemon_3500	Lisbon_Lemon
Foothill_Lisbon_lemon_3836	Lisbon_Lemon
Frost_Lisbon_lemon_3176	Lisbon_Lemon
Galligan_Lisbon_lemon_3835	Lisbon_Lemon
Kaweah_#1_Lisbon_lemon_3010	Lisbon_Lemon
Limoneira_8A_IR1	Lisbon_Lemon
Limoneira_8A_Lisbon_lemon_3501	Lisbon_Lemon
Lupe_Lisbon_lemon_3013	Lisbon_Lemon
Monroe_Lisbon_lemon_3839	Lisbon_Lemon
Prior_Lisbon_lemon_3505	Lisbon_Lemon
Rosenberger_Lisbon_lemon_3840	Lisbon_Lemon
Avana_Apireno_mandarin_4163	Mand
Baiju_nucellar_RRUT76	Mand
Batangas_mandarin_3659	Mand
Belady_mandarin_3363	Mand
Bower_mandarin_3649	Mand
Citrus_amblycarpa_2485	Mand
Citrus_kinokuni_3887	Mand
Citrus_leiocarpa_3147	Mand
Citrus_lycopersicaformis_3564	Mand
Citrus_nippokoreana_3228	Mand
Citrus_reticulata_3813	Mand
Citrus_tachibana_3150	Mand
Citrus_tardiva_3297	Mand
Citrus_yatsushiro_3880	Mand
Canton_mandarin_3576	Mand
Changsha_mandarin_3577	Mand
Somkeowan_mandarin_3752	Mand
Cleopatra_mandarin_3844	Mand
Corsica_#1_clemetine_4046	Mand
Citrus_sunki_3143	Mand
Daisy_mandarin_3910	Mand
Daisy_SL_mandarin_4203	Mand
Dancy_Weshart_602	Mand
Dancy_Frost_mandarin_3026	Mand
Empress_mandarin_3613	Mand
Encore_mandarin_3569	Mand
Fairchild_mandarin_3559	Mand
Fairchild_LS_mandarin_4205	Mand
Fallglo_mandarin_3990	Mand

Fortune_mandarin_3560	Mand
Fremont_mandarin_3558	Mand
Frua_mandarin_3022	Mand
Gold_Nugget_mandarin_3913	Mand
Hansen_mandarin_4152	Mand
Hickson_mandarin_4161	Mand
Hill_mandarin_3847	Mand
Honey_mandarin_3177	Mand
Huang_Yen_Man_Chie_3897	Mand
Hung_Kat_mandarin_2331	Mand
Imperial_mandarin_4160	Mand
Iyo_San_Ponkan_4118	Mand
Kara_mandarin_3019	Mand
King_mandarin_303	Mand
King_mandarin_3845	Mand
Kinnow_LS_mandarin_4236	Mand
Kinokuni_mandarin_696	Mand
Kobayashi_Mikan_mandarin_3817	Mand
Kunembo_mandarin_3346	Mand
Laranja_Cravo_mandarin_2893	Mand
Lee_mandarin_3851	Mand
Mandarine_sanguine_mandarin_3367	Mand
Michal_mandarin_4049	Mand
Mihrio_mandarin_4248	Mand
Murcott_mandarin_3846	Mand
Nova_IR10_mandarin_4206	Mand
Palazzelli_mandarin_4244	Mand
Parson's_Special_mandarin_300	Mand
Pixie_mandarin_3568	Mand
Richard's_Special_mandarin_4117	Mand
Rubidoux_mandarin_4031	Mand
Scarlet_Emperor_mandarin_3226	Mand
Seedless_kishu_mandarin_3906	Mand
Shiranui_VI860	Mand
Shiranui_Dekopan_4113	Mand
Soh_Himtra_mandarin_4095	Mand
Soh_Niamtra_mandarin_3260	Mand
Solid_Scarlet_mandarin_3328	Mand
Som_Kaeo_mandarin_3852	Mand
Som_Saa_mandarin_3853	Mand

SRA_337_RRUT223	Mand
Sunburst_mandarin_3809	Mand
Sun_Chu_Sha_mandarin_4003	Mand
Szinkom_mandarin_3085	Mand
Tango_mandarin_4183	Mand
Tien_Chieh_mandarin_2376	Mand
Tien_Chieh_mandarin_2590	Mand
Tieu_nucellar_mandarin_RRUT87	Mand
Tim_Kat_mandarin_2692	Mand
Unnamed_Chinese_RRUT223	Mand
US_Superna_USDA_88_2_mandarin_hybrid_3991	Mand
USDA_88_3_mandarin_hybrid_3992	Mand
Citrus_indica_3163	Mand
Fuju_nucellar_RRUT82	Mand
Emerald_mandarin_4123	Mand
W_Murcott_Afourer_mandarin_3953	Mand
Necked_mandarin_3738	Mand
Kincy_mandarin_4144	Mand
Kinnow_mandarin_3021	Mand
Wilking_mandarin_3020	Mand
Nova_mandarin_3615	Mand
US_Furr_SRA337_mandarin_4238	Mand
USDA_6-15-150_mandarin_4165	Mand
Robinson_mandarin_3850	Mand
Page_mandarin_3616	Mand
Yellow_Rind_mandarin_3895	Mand
Nagpur_orange_3727	Mand
Mandarinette_3405	Mand
Willowleaf_mandarin_3843	Mand
Ponkan_mandarin_3849	Mand
Richard's_Special_mandarin_3329	Mand
Citrus_keraji_3144	Mand
Citrus_reticulata_3239	Mand
Citrus_reticulata_3812	Mand
Citrus_depressa_2448	Mand_depressa
Citrus_tamura_3092	Mand_Hyb
Calushu_calamondin_2867	Mand_Hyb
Kinkoji_Unshiu_mandarin_graft_hybrid_3816	Mand_Hyb
NISSV_E_mandarin_hybrid_3965	Mand_Hyb
Novelty_X_Ellendale_mandarin_hybrid_3956	Mand_Hyb

Australian_red_Rangpur_2319	Mand_Hyb_Rangpur
Bishop_red_Rangpur_2451	Mand_Hyb_Rangpur
Kusaie_Rangpur_452	Mand_Hyb_Rangpur
Othaheite_acidless_Rangpur_2709	Mand_Hyb_Rangpur
Phillipine_red_Rangpur_2318	Mand_Hyb_Rangpur
Citrus_erythrosa_3292	Mand-erythrosa
Citrus_megaloxycarpa_hybrid_3241	Megaloxycarpa
Microcitrus_australasica_3672	Microcitrus
Microcitrus_australis_3665	Microcitrus
Microcitrus_australasica_3670	Microcitrus
Sydney_hybrid_1485	Microcitrus_Hyb
Apopka_navel_4151	Navel
Atwood_navel_3791	Navel
Australian_nucellar_navel_4175	Navel
Autumn_Gold_navel_4037	Navel
Bahianinha_SRA513_navel_4167	Navel
Barnfield_navel_4038	Navel
Bey_navel_3406	Navel
Bitters_Newhall_navel_4228	Navel
California_Rojo_navel_4201	Navel
Cara_cara_navel_3994	Navel
Carter_navel_2008	Navel
Ceridwen_navel_4065	Navel
Chislett_navel_4039	Navel
Cluster_Sport_navel_3732	Navel
Cogan_navel_4050	Navel
Corrugated_Thomson_navel_1381	Navel
Cpb44944A_Sdlg_navel_3305	Navel
Cpb44944B_Sdlg_navel_3306A	Navel
Cukurova_Early_navel_4086	Navel
Dream_navel_3307	Navel
Dry_navel_3354	Navel
Everhard_navel_4200	Navel
Fisher_navel_3135	Navel
Fisher_navel_3645	Navel
Frost_Washington_navel_3033	Navel
Fukumoto_navel_3903	Navel
Fulwood_navel_4112	Navel
Gillemberg_navel_4192	Navel
Gillette_navel_3017	Navel

Golden_Buckeye_avel_588	Navel
Golden_Nugget_avel_1379	Navel
Johnson_avel_4059	Navel
Leng_avel_3808	Navel
McFadden_Ribbed_avel_609	Navel
Navelate_avel_4022	Navel
Navelina_avel_4015	Navel
Newhall_avel_3014	Navel
Olympic_Gold_avel_4209	Navel
Palmer_avel_3964	Navel
Palmer_avel_4013	Navel
Paradise_avel_2853	Navel
Parent_Washington_avel_1241A	Navel
Pollock_T4_avel_4193	Navel
Powell_avel_4040	Navel
Ricalate_avel_4058	Navel
Rio_Grande_avel_3315	Navel
Robertson_avel_4023	Navel
Robyn_avel_4063	Navel
Rohde_avel_4084	Navel
Secret_Garden_avel_4194	Navel
Seedy_avel_65	Navel
Shahani_blood_avel_RRUT477	Navel
Skaggs_Bonanza_avel_3971	Navel
Spring_avel_4024	Navel
Summer_Gold_avel_4041	Navel
Thomson_avel_969	Navel
Thomson_avel_983	Navel
UCR_Early_avel_4087	Navel
Variegated_avel_1203	Navel
Whitley_avel_4195	Navel
Wiffen_avel_4083	Navel
Workman_avel_3320	Navel
Lane_Late_avel_3803	Navel
Tibbets_Parent_Washington_avel_4051	Navel
Smith's_Early_avel_574	Navel
Washington_seedy_avel_956	Navel
Parent_Washington_avel_1241B	Navel
Citrus_latipes_Khasi_3052	Papeda
Citrus_hongheesis_3797	Papeda

Citrus_hystrix_3103	Papeda
Citrus_hystrix_cabuyo_RRUT361	Papeda
Citrus_hystrix_papeda_2454	Papeda
Citrus_ichangensis_2327	Papeda
Cabuyo_papeda_432	Papeda
Ota_Yuzu_RRUT_228	Papeda
Soh_Niangrang_papeda_3203	Papeda
Unnamed_papeda_3056	Papeda
Citrus_ichangensis_2431	Papeda
Citrus_ichangensis_3931	Papeda
Citrus_micrantha_3605	Papeda
Citrus_webberii_Kalpi_1456	Papeda
Yuzu_ichandarin_1216	Papeda
Kabosu_papeda_hybrid_3943	Papeda
Citrus_yuko_3146	Papeda_Hyb
Citrus_macrophylla_papeda_hybrid_2446	Papeda_Hyb
Ichang_lemon_1219	Papeda_Hyb
Kulobot_papeda_hybrid_3612	Papeda_Hyb
Citrus_webberii_var_Montana_767	Papeda_Hyb
Ichang_lemon_1215	Papeda_Hyb
Citrus_sudachi_3471	papeda_Hyb
African_pummelo_2346	Pum
Alemoen_pummelo_2487	Pum
Banokan_pummelo_3950	Pum
Banpeiyu_pummelo_3961	Pum
BuOl_1_acidless_yellow_RRUT92	Pum
Citrus_grandis_3282	Pum
Citrus_grandis_2340	Pum
Citrus_grandis_2453	Pum
Citrus_hiroshimana_Natsuzabon_3275	Pum
Chandler_pummelo_3244	Pum
Chinese_pummelo_1224	Pum
Fleming's_Shaddock_pummelo_578	Pum
Haiku_B_pummelo_3940	Pum
Hunan_pummelo_1225	Pum
Itoshima_3928	Pum
Kao_Pan_pummelo_3948	Pum
Kao_Pan_pummelo_2242	Pum
Kao_Pan_pummelo_2243	Pum
Kao_Pan_pummelo_2249	Pum

Kao_Pan_pummelo_2349	Pum
Kao_Pan_pummelo_2356	Pum
Kao_Phuang_pummelo_3926	Pum
Kao_Ruan_Tia_pummelo_2350	Pum
Kao_Ruan_Tia_pummelo_2351	Pum
Karn_Lau_Yau_pummelo_2341	Pum
Kawachi_Bankan_3944	Pum
Kinokawa_Buntan_pummelo_3978	Pum
Moli_kurikuri_pummelo_hybrid_579	Pum
Nakhon_Chaisi_pummelo_2353	Pum
Pan_Dan_pummelo_2752	Pum
Pauthel_pummelo_3949	Pum
Pink_pummelo_2246	Pum
Pong_yau_pummelo_2342	Pum
Red_pummelo_2245	Pum
Red_pummelo_2338	Pum
Reinking_pummelo_3805	Pum
Roeding's_Pink_pummelo_1208	Pum
Rubidoux_pummelo_hybrid_#1_4028	Pum
Sarawak_pummelo_4042	Pum
Siamese_Acidless_pummelo_2240	Pum
Sour_pummelo_3066	Pum
Suisho_Buntan_pummelo_3947	Pum
Sunshine_pummelo_2236	Pum
Sweet_pummelo_3067	Pum
Tahitian_pummelo_3806	Pum
Tau_Yau_pummelo_2583	Pum
Thong_Deep_pummelo_4027	Pum
Thong_Deep_pummelo_3927	Pum
Webber_shaddock_pummelo_4233	Pum
Philippine_pummelo_hybrid_644	Pum
Kao_Phuang_pummelo_2352	Pum
Siamese_pummelo_2421	Pum
Kao_Pan_pummelo_2355	Pum
Egami_Buntan_pummelo_3959	Pum
Pin_Shan_Kong_Yan_pummelo_2348	Pum
Citrus_pennivesiculata_2434	Pum
Avana_Tardivo_di_Ciaculli_mandarin_4162	Pum
Deep_Red_pummelo_2347	Pum
Chinese_pummelo_1212	Pum

Pink_pummelo_2244	Pum
Kao_pan_pummelo_2248	Pum
Mato_Buntan_pummelo_3945	Pum
Tresca_pummelo_302	Pum
Anseikan_pummelo_hybrid_3979	Pum
Arajon_pummelo_2596	Pum
Citrus_sino-grandis_3148	Pum
Hirado_Buntan_pummelo_4122	Pum
Asahikan_pummelo_hybrid_3905	Pum_Hyb
Hassaku_Beni_pummelo_hybrid_3907	Pum_Hyb
Citrus_obovoidea_3465	Pum_Hyb
Citrus_shunkokan_3476	Pum_Hyb
Citrus_tengu_3464	Pum_Hyb
Cocktail_pummelo_hybrid_3555	Pum_Hyb
Cuban_shaddock_pummelo_hybrid_1462	Pum_Hyb
Hassaku_pummelo_hybrid_3942	Pum_Hyb
Lemelo_pummelo_hybrid_1481	Pum_Hyb
Lemelo_pummelo_hybrid_1775	Pum_Hyb
Moanalua_pummelo_hybrid_448	Pum_Hyb
Pomelit_pummelo_hybrid_4026	Pum_Hyb
Puma_pummelo_hybrid_3941	Pum_Hyb
Red_Aranyan_pummelo_2608	Pum_Hyb
Shaddock_X_St_michael_orange_hybrid_42	Pum_Hyb
Tahitian_pummelo_X_Star_Ruby_pummelo_hybrid_3781	Pum_Hyb
Yuma_Ponderosa_lemon_pummelo_hybrid_3488	Pum_Hyb
Philippine_pummelo_hybrid_2343	Pum_Hyb
Baishaishu_Rangpur_4129	Rangpur
Hangleson_Rangpur_3932	Rangpur
Kirumakki_Rangpur_4131	Rangpur
Knorr_Rangpur_4132	Rangpur
Srirampur_Poona_Rangpur_4136	Rangpur
Poona_Rangpur_4135	Rangpur
Srirampur_Rangpur_4137	Rangpur
Tuningmeng_Rangpur_4139	Rangpur
Borneo_Rangpur_2424	Rangpur
Santa_Barbara_Red_Rangpur_712	Rangpur
Japansche_citroen_Rangpur_2875	Rangpur
Feroniella_oblata_3510	Rel
Cleara_F2_4109	Rel
Oxanthera_neo-caledonica_3793	Rel

Eremolemon_2439	Rel_Hyb
Unnamed_rough_lemon_3060	Rough_Lemon
Florida_A_rough_lemon_3385	Rough_Lemon
Gomiri_rough_lemon_2557	Rough_Lemon
Jullander_Khatti_rough_lemon_4173	Rough_Lemon
Jullundri_Khatti_rough_lemon_3063	Rough_Lemon
Limoneira_rough_lemon_3834	Rough_Lemon
Mazoe_rough_lemon_1222	Rough_Lemon
South_African_rough_lemon_2325	Rough_Lemon
Vangasay_807_rough_lemon_3996	Rough_Lemon
Florida_rough_lemon_400	Rough_Lemon
Milam_rough_lemon_3396	Rough_Lemon
Schaub_rough_lemon_3879	Rough_Lemon
Stowe_red_rough_lemon_3185	Rough_Lemon
Aguzdera_Satsuma_4090	Satsuma
Aoshima_Satsuma_4066	Satsuma
Armstrong_Satsuma_4064	Satsuma
China_S_1_Satsuma_4158	Satsuma
China_S_11_Satsuma_4159	Satsuma
China_S_12_Satsuma_4154	Satsuma
China_S_15_Satsuma_4178	Satsuma
China_S_17_Satsuma_4179	Satsuma
China_S_2_Satsuma_4060	Satsuma
China_S_20_Satsuma_4198	Satsuma
China_S_5_Satsuma_4153	Satsuma
China_S_8_Satsuma_4177	Satsuma
China_S_9_Satsuma_4061	Satsuma
China_Satsuma_S-3_4115	Satsuma
China_Satsuma_S-6_4088	Satsuma
China_Satsuma_S-7_4089	Satsuma
China_6_15_Satsuma_4180	Satsuma
China_6_18_Satsuma_4199	Satsuma
China_6_21_Satsuma_4213	Satsuma
China_6_22_Satsuma_4214	Satsuma
Clausellina_Satsuma_4012	Satsuma
Dart_North_Satsuma_4100	Satsuma
Dart_South_Satsuma_4099	Satsuma
Dobashi_Beni_Satsuma_3794	Satsuma
Dungan_Satsuma_4101	Satsuma
Iveriya_Satsuma_4091	Satsuma

Iwasaki_Satsuma_4218	Satsuma
Kawano_Wase_Satsuma_1851	Satsuma
Kuno_Wase_Satsuma_4030	Satsuma
Lange_#3_Satsuma_4103	Satsuma
Louisiana_Early_Satsuma_4225	Satsuma
McEwen_satsuma_4098	Satsuma
Miyagawa_Satsuma_4074	Satsuma
Miyamoto_Satsuma_4216	Satsuma
Nepolitana_Satsuma_3848	Satsuma
Nichigan_Ichigo_Satsuma_4217	Satsuma
Road_164_satsuma_4102	Satsuma
Selma_Satsuma_4119	Satsuma
Shirokolistvennyi_Satsuma_4091	Satsuma
Silverhill_Satsuma_4072	Satsuma
Ueno_Satsuma_4223	Satsuma
Xie_Shan_Satsuma_4081	Satsuma
Miho_Wase_satsuma_4047	Satsuma
Okitsu_wase_Satsuma_3820	Satsuma
Frost_Owari_satsuma_3178	Satsuma
Bouquetier_de_Nice_sour_orange_4157	Sour_Orange
Citrus_myrtifolia_sour_orange_4164	Sour_Orange
Chinotto_sour_orange_hybrid_3728	Sour_Orange
Chinotto_sour_orange_3857	Sour_Orange
Daidai_double_calyx_sour_orange_656	Sour_Orange
Granitos_sour_orange_2715	Sour_Orange
Keen_sour_orange_2624	Sour_Orange
Mytrifoliataia_sour_orange_3290	Sour_Orange
Nansho_Daidai_sour_orange_2588	Sour_Orange
Rubidoux_sour_orange_3855	Sour_Orange
Seville_sour_orange_3856	Sour_Orange
Smooth_Flat_Seville_sour_orange_4079	Sour_Orange
Citrus_aurantium_var_salicifolia_3289	Sour_Orange
Stow_#20_sour_orange_2192	Sour_Orange
Tosu_sour_orange_hybrid_3607	Sour_Orange
Tunisian_sour_orange_2438	Sour_Orange
Variegated_sour_orange_622	Sour_Orange
Zhuluan_sour_orange_3930	Sour_Orange
Zhuluan_sour_orange_3981	Sour_Orange
Standard_sour_orange_628	Sour_Orange
Chinotto_sour_orange_hybrid_2375	Sour_Orange

Olivelands_sour_orange_2717	Sour_Orange
Citrus_canaliculata_3565	Sour_orange
Brazilian_sour_orange_1689	Sour_Orange
Bergamot_sour_orange_hybrid_2881	Sour_Orange
Bouquet_de_Fleurs_sour_orange_571	Sour_Orange
Citrus_intermedia_3474	Sour_Orange_Hyb
Citrus_maderaspatana_3225	Sour_Orange_Hyb
Citrus_natsudaiddai_3235	Sour_Orange_Hyb
Citrus_rokugatsu_3473	Sour_Orange_Hyb
Citrus_sulcata_3257	Sour_Orange_Hyb
Chinotto_Broadleaf_sour_orange_hybrid_4104	Sour_Orange_Hyb
Citrus_vulgaris_sour_orange_hybrid_760	Sour_Orange_Hyb
Gabbuchinee_sour_orange_hybrid_3079	Sour_Orange_Hyb
Goutoucheng_sour_orange_4004	Sour_Orange_Hyb
Konejime_sour_orange_hybrid_3611	Sour_Orange_Hyb
Orogold_sour_orange_hybrid_693	Sour_Orange_Hyb
Indian_sour_orange_hybrid_3175	Sour_Orange_Hyb
Pursha_lime_x_Chinotto_3578	Sour_Orange_Hyb
Citrus_miaray_3574	Sour_Orange_Hyb
Bahman_Persian#2_Sweet_lime_4187	Sw_Lemon
Mitha_Tulia_limetta_3051	Sw_Lemon
Pomona_sweet_lemon_4068	Sw_Lemon
Sweet_lemon_3093	Sw_Lemon
Bahman_Persian#1_Sweet_lime_4174	Sw_Lime
Citrus_limettioides_Sweet_lime_919	Sw_Lime
Citrus_limettioides_Sweet_lime_921	Sw_Lime
Mary_Ellen_sweet_lime_4053	Sw_Lime
Sniff_acidless_Rangpur_2484	Sw_Lime
Unnamed_sweet_lime_363	Sw_Lime
Weirick_acidless_rangpur_1684	Sw_Lime
Palestine_sweet_lime_1482	Sw_Lime
Agua_Tibia_sweet_orange_RRUT366	Swo
Akcay_Sekeri_sweet_orange_3802	Swo
Ambersweet_sweet_orange_4002	Swo
Argentina_sweet_orange_2802	Swo
Aziza_sweet_orange_3403	Swo
Bahianinha_Araras_navel_4168	Swo
Berna_sweet_orange_3272	Swo
Bessie_sweet_orange_1693	Swo
Bessie_sweet_orange_245	Swo

Bidwell's_Bar_sweet_orange_1512	Swo
Biondo_Comune_sweet_orange_3866	Swo
Biondo_Riccio_sweet_orange_3630	Swo
Blanca_Macetera_sweet_orange_3898	Swo
Boukhobza_sweet_orange_4156	Swo
Citrus_sinensis_433933_sweet_orange_3814	Swo
Citrus_sinensis_China_sweet_orange_3933	Swo
Citrus_sinensis_China_sweet_orange_3934	Swo
Cadena_Punchosa_sweet_orange_3246	Swo
Cadenera_Fina_sweet_orange_2856	Swo
Cadenera_sweet_orange_3632	Swo
Calabrese_Ovale_sweet_orange_3921	Swo
Capuchin_sweet_orange_3245	Swo
Catlin_sweet_orange_3860	Swo
Cipo_sweet_orange_3894	Swo
Curry_Early_sweet_orange_3347	Swo
Dalandan_sweet_orange_2373	Swo
Diller_sweet_orange_2602	Swo
East_Highlands_sweet_orange_2369	Swo
Espagnole_Sans_Pepins_sweet_orange_3249	Swo
Finike_sweet_orange_3870	Swo
Fuya_Menuda_sweet_orange_3787	Swo
Hamlin_sweet_orange_3862	Swo
Homosassa_sweet_orange_1696	Swo
Homosassa_sweet_orange_292	Swo
Imperial_variegated_sweet_orange_321	Swo
Indian_river_sweet_orange_3859	Swo
Jaffa_sweet_orange_71	Swo
Jincheng_sweet_orange_3983	Swo
Joppa_sweet_orange_281	Swo
Khalily_sweet_orange_3251	Swo
Koethen_sweet_orange_1106	Swo
Selecta_sweet_orange_1045	Swo
Macetera_sweet_orange_3871	Swo
Madam_Vinous_sweet_orange_3861	Swo
Maltese_Oval_sweet_orange_274	Swo
Marr's_Early_sweet_orange_3865	Swo
Mediterranean_nucellar_sweet_orange_4176	Swo
Navelencia_sweet_orange_287	Swo
Olivelands_sweet_orange_3863	Swo

Orange_a'_Feuilles_Lachiniees_sweet_orange_4202	Swo
Orange_de_Nice_sweet_orange_3624	Swo
Orange_of_Heaven_acideless_sweet_orange_371	Swo
Ovale_sweet_orange_3625	Swo
Parson_Brown_sweet_orange_3873	Swo
Pera_sweet_orange_3864	Swo
Perao_Valencia_3599	Swo
Rio_Farms_Vaniglia_sweet_orange_4128	Swo
Rotuma_Island_sweet_orange_3867	Swo
Rufert_A_sweet_orange_4141	Swo
Ruvel_sweet_orange_4142	Swo
Salustiana_sweet_orange_3868	Swo
Sarah_sweet_orange_3584	Swo
Sathgudi_sweet_orange_3083	Swo
Selecta_sweet_orange_3754	Swo
Shamouti_sweet_orange_3746	Swo
Star_sweet_orange_366	Swo
St_Micheal_Paperrind_sweet_orange_591	Swo
Tahiti_sweet_orange_3869	Swo
Trovita_Strain_A_sweet_orange_3182	Swo
Vaniglia_Sanguigno_acidless_sweet_orange_3801	Swo
Variegated_Cara_cara_navel_4207	Swo
Verna_sweet_orange_3807	Swo
Wetumpka_sweet_orange_3550	Swo
Xianfengcheng_sweet_orange_3982	Swo
Xinhuicheng_sweet_orange_3985	Swo
Xuegan_sweet_orange_3984	Swo
Pineapple_sweet_orange_3858	Swo
Lima_acidless_sweet_orange_950	Swo
Ruvel_nucellar_navel_3181	Swo
Unnamed_sweet_orange_hybrid_648	Swo_Hyb
US_119_3998	Swo_Hyb
Allspice_tangelo_4143	Tangelo
Baygold_tangelo_4186	Tangelo
Citrus_funadoko_3274	Tangelo
Clement_tangelo_2603	Tangelo
Early_tangelo_2560	Tangelo
Mapo_tangelo_4222	Tangelo
Marlow_tangelo_4145	Tangelo
Pearl_tangelo_2849	Tangelo

Sacaton_tangelo_2609	Tangelo
Sacaton_tangelo_3331	Tangelo
Sampson_tangelo_2418	Tangelo
San_Jacinto_tangelo_2011	Tangelo
Santa_Barbara_tangelo_4073	Tangelo
Seminole_tangelo_2604	Tangelo
Sexton_tangelo_2543	Tangelo
Siamelo_tangelo_2586	Tangelo
Sunrise_tangelo_2606	Tangelo
Sunshine_tangelo_2788	Tangelo
Thorton_tangelo_2013	Tangelo
Ugli_tangelo_2780	Tangelo
Ugli_tangelo_4171	Tangelo
UC_Seedless_Surprise_USDA_1_77_19_tangelo_4208	Tangelo
Webber_tangelo_2746	Tangelo
Wekiwa_tangelo_2012	Tangelo
Willial_tangelo_4147	Tangelo
Williams_tangelo_2787	Tangelo
Yalaha_tangelo_2559	Tangelo
Orlando_tangelo_3874	Tangelo
Minneola_tangelo_3340	Tangelo
Ortanique_tangor_3368	Tangelo
Altoona_tangor_2792	Tangor
Citrus_Benikoji_3149	Tangor
Dweet_tangor_3018	Tangor
Ellendale_tangor_3960	Tangor
H_56_tangor_3096	Tangor
Iyomikan_tangor_3980	Tangor
Ugli_tangelo_4127	Tangor
Koster_tangor_3958	Tangor
Mency_tangor_3183	Tangor
Miyauchi_Iyo_tangor_3946	Tangor
Ruddy_tangor_4146	Tangor
Sue_Linda_Temple_tangor_3810	Tangor
Tankan_tangor_2224	Tangor
Tankan_tangor_3875	Tangor
Kiyomi_tangor_3952	Tangor
Wincia_tangor_4148	Tangor
Temple_tangor_2598	Tangor
Argentine_trifoliate_3206	Trifol

Benoit_trifoliolate_3547	Trifol
Florida_trifoliolate_2862	Trifol
Hiryu_Flying_dragon_trifoliolate_3882	Trifol
Jacobson_trifoliolate_3411	Trifol
Marks_trifoliolate_3588	Trifol
#22_trifoliolate_4172	Trifol
P_trifoliolateiata_4007	Trifol
P_trifoliolateiata_4009	Trifol
Rich_16_6_trifoliolate_3485	Trifol
Ronnse_trifoliolate_3570	Trifol
Texas_trifoliolate_2861	Trifol
Usda_trifoliolate_1498	Trifol
Australian_trifoliolate_3151	Trifol
Barnes_trifoliolate_2554	Trifol
Benecke_trifoliolate_3338	Trifol
Christiansen_trifoliolate_3345	Trifol
English_Dwarf_trifoliolate_3876	Trifol
English_Large_trifoliolate_3548	Trifol
Fairhope_trifoliolate_3351	Trifol
Flying_dragon_trifoliolate_3330A	Trifol
Flying_dragon_trifoliolate_3330B	Trifol
Frost_trifoliolate_3484	Trifol
Hiryu_Flying_dragon_trifoliolate_3795	Trifol
Kryder_Medium_trifoliolate_3212	Trifol
Kryder_16_6_trifoliolate_3210	Trifol
Kryder_28_3_trifoliolate_3219	Trifol
Kryder_5_5_trifoliolate_3586	Trifol
Kryder_55_1_trifoliolate_3486	Trifol
Kryder_55_5_trifoliolate_3215	Trifol
Kryder_60_2_trifoliolate_3213	Trifol
Kryder_8_5_trifoliolate_3218	Trifol
Monoembryonic_trifoliolate_3888	Trifol
Pomeroy_trifoliolate_1717	Trifol
#26_trifoliolate_3939	Trifol
#27_trifoliolate_3938	Trifol
P_trifoliolateiata_4006	Trifol
Rich_12_2_trifoliolate_3209	Trifol
Rich_22_2_trifoliolate_3211	Trifol
Rich_7_5_trifoliolate_3587	Trifol
Rubidoux_trifoliolate_838	Trifol

Simmons_trifoliolate_3549	Trifol
Small_Leaf_4017	Trifol
Taylor_trifoliolate_3571	Trifol
Towne_F_trifoliolate_3572	Trifol
Towne_G_trifoliolate_3207	Trifol
Webber_Fawcette_#22_trifoliolate_2552	Trifol
Yamaguchi_trifoliolate_3412	Trifol
Citradarin_trifoliolate_hybrid_2618	Trifol_Hyb
Citradia_trifoliolate_hybrid_1436	Trifol_Hyb
Citradia_trifoliolate_hybrid_1437	Trifol_Hyb
Citradia_trifoliolate_hybrid_1438	Trifol_Hyb
Citraldin_trifoliolate_hybrid_2629	Trifol_Hyb
Citrangor_trifoliolate_hybrid_1447	Trifol_Hyb
Citremon_trifoliolate_hybrid_1448	Trifol_Hyb
Citremon_trifoliolate_hybrid_1449	Trifol_Hyb
Glen_citrangedin_trifoliolate_hybrid_3573	Trifol_Hyb
Hall_grapefruit_X_Ptrifoliolate_3821	Trifol_Hyb
Microcitrus_X_trifoliolate_3566	Trifol_Hyb
S281_citrangelo_3552	Trifol_Hyb
S302_citranguma_3415	Trifol_Hyb
Trifeola_trifoliolate_hybrid_3954	Trifol_Hyb
X639_trifoliolate_hybrid_3957	Trifol_Hyb
Citrus_ujukitsu_3467	ujukitsu
Campbell_Valencia_3032	Valencia
Chapman_Valencia_4229	Valencia
Coarse_Valencia_1518	Valencia
Cutter_Valencia_3030	Valencia
Dom_Joao_Valencia_3872	Valencia
Frost_Valencia_3031	Valencia
Fuzzy_Valencia_1240	Valencia
Hart's_late_Valencia_314	Valencia
Hart's_Tardiff_Valencia_570	Valencia
Lue_Gim_Gong_Valencia_3025	Valencia
Midknight_Valencia_3955	Valencia
Olinda_McEwen_Valencia_4230	Valencia
Olinda_Mulholland_Valencia_4231	Valencia
Pehrson_#3_Valencia_4196	Valencia
Pehrson_#4_Valencia_4197	Valencia
Precoce_de_Valence_3373	Valencia
Rocky_Hill_Valencia_4232	Valencia

San_Marino_Valencia_2689	Valencia
Seedless_Valencia_2776	Valencia
Variegated_Valencia_4001	Valencia
Werley_Valencia_3339	Valencia
Olinda_Valencia_2750	Valencia
Rhode_Red_Valencia_3395	Valencia
Valencia_late_4077	Valencia
Delta_Valencia_3963	Valencia

Table. 2 Accessions in CVC that were represented in the 382 UCR diversity set
382 accessions genotyped at 56,367 PHR and NMH loci for the study. Members belonging to clonal groups and identical/near identical accessions have been removed based on percent identity

Sample Name (CVC) and CRC	Type (new classification)
Orlando_bittersweet_orange_1588	BitterSweet_Orange
Amoa_8_Blood_mandarin_4237	Blood_Mand
Bitters_C22_4219	Citrandarin
Furr_C57_citrange_4221	Citrandarin
US_852_4245	Citrandarin
X639_trifoliate_hybrid_3957	Citrandarin
Benton_citrange_3908	Citrange
C32_citrange_3911	Citrange
C35_citrange_3912	Citrange
Carrizo_citrange_2863	Citrange
Citrange_#1416_citrange_2866	Citrange
Cunningham_citrange_271	Citrange
Morton_citrange_1463	Citrange
Rusk_citrange_1441	Citrange
Sanford_citrange_276	Citrange
Savage_citrange_275	Citrange
Assads_citron_RSD2008001	Citron
Buddha's_Hand_citron_3768	Citron
Citron_of_Commerce_3518	Citron
Corsican_citron_1795	Citron
Diamante_citron_3523	Citron
Dulcia_citron_3654	Citron
Etrog_citron_3891	Citron
GouTouXiangYuan_RRUT395	Citron
GouTouXiangYuan_RRUT396	Citron
GouTouXiangYuan_RRUT398	Citron
Indian_citron_138	Citron
Italian_citron_3530	Citron
Limau_Mata_Susu_citron_hybrid_3755	Citron
Mexican_citron_3531	Citron
Morningsong_Temoni_citron-Karp	Citron
Odorata_citron_3655	Citron
Papuan_citron_3532	Citron
Philippine_citron_3533	Citron
Sicily_citron_3534	Citron
South_Coast_Field_Station_citron_3546	Citron

Unnamed_citron_hybrid_3174	Citron
Yemen_citron_2847	Citron
Yunnanese_citron_3798	Citron
Badhri_citron_hybrid_3723	Citron_Hyb
Bengal_citron_3055	Citron_Hyb
Dabbe_citron_hybrid_3190	Citron_Hyb
Djerok_citron_hybrid_2456	Citron_Hyb
Indian_citron_661	Citron_Hyb
Paak_Ling_Mung_lemon_hybrid_3193	Citron_Mand_Hyb
Unnamed_lemon_hybrid_3172	Citron_Mand_Hyb
Citrus_pennivesiculata_2434	Citron_papeda_Hyb
Citrus_vulgaris_sour_orange_hybrid_760	Citron_papeda_Hyb
African_Shaddock_X_Rubidoux_trifoliolate_3969	Citrumelo
Citrumelo_1452	Citrumelo
Citrumelo_3348	Citrumelo
Hall_grapefruit_X_Rubiboux_trifoliolate_3889	Citrumelo
Sacaton_citrumelo_3337	Citrumelo
Sacaton_citrumelo_3417	Citrumelo
Swingle_citrumelo_3767	Citrumelo
Algerian_clementine_279	Clem
Clementine_X_Pearl_mandarin_hybrid_4020	Clem_Hyb
Clementine_X_Silverhill_mandarin_hybrid_3731	Clem_Hyb
Duncan_grapefruit_3832	Gft
Royal_grapefruit_248	Gft
Citrus_otachibana_grapefruit_hybrid_3470	Gft_Hyb
Citrus_rugulosa_pummelo_hybrid_3556	Gft_Hyb
Chironja_grapefruit_hybrid_3909	Gft_Hyb
Citrus_hiroshimana_Natsuzabon_3904	Gft_Hyb
Rex_Union_grapefruit_hybrid_3962	Gft_Hyb
Rojo_Blanco_pummelo_hybrid_4029	Gft_Hyb
Sukega_grapefruit_hybrid_4149	Gft_Hyb
Triumph_grapefruit_297	Gft_Hyb
Valentine_pummelo_hybrid_4182	Gft_Hyb
Wheeny_grapefruit_hybrid_2885	Gft_Hyb
Fortunella_obovata_3475	Kumquat
Fortunella_polyandra_3901	Kumquat
Marumi_SRA489_kumquat_4125	Kumquat
Nordmann_Seedless_kumquat_4070	Kumquat
Citrangquat_19_15_7_3259	Kumquat_Hyb
Eustis_limequat_3603	Kumquat_Hyb

Excalibur_red_lime_hybrid_4239	Kumquat_Hyb
Indio_mandarinquat_3759	Kumquat_Hyb
Procimequat_kumquat_hybrid_3295	Kumquat_Hyb
Rio_Grande_Valley_lemonquat_4097	Kumquat_Hyb
Sinton_citranguequat_3642	Kumquat_Hyb
Tavares_kumquat_hybrid_4048	Kumquat_Hyb
Thomasville_citranguequat_1440	Kumquat_Hyb
Warren_limequat_3069	Kumquat_Hyb
Citrus_kharna_3059	Lemon
Interdonato_lemon_hybrid_3593	Lemon
Iran_lemon_3162	Lemon
Khoubs_el_arsa_lemon_2477	Lemon
Khoubs_el_ersa_lemon_2489	Lemon
Kulu_lemon_3045	Lemon
Lapithiokii_lemon_4005	Lemon
Primofiore_lemon_3491	Lemon
Soh_Synteng_lemon_hybrid_3216	Lemon
Unnnamed_lemon_hybrid_3155	Lemon
Bitrouni_lemon_3265	Lemon_Hyb
Citrus_succosa_3280	Lemon_Hyb
Citrus_volkameriana_3050	Lemon_Hyb
Improved_Meyer_lemon_3737	Lemon_Hyb
Iraq_lemon_limetta_3492	Lemon_Hyb
Limonette_De_Marrakech_lemon_hybrid_3989	Lemon_Hyb
Lumia_lemon_hybrid_3925	Lemon_Hyb
New_Zealand_lemonade_lemon_4181	Lemon_Hyb
Orban_lemon_hybrid_4140	Lemon_Hyb
Peretta_lemon_hybrid_3924	Lemon_Hyb
Unnamed_lemon_hybrid_3154	Lemon_Hyb
Unnamed_lemon_3885	Lemon_Hyb
Citrus_excelsa_2316	Lime
Davao_papeda_hybrid_2427	Lime
Mexican_lime_1710	Lime
Lemonime_lemon_hybrid_2459	Lime_Hyb
Nicaraguan_lemon_3841	Lime_Hyb
Perrine_lime_hybrid_2458	Lime_Hyb
Rangpur_X_Troyer_hybrid_3997	Lime_Hyb
Soh_Long_lemon_hybrid_3199	Lime_Hyb
Mitha_Tulia_limetta_3051	limetta
Frost_Lisbon_lemon_3176	Lisbon_Lemon

Prior_Lisbon_lemon_3505	Lisbon_Lemon
Citrus_lycopersicaformis_3564	Mand_Type_1
Citrus_tachibana_3150	Mand_Type_1
Sun_Chu_Sha_mandrin_4003	Mand_Type_1
Yellow_Rind_mandarin_3895	Mand_Type_1
Baiju_nucellar_RRUT76	Mand_Hyb
Bower_mandarin_3649	Mand_Hyb
Citrus_amblycarpa_2485	Mand_Hyb
Citrus_depressa_2448	Mand_Hyb
Citrus_erythrosa_3292	Mand_Hyb
Citrus_indica_3163	Mand_Hyb
Citrus_keraji_3144	Mand_Hyb
Citrus_kinokuni_3887	Mand_Hyb
Citrus_leiocarpa_3147	Mand_Hyb
Citrus_nippokoreana_3228	Mand_Hyb
Citrus_sunki_3143	Mand_Hyb
Citrus_yatsushiro_3880	Mand_Hyb
Calushu_calamondin_2867	Mand_Hyb
Changsha_mandarin_3577	Mand_Hyb
Cleopatra_mandarin_3844	Mand_Hyb
Daisy_mandarin_3910	Mand_Hyb
Dancy_Weshart_602	Mand_Hyb
Emerald_mandarin_4123	Mand_Hyb
Empress_mandarin_3613	Mand_Hyb
Encore_mandarin_3569	Mand_Hyb
Fairchild_mandarin_3559	Mand_Hyb
Fallglo_mandarin_3990	Mand_Hyb
Fortune_mandarin_3560	Mand_Hyb
Fremont_mandarin_3558	Mand_Hyb
Frua_mandarin_3022	Mand_Hyb
Gold_Nugget_mandarin_3913	Mand_Hyb
Hansen_mandarin_4152	Mand_Hyb
Hickson_mandarin_4161	Mand_Hyb
Honey_mandarin_3177	Mand_Hyb
Huang_Yen_Man_Chie_3897	Mand_Hyb
Hung_Kat_mandarin_2331	Mand_Hyb
Imperial_mandarin_4160	Mand_Hyb
Iyo_San_Ponkan_4118	Mand_Hyb
Kara_mandarin_3019	Mand_Hyb
Kincy_mandarin_4144	Mand_Hyb

King_mandarin_303	Mand_Hyb
Kinnow_mandarin_3021	Mand_Hyb
Kinokuni_mandarin_696	Mand_Hyb
Kunembo_mandarin_3346	Mand_Hyb
Laranja_Cravo_mandarin_2893	Mand_Hyb
Lee_mandarin_3851	Mand_Hyb
Mandarinette_3405	Mand_Hyb
Michal_mandarin_4049	Mand_Hyb
Murcott_mandarin_3846	Mand_Hyb
Necked_mandarin_3738	Mand_Hyb
NISSV_E_mandarin_hybrid_3965	Mand_Hyb
Nova_mandarin_3615	Mand_Hyb
Novelty_X_Ellendale_mandarin_hybrid_3956	Mand_Hyb
Page_mandarin_3616	Mand_Hyb
palazzelli_mandarin_4244	Mand_Hyb
Parson's_Special_mandarin_300	Mand_Hyb
Pixie_mandarin_3568	Mand_Hyb
Richard's_Special_mandarin_4117	Mand_Hyb
Robinson_mandarin_3850	Mand_Hyb
Rubidoux_mandarin_4031	Mand_Hyb
Scarlet_Emperor_mandarin_3226	Mand_Hyb
Shiranui_Dekopan_4113	Mand_Hyb
Solid_Scarlet_mandarin_3328	Mand_Hyb
Som_Kaeo_mandarin_3852	Mand_Hyb
US_Furr_SRA337_mandarin_4238	Mand_Hyb
Sunburst_mandarin_3809	Mand_Hyb
Szinkom_mandarin_3085	Mand_Hyb
Tango_mandarin_4183	Mand_Hyb
Tien_Chieh_mandarin_2590	Mand_Hyb
Tieu_nucellar_mandarin_RRUT87	Mand_Hyb
Indian_sour_orange_hybrid_3175	Mand_Hyb
USDA_6-15-150_mandarin_4165	Mand_Hyb
US_Superna_USDA_88_2_mandarin_hybrid_3991	Mand_Hyb
USDA_88_3_mandarin_hybrid_3992	Mand_Hyb
Wilking_mandarin_3020	Mand_Hyb
Willowleaf_mandarin_3843	Mand_Hyb
Citrus_ujukitsu_3467	Mand_Hyb_pum
Phillipine_red_Rangpur_2318	Mand_Hyb_Rangpur
Citrus_hanayu_3469	Mand-Ichangensis_Hyb
Citrus_megaloxycarpa_hybrid_3241	Megaloxycarpa

Microcitrus_australasica_3670	Microcitrus
Microcitrus_australasica_3672	Microcitrus
Microcitrus_australis_3665	Microcitrus
Sydney_hybrid_1485	Microcitrus_Hyb
Parent_Washington_navel_1241A	Navel
Citrus_hystrix_3103	Papeda_micrantha
Citrus_micrantha_3605	Papeda_micrantha
Cabuyo_papeda_432	Papeda_micrantha
Citrus_webberri_Kalpi_1456	Papeda_micrantha
Citrus_ichangensis_2327	Papeda_Ichangensis
Citrus_ichangensis_3931	Papeda_Ichangensis
Citrus_hongheesis_3797	Papeda_ichangensis_pummelo_hybrid
Citrus_latipes_Khasi_3052	Papeda_ichangensis_pummelo_hybrid
Kabosu_papeda_hybrid_3943	Papeda_ichangensis_tachibana_hybrid
Yuzu_ichandarin_1216	Papeda_ichangensis_tachibana_hybrid
Citrus_macrophylla_papeda_hybrid_2446	Papeda_Hyb
Citrus_webberii_var_Montana_767	Papeda_Hyb
Citrus_yuko_3146	Papeda_Hyb
Ichang_lemon_1215	Papeda_Hyb
Kulobot_papeda_hybrid_3612	Papeda_Hyb
Soh_Niangrang_papeda_3203	Papeda_Hyb
Unnamed_papeda_3056	Papeda_Hyb
African_pummelo_2346	Pum
Alemoen_pummelo_2487	Pum
BuOl_1_acidless_yellow_RRUT92	Pum
Citrus_grandis_2340	Pum
Citrus_grandis_2453	Pum
Citrus_grandis_3282	Pum
Chinese_pummelo_1224	Pum
Deep_Red_pummelo_2347	Pum
Fleming's_Shaddock_pummelo_578	Pum
Hirado_Buntan_pummelo_4122	Pum
Itoshima_3928	Pum
Kao_Phuang_pummelo_2352	Pum
Kao_Phuang_pummelo_3926	Pum
Kao_Ruan_Tia_pummelo_2350	Pum
Kao_Ruan_Tia_pummelo_2351	Pum
Karn_Lau_Yau_pummelo_2341	Pum

Mato_Buntan_pummelo_3945	Pum
Nakhon_Chaisi_pummelo_2353	Pum
Pan_Dan_pummelo_2752	Pum
Pauthel_pummelo_3949	Pum
Philippine_pummelo_hybrid_644	Pum
Pin_Shan_Kong_Yan_pummelo_2348	Pum
Pink_pummelo_2246	Pum
Pomelit_pummelo_hybrid_4026	Pum
Reinking_pummelo_3805	Pum
Roeding's_Pink_pummelo_1208	Pum
Siamese_Acidless_pummelo_2240	Pum
Sunshine_pummelo_2236	Pum
Sweet_pummelo_3067	Pum
Tahitian_pummelo_3806	Pum
Tau_Yau_pummelo_2583	Pum
Thong_Deep_pummelo_3927	Pum
Thong_Deep_pummelo_4027	Pum
Unnamed_Chinese_RRUT223	Pum
Kao_pan_pummelo_2248	Pum
Kao_Pan_pummelo_2355	Pum
Webber_shaddock_pummelo_4233	Pum
Cuban_shaddock_pummelo_hybrid_1462	Pum_Citron_Hyb
Lemelo_pummelo_hybrid_1481	Pum_Citron_Hyb
Lemelo_pummelo_hybrid_1775	Pum_Citron_Hyb
Moli_kurikuri_pummelo_hybrid_579	Pum_Citron_Hyb
Ponderosa_lemon_hybrid_294	Pum_Citron_Hyb
Sour_pummelo_3066	Pum_Citron_Hyb
Anseikan_pummelo_hybrid_3979	Pum_Hyb
Arajon_pummelo_2596	Pum_Hyb
Asahikan_pummelo_hybrid_3905	Pum_Hyb
Banokan_pummelo_3950	Pum_Hyb
Hassaku_Beni_pummelo_hybrid_3907	Pum_Hyb
Citrus_hiroshimana_Natsuzabon_3275	Pum_Hyb
Citrus_obovoidea_3465	Pum_Hyb
Citrus_shunkokan_3476	Pum_Hyb
Citrus_sino-grandis_3148	Pum_Hyb
Citrus_tengu_3464	Pum_Hyb
Chinese_pummelo_1212	Pum_Hyb
Cocktail_pummelo_hybrid_3555	Pum_Hyb
Kawachi_Bankan_3944	Pum_Hyb

Kinokawa_Buntan_pummelo_3978	Pum_Hyb
Philippine_pummelo_hybrid_2343	Pum_Hyb
Puma_pummelo_hybrid_3941	Pum_Hyb
Rubidoux_pummelo_hybrid_#1_4028	Pum_Hyb
Shaddock_X_St_michael_orange_hybrid_42	Pum_Hyb
Siamese_pummelo_2421	Pum_Hyb
Suisho_Buntan_pummelo_3947	Pum_Hyb
Tresca_pummelo_302	Pum_Hyb
Banpeiyu_pummelo_3961	Pum-Allele-F-admixed
Chandler_pummelo_3244	Pum-Allele-F-admixed
Egami_Buntan_pummelo_3959	Pum-Allele-F-admixed
Hunan_pummelo_1225	Pum-Allele-F-admixed
Moanalua_pummelo_hybrid_448	Pum-Allele-F-admixed
Pink_pummelo_2244	Pum-Allele-F-admixed
Pong_yau_pummelo_2342	Pum-Allele-F-admixed
Red_pummelo_2245	Pum-Allele-F-admixed
Santa_Barbara_Red_Rangpur_712	Rangpur
Sniff_acidless_Rangpur_2484	Rangpur
India_lemon_2323	Rough_lemon
Indian_rough_lemon_2544	Rough_lemon
Jullander_Khatti_rough_lemon_4173	Rough_Lemon
Schaub_rough_lemon_3879	Rough_Lemon
South_African_rough_lemon_2325	Rough_Lemon
Vangasay_807_rough_lemon_3996	Rough_Lemon
Milam_rough_lemon_3396	Rough_lemon_Hyb
Frost_Owari_satsuma_3178	Satsuma
Mytrifoliata_sour_orange_3290	Sour_Orange
Nansho_Daidai_sour_orange_2588	Sour_Orange
Smooth_Flat_Seville_sour_orange_4079	Sour_Orange
Citrus_aurantium_var_salicifolia_3289	Sour_Orange
Standard_sour_orange_628	Sour_Orange
Zhuluan_sour_orange_3930	Sour_Orange
Citrus_intermedia_3474	Sour_Orange_Hyb
Citrus_maderaspatana_3225	Sour_Orange_Hyb
Citrus_natsudaiddai_3235	Sour_Orange_Hyb
Citrus_rokugatsu_3473	Sour_Orange_Hyb
Goutoucheng_sour_orange_4004	Sour_Orange_Hyb
Orogold_sour_orange_hybrid_693	Sour_Orange_Hyb
Pursha_lime_x_Chinotto_3578	Sour_Orange_Hyb
Bergamot_sour_orange_hybrid_2881	Sour_Orange_Hyb_citron

Citrus_canaliculata_3565	Sour_Orange_Hyb_citron
Chinotto_sour_orange_hybrid_3728	Sour_Orange_Hyb_citron
Bahman_Persian#1_Sweet_lime_4174	Sw_Lemon
Bahman_Persian#2_Sweet_lime_4187	Sw_Lemon
Pomona_sweet_lemon_4068	Sw_Lemon
Sweet_lemon_3093	Sw_Lemon
Ambersweet_sweet_orange_4002	Swo
Unnamed_sweet_orange_hybrid_648	Swo_Hyb
US_119_3998	Swo_Hyb
Allspice_tangelo_4143	Tangelo
Baygold_tangelo_4186	Tangelo
Citrus_funadoko_3274	Tangelo
Clement_tangelo_2603	Tangelo
Early_tangelo_2560	Tangelo
Mapo_tangelo_4222	Tangelo
Marlow_tangelo_4145	Tangelo
Minneola_tangelo_3340	Tangelo
Orlando_tangelo_3874	Tangelo
Ortanique_tangor_3368	Tangelo
Pearl_tangelo_2849	Tangelo
Sacaton_tangelo_2609	Tangelo
Sacaton_tangelo_3331	Tangelo
Sampson_tangelo_2418	Tangelo
San_Jacinto_tangelo_2011	Tangelo
Santa_Barbara_tangelo_4073	Tangelo
Seminole_tangelo_2604	Tangelo
Sexton_tangelo_2543	Tangelo
Siamelo_tangelo_2586	Tangelo
Sunrise_tangelo_2606	Tangelo
Sunshine_tangelo_2788	Tangelo
Thorton_tangelo_2013	Tangelo
UC_Seedless_Surprise_USDA_1_77_19_tangelo_4208	Tangelo
Webber_tangelo_2746	Tangelo
Willial_tangelo_4147	Tangelo
Williams_tangelo_2787	Tangelo
Yalaha_tangelo_2559	Tangelo
Altoona_tangor_2792	Tangor
Citrus_Benikoji_3149	Tangor
Dweet_tangor_3018	Tangor
H_56_tangor_3096	Tangor

Iyomikan_tangor_3980	Tangor
Ugli_tangelo_4127	Tangor
Kiyomi_tangor_3952	Tangor
Koster_tangor_3958	Tangor
Mency_tangor_3183	Tangor
Ruddy_tangor_4146	Tangor
Tankan_tangor_2224	Tangor
Temple_tangor_2598	Tangor
Wincia_tangor_4148	Tangor
#26_trifoliate_3939	Trifol
#27_trifoliate_3938	Trifol
Argentine_trifoliate_3206	Trifol
Flying_dragon_trifoliate_3330A	Trifol
Monoembryonic_trifoliate_3888	Trifol
Simmons_trifoliate_3549	Trifol
Small_Leaf_4017	Trifol
Citradia_trifoliate_hybrid_1437	Trifol_Hyb
Citradia_trifoliate_hybrid_1438	Trifol_Hyb
Citraldin_trifoliate_hybrid_2629	Trifol_Hyb
Citradarin_trifoliate_hybrid_2618	Trifol_Hyb
Citrangor_trifoliate_hybrid_1447	Trifol_Hyb
Citremon_trifoliate_hybrid_1448	Trifol_Hyb
Citremon_trifoliate_hybrid_1449	Trifol_Hyb
Glen_citrangedin_trifoliate_hybrid_3573	Trifol_Hyb
Microcitrus_X_trifoliate_3566	Trifol_Hyb
S281_citrangelo_3552	Trifol_Hyb
S302_citranguma_3415	Trifol_Hyb
Trifeola_trifoliate_hybrid_3954	Trifol_Hyb

Table 3. 85 accessions used in maximum likelihood tree.

85 accession identified to be non-admixed belonging to citrus ancestral progenitor species and used to infer phylogeny using maximum likelihood tree.

Sample Name CRC
Assads_citron_RSD2008001
Buddhas_Hand_citron_3768
Citron_of_Commerce_3518
Corsican_citron_1795
Dulcia_citron_3654
Etrog_citron_3891
GouTouXiangYuan_citron_RRUT395
GouTouXiangYuan_citron_RRUT396
GouTouXiangYuan_citron_RRUT398
Indian_citron_138_
Italian_citron_3530
Mexican_citron_3531
Morningsong_Temoni_citron
Papuan_citron_3532
Philippine_citron_3533
Sicily_citron_3534_
Yemen_citron_2847
Diamante_citron_3523
Yunnanese_citron_3798
Unnamed_citron_3174
South_coast_field_station_citron_3546
Limau_Matu_citron_3755
Odorata_citron_3655
Fortunella_polyandra_3901
Marumi_SRA489_kumquat_4125
Nordmann_kumquat_4070
C_lycopersicaformis_mandarin_3564
Sun_Chua_Sha_mandarin_4003
Yellow_Rind_mandarin_3895
C_hystrix_3103
C_hystrix_Cabuyo_432
C_micrantha_3605
C_webberii_Kalpi_1456
African_pummelo_2346
Alemoen_pummelo_2487

C_grandis_pummelo_3282
C_grandis_pummelo_2340
C_grandis_pummelo_2453
Chinese_pummelo_1224
Flemings_Shaddock_pummelo_578
Itoshima_3928
Kao_Phuang_pummelo_3926
Pan_Dan_pummelo_2752
Pauthel_pummelo_3949
Pink_pummelo_2246
Pomelit_pummelo_4026
Reinking_pummelo_3805
Roedings_Pink_pummelo_1208
Sunshine_pummelo_2236
Sweet_pummelo_3067
Tahitian_pummelo_3806
Thong_Deep_pummelo_4027
Thong_Deep_pummelo_3927
Webber_Shaddock_pummelo_4233
Philippine_pummelo_644
Kao_Phuang_pummelo_2352
Deep_Red_pummelo_2347
Hirado_Buntan_4122
Kao_Ruan_Tia_pummelo_2350
Karn_Lau_Yau_pummelo_2341
Nakhon_Chaisi_pummelo_2353
Siamese_Acidless_pummelo_2240
Tau_Yau_pummelo_2583
Kao_Pan_pummelo_2355
Mato_Buntan_3945
Kao_Ruan_Tia_pummelo_2351
Unnamed_Chinese_pummelo_RRUT223
Pin_Shan_Kong_Yan_pummelo_2348
BuOl_1_acidless_yellow_pummelo_RRUT92
Kao_Pan_pummelo_2248
Argentine_trifoliolate_3206
26_Nanjing_trifoliolate_3939
27_trifoliolate_3938
Small_Leaf_trifoliolate_4017
Monoembryonic_trifoliolate_3888

Flying_Dragon_trifoliate_3330A
Simmons_trifoliate_3549
M_australasica_3672
M_australasica_3670
Sydney_hybrid_1485
M_australis_3665
C_tachibana_mandarin_3150
C._ichangensis_2327
C._ichangensis_3931
C._hongheensis_3797

Table 4. Diversity statistics

Genetic parameters were estimated using observed heterozygosity (Ho), expected heterozygosity (He), polymorphic information content (PIC) and percent polymorphic loci with genotyping calls at PHR 56,367 loci for 924 and the 382 UCR diversity panel accessions

Statistics	382 UCR Diversity Set	924 CVC accessions
Minor Allele Frequency	0.19	0.189
Observed Heterozygosity (Ho)	0.198	0.241
Expected Heterozygosity (He)	0.268	0.262
Mean PIC	0.251	0.223
% Polymorphic markers (q>0.05)	0.831	0.802
% Polymorphic markers (q>0.01)	0.979	0.963
% Polymorphic markers (q>0.0)	1	1

Table 5. ADMIXTURE analysis with UCR 382 diversity set
Estimated Q from ADMIXTURE results at K=6
Classification is based on admixture analysis and introgression analysis

pum	pap	trifol	kum/ Micro	mand	citron	Sample Name_CRC	Type
0.442 202	0.000 01	0.000 01	0.0000 1	0.557 758	0.000 01	Orlando_bittersweet_orange_1588	BitterSweet_Orange
0.111 554	0.000 01	0.000 01	0.0000 1	0.888 406	0.000 01	Amoa_8_Blood_mandarin_4237	Blood_Mand
0.000 01	0.000 01	0.477 238	0.0000 1	0.522 722	0.000 01	Bitters_C22_4219	Citrandarin
0.000 01	0.000 01	0.465 816	0.0000 1	0.534 144	0.000 01	Furr_C57_citrange_4221	Citrandarin
0.000 01	0.000 01	0.494 028	0.0000 1	0.505 932	0.000 01	US_852_4245	Citrandarin
0.000 01	0.000 01	0.463 488	0.0000 1	0.536 472	0.000 01	X639_trifoliate_hybrid_3957	Citrandarin
0.030 636	0.000 01	0.524 441	0.0000 1	0.444 894	0.000 01	Benton_citrange_3908	Citrange
0.150 195	0.000 01	0.553 612	0.0000 1	0.296 163	0.000 01	C32_citrange_3911	Citrange
0.244 042	0.000 01	0.546 44	0.0000 1	0.209 488	0.000 01	C35_citrange_3912	Citrange
0.202 643	0.000 01	0.529 686	0.0000 1	0.267 641	0.000 01	Carrizo_citrange_2863	Citrange
0.024 341	0.000 01	0.526 458	0.0000 1	0.449 17	0.000 01	Citrange_#1416_citrange_2866	Citrange
0.184 982	0.000 01	0.513 625	0.0000 1	0.301 363	0.000 01	Cunningham_citrange_271	Citrange
0.165 296	0.000 01	0.520 646	0.0000 1	0.314 028	0.000 01	Morton_citrange_1463	Citrange
0.047 472	0.000 01	0.526 001	0.0000 1	0.426 496	0.000 01	Rusk_citrange_1441	Citrange
0.117 997	0.000 01	0.519 403	0.0000 1	0.362 57	0.000 01	Sanford_citrange_276	Citrange
0.196 488	0.000 01	0.531 131	0.0000 1	0.272 352	0.000 01	Savage_citrange_275	Citrange
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Assads_citron_RSD2008001	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Buddha's_Hand_citron_3768	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Citron_of_Commerce_3518	Citron
0.000 01	0.000 011	0.000 011	0.0000 1	0.000 01	0.999 948	Corsican_citron_1795	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Diamante_citron_3523	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Dulcia_citron_3654	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	Etrog_citron_3891	Citron
0.000 01	0.000 01	0.000 01	0.0000 1	0.000 01	0.999 95	GouTouXiangYuan_RRUT395	Citron

0.000	0.000	0.000	0.0000	0.000	0.999	GouTouXiangYuan_RRUT396	Citron
01	017	01	1	01	943		
0.000	0.000	0.000	0.0000	0.000	0.999	GouTouXiangYuan_RRUT398	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Indian_citron_138	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Italian_citron_3530	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Limau_Mata_Susu_citron_hybrid_3755	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Mexican_citron_3531	Citron
01	01	01	11	01	949		
0.000	0.000	0.000	0.0000	0.000	0.999	Morningsong_Temoni_citron-Karp	Citron
014	01	01	1	01	946		
0.000	0.000	0.000	0.0000	0.000	0.999	Odorata_citron_3655	Citron
01	013	01	1	01	947		
0.000	0.000	0.000	0.0000	0.000	0.999	Papuan_citron_3532	Citron
01	011	01	1	01	949		
0.000	0.000	0.000	0.0000	0.000	0.999	Philippine_citron_3533	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Sicily_citron_3534	Citron
011	01	01	1	01	948		
0.000	0.000	0.000	0.0000	0.000	0.999	South_Coast_Field_Station_citron_3546	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Unnamed_citron_hybrid_3174	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Yemen_citron_2847	Citron
01	01	01	1	01	95		
0.000	0.000	0.000	0.0000	0.000	0.999	Yunnanese_citron_3798	Citron
01	019	01	1	01	941		
0.348	0.000	0.000	0.0000	0.107	0.543	Badhri_citron_hybrid_3723	Citron_Hyb
965	01	01	1	462	542		
0.671	0.000	0.000	0.0000	0.000	0.328	Bengal_citron_3055	Citron_Hyb
091	01	01	1	01	869		
0.295	0.000	0.000	0.0000	0.181	0.523	Dabbe_citron_hybrid_3190	Citron_Hyb
16	01	01	1	417	393		
0.504	0.000	0.000	0.0000	0.000	0.495	Djerok_citron_hybrid_2456	Citron_Hyb
56	01	01	1	01	4		
0.000	0.459	0.000	0.0000	0.000	0.540	Indian_citron_661	Citron_Hyb
01	894	01	1	01	066		
0.000	0.000	0.000	0.0000	0.508	0.491	Paak_Ling_Mung_lemon_hybrid_3193	Citron_Mand_Hyb
01	01	01	1	734	226		
0.000	0.000	0.000	0.0000	0.507	0.492	Unnamed_lemon_hybrid_3172	Citron_Mand_Hyb
01	01	01	1	613	347		
0.000	0.450	0.000	0.0000	0.000	0.549	Citrus_pennivesiculata_2434	Citron_papeda_Hyb
01	099	01	1	01	861		
0.000	0.468	0.000	0.0000	0.000	0.531	Citrus_vulgaris_sour_orange_hybrid_760	Citron_papeda_Hyb
01	118	01	1	01	841		
0.544	0.000	0.455	0.0000	0.000	0.000	African_Shaddock_X_Rubidoux_trifoliolate_3969	Citrumelo
345	01	614	11	01	01		
0.435	0.000	0.497	0.0000	0.067	0.000	Citrumelo_1452	Citrumelo
196	01	001	1	774	01		
0.292	0.000	0.510	0.0000	0.197	0.000	Citrumelo_3348	Citrumelo
071	01	637	1	262	01		
0.300	0.000	0.517	0.0000	0.181	0.000	Hall_grapefruit_X_Rubiboux_trifoliolate_3889	Citrumelo
581	01	899	1	49	01		
0.372	0.000	0.508	0.0000	0.119	0.000	Sacaton_citrumelo_3337	Citrumelo
106	01	514	1	35	01		

0.341	0.000	0.549	0.0000	0.109	0.000	Sacaton_citrumelo_3417	Citrumelo
088	01	042	1	84	01		
0.289	0.000	0.506	0.0000	0.203	0.000	Swingle_citrumelo_3767	Citrumelo
973	01	042	1	956	01		
0.052	0.000	0.000	0.0000	0.947	0.000	Algerian_clementine_279	Clem
082	01	01	1	878	01		
0.277	0.000	0.000	0.0000	0.722	0.000	Clemetine_X_Pearl_mandarin_hy	Clem_Hyb
59	01	01	1	37	01	brid_4020	
0.158	0.000	0.000	0.0000	0.841	0.000	Clemetine_X_Silverhill_mandarin	Clem_Hyb
262	01	01	1	698	01	_hybrid_3731	
0.609	0.000	0.000	0.0000	0.390	0.000	Duncan_grapefruit_3832	Gft
267	01	01	1	693	01		
0.623	0.000	0.000	0.0000	0.376	0.000	Royal_grapefruit_248	Gft
734	01	01	1	226	01		
0.593	0.000	0.000	0.0000	0.406	0.000	Citrus_otachibana_grapefruit_hyb	Gft_Hyb
911	01	01	16	043	01	rid_3470	
0.479	0.000	0.000	0.0000	0.520	0.000	Citrus_rugulosa_pummelo_hybrid	Gft_Hyb
571	01	01	19	38	01	_3556	
0.479	0.000	0.000	0.0000	0.520	0.000	Chironja_grapefruit_hybrid_3909	Gft_Hyb
467	01	01	1	493	01		
0.656	0.000	0.000	0.0000	0.343	0.000	Citrus_hiroshimana_Natsuzabon_	Gft_Hyb
781	01	01	1	179	01	3904	
0.580	0.000	0.000	0.0000	0.419	0.000	Rex_Union_grapefruit_hybrid_39	Gft_Hyb
232	01	01	1	728	01	62	
0.784	0.000	0.000	0.0000	0.215	0.000	Rojo_Blanco_pummelo_hybrid_4	Gft_Hyb
217	01	01	1	743	01	029	
0.524	0.000	0.000	0.0000	0.475	0.000	Sukega_grapefruit_hybrid_4149	Gft_Hyb
673	01	01	1	287	01		
0.608	0.000	0.000	0.0000	0.391	0.000	Triumph_grapefruit_297	Gft_Hyb
378	01	01	1	582	01		
0.580	0.000	0.000	0.0000	0.419	0.000	Valentine_pummelo_hybrid_4182	Gft_Hyb
618	01	01	1	342	01		
0.478	0.000	0.000	0.0000	0.521	0.000	Wheeny_grapefruit_hybrid_2885	Gft_Hyb
41	01	01	1	55	01		
0.000	0.000	0.000	0.7705	0.229	0.000	Fortunella_obovata_3475	Kumquat
01	02	01	55	396	01		
0.000	0.000	0.000	0.9999	0.000	0.000	Fortunella_polyandra_3901	Kumquat
01	01	01	5	01	01		
0.000	0.000	0.000	0.9999	0.000	0.000	Marumi_SRA489_kumquat_4125	Kumquat
01	01	01	5	01	01		
0.000	0.000	0.000	0.9999	0.000	0.000	Nordmann_Seedless_kumquat_40	Kumquat
01	013	01	47	01	01	70	
0.036	0.000	0.000	0.5440	0.419	0.000	Citrangquat_19_15_7_3259	Kumquat_Hyb
099	01	01	66	805	01		
0.000	0.119	0.000	0.5225	0.000	0.358	Eustis_limequat_3603	Kumquat_Hyb
01	069	01	14	01	387		
0.145	0.000	0.000	0.5262	0.328	0.000	Excalibur_red_lime_hybrid_4239	Kumquat_Hyb
542	01	01	59	169	01		
0.000	0.000	0.000	0.5111	0.488	0.000	Indio_mandarinquat_3759	Kumquat_Hyb
01	01	01	19	841	01		
0.000	0.047	0.000	0.8441	0.000	0.108	Procimequat_kumquat_hybrid_32	Kumquat_Hyb
01	22	01	06	01	644	95	
0.102	0.000	0.000	0.5234	0.373	0.000	Rio_Grande_Valley_lemonquat_4	Kumquat_Hyb
945	01	01	21	604	01	097	
0.011	0.000	0.175	0.5496	0.262	0.000	Sinton_citrangquat_3642	Kumquat_Hyb
931	01	532	5	868	01		
0.000	0.170	0.000	0.5321	0.000	0.297	Tavares_kumquat_hybrid_4048	Kumquat_Hyb
011	023	01	4	01	805		

0.019	0.000	0.212	0.5866	0.182	0.000	Thomasville_citrangquat_1440	Kumquat_Hyb
244	01	007	78	05	01		
0.000	0.119	0.000	0.5217	0.000	0.358	Warren_limequat_3069	Kumquat_Hyb
01	895	01	65	01	31		
0.171	0.000	0.000	0.0000	0.308	0.519	Citrus_kharna_3059	Lemon
851	01	01	1	704	415		
0.045	0.000	0.000	0.0000	0.238	0.715	Interdonato_lemon_hybrid_3593	Lemon
701	01	01	1	568	7		
0.337	0.000	0.000	0.0000	0.134	0.527	Iran_lemon_3162	Lemon
848	01	01	1	15	972		
0.223	0.000	0.000	0.0000	0.250	0.525	Khoubs_el_arsa_lemon_2477	Lemon
568	01	01	1	623	779		
0.223	0.000	0.000	0.0000	0.250	0.525	Khoubs_el_ersa_lemon_2489	Lemon
727	01	01	1	97	273		
0.187	0.000	0.000	0.0000	0.292	0.520	Kulu_lemon_3045	Lemon
641	01	01	1	287	042		
0.153	0.000	0.000	0.0000	0.335	0.510	Lapithiokii_lemon_4005	Lemon
195	01	01	1	805	97		
0.139	0.000	0.000	0.0000	0.327	0.533	Primofiore_lemon_3491	Lemon
389	01	01	1	306	275		
0.359	0.000	0.000	0.0000	0.120	0.520	Soh_Synteng_lemon_hybrid_3216	Lemon
019	01	01	1	215	735		
0.297	0.000	0.000	0.0000	0.183	0.518	Unnnamed_lemon_hybrid_3155	Lemon
848	01	01	1	953	169		
0.576	0.000	0.000	0.0000	0.024	0.398	Bitrouni_lemon_3265	Lemon_Hyb
804	01	01	1	447	718		
0.421	0.000	0.000	0.0000	0.578	0.000	Citrus_succosa_3280	Lemon_Hyb
227	01	01	1	733	01		
0.000	0.000	0.000	0.0000	0.517	0.482	Citrus_volkameriana_3050	Lemon_Hyb
012	01	01	1	894	065		
0.066	0.000	0.000	0.0000	0.429	0.503	Improved_Meyer_lemon_3737	Lemon_Hyb
933	01	01	1	144	893		
0.103	0.000	0.000	0.0000	0.383	0.512	Iraq_lemon_limetta_3492	Lemon_Hyb
785	01	01	1	784	401		
0.108	0.000	0.000	0.0000	0.363	0.528	Limonette_De_Marrakech_lemon	Lemon_Hyb
491	01	01	1	166	313	_hybrid_3989	
0.572	0.000	0.000	0.0000	0.138	0.289	Lumia_lemon_hybrid_3925	Lemon_Hyb
919	01	01	1	038	013		
0.181	0.000	0.000	0.0000	0.455	0.362	New_Zealand_lemonade_lemon_4	Lemon_Hyb
578	01	01	1	925	468	181	
0.291	0.000	0.000	0.0000	0.397	0.311	Orban_lemon_hybrid_4140	Lemon_Hyb
705	01	01	1	152	113		
0.117	0.000	0.000	0.0000	0.269	0.612	Peretta_lemon_hybrid_3924	Lemon_Hyb
464	01	01	1	698	808		
0.298	0.000	0.000	0.0000	0.183	0.517	Unnamed_lemon_hybrid_3154	Lemon_Hyb
989	01	01	1	354	627		
0.188	0.061	0.000	0.0000	0.284	0.465	Unnamed_lemon_3885	Lemon_Hyb
758	545	01	1	642	035		
0.000	0.483	0.000	0.0000	0.000	0.516	Citrus_excelsa_2316	Lime
01	888	01	1	01	072		
0.000	0.493	0.000	0.0000	0.000	0.506	Davao_papeda_hybrid_2427	Lime
01	825	01	1	01	135		
0.000	0.498	0.000	0.0000	0.000	0.501	Mexican_lime_1710	Lime
01	393	01	1	01	567		
0.069	0.335	0.000	0.0000	0.224	0.370	Lemonime_lemon_hybrid_2459	Lime_Hyb
779	348	01	1	495	359		
0.000	0.266	0.000	0.0000	0.000	0.733	Nicaraguan_lemon_3841	Lime_Hyb
01	606	012	1	01	353		

0.001	0.347	0.000	0.0000	0.103	0.547	Perrine_lime_hybrid_2458	Lime_Hyb
894	114	01	1	85	121		
0.086	0.000	0.197	0.0000	0.516	0.199	Rangpur_X_Troyer_hybrid_3997	Lime_Hyb
093	01	941	1	781	165		
0.000	0.463	0.000	0.0000	0.000	0.536	Soh_Long_lemon_hybrid_3199	Lime_Hyb
01	157	01	1	01	803		
0.000	0.000	0.000	0.0000	0.515	0.484	Mitha_Tulia_limetta_3051	limetta
011	01	01	1	921	038		
0.138	0.000	0.000	0.0000	0.328	0.533	Frost_Lisbon_lemon_3176	Lisbon_Lemon
56	01	01	1	127	282		
0.139	0.000	0.000	0.0000	0.326	0.534	Prior_Lisbon_lemon_3505	Lisbon_Lemon
095	01	01	1	562	312		
0.000	0.000	0.000	0.0000	0.999	0.000	Citrus_lycopersicaformis_3564	Mand_Type_1
01	01	01	1	95	01		
0.000	0.000	0.000	0.1152	0.884	0.000	Citrus_tachibana_3150	Mand_Type_1
01	01	01	95	665	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Sun_Chu_Sha_mandarin_4003	Mand_Type_1
01	01	01	1	95	01		
0.000	0.000	0.000	0.0092	0.990	0.000	Yellow_Rind_mandarin_3895	Mand_Type_1
01	01	01	7	69	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Baiju_nucellar_RRUT76	Mand_Hyb
01	01	011	1	949	01		
0.027	0.000	0.000	0.0000	0.972	0.000	Bower_mandarin_3649	Mand_Hyb
674	01	01	1	286	01		
0.000	0.478	0.000	0.0000	0.521	0.000	Citrus_amblycarpa_2485	Mand_Hyb
01	848	01	1	112	01		
0.000	0.000	0.000	0.0741	0.925	0.000	Citrus_depressa_2448	Mand_Hyb
01	01	01	53	807	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Citrus_erythrosa_3292	Mand_Hyb
01	01	01	1	95	01		
0.000	0.281	0.000	0.1165	0.000	0.602	Citrus_indica_3163	Mand_Hyb
01	304	01	29	01	137		
0.118	0.000	0.000	0.1217	0.760	0.000	Citrus_keraji_3144	Mand_Hyb
194	01	01	66	009	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Citrus_kinokuni_3887	Mand_Hyb
01	01	01	1	95	01		
0.057	0.000	0.000	0.1119	0.830	0.000	Citrus_leiocarpa_3147	Mand_Hyb
39	01	01	12	668	01		
0.000	0.000	0.000	0.3423	0.657	0.000	Citrus_nippokoreana_3228	Mand_Hyb
01	01	01	39	621	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Citrus_sunki_3143	Mand_Hyb
01	01	01	1	95	01		
0.150	0.000	0.000	0.0000	0.849	0.000	Citrus_yatsushiro_3880	Mand_Hyb
121	01	01	1	839	01		
0.042	0.000	0.000	0.3442	0.612	0.000	Calushu_calamondin_2867	Mand_Hyb
845	015	01	31	889	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Changsha_mandarin_3577	Mand_Hyb
01	01	01	1	95	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Cleopatra_mandarin_3844	Mand_Hyb
01	01	01	1	95	01		
0.145	0.000	0.000	0.0000	0.854	0.000	Daisy_mandarin_3910	Mand_Hyb
146	01	01	1	814	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Dancy_Weshart_602	Mand_Hyb
01	01	01	1	95	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Emerald_mandarin_4123	Mand_Hyb
01	01	01	1	95	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Empress_mandarin_3613	Mand_Hyb
01	01	01	1	95	01		

0.000	0.000	0.000	0.0000	0.999	0.000	Encore_mandarin_3569	Mand_Hyb
01	01	01	1	95	01		
0.075	0.000	0.000	0.0000	0.924	0.000	Fairchild_mandarin_3559	Mand_Hyb
774	01	01	1	186	01		
0.079	0.000	0.000	0.0000	0.920	0.000	Fallglo_mandarin_3990	Mand_Hyb
743	01	01	1	217	01		
0.096	0.000	0.000	0.0000	0.903	0.000	Fortune_mandarin_3560	Mand_Hyb
126	01	01	1	834	01		
0.045	0.000	0.000	0.0000	0.954	0.000	Fremont_mandarin_3558	Mand_Hyb
828	01	01	1	132	01		
0.028	0.000	0.000	0.0000	0.971	0.000	Frua_mandarin_3022	Mand_Hyb
442	01	01	1	518	01		
0.050	0.000	0.000	0.0000	0.949	0.000	Gold_Nugget_mandarin_3913	Mand_Hyb
547	01	01	1	413	01		
0.031	0.000	0.000	0.0000	0.968	0.000	Hansen_mandarin_4152	Mand_Hyb
119	01	01	1	841	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Hickson_mandarin_4161	Mand_Hyb
01	01	01	1	95	01		
0.004	0.000	0.000	0.0000	0.995	0.000	Honey_mandarin_3177	Mand_Hyb
483	01	01	1	477	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Huang_Yen_Man_Chie_3897	Mand_Hyb
01	01	01	1	95	01		
0.015	0.000	0.000	0.0018	0.982	0.000	Hung_Kat_mandarin_2331	Mand_Hyb
172	01	01	02	996	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Imperial_mandarin_4160	Mand_Hyb
011	01	01	1	949	01		
0.420	0.000	0.000	0.0000	0.579	0.000	Iyo_San_Ponkan_4118	Mand_Hyb
772	01	01	1	188	01		
0.083	0.000	0.000	0.0000	0.916	0.000	Kara_mandarin_3019	Mand_Hyb
741	01	01	1	219	01		
0.046	0.000	0.000	0.0000	0.953	0.000	Kincy_mandarin_4144	Mand_Hyb
184	01	01	1	776	01		
0.113	0.000	0.000	0.0000	0.886	0.000	King_mandarin_303	Mand_Hyb
353	01	01	1	607	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Kinnow_mandarin_3021	Mand_Hyb
01	01	01	1	95	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Kinokuni_mandarin_696	Mand_Hyb
01	01	01	1	95	01		
0.250	0.000	0.000	0.0000	0.749	0.000	Kunembo_mandarin_3346	Mand_Hyb
816	01	01	1	144	01		
0.047	0.000	0.000	0.0000	0.952	0.000	Laranja_Cravo_mandarin_2893	Mand_Hyb
501	01	01	1	459	01		
0.059	0.000	0.000	0.0000	0.940	0.000	Lee_mandarin_3851	Mand_Hyb
858	01	01	1	102	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Mandarinette_3405	Mand_Hyb
01	015	01	1	945	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Michal_mandarin_4049	Mand_Hyb
01	01	01	1	95	01		
0.039	0.000	0.000	0.0000	0.960	0.000	Murcott_mandarin_3846	Mand_Hyb
823	01	01	1	137	01		
0.166	0.000	0.000	0.0000	0.833	0.000	Necked_mandarin_3738	Mand_Hyb
497	01	01	1	463	01		
0.065	0.000	0.000	0.0000	0.934	0.000	NISSV_E_mandarin_hybrid_3965	Mand_Hyb
339	01	01	1	621	01		
0.144	0.000	0.000	0.0000	0.855	0.000	Nova_mandarin_3615	Mand_Hyb
518	01	01	1	442	01		
0.189	0.000	0.000	0.0000	0.810	0.000	Novelty_X_Ellendale_mandarin_hybrid_3956	Mand_Hyb
099	01	01	1	861	01		

0.160	0.000	0.000	0.0000	0.838	0.000	Page_mandarin_3616	Mand_Hyb
998	01	01	11	961	01		
0.097	0.000	0.000	0.0000	0.902	0.000	palazzelli_mandarin_4244	Mand_Hyb
622	01	01	1	338	01		
0.089	0.000	0.000	0.0000	0.910	0.000	Parson's_Special_mandarin_300	Mand_Hyb
032	01	01	1	928	01		
0.047	0.000	0.000	0.0000	0.952	0.000	Pixie_mandarin_3568	Mand_Hyb
429	01	01	1	531	01		
0.013	0.000	0.000	0.0000	0.986	0.000	Richard's_Special_mandarin_4117	Mand_Hyb
027	01	01	1	933	01		
0.054	0.000	0.000	0.0000	0.945	0.000	Robinson_mandarin_3850	Mand_Hyb
056	01	01	1	904	01		
0.029	0.000	0.000	0.0000	0.970	0.000	Rubidoux_mandarin_4031	Mand_Hyb
673	01	01	1	287	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Scarlet_Emperor_mandarin_3226	Mand_Hyb
01	01	01	1	95	01		
0.148	0.000	0.000	0.0000	0.851	0.000	Shiranui_Dekopan_4113	Mand_Hyb
046	01	01	1	914	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Solid_Scarlet_mandarin_3328	Mand_Hyb
01	01	01	1	95	01		
0.033	0.000	0.000	0.0000	0.966	0.000	Som_Kaeo_mandarin_3852	Mand_Hyb
18	01	01	1	78	01		
0.083	0.000	0.000	0.0000	0.916	0.000	US_Furr_SRA337_mandarin_423	Mand_Hyb
38	01	01	1	58	01	8	
0.066	0.000	0.000	0.0000	0.933	0.000	Sunburst_mandarin_3809	Mand_Hyb
7	01	01	1	26	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Szinkom_mandarin_3085	Mand_Hyb
01	01	01	1	95	01		
0.001	0.000	0.000	0.0000	0.998	0.000	Tango_mandarin_4183	Mand_Hyb
492	01	01	1	468	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Tien_Chieh_mandarin_2590	Mand_Hyb
01	01	01	1	95	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Tieu_nucellar_mandarin_RRUT87	Mand_Hyb
01	01	01	1	95	01		
0.000	0.370	0.000	0.0000	0.629	0.000	Indian_sour_orange_hybrid_3175	Mand_Hyb
01	393	01	1	567	01		
0.101	0.000	0.000	0.0000	0.898	0.000	USDA_6-15-150_mandarin_4165	Mand_Hyb
596	01	01	1	364	01		
0.079	0.000	0.000	0.0000	0.920	0.000	US_Superna_USDA_88_2_manda	Mand_Hyb
738	01	01	1	222	01	rin_hybrid_3991	
0.000	0.000	0.000	0.0000	0.999	0.000	USDA_88_3_mandarin_hybrid_3	Mand_Hyb
01	01	01	1	95	01	992	
0.012	0.000	0.000	0.0000	0.987	0.000	Wiling_mandarin_3020	Mand_Hyb
15	01	01	1	81	01		
0.000	0.000	0.000	0.0000	0.999	0.000	Willowleaf_mandarin_3843	Mand_Hyb
01	012	01	14	944	01		
0.353	0.000	0.000	0.0000	0.646	0.000	Citrus_ujukitsu_3467	Mand_Hyb_pum
546	011	01	1	413	01		
0.000	0.000	0.000	0.0000	0.508	0.491	Phillipine_red_Rangpur_2318	Mand_Hyb_Rangpur
01	01	01	1	256	704		
0.000	0.000	0.000	0.4456	0.554	0.000	Citrus_hanayu_3469	Mand-Ichangensis_Hyb
01	01	01	32	328	01		
0.357	0.022	0.000	0.0000	0.025	0.594	Citrus_megaloxycarpa_hybrid_32	Megaloxycarpa
146	869	01	1	614	351	41	
0.000	0.000	0.000	0.9999	0.000	0.000	Microcitrus_australasica_3670	Microcitrus
01	01	01	5	01	01		
0.000	0.000	0.000	0.9999	0.000	0.000	Microcitrus_australasica_3672	Microcitrus
01	01	01	5	01	01		

0.000	0.000	0.000	0.9999	0.000	0.000	Microcitrus_australis_3665	Microcitrus
01	01	01	5	01	01		
0.000	0.000	0.000	0.9999	0.000	0.000	Sydney_hybrid_1485	Microcitrus_Hyb
01	01	01	5	01	01		
0.329	0.000	0.000	0.0000	0.670	0.000	Parent_Washington_navel_1241A	Navel
777	01	01	1	183	01		
0.000	0.999	0.000	0.0000	0.000	0.000	Citrus_hystrix_3103	Papeda_micrantha
015	945	01	1	01	01		
0.000	0.999	0.000	0.0000	0.000	0.000	Citrus_micrantha_3605	Papeda_micrantha
01	947	01	13	01	01		
0.000	0.999	0.000	0.0000	0.000	0.000	Cabuyo_papeda_432	Papeda_micrantha
01	95	01	1	01	01		
0.000	0.999	0.000	0.0000	0.000	0.000	Citrus_webberri_Kalpi_1456	Papeda_micrantha
01	947	013	1	01	01		
0.000	0.087	0.000	0.9129	0.000	0.000	Citrus_ichangensis_2327	Papeda_Ichangensis
01	044	01	16	01	01		
0.000	0.077	0.002	0.9194	0.000	0.000	Citrus_ichangensis_3931	Papeda_Ichangensis
01	777	742	51	01	01		
0.543	0.244	0.000	0.2115	0.000	0.000	Citrus_hongheesis_3797	Papeda_ichangensis_pu
677	786	01	01	016	01		mmelo_hybrid
0.501	0.271	0.000	0.2263	0.000	0.000	Citrus_latipes_Khasi_3052	Papeda_ichangensis_pu
764	838	01	68	01	01		mmelo_hybrid
0.122	0.000	0.000	0.3270	0.550	0.000	Kabosu_papeda_hybrid_3943	Papeda_ichangensis_tac
303	01	01	69	598	01		hibana_hybrid
0.000	0.000	0.000	0.5513	0.448	0.000	Yuzu_ichandarin_1216	Papeda_ichangensis_tac
01	01	01	42	618	01		hibana_hybrid
0.000	0.490	0.000	0.0000	0.000	0.509	Citrus_macrophylla_papeda_hybri	Papeda_Hyb
01	338	01	1	01	622	d_2446	
0.000	0.446	0.000	0.0000	0.553	0.000	Citrus_webberii_var_Montana_76	Papeda_Hyb
01	92	017	1	032	01	7	
0.148	0.000	0.000	0.1766	0.675	0.000	Citrus_yuko_3146	Papeda_Hyb
072	01	01	87	211	01		
0.514	0.000	0.000	0.3378	0.147	0.000	Ichang_lemon_1215	Papeda_Hyb
336	01	01	78	756	01		
0.000	0.534	0.000	0.0000	0.465	0.000	Kulobot_papeda_hybrid_3612	Papeda_Hyb
017	514	01	1	44	01		
0.245	0.115	0.000	0.0889	0.549	0.000	Soh_Niangrang_papeda_3203	Papeda_Hyb
668	756	01	15	641	01		
0.499	0.271	0.000	0.2291	0.000	0.000	Unnamed_papeda_3056	Papeda_Hyb
153	695	01	22	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	African_pummelo_2346	Pum
946	01	01	1	014	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Alemoen_pummelo_2487	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	BuOl_1_acidless_yellow_RRUT9	Pum
95	01	01	1	01	01	2	
0.999	0.000	0.000	0.0000	0.000	0.000	Citrus_grandis_2340	Pum
944	016	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Citrus_grandis_2453	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Citrus_grandis_3282	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Chinese_pummelo_1224	Pum
944	016	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Deep_Red_pummelo_2347	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Fleming's_Shaddock_pummelo_5	Pum
95	01	01	1	01	01	78	

0.999	0.000	0.000	0.0000	0.000	0.000	Hirado_Buntan_pummelo_4122	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Itoshima_3928	Pum
938	013	013	16	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_Phuang_pummelo_2352	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_Phuang_pummelo_3926	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_Ruan_Tia_pummelo_2350	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_Ruan_Tia_pummelo_2351	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Karn_Lau_Yau_pummelo_2341	Pum
944	016	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Mato_Buntan_pummelo_3945	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Nakhon_Chaisi_pummelo_2353	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pan_Dan_pummelo_2752	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pauthel_pummelo_3949	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Philippine_pummelo_hybrid_644	Pum
948	011	012	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pin_Shan_Kong_Yan_pummelo_2	Pum
95	01	01	1	01	01	348	
0.999	0.000	0.000	0.0000	0.000	0.000	Pink_pummelo_2246	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pomelit_pummelo_hybrid_4026	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Reinking_pummelo_3805	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Roeding's_Pink_pummelo_1208	Pum
948	01	01	12	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Siamese_Acidless_pummelo_224	Pum
95	01	01	1	01	01	0	
0.999	0.000	0.000	0.0000	0.000	0.000	Sunshine_pummelo_2236	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Sweet_pummelo_3067	Pum
944	016	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Tahitian_pummelo_3806	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Tau_Yau_pummelo_2583	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Thong_Deep_pummelo_3927	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Thong_Deep_pummelo_4027	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Unnamed_Chinese_RRUT223	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_pan_pummelo_2248	Pum
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Kao_Pan_pummelo_2355	Pum
936	016	01	14	014	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Webber_shaddock_pummelo_423	Pum
95	01	01	1	01	01	3	
0.445	0.000	0.000	0.0000	0.059	0.494	Cuban_shaddock_pummelo_hybri	Pum_Citron_Hyb
254	01	01	1	978	737	d_1462	

0.556	0.000	0.000	0.0000	0.151	0.292	Lemelo_pummelo_hybrid_1481	Pum_Citron_Hyb
045	01	01	1	393	532		
0.778	0.000	0.000	0.0000	0.026	0.194	Lemelo_pummelo_hybrid_1775	Pum_Citron_Hyb
423	01	01	1	77	776		
0.574	0.000	0.000	0.0000	0.124	0.300	Moli_kurikuri_pummelo_hybrid_	Pum_Citron_Hyb
285	01	01	1	975	71	579	
0.512	0.000	0.000	0.0000	0.000	0.487	Ponderosa_lemon_hybrid_294	Pum_Citron_Hyb
633	01	01	1	01	327		
0.904	0.033	0.000	0.0000	0.000	0.062	Sour_pummelo_3066	Pum_Citron_Hyb
339	288	01	1	01	343		
0.768	0.000	0.000	0.0000	0.230	0.000	Anseikan_pummelo_hybrid_3979	Pum_Hyb
982	01	01	1	978	01		
0.742	0.000	0.000	0.0000	0.256	0.000	Arajon_pummelo_2596	Pum_Hyb
965	01	01	1	995	01		
0.527	0.000	0.000	0.0000	0.472	0.000	Asahikan_pummelo_hybrid_3905	Pum_Hyb
062	01	01	1	897	01		
0.561	0.000	0.000	0.0000	0.438	0.000	Banokan_pummelo_3950	Pum_Hyb
042	01	01	1	918	01		
0.552	0.000	0.000	0.0000	0.447	0.000	Hassaku_Beni_pummelo_hybrid_	Pum_Hyb
401	01	01	1	559	01	3907	
0.812	0.000	0.000	0.0000	0.187	0.000	Citrus_hiroshimana_Natsuzabon_	Pum_Hyb
815	01	01	1	145	01	3275	
0.542	0.000	0.000	0.0000	0.456	0.000	Citrus_obovoidea_3465	Pum_Hyb
998	01	01	1	962	01		
0.376	0.000	0.000	0.0000	0.623	0.000	Citrus_shunkokan_3476	Pum_Hyb
019	01	01	1	941	01		
0.786	0.000	0.000	0.0000	0.213	0.000	Citrus_sino-grandis_3148	Pum_Hyb
194	01	01	1	766	01		
0.428	0.000	0.000	0.0000	0.571	0.000	Citrus_tengu_3464	Pum_Hyb
031	01	01	1	929	01		
0.787	0.000	0.000	0.0000	0.212	0.000	Chinese_pummelo_1212	Pum_Hyb
218	01	01	1	742	01		
0.485	0.000	0.000	0.0000	0.514	0.000	Cocktail_pummelo_hybrid_3555	Pum_Hyb
321	01	01	1	639	01		
0.653	0.000	0.000	0.0000	0.346	0.000	Kawachi_Bankan_3944	Pum_Hyb
228	01	01	1	732	01		
0.411	0.000	0.000	0.0000	0.588	0.000	Kinokawa_Buntan_pummelo_397	Pum_Hyb
93	01	01	1	03	01	8	
0.461	0.000	0.000	0.0000	0.538	0.000	Philippine_pummelo_hybrid_2343	Pum_Hyb
603	01	01	1	357	01		
0.789	0.000	0.000	0.0000	0.210	0.000	Puma_pummelo_hybrid_3941	Pum_Hyb
603	01	01	1	357	01		
0.784	0.000	0.000	0.0000	0.215	0.000	Rubidoux_pummelo_hybrid_#1_4	Pum_Hyb
259	01	01	1	701	01	028	
0.603	0.000	0.000	0.0000	0.396	0.000	Shaddock_X_St_michael_orange_	Pum_Hyb
041	01	01	1	919	01	hybrid_42	
0.815	0.000	0.000	0.0000	0.000	0.184	Siamese_pummelo_2421	Pum_Hyb
957	01	01	1	01	003		
0.865	0.000	0.000	0.0000	0.133	0.000	Suisho_Buntan_pummelo_3947	Pum_Hyb
973	01	01	1	987	01		
0.671	0.000	0.000	0.0000	0.328	0.000	Tresca_pummelo_302	Pum_Hyb
527	01	01	1	433	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Banpeiyu_pummelo_3961	Pum-Allele-F-admixed
947	013	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Chandler_pummelo_3244	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Egami_Buntan_pummelo_3959	Pum-Allele-F-admixed
95	01	01	1	01	01		

0.999	0.000	0.000	0.0000	0.000	0.000	Hunan_pummelo_1225	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Moanalua_pummelo_hybrid_448	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pink_pummelo_2244	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Pong_yau_pummelo_2342	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.999	0.000	0.000	0.0000	0.000	0.000	Red_pummelo_2245	Pum-Allele-F-admixed
95	01	01	1	01	01		
0.000	0.000	0.000	0.0000	0.510	0.489	Santa_Barbara_Red_Rangpur_712	Rangpur
01	01	01	1	304	656		
0.000	0.000	0.000	0.0000	0.508	0.491	Sniff_acidless_Rangpur_2484	Rangpur
01	01	01	1	049	911		
0.000	0.000	0.000	0.0000	0.515	0.484	India_lemon_2323	Rough_lemon
01	011	01	1	549	41		
0.000	0.000	0.000	0.0000	0.518	0.481	Indian_rough_lemon_2544	Rough_lemon
01	01	01	1	896	064		
0.000	0.000	0.000	0.0000	0.520	0.479	Jullander_Khatti_rough_lemon_41	Rough_Lemon
01	01	01	1	301	659	73	
0.000	0.000	0.000	0.0000	0.519	0.480	Schaub_rough_lemon_3879	Rough_Lemon
012	01	01	1	739	219		
0.000	0.000	0.000	0.0000	0.517	0.482	South_African_rough_lemon_232	Rough_Lemon
01	01	01	1	512	448	5	
0.000	0.000	0.000	0.0000	0.518	0.481	Vangasay_807_rough_lemon_399	Rough_Lemon
01	01	01	1	48	48	6	
0.115	0.000	0.000	0.0000	0.659	0.224	Milam_rough_lemon_3396	Rough_lemon_Hyb
726	01	01	1	246	998		
0.150	0.000	0.000	0.0000	0.849	0.000	Frost_Owari_satsuma_3178	Satsuma
057	01	01	1	903	01		
0.428	0.000	0.000	0.0000	0.571	0.000	Mytrifoliataia_sour_orange_3290	Sour_Orange
714	01	01	1	246	01		
0.422	0.000	0.000	0.0150	0.562	0.000	Nansho_Daidai_sour_orange_258	Sour_Orange
33	013	01	53	583	01	8	
0.462	0.000	0.000	0.0000	0.537	0.000	Smooth_Flat_Seville_sour_orange	Sour_Orange
093	01	01	1	867	01	_4079	
0.522	0.000	0.000	0.0036	0.473	0.000	Citrus_aurantium_var_salicifolia_	Sour_Orange
705	01	01	99	567	01	3289	
0.442	0.000	0.000	0.0000	0.557	0.000	Standard_sour_orange_628	Sour_Orange
688	01	01	1	272	01		
0.397	0.000	0.000	0.0000	0.602	0.000	Zhuluan_sour_orange_3930	Sour_Orange
046	01	01	1	914	01		
0.510	0.000	0.000	0.0000	0.489	0.000	Citrus_intermedia_3474	Sour_Orange_Hyb
907	01	01	1	053	01		
0.000	0.449	0.000	0.0000	0.550	0.000	Citrus_maderaspatana_3225	Sour_Orange_Hyb
01	315	01	1	645	01		
0.466	0.000	0.000	0.0000	0.533	0.000	Citrus_natsudaiddai_3235	Sour_Orange_Hyb
784	011	01	1	175	01		
0.209	0.000	0.000	0.1334	0.656	0.000	Citrus_rokugatsu_3473	Sour_Orange_Hyb
61	01	01	6	9	01		
0.433	0.000	0.000	0.0164	0.549	0.000	Goutoucheng_sour_orange_4004	Sour_Orange_Hyb
056	552	011	84	887	01		
0.559	0.000	0.000	0.0000	0.440	0.000	Orogold_sour_orange_hybrid_693	Sour_Orange_Hyb
37	01	01	1	59	01		
0.000	0.000	0.000	0.0494	0.687	0.262	Pursha_lime_x_Chinotto_3578	Sour_Orange_Hyb
01	01	01	2	609	942		
0.362	0.000	0.000	0.0000	0.313	0.324	Bergamot_sour_orange_hybrid_28	Sour_Orange_Hyb_citr
171	01	01	1	42	379	81	on

0.684	0.000	0.000	0.0717	0.148	0.094	Citrus_canaliculata_3565	Sour_Orange_Hyb_citr
896	01	01	56	435	893		on
0.224	0.000	0.000	0.0283	0.565	0.181	Chinotto_sour_orange_hybrid_37	Sour_Orange_Hyb_citr
432	01	01		663	585	28	on
0.100	0.000	0.000	0.0000	0.383	0.515	Bahman_Persian#1_Sweet_lime_4	Sw_Lemon
708	01	01	1	404	858	174	
0.100	0.000	0.000	0.0000	0.382	0.516	Bahman_Persian#2_Sweet_lime_4	Sw_Lemon
914	01	01	1	215	841	187	
0.111	0.000	0.000	0.0000	0.363	0.525	Pomona_sweet_lemon_4068	Sw_Lemon
119	01	01	1	501	35		
0.101	0.000	0.000	0.0000	0.382	0.516	Sweet_lemon_3093	Sw_Lemon
206	01	01	1	699	065		
0.336	0.000	0.000	0.0000	0.663	0.000	Ambersweet_sweet_orange_4002	Swo
311	01	01	1	649	01		
0.245	0.000	0.000	0.0000	0.445	0.308	Unnamed_sweet_orange_hybrid_6	Swo_Hyb
882	01	01	1	132	956	48	
0.286	0.000	0.299	0.0000	0.413	0.000	US_119_3998	Swo_Hyb
919	01	495	1	556	01		
0.323	0.000	0.000	0.0000	0.676	0.000	Allspice_tangelo_4143	Tangelo
786	01	01	1	174	01		
0.255	0.000	0.000	0.0000	0.744	0.000	Baygold_tangelo_4186	Tangelo
91	011	01	1	049	01		
0.343	0.000	0.000	0.0000	0.656	0.000	Citrus_funadoko_3274	Tangelo
714	01	01	1	246	01		
0.375	0.000	0.000	0.0000	0.624	0.000	Clement_tangelo_2603	Tangelo
352	01	01	1	608	01		
0.242	0.000	0.000	0.0000	0.757	0.000	Early_tangelo_2560	Tangelo
355	01	01	1	605	01		
0.226	0.000	0.000	0.0000	0.773	0.000	Mapo_tangelo_4222	Tangelo
435	01	01	1	525	01		
0.300	0.000	0.000	0.0000	0.699	0.000	Marlow_tangelo_4145	Tangelo
222	01	01	1	738	01		
0.179	0.000	0.000	0.0000	0.820	0.000	Minneola_tangelo_3340	Tangelo
87	01	01	1	09	01		
0.207	0.000	0.000	0.0000	0.792	0.000	Orlando_tangelo_3874	Tangelo
736	01	01	1	224	01		
0.222	0.000	0.000	0.0000	0.777	0.000	Ortanique_tangor_3368	Tangelo
544	01	01	1	416	01		
0.348	0.000	0.000	0.0000	0.651	0.000	Pearl_tangelo_2849	Tangelo
479	01	01	1	481	01		
0.253	0.000	0.000	0.0000	0.746	0.000	Sacaton_tangelo_2609	Tangelo
185	01	01	1	775	01		
0.471	0.000	0.000	0.0000	0.528	0.000	Sacaton_tangelo_3331	Tangelo
144	01	01	1	816	01		
0.240	0.000	0.000	0.0000	0.759	0.000	Sampson_tangelo_2418	Tangelo
875	01	01	1	085	01		
0.189	0.000	0.000	0.0000	0.810	0.000	San_Jacinto_tangelo_2011	Tangelo
676	01	01	1	284	01		
0.353	0.000	0.000	0.0000	0.646	0.000	Santa_Barbara_tangelo_4073	Tangelo
719	01	01	1	241	01		
0.224	0.000	0.000	0.0000	0.775	0.000	Seminole_tangelo_2604	Tangelo
534	01	01	1	426	01		
0.555	0.000	0.000	0.0000	0.444	0.000	Sexton_tangelo_2543	Tangelo
001	01	01	1	959	01		
0.207	0.000	0.000	0.0000	0.791	0.000	Siamelo_tangelo_2586	Tangelo
966	01	01	1	994	01		
0.238	0.000	0.000	0.0000	0.761	0.000	Sunrise_tangelo_2606	Tangelo
207	01	01	1	753	01		

0.275	0.000	0.000	0.0000	0.724	0.000	Sunshine_tangelo_2788	Tangelo
421	01	01	1	539	01		
0.212	0.000	0.000	0.0000	0.787	0.000	Thorton_tangelo_2013	Tangelo
643	01	01	1	317	01		
0.575	0.000	0.000	0.0000	0.424	0.000	UC_Seedless_Surprise_USDA_1_	Tangelo
067	01	01	1	893	01	77_19_tangelo_4208	
0.301	0.000	0.000	0.0000	0.698	0.000	Webber_tangelo_2746	Tangelo
609	01	01	1	351	01		
0.240	0.000	0.000	0.0000	0.759	0.000	Willial_tangelo_4147	Tangelo
582	01	01	1	378	01		
0.274	0.000	0.000	0.0000	0.725	0.000	Williams_tangelo_2787	Tangelo
773	01	01	1	187	01		
0.206	0.000	0.000	0.0000	0.793	0.000	Yalaha_tangelo_2559	Tangelo
612	01	01	1	348	01		
0.203	0.000	0.000	0.0000	0.796	0.000	Altoona_tangor_2792	Tangor
485	01	01	1	475	01		
0.227	0.001	0.000	0.0053	0.766	0.000	Citrus_Benikoji_3149	Tangor
282	203	01	36	159	01		
0.197	0.000	0.000	0.0000	0.802	0.000	Dweet_tangor_3018	Tangor
713	01	01	1	247	01		
0.163	0.000	0.000	0.0000	0.836	0.000	H_56_tangor_3096	Tangor
792	01	01	1	168	01		
0.315	0.000	0.000	0.0000	0.684	0.000	Iyomikan_tangor_3980	Tangor
664	01	01	1	296	01		
0.327	0.000	0.000	0.0000	0.672	0.000	Ugli_tangelo_4127	Tangor
083	01	01	1	877	01		
0.277	0.000	0.000	0.0000	0.722	0.000	Kiyomi_tangor_3952	Tangor
22	01	01	1	739	01		
0.062	0.000	0.000	0.0000	0.937	0.000	Koster_tangor_3958	Tangor
668	01	01	1	292	01		
0.070	0.000	0.000	0.0000	0.929	0.000	Mency_tangor_3183	Tangor
27	01	012	1	688	01		
0.214	0.000	0.000	0.0000	0.785	0.000	Ruddy_tangor_4146	Tangor
076	01	01	1	884	01		
0.100	0.000	0.000	0.0000	0.899	0.000	Tankan_tangor_2224	Tangor
564	01	01	1	396	01		
0.203	0.000	0.000	0.0000	0.796	0.000	Temple_tangor_2598	Tangor
797	017	01	1	157	01		
0.117	0.000	0.000	0.0000	0.882	0.000	Wincia_tangor_4148	Tangor
428	01	01	1	532	01		
0.000	0.000	0.999	0.0000	0.000	0.000	#26_trifoliate_3939	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	#27_trifoliate_3938	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	Argentine_trifoliate_3206	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	Flying_dragon_trifoliate_3330A	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	Monoembryonic_trifoliate_3888	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	Simmons_trifoliate_3549	Trifol
01	01	95	1	01	01		
0.000	0.000	0.999	0.0000	0.000	0.000	Small_Leaf_4017	Trifol
01	01	95	1	01	01		
0.174	0.000	0.511	0.0000	0.314	0.000	Citradia_trifoliate_hybrid_1437	Trifol_Hyb
151	01	49	13	326	01		
0.128	0.000	0.505	0.0000	0.366	0.000	Citradia_trifoliate_hybrid_1438	Trifol_Hyb
081	01	562	1	327	01		

0.000	0.000	0.468	0.2542	0.277	0.000	Citraldin_trifoliolate_hybrid_2629	Trifol_Hyb
01	01	182	23	566	01		
0.000	0.000	0.490	0.0000	0.509	0.000	Citradarin_trifoliolate_hybrid_2618	Trifol_Hyb
01	01	632	1	328	01		
0.209	0.000	0.535	0.0000	0.254	0.000	Citrangor_trifoliolate_hybrid_1447	Trifol_Hyb
892	01	333	1	746	01		
0.079	0.000	0.535	0.0000	0.164	0.220	Citremon_trifoliolate_hybrid_1448	Trifol_Hyb
786	01	437	1	326	431		
0.099	0.000	0.534	0.0000	0.232	0.133	Citremon_trifoliolate_hybrid_1449	Trifol_Hyb
339	01	565	1	961	116		
0.000	0.000	0.264	0.3434	0.392	0.000	Glen_citrangedin_trifoliolate_hybrid	Trifol_Hyb
01	01	095	36	44	01	_3573	
0.000	0.000	0.488	0.5111	0.000	0.000	Microcitrus_X_trifoliolate_3566	Trifol_Hyb
01	01	858	02	01	01		
0.355	0.000	0.296	0.0000	0.348	0.000	S281_citrangelo_3552	Trifol_Hyb
12	01	678	1	171	01		
0.300	0.000	0.506	0.0000	0.193	0.000	S302_citranguma_3415	Trifol_Hyb
113	01	677	1	181	01		
0.132	0.000	0.506	0.0000	0.360	0.000	Trifeola_trifoliolate_hybrid_3954	Trifol_Hyb
629	01	777	1	564	01		

Table 6. Proportion of genomic composition for F1 hybrid
Average, mean and standard deviation for the proportion of genomic composition observed for known F1 hybrids (26 accessions) were calculated after ADMIXTURE results of suspected F1 hybrids were compared to results obtained by chromosomal level introgression analysis.

Statistics	Genomic composition 1	Genomic composition 2
Average	0.478322	0.521638
Median	0.481964	0.517996
Stdv	0.015336	0.015335

Table 7. ID for the accessions in the 382 UCR diversity panel used in phylogeny network analysis

ID	Sample Name CRC
1	African pummelo 2346
2	African Shaddock x Rubidoux 3969
3	Alemoen pummelo 2487
4	Algerian clementine 279
5	Allspice tangelo 4143
6	Altoona tangor 2792
7	Amberswee sweet orange hybrid 4002
8	Amoa 8 blood mandarin 4237
9	Argentine trifoliolate 3206
10	Asahikan pummelo hybrid 3905
11	Assads citron RSD2008001
12	Badhri citron 3723
13	Bahman Persian 1 sweet lime 4174
14	Bahman Persian 2 sweet lime 4187
15	Baiju nucellar RRUT76
16	Banokan pummelo 3950
17	Banpeiyu pummelo 3961
18	Baygold tangelo 4186
19	Bengal citron 3055
20	Hassaku Beni pummelo hybrid 3907
21	Benton citrange 3908
22	Bitrouni lemon 3265
23	Bitters C22 4219
24	Bower mandarin 3649
25	Buddha's Hand citron 3768
26	BuOl 1 acidless yellow pummelo RRUT92
27	C. grandis 3282
28	C. latipes 3052
29	C. amblycarpa 2485
30	C. benikoji 3149
31	C. depressa 2448
32	C. erythrosa 3292
33	C. funadoko3274
34	C. grandis 2340
35	C. grandis 2453
36	C. hiroshimana Natsuzabon 3275
37	C. hongheesis 3797

38	<i>C. hystrix</i> 3103
39	<i>C. ichangensis</i> 2327
40	<i>C. intermedia</i> 3474
41	<i>C. kinokuni</i> 3887
42	<i>C. leiocarpa</i> 3147
43	<i>C. lycopersicaformis</i> 3564
44	<i>C. maderaspatana</i> papeda hybrid 3225
45	<i>C. natsudaikai</i> 3235
46	<i>C. Nippokoreana</i> 3228
47	<i>C. obovoidae</i> 3465
48	<i>C. otachibana</i> 3470
49	<i>C. rokugatsu</i> 3473
50	<i>C. rugulosa</i> 3556
51	<i>C. shunkokan</i> 3476
52	<i>C. tachibana</i> 3150
53	<i>C. tengu</i> 3464
54	<i>C. ujukitsu</i> 3467
55	<i>C. volkameriana</i> 3050
56	<i>C. yatsushiro</i> 3880
57	<i>C. yuko</i> 3146
58	C32 citrange 3911
59	C35 citrange 3912
60	Cabuyo Papeda 432
61	Calushu calmondin 2867
62	Chandler pummelo 3244
63	Changsha mandarin 3577
64	Chinese pummelo 1224
65	Chinotto sour orange 3728
66	Chironja grapefruit hybrid 3909
67	Citradarin trifoliolate hybrid 2618
68	Citrardia trifoliolate hybrid 1437
69	Citrardia trifoliolate hybrid 1438
70	Citraldin trifoliolate hybrid 2629
71	Citrange #1416 citrange 2866
72	Citrangquat 19 15 7 3259
73	Citrangor trifoliolate hybrid 1447
74	Citremon trifoliolate hybrid 1448
75	Citremon trifoliolate hybrid 1449
76	Citron of Commerce 3518
77	Citrumelo 1452

78	Citrumelo 3348
79	C. macrophylla papeda hybrid 2446
80	C. vulgaris sour orange hybrid 760
81	Clement tangelo 2603
82	Clementine x Pearl mandarin hybrid 4020
83	Clementine x Silverhill mandarin hybrid 3731
84	Cleopatra mandarin 3844
85	Cocktail pummelo hybrid 3555
86	Corsican citron 1795
87	C. sunki mandarin 3143
88	Cuban Shaddock pummelo hybrid 1462
89	Cunningham citrange 271
90	Dabbe citron hybrid 3190
91	Daisy mandarin 3910
92	Dancy Weshart 602
93	Davao papeda hybrid 2427
94	Djerok citron hybrid 2456
95	Dulcia citron 3654
96	Dweet tangor 3018
97	Early tangelo 2560
98	Empress mandarin 3613
99	Encore mandarin 3569
100	Etrog citron 3891
101	Eustis Limequat 3603
102	Excalibur red lime hybrid 4239
103	Fairchild mandarin 3559
104	Fallglo mandarin 3990
105	Fleming's Shaddock pummelo 578
106	Fortune mandarin 3560
107	Fortunella polyandra 3901
108	Fremont mandarin 3558
109	Frost Lisbon lemon 3176
110	Frua mandarin 3022
111	Furr C57 citrange 4221
112	Glen citrangedin trifoliolate hybrid 3573
113	Gold Nugget mandarin 3913
114	Goutoucheng sour orange hybrid 4004
115	GouTouXiangYuan RRUT 395
116	GouTouXiangYuan RRUT 396
117	GouTouXiangYuan RRUT 398

118	H 56 tangor 3096
119	Hall grapefruit X Rubidoux 3889
120	Hansen mandarin 4152
121	Hickson mandarin 4161
122	Citrus hiroschimana Natsuzabon 3904
123	Honey mandarin 3177
124	Huang Yen Man Chie mandarin 3897
125	Hung Kat mandarin 2331
126	Hunan pummelo 1225
127	Imperial mandarin 4160
128	India lemon 2323
129	Indian citron 138
130	Indian citron 661
131	Indian rough lemon 2544
132	Indio mandarinat 3759
133	Interdonato lemon hybrid 3593
134	Iran lemon 3162
135	Italian citron 3530
136	Itoshima 3928
137	Iyo San Ponkan 4118
138	Iyomikan tangor 3980
139	uUgli 4127
140	Jullander Khatti rough lemon 4173
141	Kao Phuang pummelo 3926
142	Kao Ruan Tia pummelo 2350
143	Kao Ruan Tia pummelo 2351
144	Kara mandarin 3019
145	Karn Lau Yau pummelo 2341
146	Kawachi Bankan 3944
147	Khoubs El Arsa lemon 2477
148	Khoubs El Ersal lemon 2489
149	King mandarin 303
150	Kinokawa Buntan pummelo 3978
151	Kinokuni mandarin 696
152	Koster tangor 3958
153	kulobot papeda hybrid 3612
154	Kulu lemon 3045
155	Kunembo mandarin 3346
156	Lapithiokii lemon 4005
157	Laranja Cravo mandarin 2983

158	Lee mandarin 3851
159	Lemelo pummelo hybrid 1481
160	Lemelo pummelo hybrid 1775
161	lemonime lemon hybrid 2459
162	Limau Matu 3755
163	Limonette De Marrakech lemon hybrid 3989
164	Lumina lemon 3925
165	Mapo tangelo 4222
166	Marlow tangelo 4145
167	Marumi SRA489 kumquat 4125
168	Mency tangor 3183
169	Mexican citron 3531
170	Michal mandarin 4049
171	Microcitrus x trifoliae 3566
172	Mitha Tulia limetta 3051
173	Moanalua pummelo hybrid 448
174	Moli kurikuri pummelo hybrid 579
175	Morningsong Temoni citron-Karp
176	Morton citrange 1463
177	Murcott mandarin 3846
178	Mytrifolia sour orange 3290
179	Nakhon Chaisi pummelo 2353
180	Nansho Daidai sour orange 2588
181	New Zealand lemonade lemon 4181
182	Nicaraguan lemon 3841
183	NISSV E mandarin hybrid 3965
184	Nordmann seedless kumquat 4070
185	Novelty x Ellendale mandarin hybrid 3956
186	Orban lemon hybrid 4140
187	Orlando bittersweet orange 1588
188	Orogold sour orange hybrid 693
189	Paak Ling Mung lemon hybrid 3193
190	Pan Dan pummelo 2752
191	Papuan citron 3532
192	Palazzelli mandarin 4244
193	Parent Washington 1241A
194	Parson's Special mandarin 300
195	Pauthel pummelo 3949
196	Pearl tangelo 2849
197	Peretta lemon hybrid 3924

198	Perrine lemon hybrid 2458
199	Philippine citron 3533
200	Phillipine red rangpur 2318
201	Pink pummelo 2246
202	Pixie mandarin 3568
203	Pomelit pummelo hybrid 4026
204	Pomona sweet lemon 4068
205	Ponderosa lemon hybrid 294
206	Pong yau pummelo 2342
207	Primofiore lemon 3491
208	Prior Lisbon lemon 3505
209	Procimequat kumquat hybrid 3295
210	Puma pummelo hybrid 3941
211	Rangpur x Troyer 3997
212	Red pummelo 2245
213	Reinking pummelo 3805
214	Richard's Special mandarin 4117
215	Rio Grande Valley lemonquat 4097
216	Roeding's Pink pummelo 1208
217	Rojo Blanco pummelo hybrid 4029
218	Rubidoux mandarin 4031
219	Rubidoux pummelo hybrid #1 4028
220	Ruddy tangor 4146
221	Rusk citrange 1441
222	S281 citrangelo trifoliolate hybrid 3552
223	S302 Citranguma trifoliolate hybrid 3415
224	Sacaton citrumelo 3417
225	Sacaton tangelo 2609
226	Sacaton tangelo 3331
227	Sampson tangelo 2418
228	San Jacinto tangelo 2011
229	Sanford citrange 276
230	Santa Barbara tangelo 4073
231	Savage citrange 275
232	Scarlet Emperor Ponkan mandarin 3226
233	Seminole tangelo 2604
234	Sexton tangelo 2543
235	shaddock X St. Michael orange hybrid 42
236	Siamelo tangelo 2586
237	Siamese Acidless pummelo 2240

238	Sicily citron 3534
239	Sinton_citrangquat_3642
240	Siranui Dekopan 4113
241	smooth flat seville sour orange 4079
242	Sniff_acidless_Rangpur_2484
243	Soh Long lemon hybrid 3199
244	Soh Niangrang Papeda 3203
245	Soh Synteng lemon hybrid 3261
246	Solid Scarlet mandarin 3328
247	Som Kaeo mandarin 3852
248	Sour pummelo 3066
249	C. aurantium var salicifolia sour orange 3289
250	South African rough lemon 2325
251	Suisho Buntan pummelo 3947
252	Sukega grapefruit 4149
253	Sunburst mandarin 3809
254	Sun Chu Sha mandarin 4003
255	Sunrise tangelo 2606
256	Sunshine pummelo 2236
257	Sunshine tangelo 2788
258	Sweeteet lemon 3093
259	Sweet pummelo 3067
260	Swingle citrumelo 3767
261	Sydney hybrid 1485
262	Szinkom mandarin 3085
263	Tahitian pummelo 3806
264	Tango mandarin 4183
265	Tankan tangor 2224
266	Tau Yau pummelo 2583
267	Tavares kumquat hybrid 4048
268	Duncan grapefruit 3832
269	Thomasville citrangequat 1440
270	Thong Dee pummelo 4027
271	Thong Dee pummelo 3927
272	Thorton tangelo 2013
273	Tien Chieh mandarin 2590
274	Tieu nucellar mandarin RRUT 87
275	Trifeola trifoliolate hybrid 3954
276	Unknown lemon hybrid 3172
277	Unnamed Chinese RRUT 223

278	India sour orange hybrid 3175
279	Unnamed lemon hybrid 3154
280	Unnamed papeda 3056
281	Unnamed sweet orange hybrid 648
282	Unnnamed lemon hybrid 3155
283	US 852 4245
284	UC Seedless Surprise USDA 1-77-19 tangelo 4208
285	US Superna USDA 88-2 mandarin hybrid 3991
286	USDA 88 3 mandarin 3992
287	Vangasay 807 rough lemon 3996
288	Warren limequat 3069
289	Webber Shaddock pummelo 4233
290	Webber tangelo 2746
291	Wheeny grapefruit hybrid 2885
292	Willial tangelo 4147
293	Williams tangelo 2787
294	X639 trifoliate hybrid 3957
295	Yalaha tangelo 2559
296	Yemen citron 2847
297	Zhuluan sour orange 3930
298	<i>C. ichangensis</i> 3931
299	<i>C. indica</i> 3163
300	<i>C. micrantha</i> 3605
301	Flying dragon trifoliate 3330A
302	<i>C. webberri</i> kalpi 1456
303	<i>M. australasica</i> 3672
304	<i>M. australis</i> 3665
305	<i>M. australasica</i> 3670
306	Monoembryonic trifoliate 3888
307	#26 Nanjing 3939
308	#27 trifoliate 3938
309	Simmons trifoliate 3549
310	Small Leaf 4017
311	Emerald mandarin 4123
312	Philippine pummelo hybrid 2343
313	Standard sour orange 628
314	Yuzu ichandarin 1216
315	Triumph grapefruit 297
316	Kabosu papeda hybrid 3943
317	<i>C. canaliculata</i> 3565

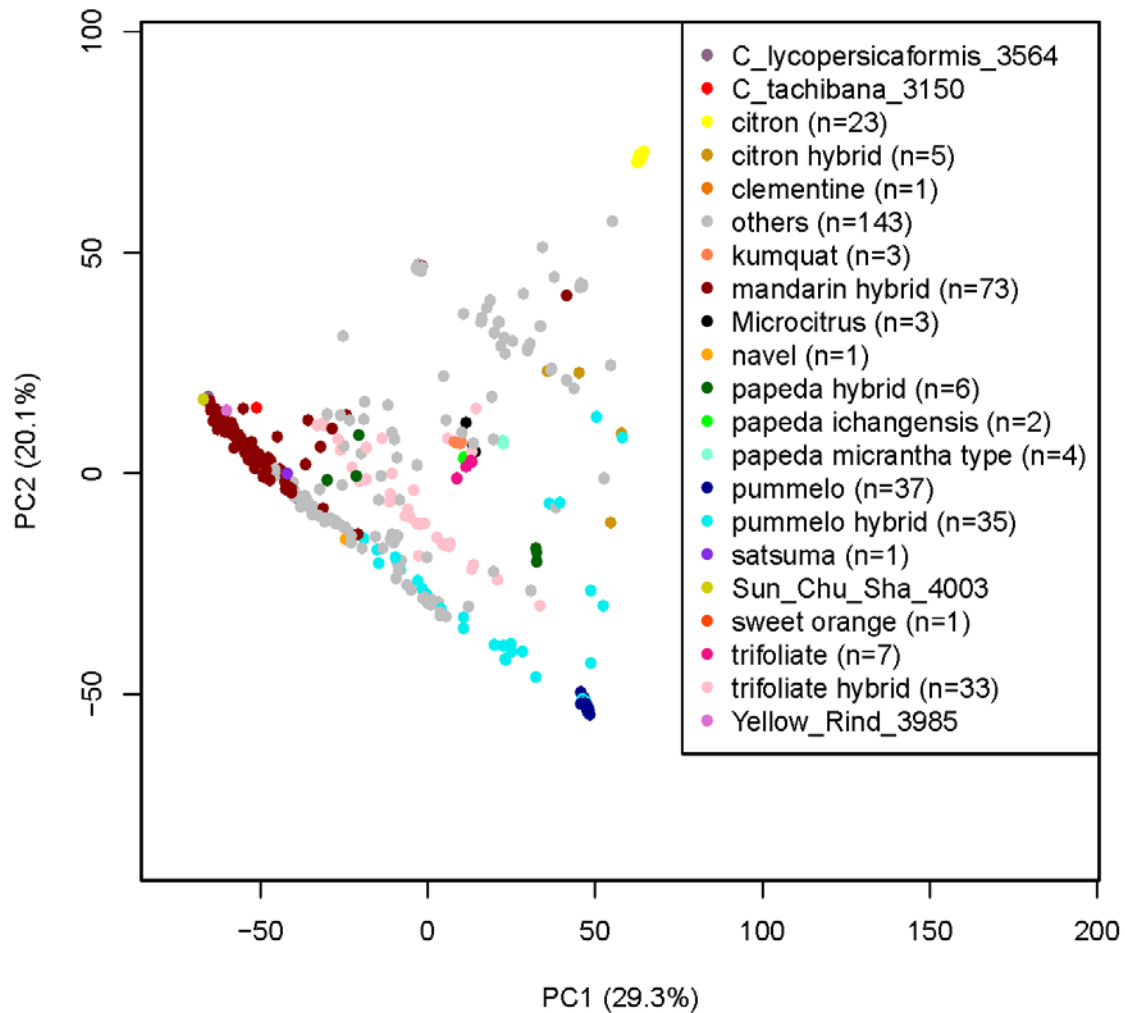
318	Philippine pummelo 644
319	Santa Barbara red rangpur lime 712
320	Kiyomi tangor 3952
321	Kao Phuang pummelo 2352
322	Rex Union grapefruit hybrid 3962
323	Odorata citron 3655
324	Citrus webberii var Montana 767
325	Bergamot sour orange hybrid 2881
326	Fortunella obovata 3475
327	Necked mandarin 3738
328	Carrizo citrange 2863
329	Ichang lemon 1215
330	C. excelsa 2316
331	Siamese pummelo 2421
332	Royal grapefruit 248
333	Kao Pan pummelo 2355
334	Kincy mandarin 4144
335	Egami Buntan pummelo 3959
336	Iraq lemon limetta 3492
337	Milam rough lemon 3396
338	Schaub rough lemon 3879
339	Pin Shan Kong Yan pummelo 2348
340	C. pennivesiculata 2434
341	Orlando tangelo 3874
342	Minneola tangelo 3340
343	Kinnow mandarin 3021
344	Wilking mandarin 3020
345	Sacaton citrumelo 3337
346	Valentine pummelo hybrid 4182
347	Deep Red pummelo 2347
348	Nova mandarin 3615
349	US 119 3998
350	Wincia tangor 4148
351	US Furr SRA337 mandarin 4238
352	Chinese pummeo 1212
353	Temple tangor 2598
354	USDA 6-15-150 manarin 4165
355	Robinson mandarin 3850
356	Pursha lime x Chinotto 3578
357	Ortanique tangelo 3368

358	Page mandarin 3616
359	Pink pummelo 2244
360	Yellow Rind mandarinarin 3895
361	Unnamed pummelo 2248
362	Mato Buntan pummelo 3945
363	Diamante citron 3523
364	Yunnanese citron 3798
365	Mandarinette 3405
366	Frost Owari satsuma 3178
367	C. kharna 3059
368	Unnamed citron hybrid 3174
369	Mexican lime 1710
370	Tresca pummelo 302
371	C. megaloxycarpa hybrid 3241
372	Anseikan pummelo 3979
373	C. hanayu 3469
374	Arajon pummelo 2596
375	C. succosa 3280
376	Improved Meyer lemon 3737
377	Willowleaf mandarinarin nucellar 3843
378	C. sino-grandis 3148
379	Unnamed lemon 3885
380	C. keraji 3144
381	Hirado Buntan pummelo 4122
382	South Coast Field Station citron 3546

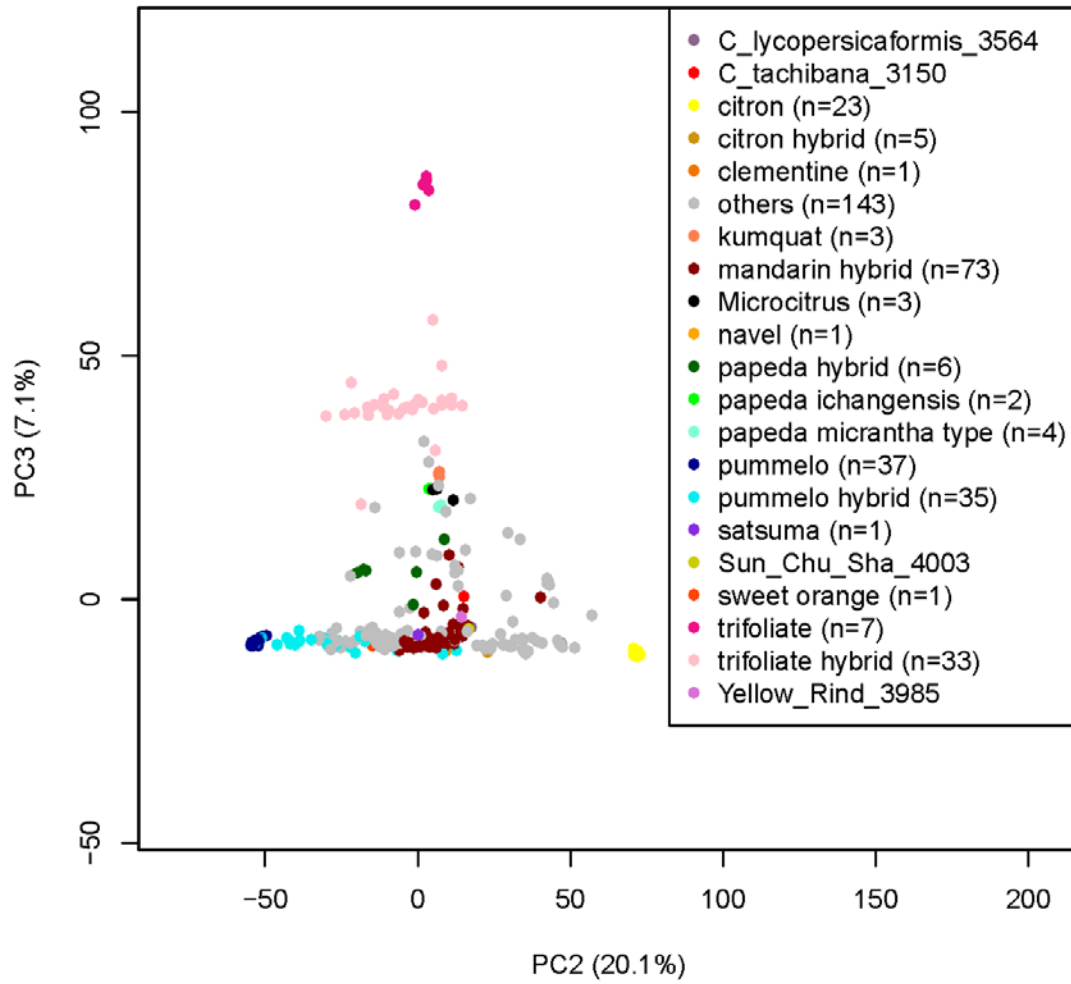
Figure 1. PCA of 382 accessions in UCR diversity panel

PCA was carried out with 382 genomic DNA samples from the 382 UCR diversity panel genotyped at 56,367 loci with Axiom® Citrus Genotyping Array. UCR diversity panel included type-1 mandarins *C. lycopersicaformis* CRC 3564 (plum), *C. tachibana* CRC 3150 (red), Sun Chu Sha CRC 4003 (darkyellow), Yellow Rind CRC 3895 (orchid), 23 citrons (yellow), 5 citron hybrids (gold), 1 clementine (dark orange), 3 kumquats (coral), 73 mandarin hybrids (darkred), 3 *Microcitrus* accessions (black), 1 navel (orange), 6 papeda hybrids (green), 2 *ichangensis* accessions (darkgreen), 4 *micrantha* type papedas (aquamarine), 37 pummelos (darkblue), 35 pummelo hybrids (cyan), 1 satsuma (purple), 1 sweet orange (orangered), 7 trifoliate (deeppink), and 33 trifoliate hybrid (pink). PC1, 2 and 3 explained 56.4% of the total variation that exists in the data. (a) PC1(29.3%) vs PC2(20.1%) (b) PC2 (20.1%) vs PC3 (7.1%) (c) PC1 (29.3%) vs PC3 (7.1%)

(a)



(b)



(c)

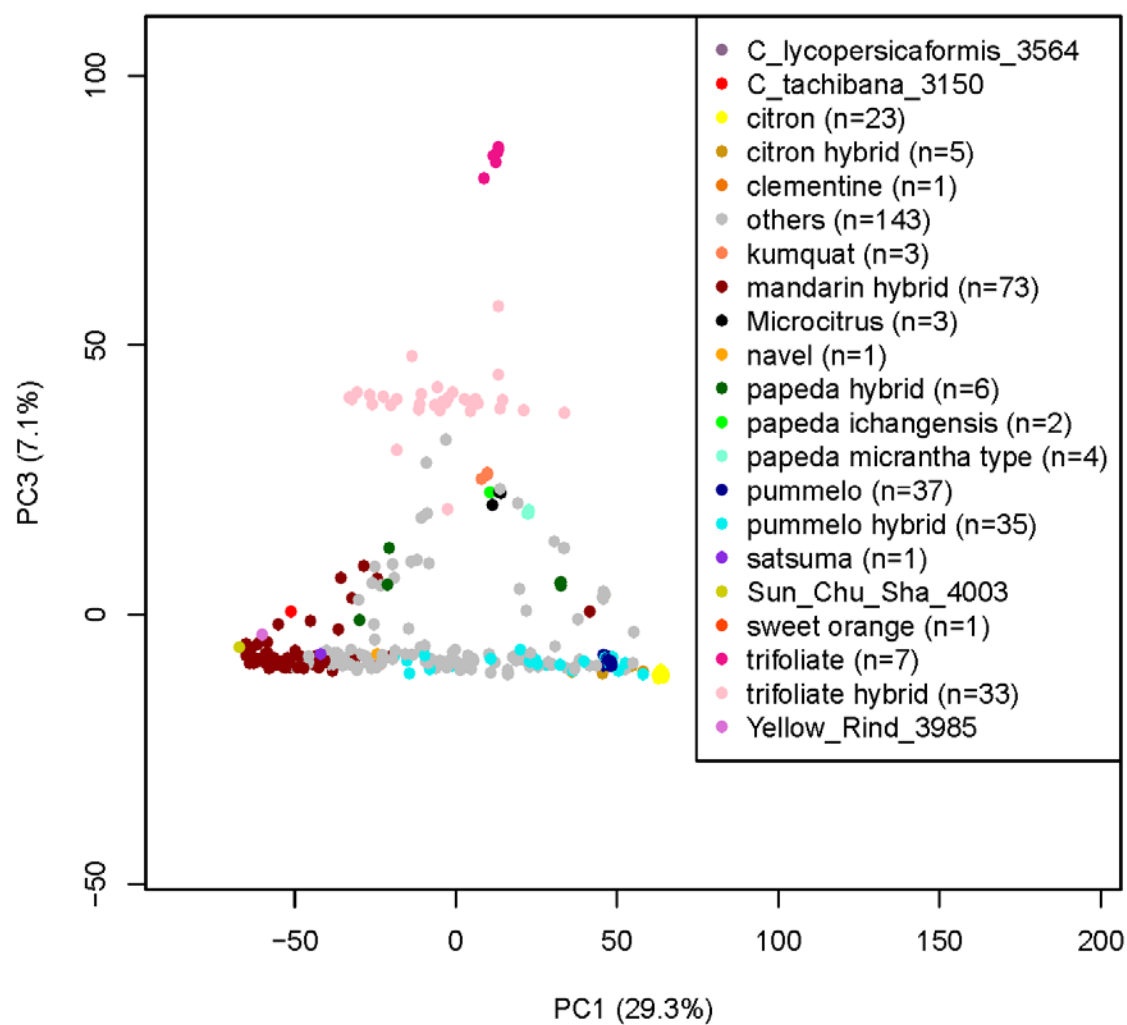


Figure 2 Cross-validation error plot for admixture analysis with the 382 UCR diversity panel

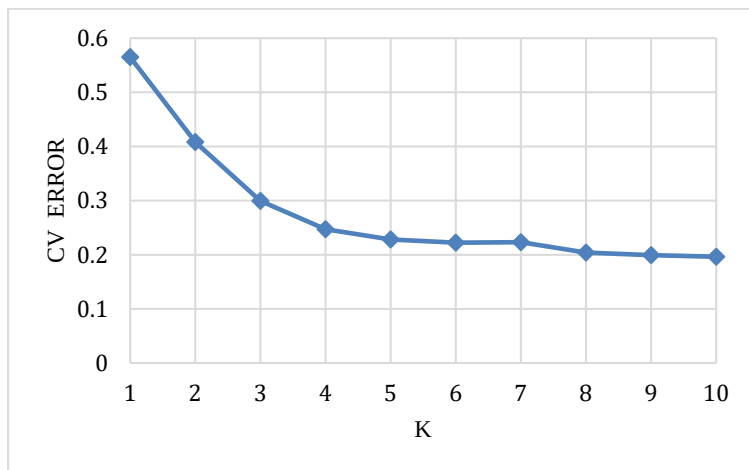
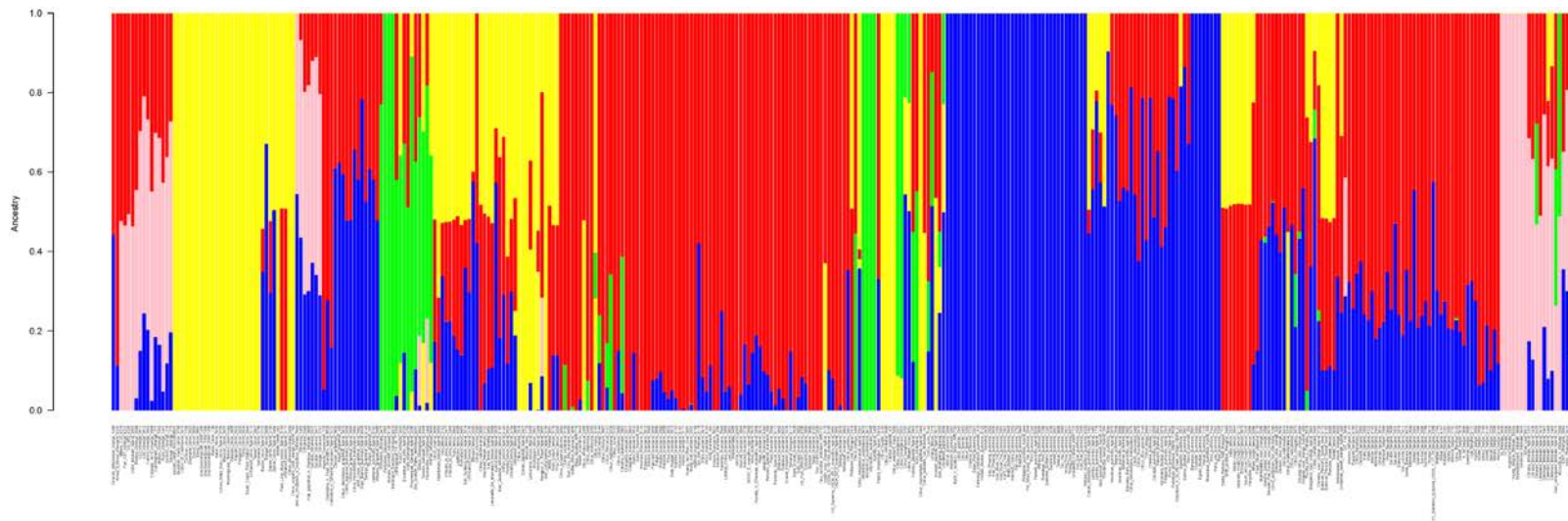


Figure 3. Admixture analysis with the 382 UCR diversity panel

Admixture analysis conducted with 56,367 PHR and NMH autosomal loci with Axiom® Citrus Genotyping Array using software ADMIXTURE (a) K=6 showing trifoliate (pink), citron (yellow), kumquat/papeda/Micrositrus (orange), mandarin (red), pummelo (blue), and papeda-micrantha type (green). (b) K=5 showing trifoliate (pink), citron (yellow), kumquat/papeda/Micrositrus (orange), mandarin (red), and pummelo (blue). (c) the error rate plot estimated by CV test for each K in the ADMIXTURE analysis from K=1 through K=10

(a)



(b)

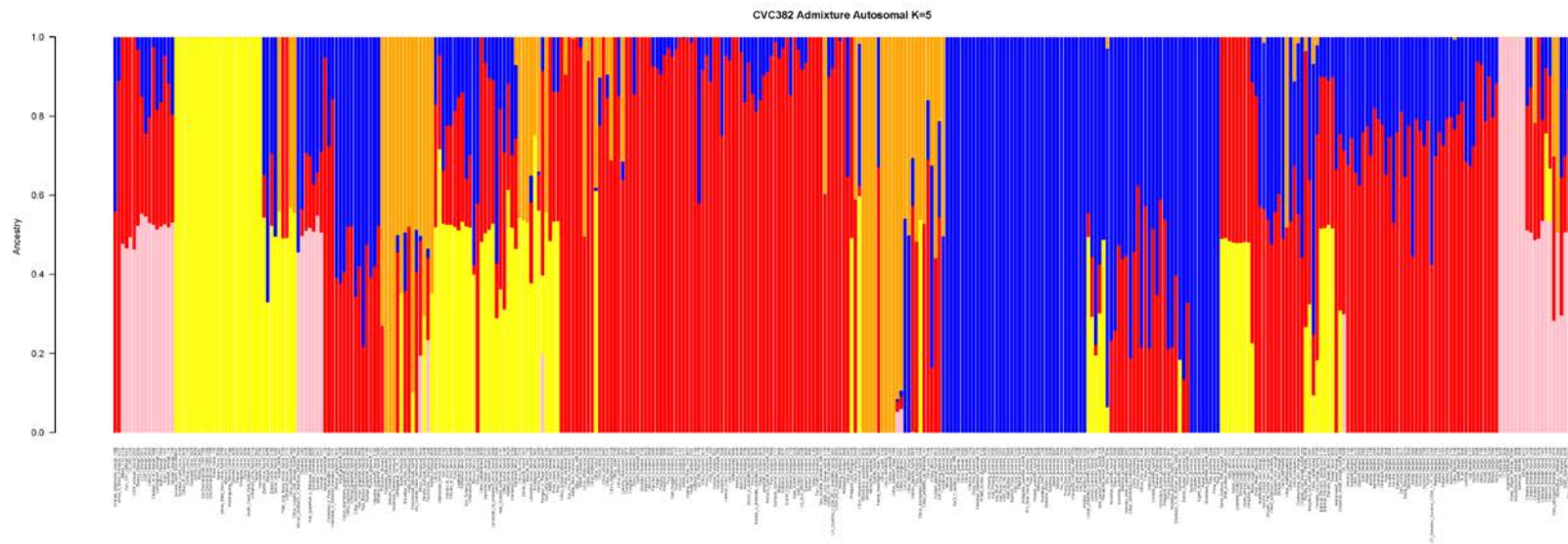


Figure 4. Maximum likelihood phylogeny tree of 85 accessions in *Citrus* and close relatives. Phylogenetic tree was generated with Axiom® Citrus Genotyping Array data based on 41,524 concatenated SNPs generated by SNPhylo. 85 accessions considered for the analysis includes non-hybrid and non-admixed accessions in CVC based on heterozygosity and admixture results. Maximum likelihood trees were generated by pipeline SNPhylo and visualized by FigTree v1.4.3. Bootstrap values based on 1000 replications are indicated on the branch nodes. Bootstrap value of above 90% is indicated in bold line.

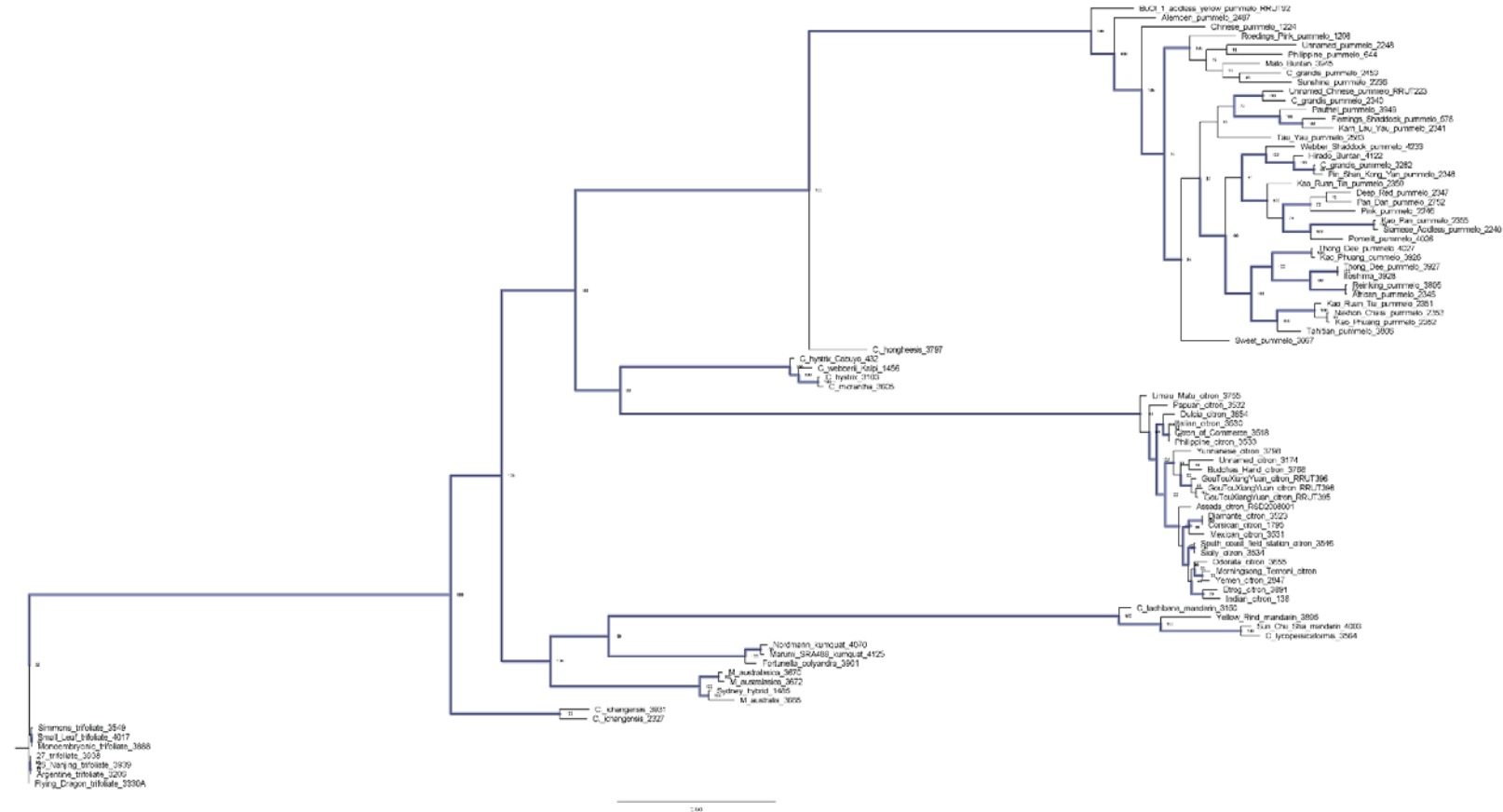


Figure 5. Phylogeny network analysis of the 382 UCR diversity set.

Genotyping calls of 12,183 loci were used for the Neighbor-net analysis which was carried out with default setting for the 382 UCR diversity set using SplitsTree v4.14.6 to investigate a reticulated structure in citrus. The numeric codes for the accession are identified in Table 7.

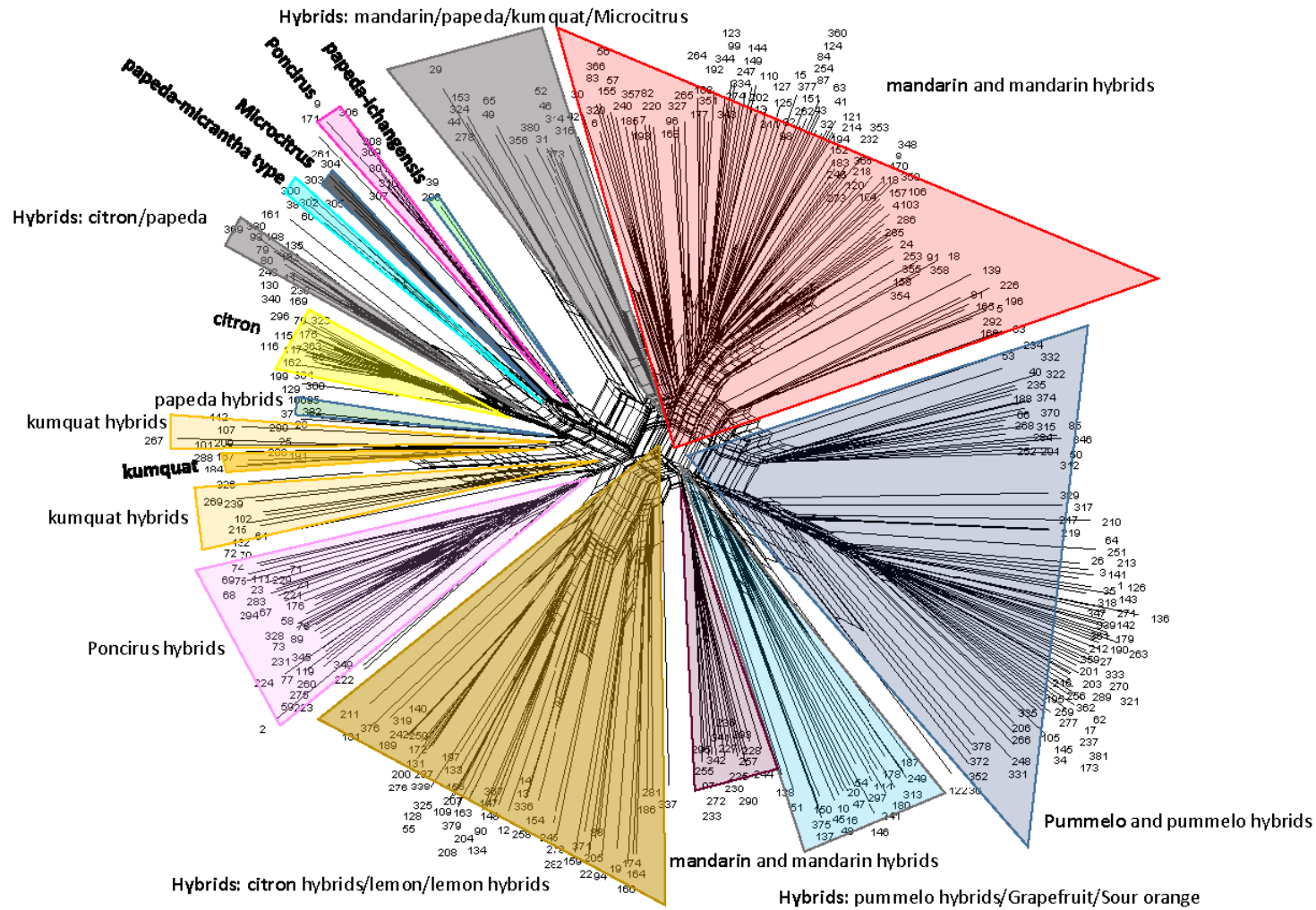


Figure 6. Graphical genotyping of Koster CRC 3958 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

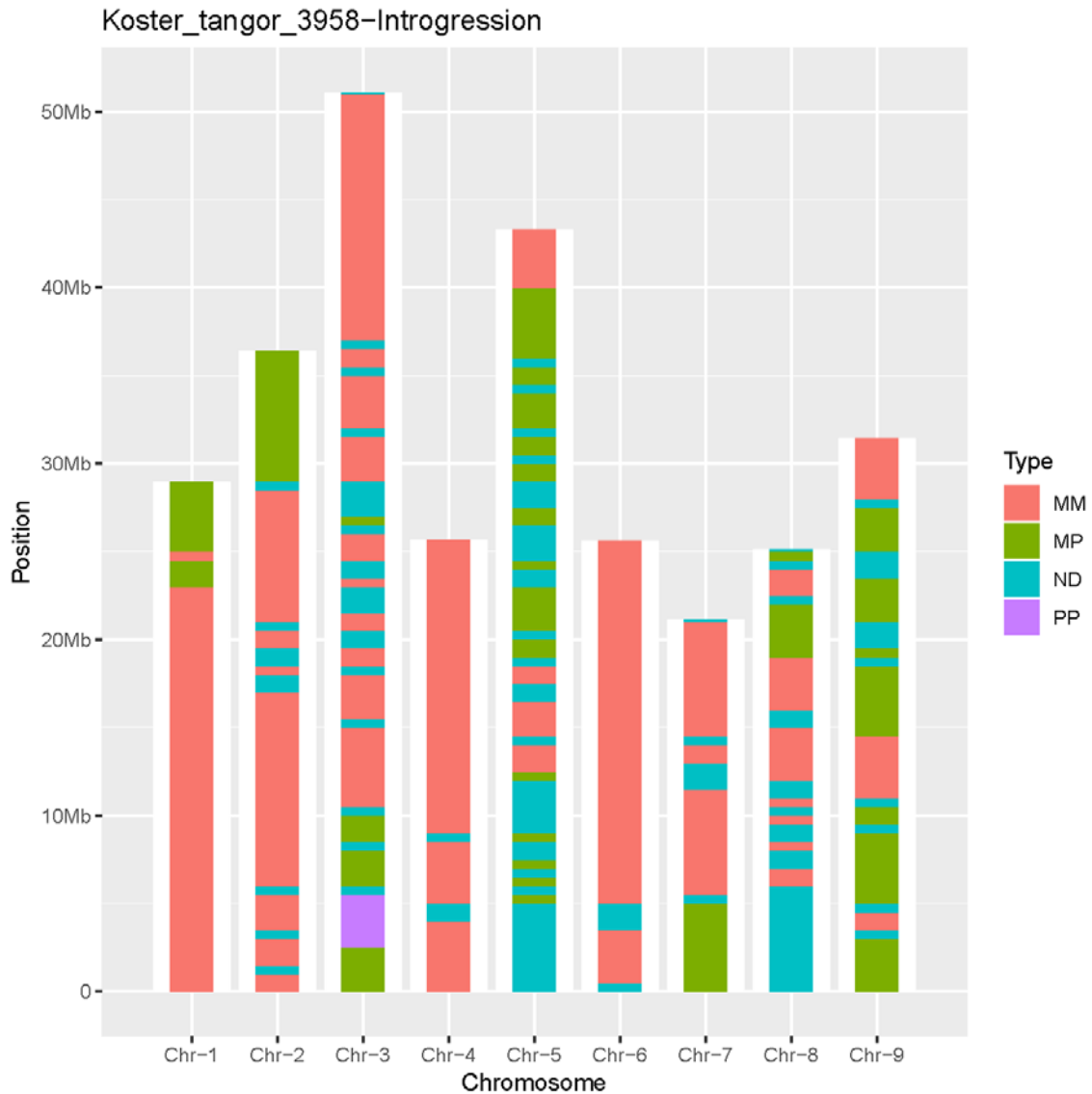


Figure 7. Graphical genotyping of Ellendale CRC 3960 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

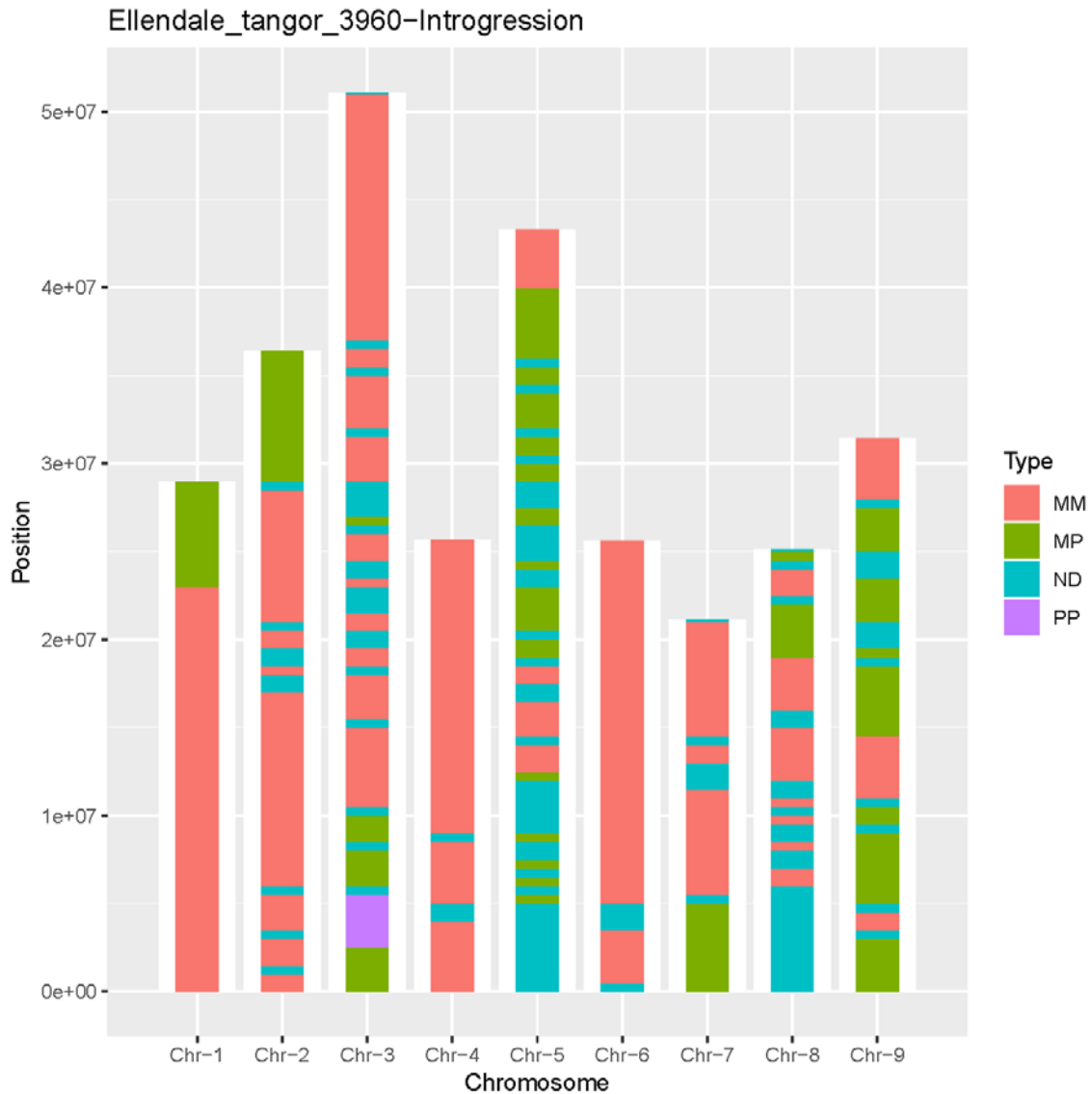


Figure 8. Graphical genotyping of *C. amblycarpa* CRC 2485 showing the pattern of introgression between papeda and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and papeda allele (green), homozygous for papeda allele (purple), and undetermined (teal) based on the frequency of SNP loci in each genotype class.

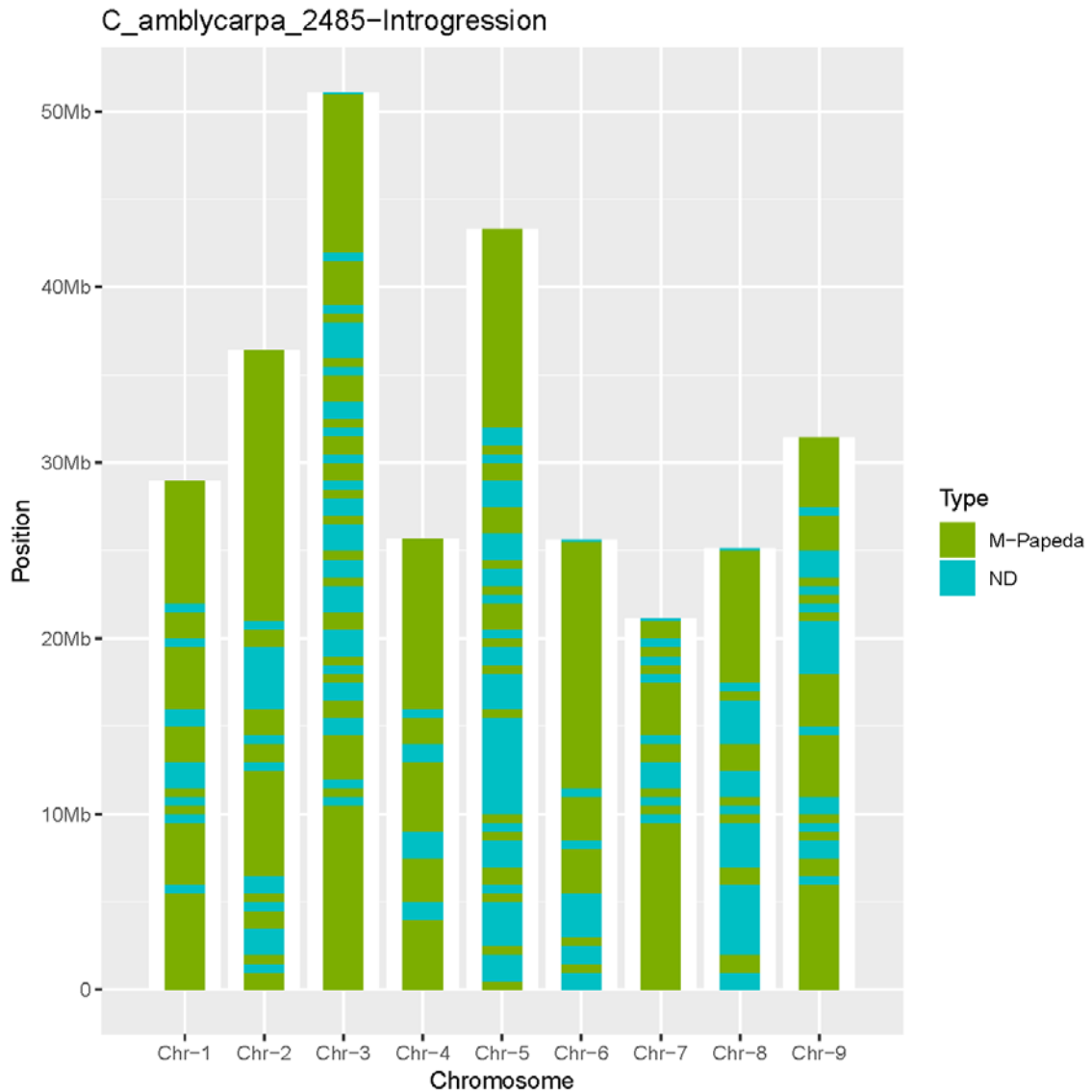


Figure 9. Graphical genotyping of *C. leiocarpa* CRC 3147 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

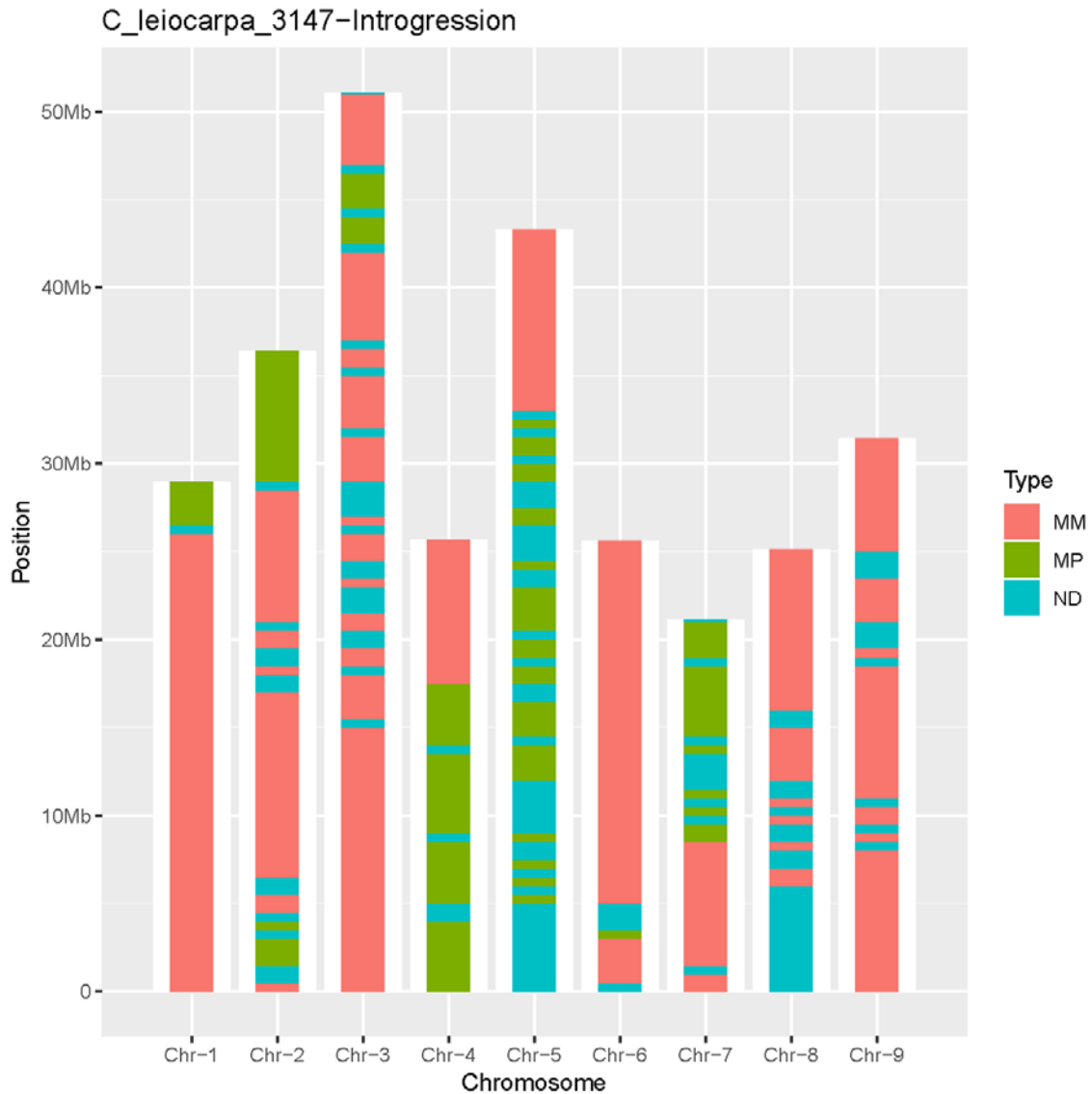


Figure 10. Graphical genotyping of *C. keraji* CRC 3144 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

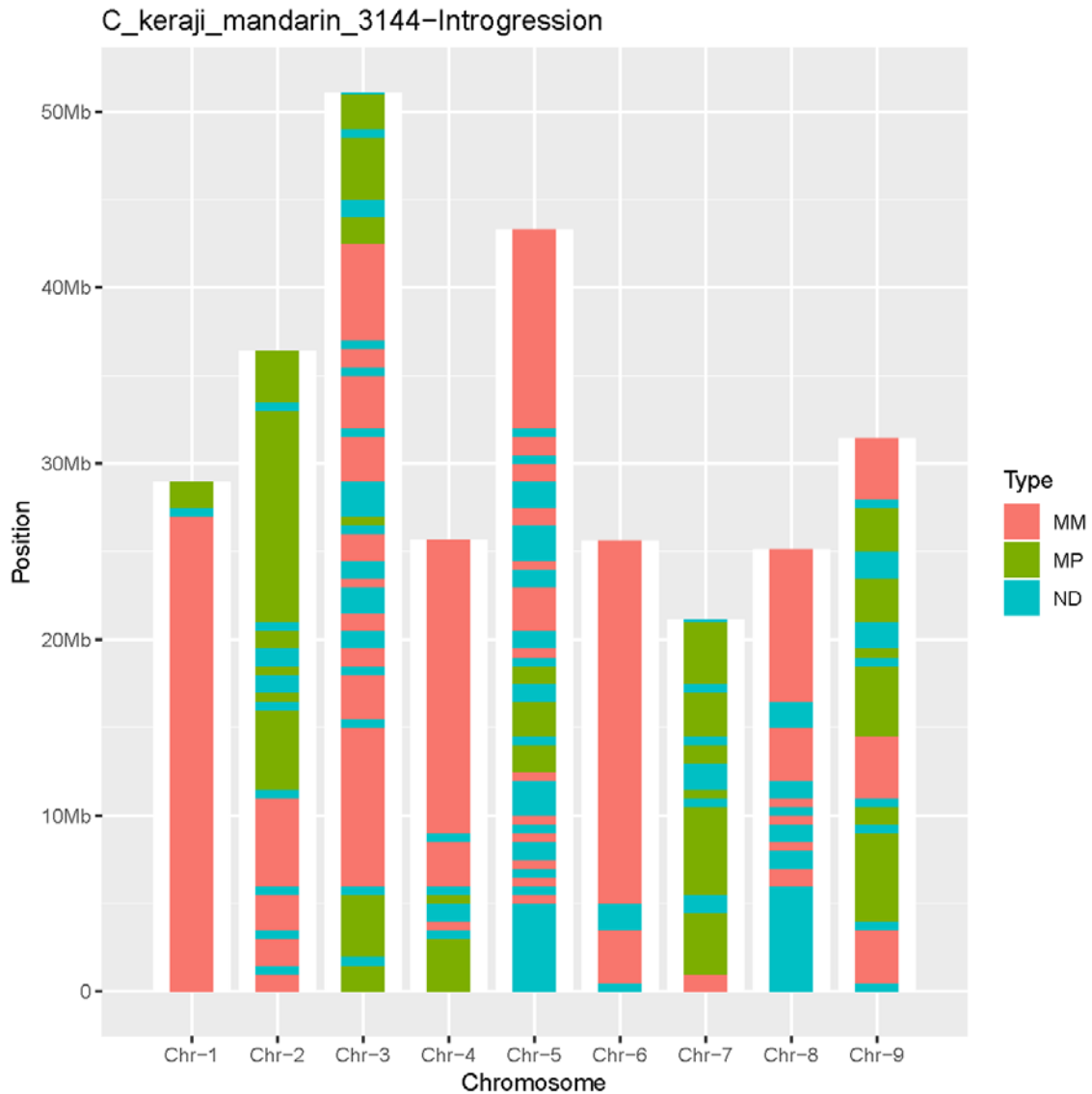


Figure 11. Graphical genotyping of *C. tachibana* CRC 3150 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

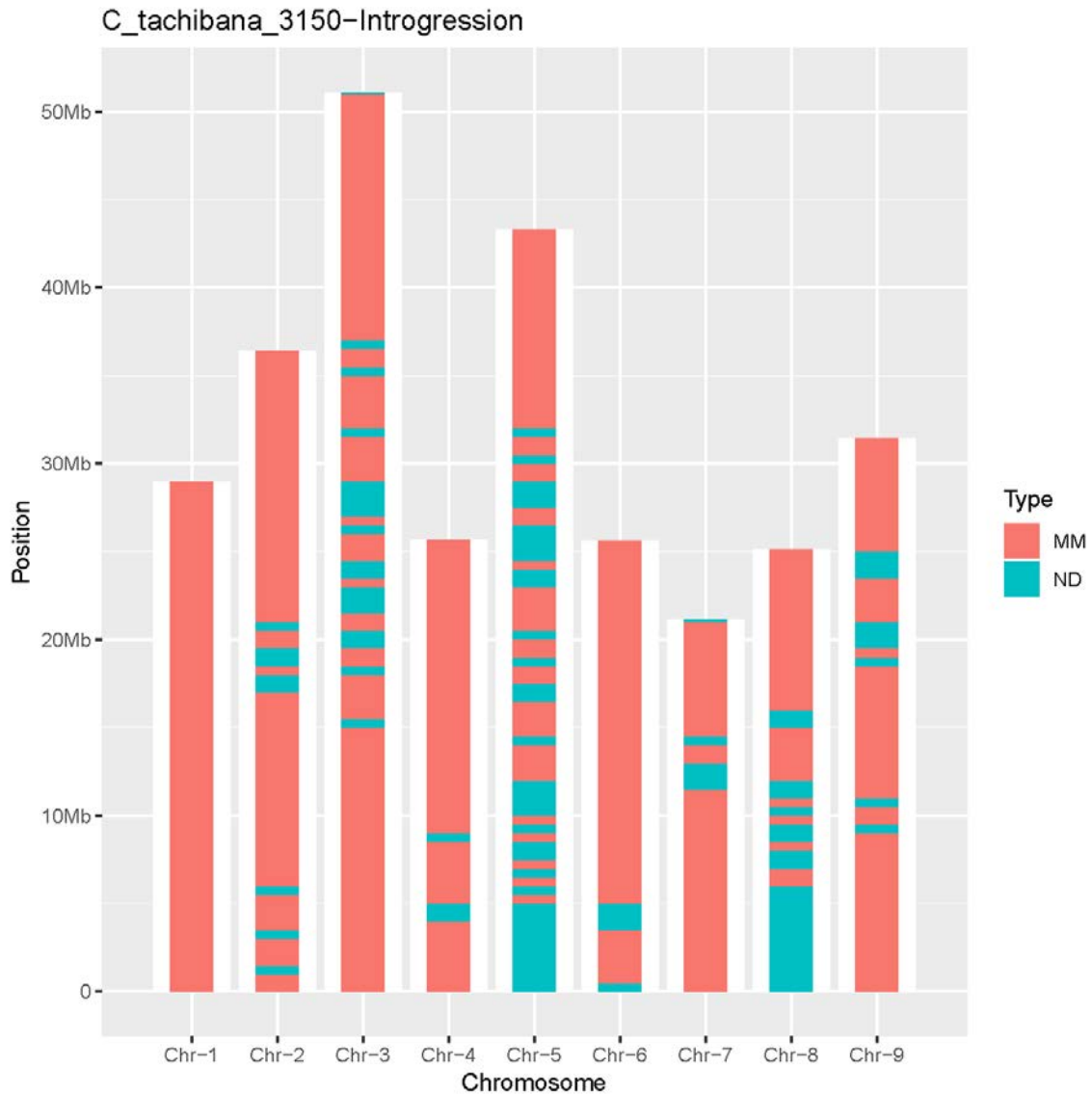


Figure 12. Graphical genotyping of *C. benikoji* CRC 3149 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.



Figure 13. Graphical genotyping of Kunembo CRC 3346 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

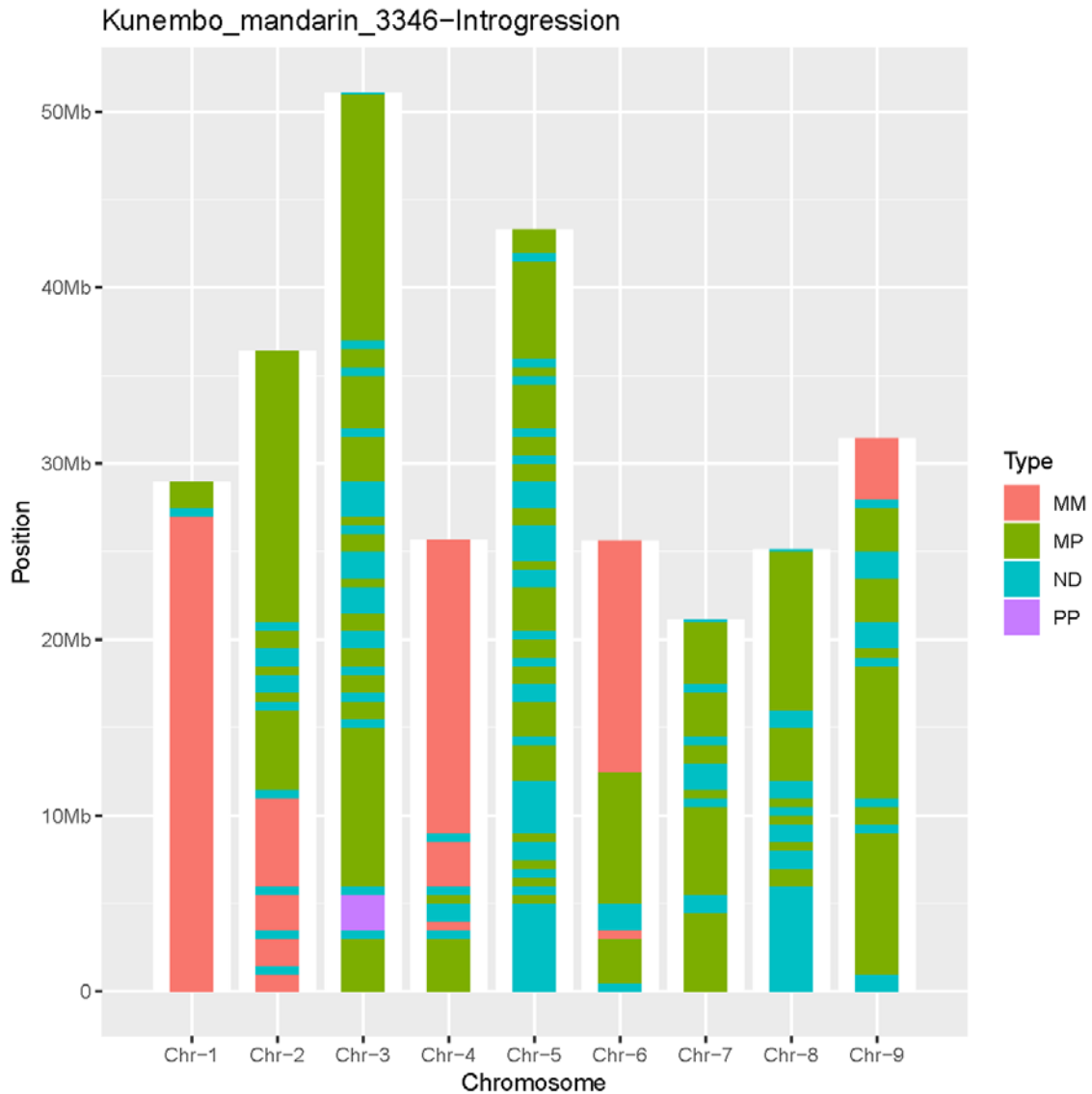


Figure 14. Graphical genotyping of *C. depressa* CRC 2448 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

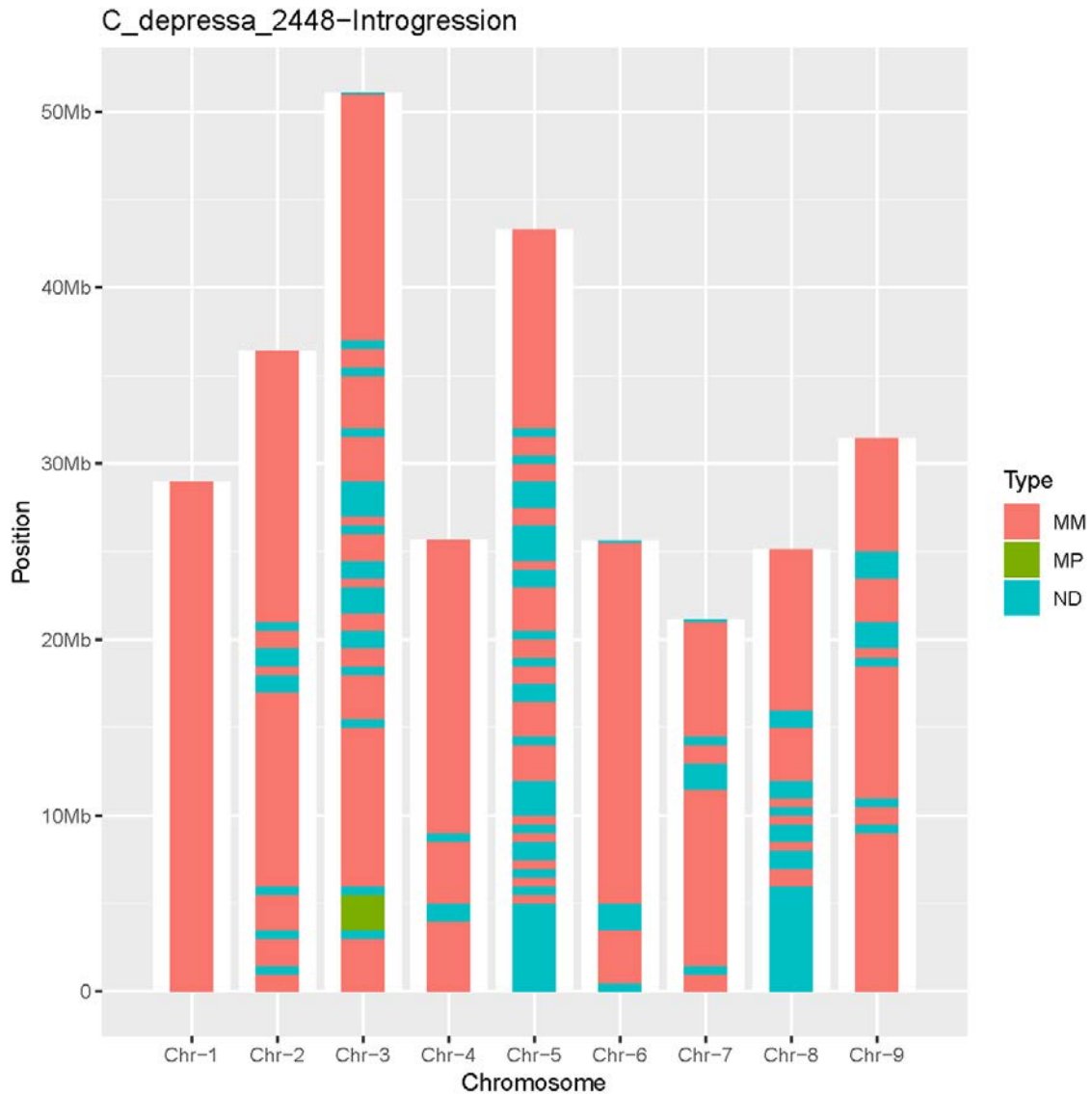


Figure 15. Graphical genotyping of Cleopatra CRC 3844 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.



Figure 16. Graphical genotyping of *C. sunki* CRC 3143 showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

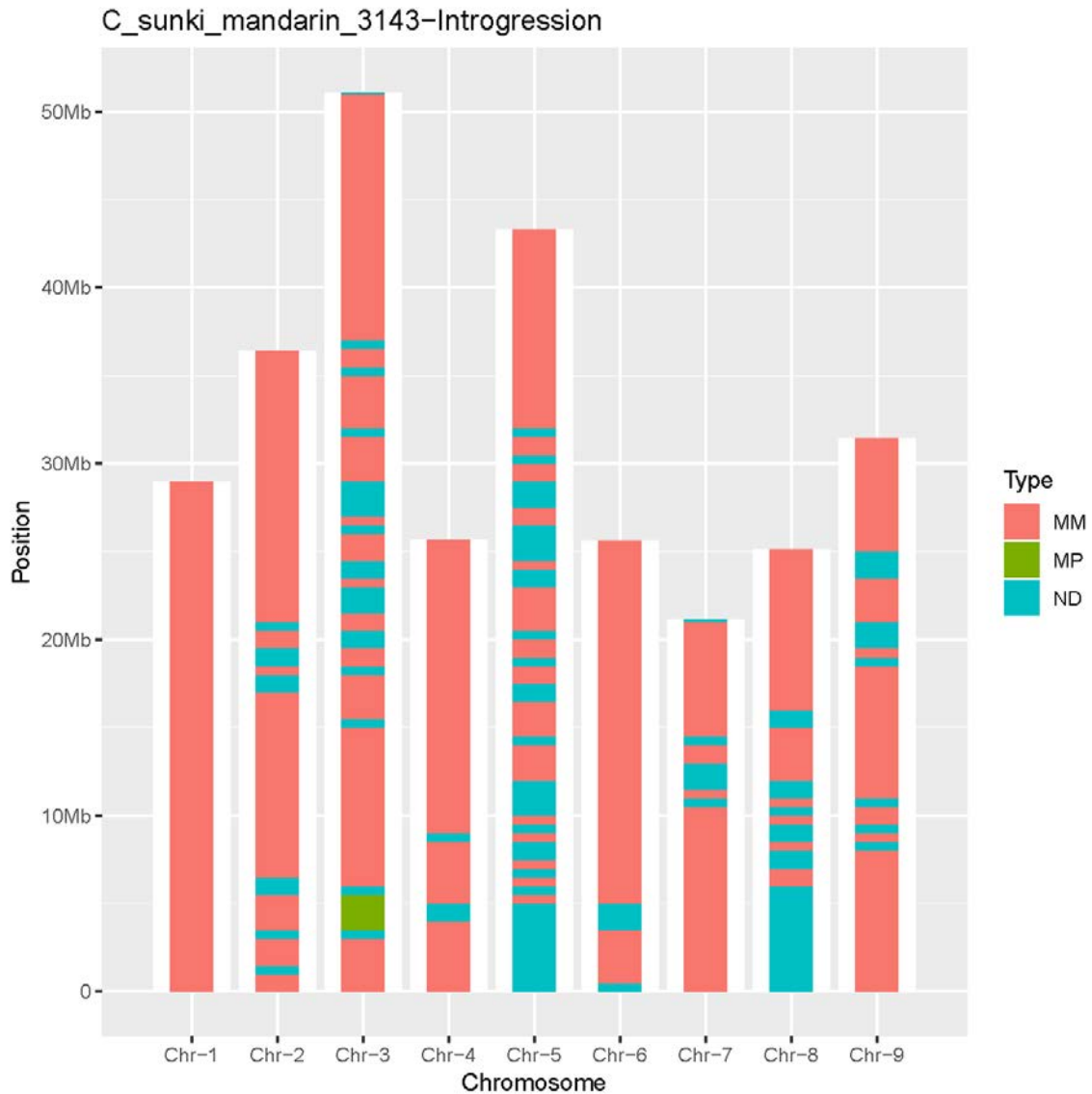


Figure 17. Graphical genotyping of candidate type-1 mandarins showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for a) Yellow Rind mandarin CRC 3895 b) *C. lycopersicaformis* CRC 3564 and c) *C. reticulata* CRC 3239

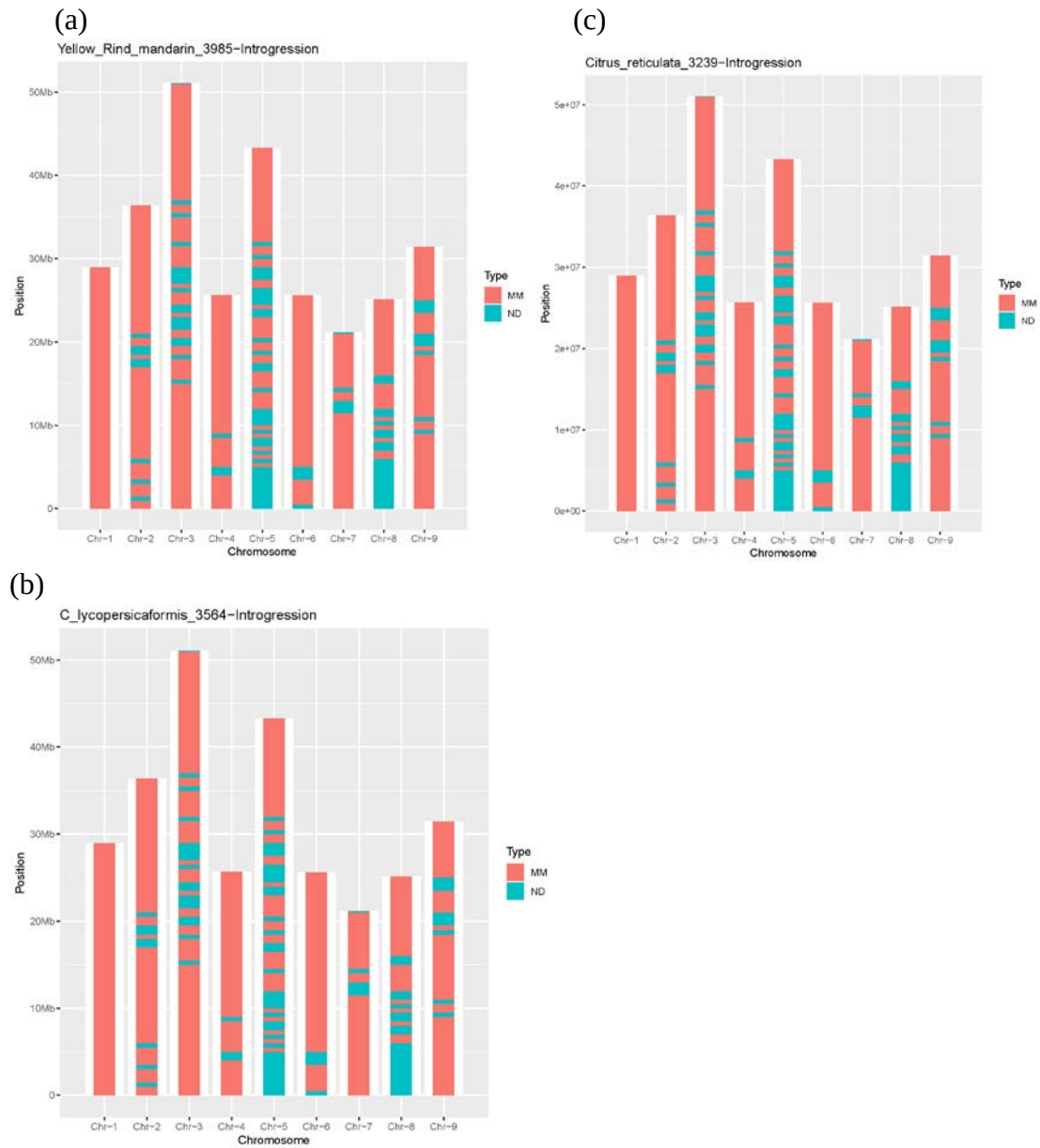


Figure 18. Graphical genotyping of candidate type-2 mandarins showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for a) Canton CRC 3576, and (b) Mihrio CRC 4248.

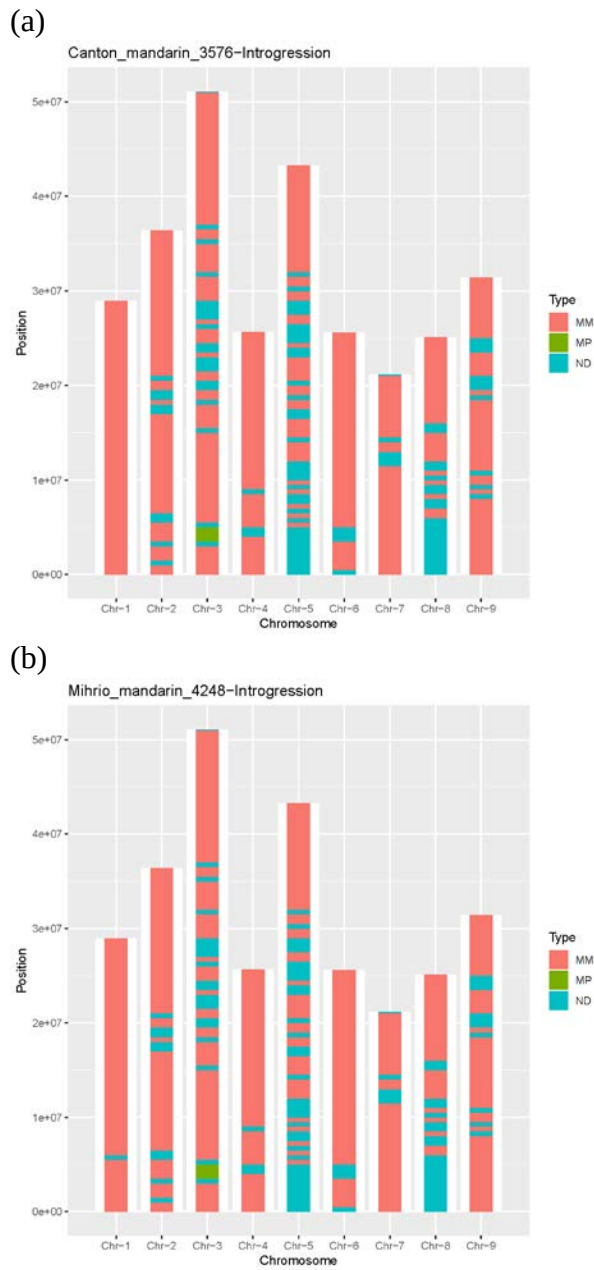


Figure 19. Graphical genotyping of Djerok CR 2456 showing the pattern of introgression between pummelo and citron at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for citron allele (red), admixed (heterozygous) for citron and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

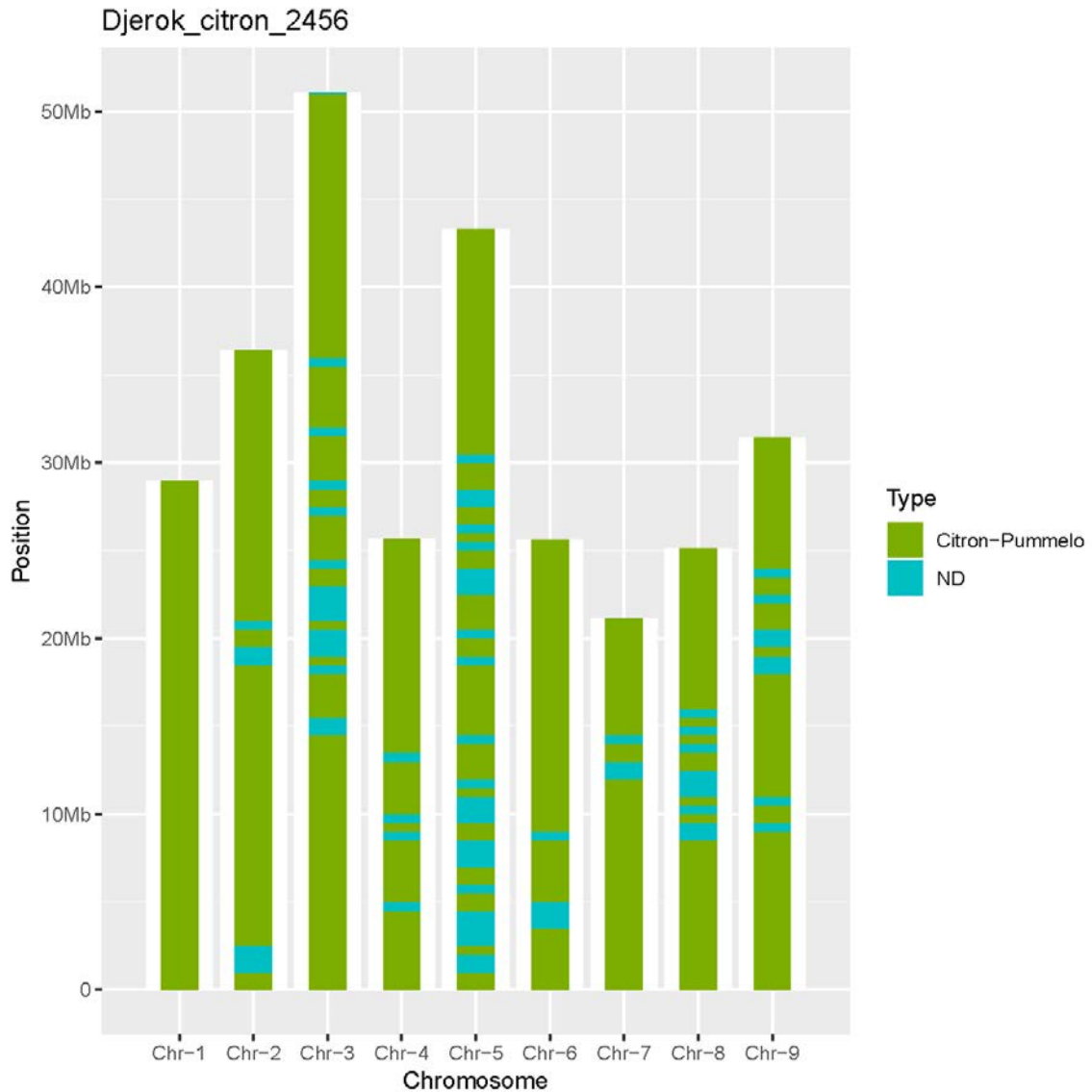


Figure 20 Graphical genotyping of Bengal CRC 3055 showing the pattern of introgression between pummelo and citron at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for citron allele (red), admixed (heterozygous) for citron and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

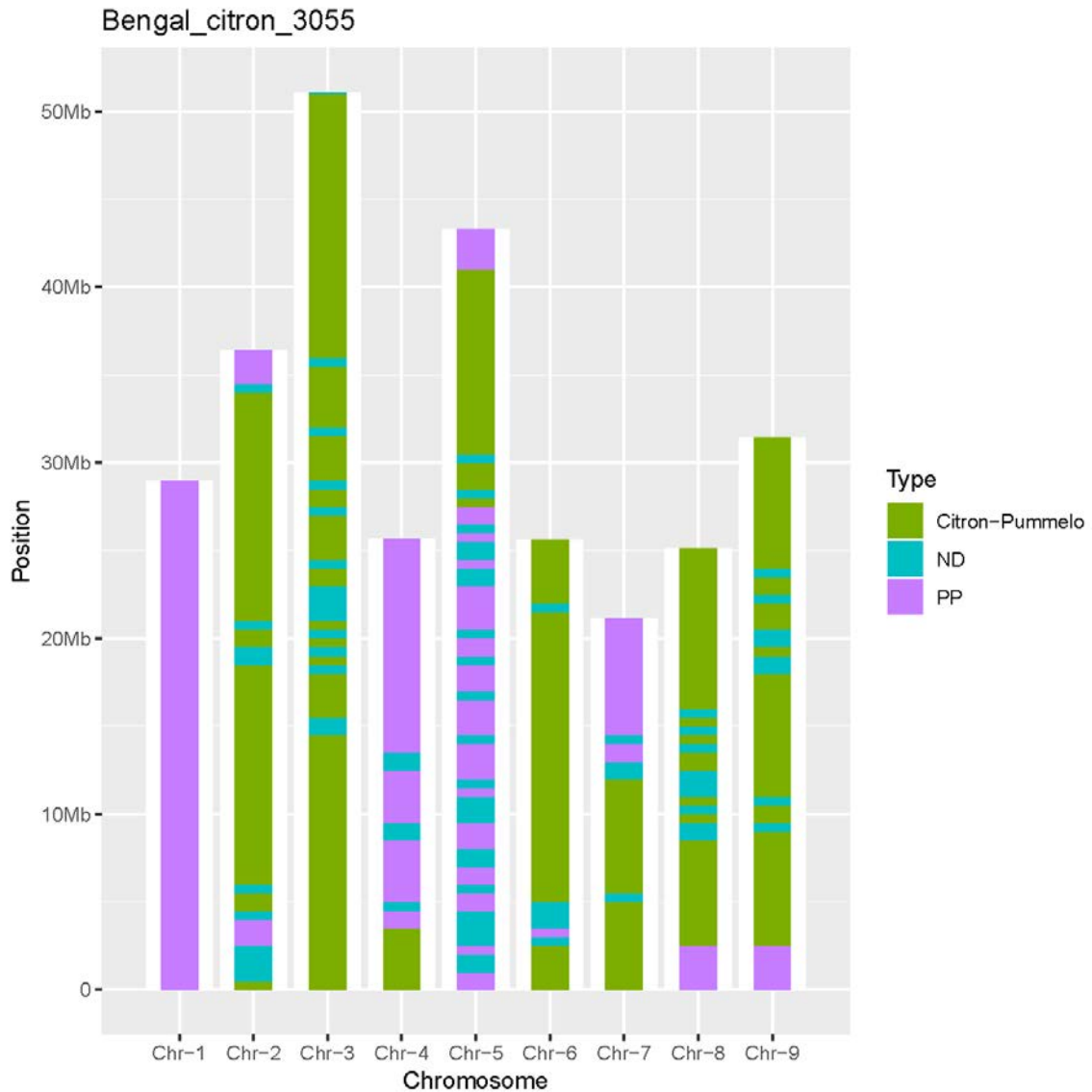
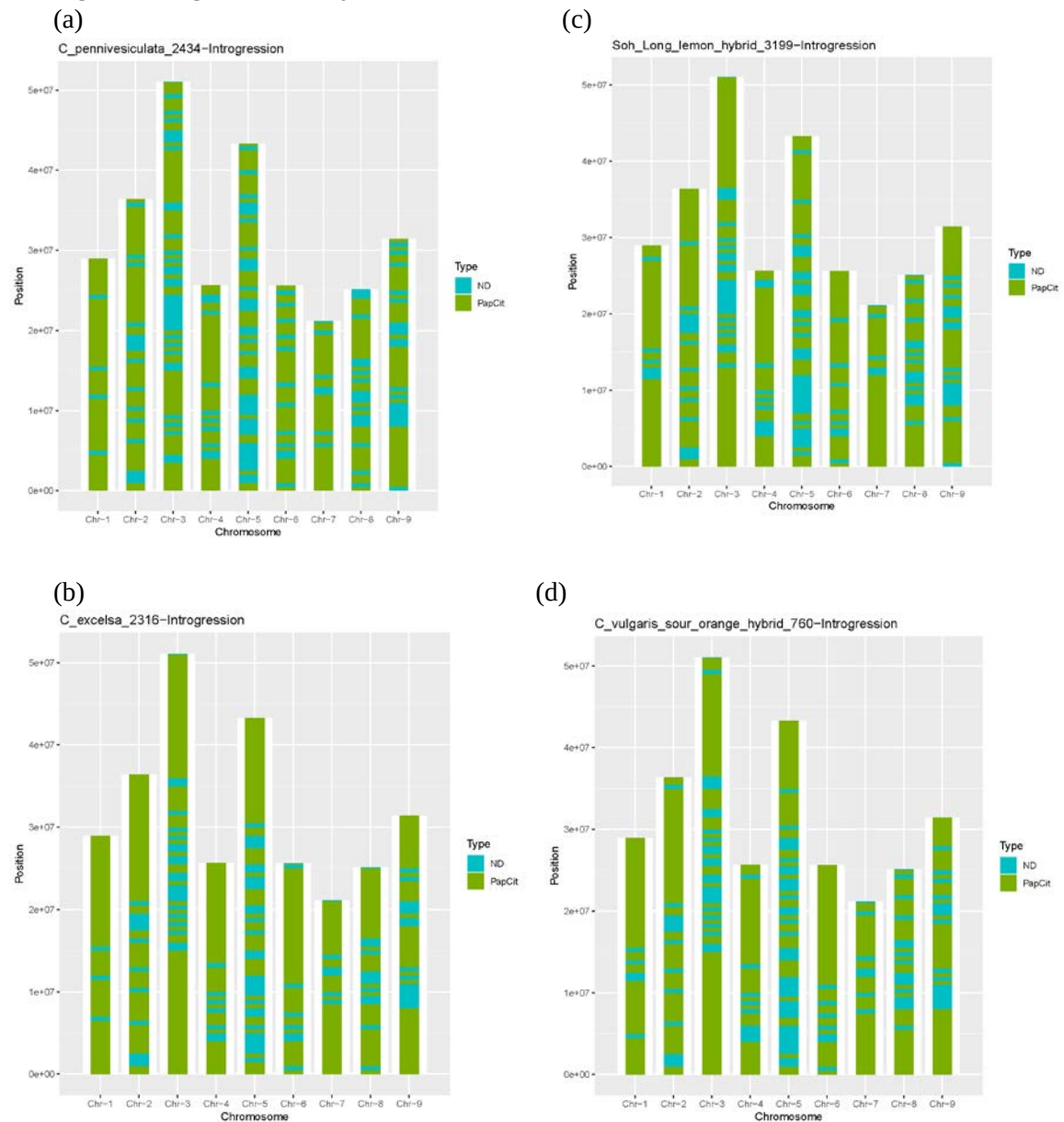
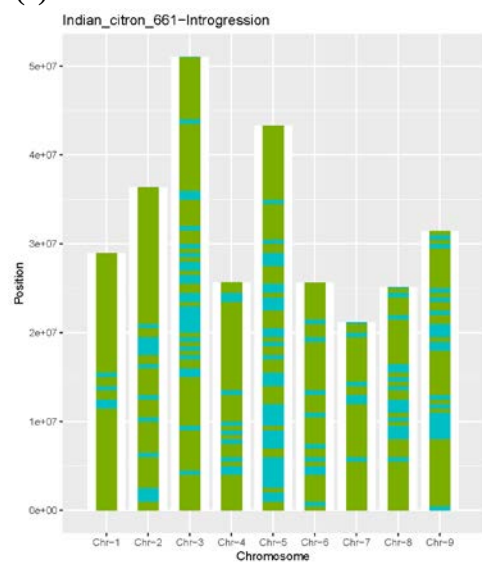


Figure 21 Graphical genotyping of *C. pennivesiculata* CRC 2434, *C. excelsa* CRC 2316, Soh Long CRC 3199, *C. vulgaris* CRC 760, India CRC 661, and *C. macrophylla* CRC 2446 showing the pattern of introgression between papada and citron at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for papada allele (red), admixed (heterozygous) for citron and papada allele (green), homozygous for citron allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for (a) *C. pennivesiculata* CRC 2434, (b) *C. excelsa* CRC 2316, (c) Soh Long CRC 3199, (d) *C. vulgaris* CRC 760, (e) India CRC 661, and (f) *C. macrophylla* CRC 2446. Citron homozygous region on chromosome 5 (9000000-9500000) of *C. macrophylla* CRC 2446 could be an error resulting from this region having a low density of markers



(e)



(f)

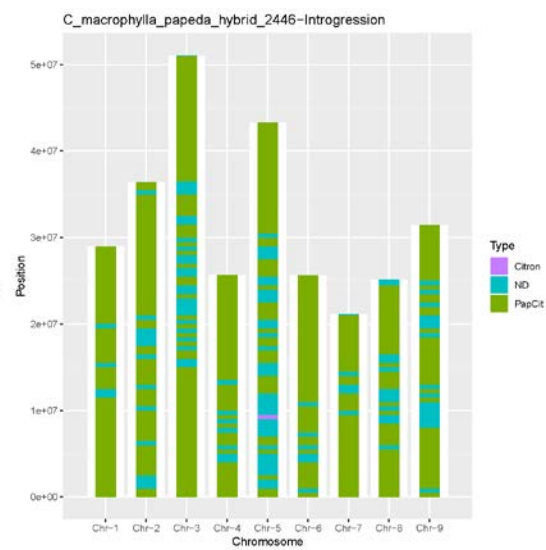


Figure 22 Graphical genotyping of Siamese pummelo CRC 2421 showing the pattern of introgression between pummelo and citron at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for citron allele (red), admixed (heterozygous) for citron and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

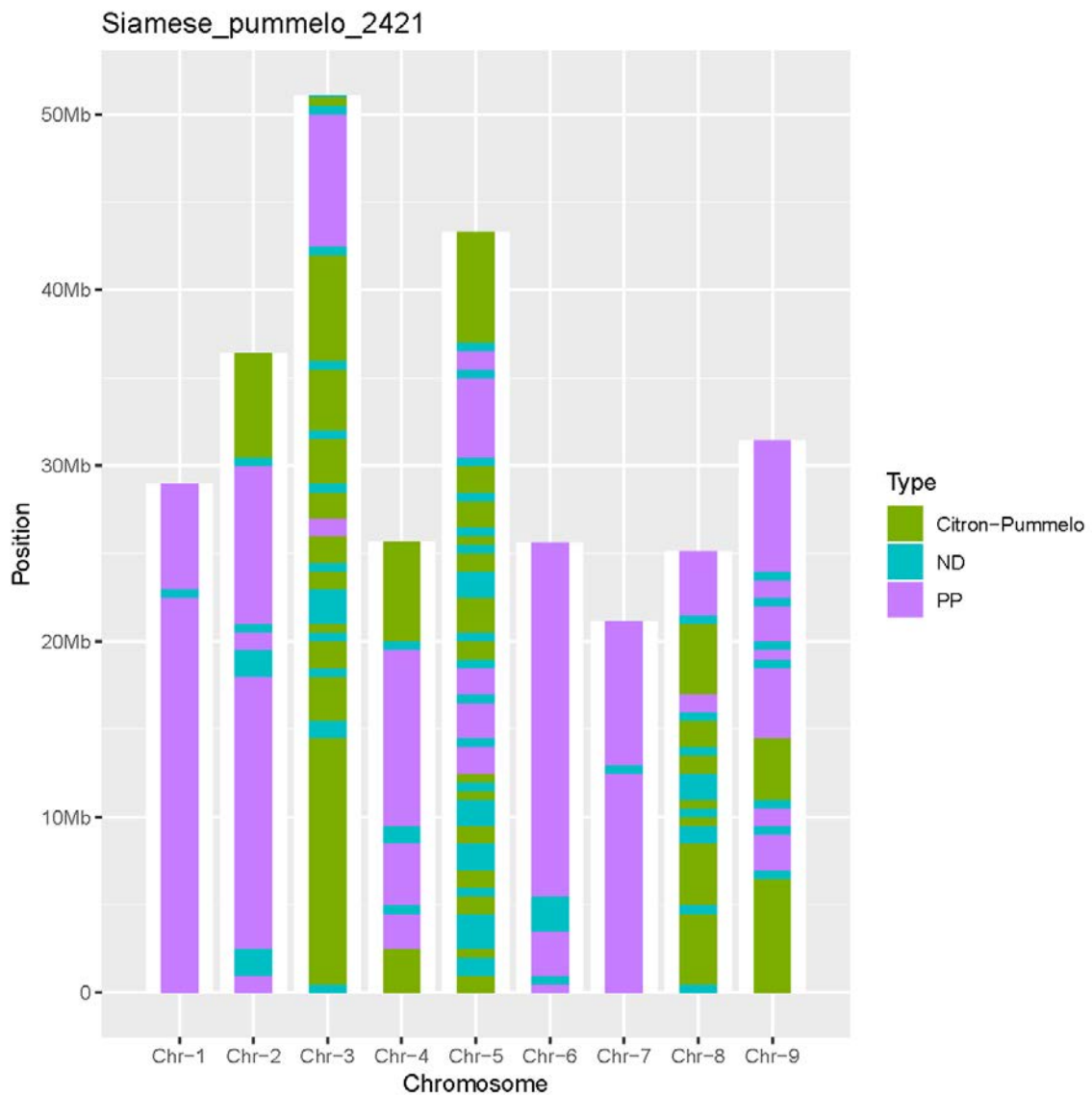
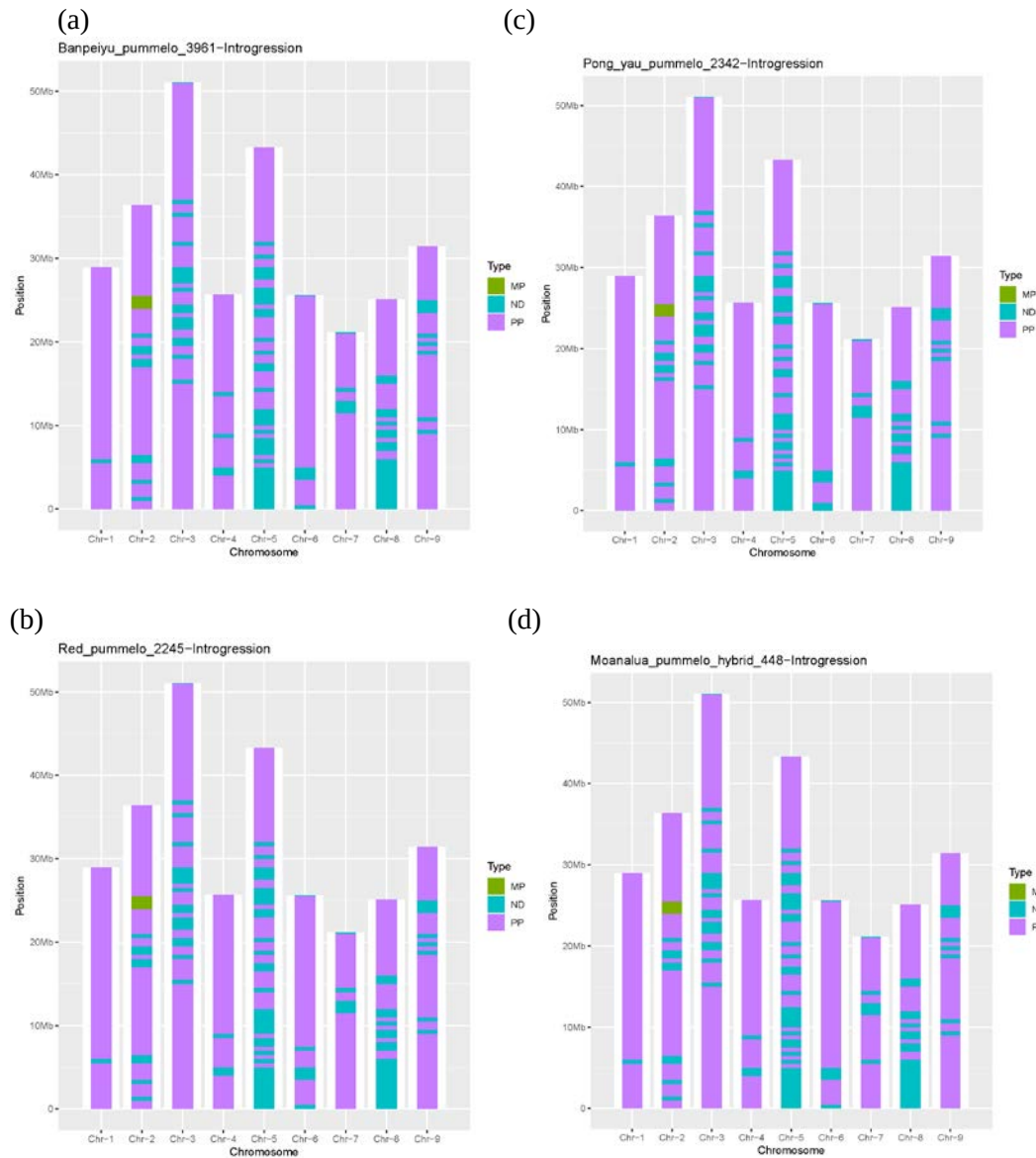


Figure 23. Graphical genotyping for Chandler type pummelo accessions showing the pattern of introgression between mandarin and pummelo at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for a) Banpeiyu CRC 3961, (b) Red pummelo CRC 2245, (c) Pong Yau CRC 2342, (d) Moanalua CRC 448, and (e) Pink pummelo CRC 2244



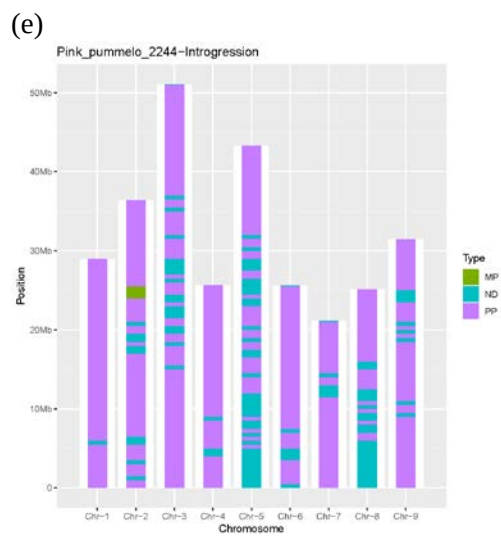


Figure 24. Graphical genotyping for Washington navel CRC 1241A showing the pattern of introgression between pummelo and mandarin at the chromosome level

Chromosome level introgression was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

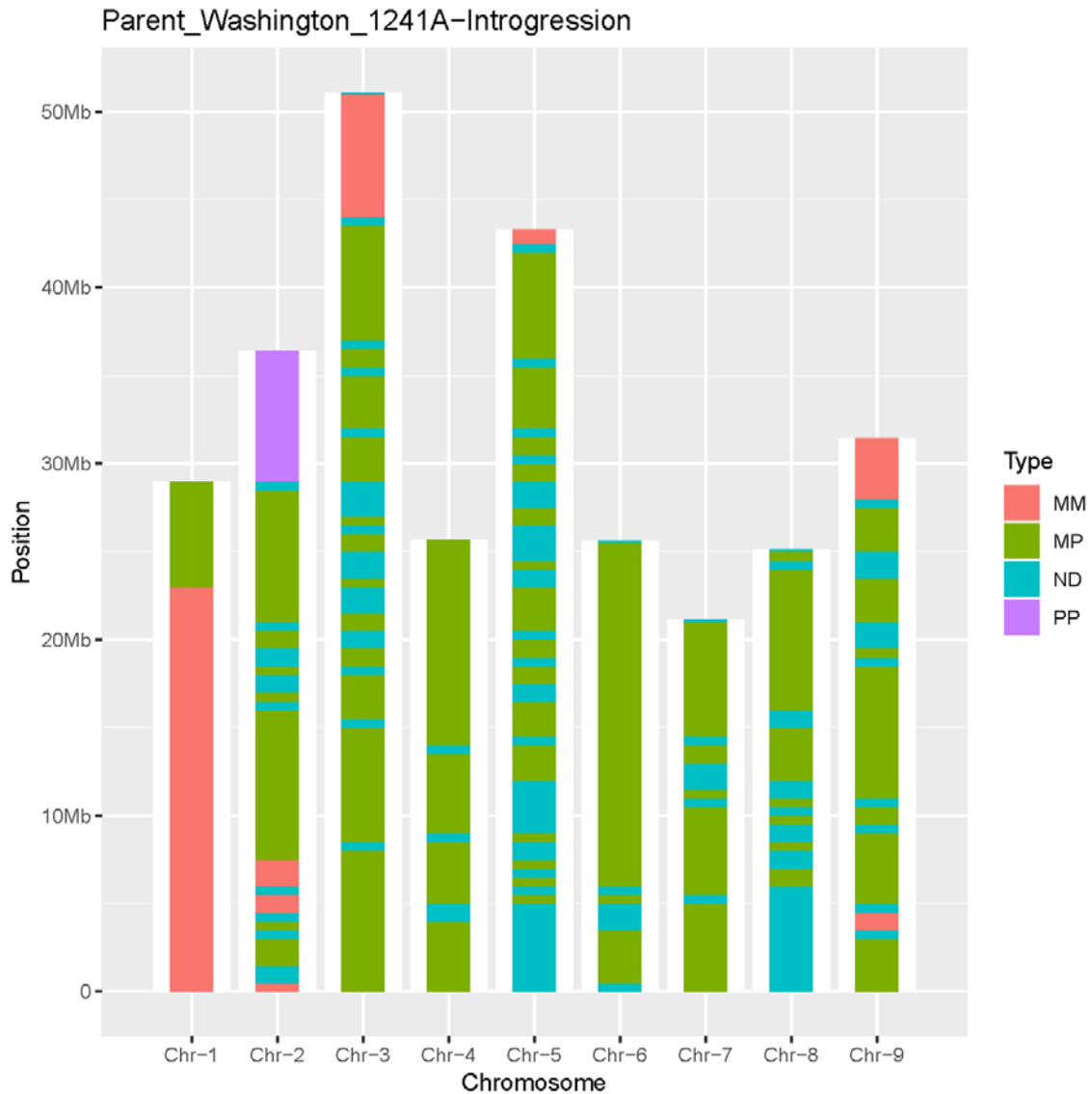


Figure 25. Graphical genotyping for Ambersweet CRC 4002 showing the pattern of introgression between pummelo and mandarin at the chromosome level

The pattern of introgression between pummelo and mandarin at the chromosome level was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class.

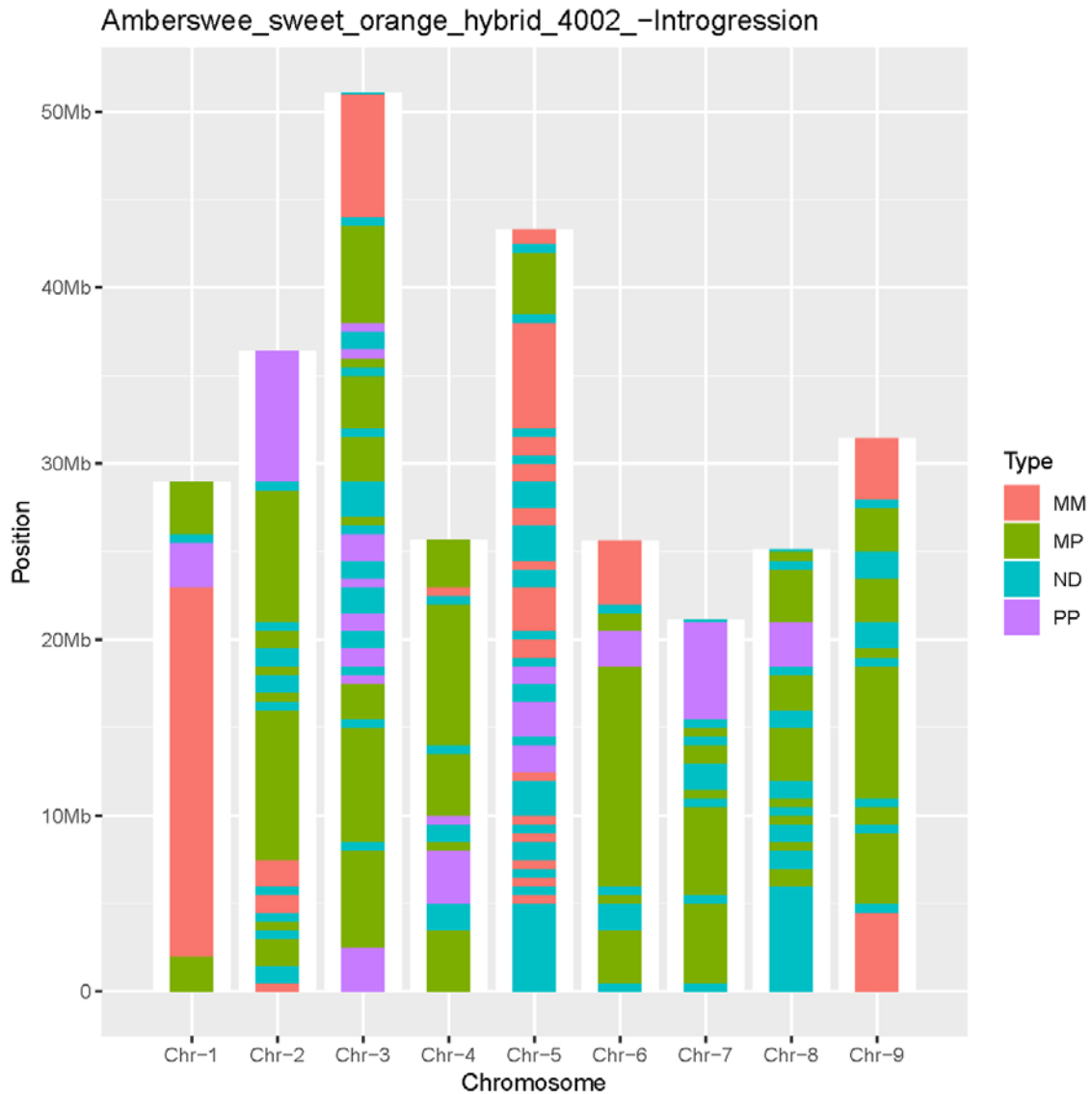


Figure 26. Graphical genotyping for Indian CRC 3175 and *C. maderaspatana* CRC 3225 showing the pattern of introgression between papeda and mandarin at the chromosome level

The pattern of introgression between pummelo and mandarin at the chromosome level was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and papeda allele (green), homozygous for papeda allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for (a) Indian CRC 3175 and (b) *C. maderaspatana* CRC 3225

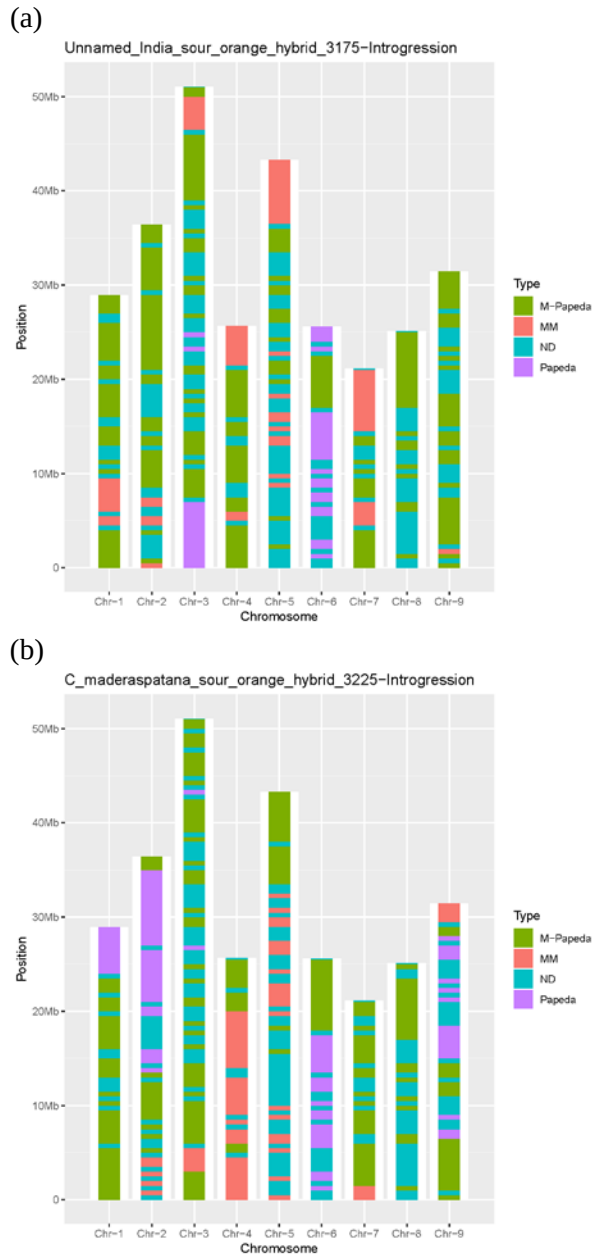


Figure 27. Graphical genotyping for *C. intermedia* CRC 3474 and Orogold CRC 693 showing the pattern of introgression between pummelo and mandarin at the chromosome level

The pattern of introgression between pummelo and mandarin at the chromosome level was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for (a) *C. intermedia* CRC 3474 and (b) Orogold CRC 693

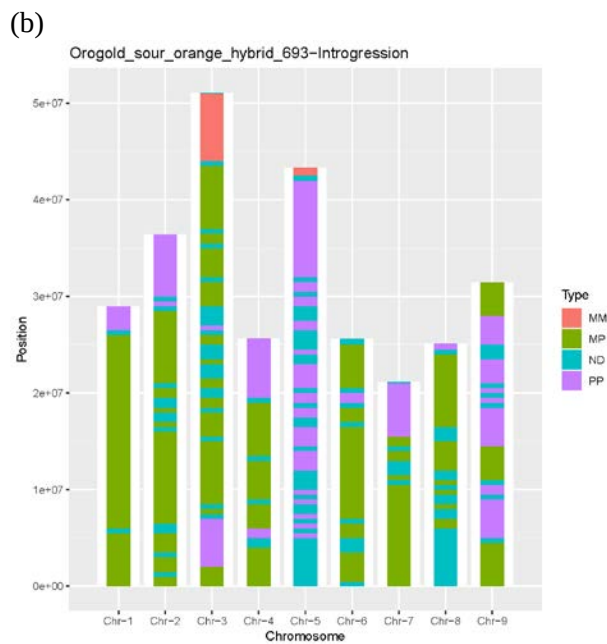
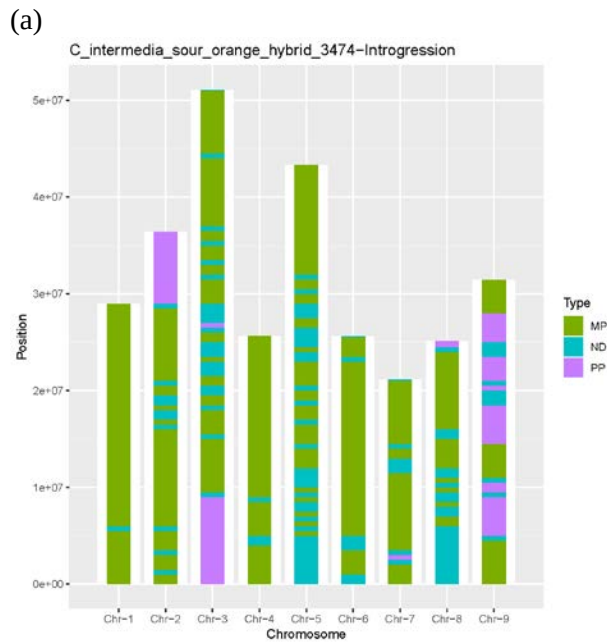
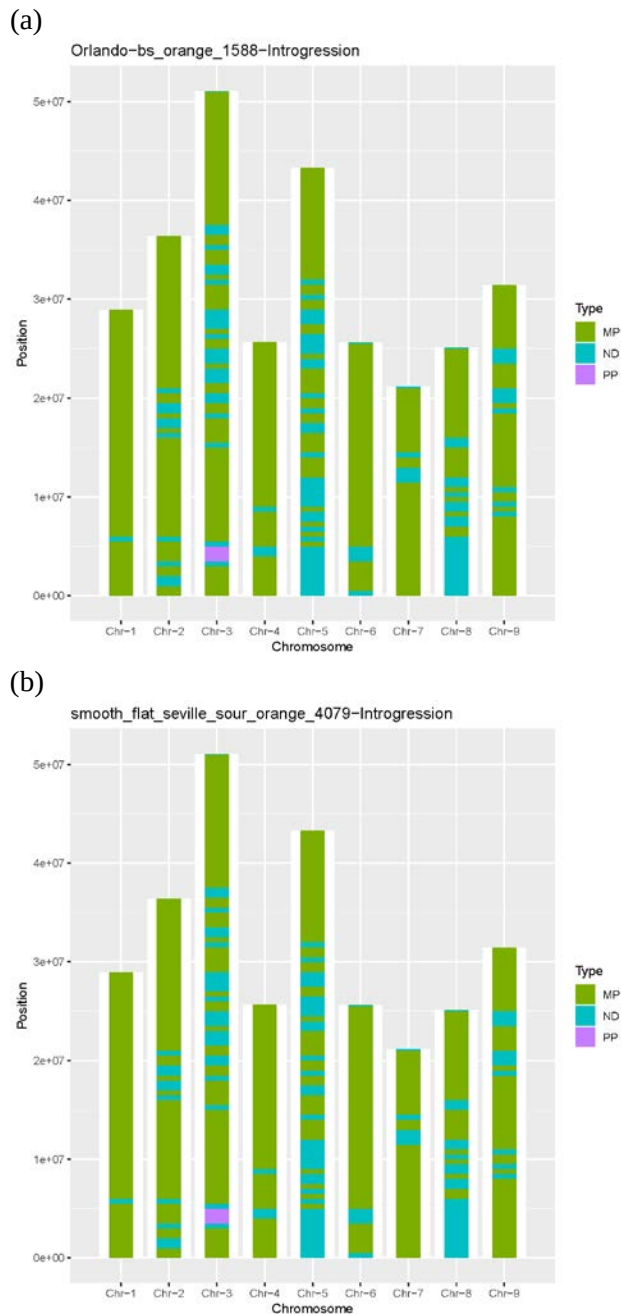


Figure 28. Graphical genotyping for Orlando CRC 1588, Smooth flat Seville CRC 4079 showing the pattern of introgression between pummelo and mandarin at the chromosome level

The pattern of introgression between pummelo and mandarin at the chromosome level was plotted by using species specific allele found in putative ancestral taxa and each bin (0.5 Mb) was assigned homozygous for mandarin allele (red), admixed (heterozygous) for mandarin and pummelo allele (green), homozygous for pummelo allele (purple) and undetermined (teal) based on the frequency of SNP loci in each genotype class for (a) Orlando CRC 1588 and (b) Smooth flat Seville CRC 4079



Supplemental Table 1

Pairwise Fst values for 7 populations studied in the citrus diversity study

The level of genetic divergence estimated within non-admixed citron, kumquat, mandarin, Microcitrus, papeda (micrantha type), pummelo and trifoliate (a) mean Fst (b)weighted Fst

(a)

mean Fst	citron	kumquat	mandarin	microcitrus	papeda	pummelo	trifoliate
citron	+	+	+	+	+	+	+
kumquat	0.809629	+	+	+	+	+	+
mandarin	0.826847	0.5848	+	+	+	+	+
microcitrus	0.782221	0.514582	0.575175	+	+	+	+
papeda	0.77866	0.574773	0.620023	0.531792	+	+	+
pummelo	0.640989	0.530565	0.666571	0.519779	0.507857	+	+
trifoliate	0.844933	0.741536	0.722331	0.706593	0.72608	0.598577	+

(b)

weighted Fst	citron	kumquat	mandarin	microcitrus	papeda	pummelo	trifoliate
citron	+	+	+	+	+	+	+
kumquat	0.922384	+	+	+	+	+	+
mandarin	0.93381	0.774507	+	+	+	+	+
microcitrus	0.918474	0.758414	0.775258	+	+	+	+
papeda	0.910421	0.774089	0.805703	0.766427	+	+	+
pummelo	0.831869	0.708717	0.798419	0.702213	0.684647	+	+
trifoliate	0.940793	0.883116	0.876448	0.880475	0.877535	0.773876	+

Supplemental Table 2. Sample ID and Sample names for accessions in citrus variant discovery panel used for network analysis

ID	Sample Name CRC
1	Sun Chu Sha mandarin 4003
2	Triumph grapefruit 297
3	Citrus tachibana mandarin 3150
4	Ckinokuni mandarin 3887
5	Citrus sunki mandarin 3143
6	Cleopatra 3844
7	W. Murcott mandarin 3953
8	Clementine de nules 4056
9	Pineapple sweet orange 3858
10	Willowleaf mandarin 3843
11	Changsha mandarin 3577
12	Dancy Frost mandarin 3026
13	Kinokuni mandarin 696
14	Buddha's hand Citron 3768
15	Yunnanese citron 3798
16	Flying Dragon trifoliolate 3330A
17	Siamese Acidless pummelo 2240
18	Chandler pummelo 3224
19	Kao Phuang pummelo 2352
20	Thong Dee pummelo 3927
21	Mato Buntan pummelo 3945
22	Hunnan pummelp 1225
23	Duncan grapefruit 3832
24	Citrus hystrix 3103
25	Mexican lime 1710
26	Schaub rough lemon 3879
27	Frost Lisbon lemon 3176
28	Etrog citron 3891
29	Tahitian pummelo 3806
30	Citrus microcarpa 3605
31	Pomeroy trifoliolate 1717
32	Standard sour orange 628
33	Eremocitrus Glauca 3463
34	Microcitrus australis 3673
35	Nagami kumquat 3877
36	Microcitrus australasica 484

Supplemental Table 3. Delta and Qresidual scores

The amount of reticulate (non-tree like) signal in the 36 accessions in variant discovery panel

CVC Sample Name	Delta score	Q-residual
Clementine de Nules 4056	0.27832	0.02364
W. Murcott mandarin 3953	0.24504	0.018234
Citrus tachibana mandarin 3150	0.22005	0.017077
Changsha mandarin 3577	0.21883	0.016481
Dancy Frost mandarin 3026	0.22252	0.016628
Citrus kinokuni mandarin 3887	0.20776	0.016441
Willowleaf mandarin 3843	0.21976	0.016763
Kinokuni mandarin 696	0.20156	0.016207
Citrus sunki mandarin 3143	0.20371	0.016833
Sun Chu Sha mandarin 4003	0.20528	0.017439
Cleopatra 3844	0.20611	0.016929
Standard sour orange 628	0.27713	0.039835
Pineapple sweet orange 3858	0.2764	0.037614
Triumph grapefruit 297	0.28766	0.04518
Duncan grapefruit 3832	0.28743	0.046413
Hunan pummelp 1225	0.19134	0.021018
Mato Buntan pummelo 3945	0.19041	0.021381
Siamese Acidless pummelo 2240	0.19164	0.021097
Chandler pummelo 3244	0.18693	0.021617
Thong Dee pummelo 3927	0.18797	0.021491
Kao Phuang pummelo 2352	0.18839	0.021768
Tahitian pummelo 3806	0.18667	0.021444
Schaub rough lemon 3879	0.28038	0.061097
Frost Lisbon lemon 3176	0.28634	0.03836
Flying dragon trifoliolate 3330A	0.24917	0.016986
Pomeroy trifoliolate 1717	0.2487	0.016844
Citrus hystrix 3103	0.22538	0.015908
Citrus micrantha 3605	0.21549	0.017275
Nagami kumquat 3877	0.27577	0.014271
Eremocitrus Glauca 3463	0.22949	0.015639
Microcitrus australasica 484	0.21473	0.016288
Microcitrus australis 3673	0.22734	0.016545
Mexican lime 1710	0.19765	0.030907
Yunnanese citron 3798	0.21096	0.026303
Buddha's hand Citron 3768	0.2156	0.027009
Etrog citron 3891	0.21484	0.026824

Supplemental Table 4. Delta and Qresidual scores
The amount of reticulate (non-tree like) signal in the 382 UCR diversity panel

382 CVC diversity panel	Delta Score	Q-residual
African pummelo 2346	0.26569	0.013006
African Shaddock x Rubidoux citrumelo 3969	0.32667	0.019526
Alemoen pummelo 2487	0.27071	0.012152
Algerian clementine 279	0.29026	0.010107
Allspice tangelo 4143	0.33516	0.011817
Altoona tangor 2792	0.32741	0.010873
Ambersweet sweet orange hybrid 4002	0.34297	0.011476
Amoa 8 blood mandarin 4237	0.29714	0.009573
Anseikan pummelo 3979	0.29944	0.011248
Arajon pummelo 2596	0.32287	0.013177
Argentine trifoliate 3206	0.31814	0.016526
Asahikan pummelo hybrid 3905	0.33183	0.014073
Assads citron RSD2008001	0.28458	0.014814
Badhri citron 3723	0.3313	0.013379
Bahman Persian 1 sweet lime 4174	0.31213	0.011861
Bahman Persian 2 sweet lemon 4187	0.31217	0.011851
Baiju nucellar RRUT76	0.29326	0.011495
Banokan pummelo 3950	0.3224	0.012969
Banpeiyu pummelo 3961	0.26761	0.012871
Baygold tangelo 4186	0.32541	0.009901
Bengal citron 3055	0.3273	0.017783
Hassaku Beni pummelo hybrid 3907	0.32625	0.013777
Benton citrange 3908	0.30631	0.011803
Bergamot sour orange 2881	0.3453	0.012973
Bitrouni lemon 3265	0.33411	0.019781
Bitters C22 citrange 4219	0.31685	0.015638
Bower mandarin 3649	0.28651	0.009797
Buddha's Hand citron 3768	0.28235	0.015177
BuOl 1 acidless yellow pummelo RRUT92	0.27249	0.011836
C. ichangensis 3931	0.323	0.012709
C. indica 3163	0.30076	0.013344
C32 citrange 3911	0.30648	0.010931
C35 citrange 3912	0.31855	0.012607
Cabuyo Papeda 432	0.31341	0.013928
Calushu mandarin hybrid 2867	0.33325	0.010679
Carrizo citrange 2863	0.3099	0.011328

Chandler pummelo 3244	0.26819	0.01343
Changsha mandarin 3577	0.30588	0.012054
Chinese pummelo 1212	0.30196	0.010458
Chinese pummelo 1224	0.27169	0.012046
Chinotto sour orange 3728	0.35438	0.011665
Chironja grapefruit hybrid 3909	0.34355	0.01382
Citradarin trifoliolate hybrid 2618	0.31713	0.013545
Citradia trifoliolate hybrid 1437	0.31742	0.012672
Citradia trifoliolate hybrid 1438	0.31568	0.012474
Citraldin trifoliolate hybrid 2629	0.30996	0.012865
Citrangle #1416 2866	0.30271	0.011763
Citranglequat 19 15 7 3259	0.33589	0.011251
Citrangor trifoliolate hybrid 1447	0.31379	0.011335
Citremon trifoliolate hybrid 1448	0.3326	0.014786
Citremon trifoliolate hybrid 1449	0.32729	0.014343
Citron of Commerce 3518	0.2812	0.015444
Citrumelo 1452	0.32945	0.015907
Citrumelo 3348	0.32035	0.012662
C. amblycarpa 2485	0.33493	0.014688
C. benikoji 3149	0.33707	0.010242
C. canaliculata 3565	0.30709	0.010587
C. depressa 2448	0.31972	0.012265
C. erythrosa 3292	0.30506	0.011892
C. excelsa 2316	0.28202	0.01465
C. funadoko 3274	0.33546	0.011488
C. grandis 2340	0.26712	0.012494
C. grandis 3282	0.26462	0.013072
C. grandis 2453	0.26862	0.012335
C. hanayu 3469	0.33139	0.0102
C. hiroschimana Natzuzabon 3275	0.30609	0.011277
C. hongheensis 3797	0.32108	0.012644
C. hystrix 3103	0.31199	0.014238
C. ichangensis 2327	0.32321	0.012759
C. intermedia 3474	0.33225	0.014727
C. keraji 3144	0.31958	0.008854
C. kharna 3059	0.30985	0.01301
C. kinokuni 3887	0.31658	0.01403
C. latipes 3052	0.32101	0.012444
C. leiocarpa 3147	0.31313	0.009088
C. lycopersicaformis 3564	0.32429	0.014556

C. macrophylla 2446	0.28248	0.014632
C. maderaspatana 3225	0.33539	0.011464
C. megaloxycarpa hybrid 3241	0.32484	0.016134
C. micrantha 3605	0.31395	0.015146
C. natsudaikai 3235	0.33618	0.014389
C. nippokoreana 3228	0.33063	0.010625
C. obovoidae 3465	0.32506	0.013395
C. otachibana 3470	0.32228	0.01252
C. pennivesiculata 2434	0.28211	0.013808
C. rokugatsu 3473	0.34929	0.010379
C. rugulosa 3556	0.34832	0.015092
C. shunkokan 3476	0.33957	0.011135
C. sino-grandis 3148	0.29802	0.011053
C. succosa 3280	0.33673	0.013355
C. tachibana 3150	0.33226	0.01251
C. tengu 3464	0.34202	0.012273
C. ujukitsu 3467	0.33465	0.011763
C. volkameriana 3050	0.32016	0.018452
C. vulgaris 760	0.28297	0.014144
Citrus webberii var Montana 767	0.33502	0.014194
C. yatsushiro 3880	0.31839	0.009906
C. yuko ichandarin 3146	0.32449	0.008861
Clement tangelo 2603	0.3482	0.011427
Clementine X Pearl 4020	0.32895	0.009906
Cleopatra mandarin 3844	0.31699	0.013685
Clementine X Silverhill 3731	0.31316	0.009969
Cocktail pummelo hybrid 3555	0.3436	0.01519
Corsican citron 1795	0.28368	0.014819
C. sunki mandarin 3143	0.30969	0.013274
Cuban Shaddock pummelo hybrid 1462	0.33918	0.019941
Cunningham citrange 271	0.30833	0.010651
Dabbe citron hybrid 3190	0.32032	0.015089
Daisy mandarin 3910	0.3012	0.008908
Dancy Weshart mandarin 602	0.29925	0.012638
Davao papeda hybrid 2427	0.28197	0.014634
Deep Red pummelo 2347	0.26602	0.012886
Diamante citron 3523	0.2813	0.015499
Djerok citron hybrid 2456	0.3347	0.022685
Dulcia citron 3654	0.28494	0.014825
Duncan grapefruit 3832	0.33771	0.017144

Dweet tangor 3018	0.31268	0.010084
Early tangelo 2560	0.31609	0.010107
Egami Buntan pummelo 3959	0.27632	0.011983
Emerald mandarin 4123	0.29704	0.009498
Empress mandarin 3613	0.29407	0.011982
Encore mandarin 3569	0.28766	0.010176
Etrog citron 3891	0.28322	0.015
Eustis Limequat 3603	0.3034	0.014456
Excalibur red lime hybrid 4239	0.34186	0.011129
Fairchild mandarin 3559	0.2858	0.009175
Fallglo mandarin 3990	0.29343	0.008936
Fleming's Shaddock pummelo 578	0.26628	0.012389
Flying dragon trifoliate 3330A	0.31946	0.017465
Fortune mandarin 3560	0.28595	0.00887
Fortunella obovata 3475	0.3273	0.011559
Fortunella polyandra 3901	0.32483	0.014198
Fremont mandarin 3558	0.29362	0.009714
Frost Lisbon 3176	0.30376	0.013782
Frost Owari satsuma 3178	0.31814	0.00982
Frua mandarin 3022	0.29639	0.010275
Furr C57 citrange 4221	0.31774	0.015774
Glen citrangedin trifoliate hybrid 3573	0.32776	0.0109
Gold Nugget mandarin 3913	0.2922	0.008789
Goutoucheng sour orange hybrid 4004	0.34153	0.012486
GouTouXiangYuan RRUT395	0.28235	0.014852
GouTouXiangYuan RRUT396	0.28156	0.01504
GouTouXiangYuan RRUT398	0.28301	0.014943
H 56 tangor 3096	0.30841	0.0099
Hall grapefruit x Rubidoux 3889	0.3182	0.012118
Hansen mandarin 4152	0.29467	0.009915
Hickson mandarin 4161	0.29349	0.010665
Hirado Buntan 4122	0.26425	0.012675
C. hiroshima Natsuzabon grapefruit 3904	0.33344	0.014279
Honey mandarin 3177	0.29456	0.010792
Huang Yen Man Chie mandarin 3897	0.31347	0.013013
Hung Kat mandarin 2331	0.2973	0.011257
Hunan pummelo 1225	0.26844	0.012357
Ichang lemon 1215	0.32708	0.010444
Imperial mandarin 4160	0.28834	0.010722
Improved Meyer lemon 3737	0.32425	0.013337

India lemon 2323	0.31812	0.018355
Indian citron 138	0.28214	0.015039
Indian citron 661	0.28263	0.013862
Indian rough lemon 2544	0.31863	0.018339
Indio mandarin 3759	0.33277	0.011245
Interdonato lemon hybrid 3593	0.29707	0.012466
Iran lemon 3162	0.31712	0.016477
Iraq lemon limetta 3492	0.31224	0.011931
Italian citron 3530	0.28111	0.015381
Itoshima 3928	0.26545	0.012413
Iyo San Ponkan 4118	0.33666	0.013263
Iyomikan tangor 3980	0.33605	0.011192
Ugli 4127	0.34134	0.012116
Jullander Khatti rough lemon 4173	0.31864	0.018387
Kabosu ichandarin 3943	0.33495	0.008801
C. webberri kalpi 1456	0.31332	0.014824
Kao Phuang pummelo 3926	0.26365	0.012383
Kao Phuang pummelo 2352	0.26544	0.01293
Kao Ruan Tia pummelo 2350	0.2669	0.013038
Kao Ruan Tia pummelo 2351	0.26735	0.013186
Kara mandarin 3019	0.30197	0.00971
Karn Lau Yau pummelo 2341	0.26872	0.012167
Kawachi banokan 3944	0.31514	0.011757
Khoubs El Arsa lemon 2477	0.31011	0.013201
Khoubs El Ers lemon 2489	0.31029	0.013164
Kincy mandarin 4144	0.29173	0.009462
King mandarin 303	0.30156	0.009723
Kinnow mandarin 3021	0.28743	0.010385
Kinokawa Buntan pummelo 3978	0.33389	0.013827
Kiyomi tangor 3952	0.33287	0.010784
Kinokuni mandarin 696	0.30611	0.01261
Koster tangor 3958	0.29194	0.00961
Kulobot papeda hybrid 3612	0.33241	0.01206
Kulu lemon 3045	0.31066	0.011452
Kunembo mandarin 3346	0.32494	0.011525
Lapithiokii lemon 4005	0.30607	0.013136
Laranja Cravo Ponkan mandarin 2983	0.292	0.009982
Lee mandarin 3851	0.28649	0.009041
Lemelo pummelo hybrid 1481	0.33688	0.014172
Lemelo pummelo hybrid 1775	0.30936	0.015061

lemonime lemon hybrid 2459	0.33437	0.012548
Limau Matu 3755	0.28508	0.014687
Limonette De Marrakech lemon hybrid 3989	0.31233	0.013784
Unnamed lemon 3885	0.31243	0.013025
Lumina lemon 3925	0.33871	0.016006
M. australasica 3672	0.31463	0.013332
Mandarinette 3405	0.28864	0.009819
Mapo tangelo 4222	0.31722	0.010039
Marlow tangelo 4145	0.3326	0.011584
Marumi SRA489 kumquat 4125	0.32087	0.014687
Mato Buntan pummelo 3945	0.26797	0.012879
Mency tangor 3183	0.2921	0.009436
Mexican citron 3531	0.28584	0.014463
Mexican lime 1710	0.28181	0.015069
Michal mandarin 4049	0.28505	0.010233
M. australis 3665	0.31588	0.01372
Microcitrus x trifoliae trifoliae hybrid 3566	0.31166	0.015015
M. australasica 3670	0.32226	0.012938
Milam rough lemon 3396	0.361	0.011845
Minneola tangelo 3340	0.30501	0.01066
Mitha Tulia limetta 3051	0.31756	0.017947
Moanalua pummelo hybrid 448	0.26688	0.012729
Moli kurikuri pummelo hybrid 579	0.33946	0.016783
Monoembryonic trifoliolate 3888	0.32141	0.015458
Morningsong Temoni citron-Karp	0.28386	0.01489
Morton citrange 1463	0.31122	0.010988
Murcott mandarin 3846	0.29247	0.009373
Mytrifolia sour orange 3290	0.36087	0.012078
Nakhon Chaisi pummelo 2353	0.26546	0.013092
Nansho Daidai sour orange 2588	0.34729	0.012974
Necked mandarin 3738	0.31284	0.009262
New Zealand lemonade lemon 4181	0.35173	0.011929
Nicaraguan lemon hybrid 3841	0.28019	0.014712
NISSV E mandarin hybrid 3965	0.29181	0.009611
Nordmann kumquat 4070	0.32152	0.014734
Nova mandarin 3615	0.29801	0.009322
Novelty x Ellendale mandarin hybrid 3956	0.31511	0.009389
Odorata citron 3655	0.28445	0.014809
Orban lemon hybrid 4140	0.37535	0.013361
Orlando bittersweet orange 1588	0.3458	0.019047

Orlando tangelo 3874	0.30654	0.010217
Orogold sour orange hybrid 693	0.344	0.012784
Ortanique tangelo 3368	0.31597	0.010601
Paak Ling Mung lemon hybrid 3193	0.31554	0.01738
Page mandarin 3616	0.31011	0.009962
Pan Dan pummelo 2752	0.26613	0.012913
Papuan citron 3532	0.28435	0.014781
Palazzelli mandarin 4244	0.29995	0.009056
Parent Washington 1241A	0.34842	0.016618
Parson's Special mandarin 300	0.29903	0.009348
Pauthel pummelo 3949	0.267	0.012206
Pearl tangelo 2849	0.34269	0.013264
Peretta lemon hybrid 3924	0.30038	0.012792
Perrine lemon hybrid 2458	0.28895	0.012952
Philippine citron 3533	0.28463	0.014609
Philippine pummelo 644	0.27029	0.012548
Philippine pummelo hybrid 2343	0.34478	0.014677
Philippine mandarin hybrid 2318	0.31527	0.017332
Pin Shan Kong Yan pummelo 2348	0.26773	0.013101
Pink pummelo 2246	0.26607	0.012602
Pink pummelo 2244	0.26709	0.012678
Pixie mandarin 3568	0.30223	0.009362
Pomelit pummelo hybrid 4026	0.26761	0.012414
Pomona sweet lemon 4068	0.31146	0.014028
Ponderosa lemon hybrid 294	0.34414	0.023591
Pong yau pummelo 2342	0.27807	0.011932
Primofiore lemon 3491	0.30334	0.013944
Prior Lisbon lemon 3505	0.30362	0.013792
Procimequat kumquat hybrid 3295	0.321	0.012616
#26 trifoliate 3939	0.31982	0.017222
#27 trifoliate 3938	0.32009	0.016997
Puma pummelo hybrid 3941	0.31163	0.012894
Pursha lime x Chinotto 3578	0.35018	0.01482
Rangpur x Troyer 3997	0.35006	0.011564
Red pummelo 2245	0.26704	0.012559
Reinking pummelo 3805	0.26432	0.012735
Rex Union grapefruit hybrid 3962	0.34014	0.016424
Richard's Special mandarin 4117	0.28975	0.010021
Rio Grande Valley lemonsquat 4097	0.33588	0.010012
Robinson mandarin 3850	0.28969	0.009348

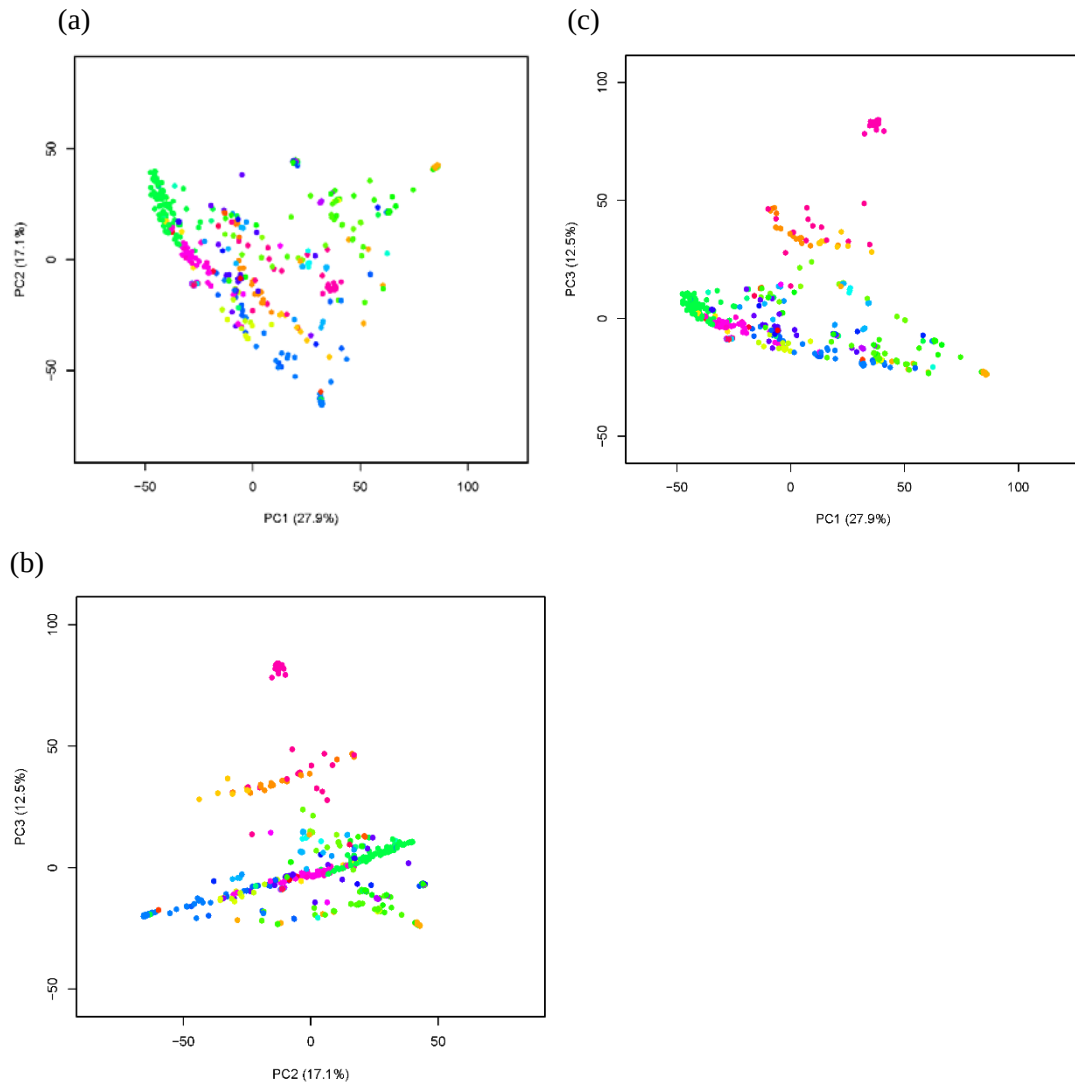
Roeding's Pink pummelo 1208	0.26885	0.012609
Rojo Blanco grapefruit hybrid 4029	0.30946	0.012879
Royal grapefruit 248	0.32282	0.013977
Rubidoux mandarin 4031	0.29063	0.009352
Rubidoux pummelo hybrid #1 4028	0.30943	0.012874
Ruddy tangor 4146	0.32907	0.009629
Rusk citrange 1441	0.30426	0.011033
S281 citrangelo 3552	0.34996	0.012034
S302 citranguma 3415	0.32166	0.012315
Sacaton citrumelo 3337	0.32107	0.013693
Sacaton citrumelo 3417	0.32049	0.012646
Sacaton tangelo 2609	0.31874	0.010829
Sacaton tangelo 3331	0.34683	0.013521
Sampson tangelo 2418	0.31854	0.010849
San Jacinto tangelo 2011	0.30669	0.009149
Sanford citrange 276	0.30772	0.010922
Santa Barbara red rangpur 712	0.31568	0.01761
Santa Barbara tangelo 4073	0.33442	0.011057
Savage citrange 275	0.31331	0.011381
Scarlet Emperor mandarin 3226	0.29949	0.012868
Schaub rough lemon 3879	0.32068	0.019024
Seminole tangelo 2604	0.31401	0.011412
Sexton tangelo 2543	0.32864	0.014096
shaddock x St. Michael pummelo hybrid 42	0.33082	0.014488
Siamelo tangelo 2586	0.30902	0.009978
Siamese Acidless pummelo 2240	0.26772	0.01279
Siamese pummelo 2421	0.3087	0.015292
Sicily citron 3534	0.28122	0.015431
Simmons trifoliolate 3549	0.32192	0.016814
Sinton citrangequat 3642	0.32817	0.011057
Siranui Dekopan 4113	0.30857	0.008991
Small Leaf trifoliolate 4017	0.32199	0.01708
smooth flat seville sour orange 4079	0.34166	0.014564
Sniff acidless Rangpur 2484	0.31519	0.017434
Soh Long lemon hybrid 3199	0.28248	0.013815
Soh Niangrang Papeda 3203	0.33585	0.008597
Soh Synteng lemon 3261	0.32468	0.015073
Solid Scarlet mandarin 3328	0.29302	0.010855
Som Kao 3852	0.2949	0.009604
Sour pummelo 3066	0.28287	0.012936

Citrus aurantium var salicifolia 3289	0.36019	0.012329
South African rough lemon 2325	0.32024	0.018839
South Coast Field Station citron 3546	0.28112	0.0156
SRA 337 RRUT223	0.29653	0.009009
Standard sour orange 628	0.34591	0.019211
Suisho Buntan 3947	0.29059	0.011506
Sukega grapefruit 4149	0.34292	0.012936
Sunburst mandarin 3809	0.29724	0.009023
Sun Chu Sha mandarin 4003	0.3126	0.014353
Sunrise tangelo 2606	0.31847	0.010082
Sunshine pummelo 2236	0.27042	0.012958
Sunshine tangelo 2788	0.32153	0.011075
Sweet lemon 3093	0.31238	0.011805
Sweet pummelo 3067	0.26885	0.012554
Swingle citrumelo 3767	0.3209	0.012388
Sydney hybrid 1485	0.31598	0.01335
Szinkom mandarin 3085	0.30242	0.012414
Tahitian pummelo 3806	0.26299	0.012751
Tango mandarin 4183	0.28994	0.009995
Tankan tangor 2224	0.30164	0.009567
Tau Yau pummelo 2583	0.27002	0.012582
Tavares kumquat hybrid 4048	0.31065	0.014178
Temple tangor 2598	0.31684	0.01089
Thomasville citrangequat 1440	0.32357	0.01084
Thong Dee pummelo 4027	0.2627	0.012717
Thong Dee pummelo 3927	0.265	0.012848
Thorton tangelo 2013	0.30867	0.010639
Tien Chieh mandarin 2590	0.30297	0.0099
Tieu nucellar mandarin RRUT87	0.29411	0.010487
Tresca pummelo 302	0.33171	0.014897
Trifeola trifoliolate hybrid 3954	0.30896	0.010597
Triumph grapefruit 297	0.32509	0.014697
Unknown lemon hybrid 3172	0.31482	0.017205
Unnamed Chinese pummelo RRUT223	0.26722	0.012703
Unnamed citron 3174	0.28491	0.014909
India sour orange hybrid 3175	0.33431	0.012056
Unnamed lemon hybrid 3154	0.31704	0.013178
Unnamed papeda 3056	0.32143	0.01241
Kao Pan pummelo 2248	0.27417	0.012264
Kao Pan pummelo 2355	0.26415	0.012732

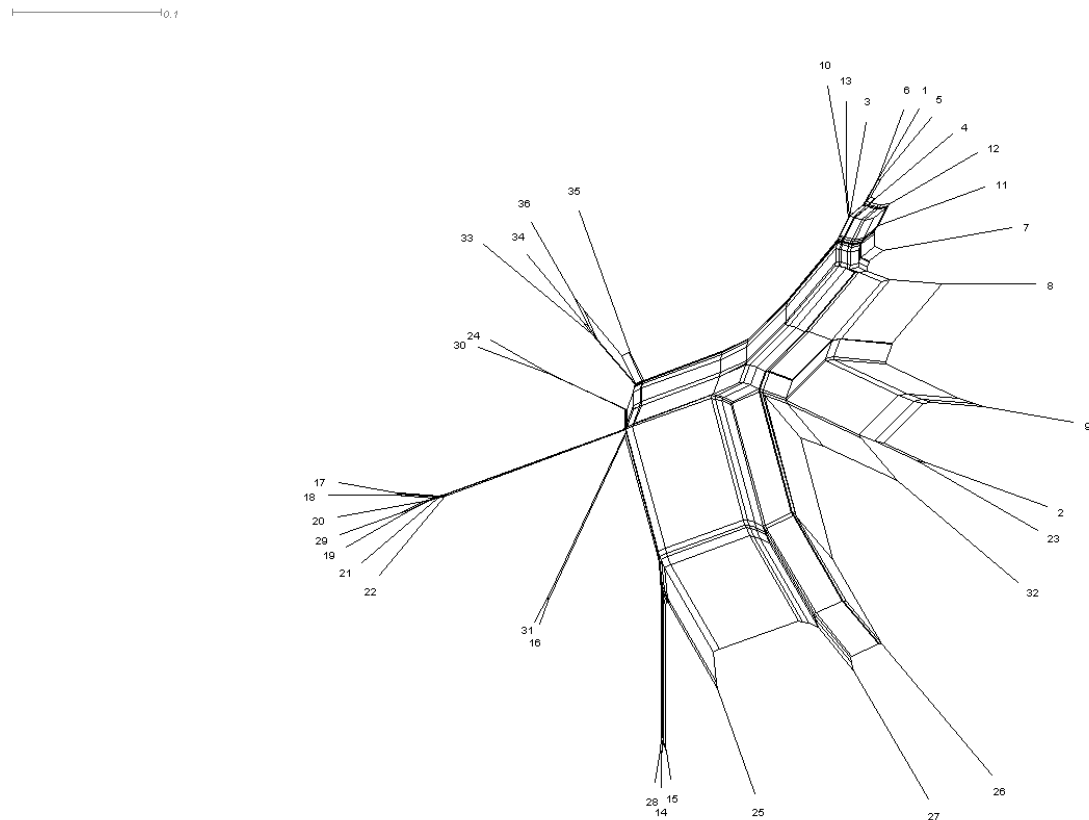
Unnamed sweet orange hybrid 648	0.36983	0.012696
Unnamed lemon hybrid 3155	0.3174	0.013056
US 119 3998	0.34822	0.011837
US 852 4245	0.30916	0.013613
UC Seedless Surprise USDA 17 7 19 4108	0.33555	0.012374
USDA 6-15-150 mandarin 4165	0.28962	0.008885
US Superna USDA 88 2 3991	0.29509	0.008844
USDA 88 3 mandarin hybrid 3992	0.29318	0.009648
Valentine pummelo hybrid 4182	0.32946	0.013643
Vangasay 807 rough lemon 3996	0.32105	0.018656
Warren limequat 3069	0.30345	0.014411
Webber Shaddock pummelo 4233	0.26562	0.012884
Webber tangelo 2746	0.32583	0.011252
Wheeny grapefruit hybrid 2885	0.33935	0.014482
Wilking mandarin 3020	0.28562	0.009646
Willial tangelo 4147	0.32463	0.010554
Williams tangelo 2787	0.32003	0.01092
Willowleaf mandarin 3843	0.29411	0.012084
Wincia tangor 4148	0.2964	0.009495
X639 trifoliate hybrid 3957	0.31689	0.01636
Yalaha tangelo 2559	0.31203	0.009769
Yellow Rind mandarin 3895	0.32163	0.012975
Yemen citron 2847	0.28392	0.014727
Yunnanese citron 3798	0.282	0.015141
Yuzu ichandarin 1216	0.33145	0.009991
Zhuluan sour orange 3930	0.34403	0.012873

Supplemental Figure 1. CVC924 PCA

PCA was carried out with 924 samples genotyped at PHR and NMH loci with Axiom® Citrus Genotyping Array. PC1(27.9%), PC2(17.1%), and PC3(12.5%) explained 57.5 % of the total variation that exists in the data. (a) PCA PC1 vs PC2. (b) PCA PC2 vs PC3. (c) PCA PC1 vs PC3.



Supplemental Figure 2. Network analysis of 36 accessions from citrus variant discovery panel
 Genotyping calls of 43,614 loci were used for the Neighbor-net analysis which was carried out with default setting for the 36 accessions from citrus variant discovery panel using SplitsTree v4.14.6 to investigate a reticulated structure in citrus. The numeric codes for the accession are identified in Supplemental Table 2.



CHAPTER 3

Distribution and Inheritance of Nucellar Embryony Loci in Citrus and its close relatives

Abstract

Nucellar embryony is an important trait in citrus because it significantly influences strategies for both scion and rootstock breeding. Some of seedlings arising from a seed of a genotype with nucellar embryony are genetically identical to the seed parent. Citrus production relies on uniform vegetative propagation by grafting a scion onto a rootstock grown from seed, making nucellar embryony a desired trait. However, nucellar embryony also hinders breeding efforts by producing few or no seeds of hybrid origin. A current hypothesis states that the induction of nucellar embryony is under the control of a single dominant locus, while its expression is controlled by minor affecting loci. To better understand the genetic basis of nucellar embryony, a genome-wide association study (GWAS) was conducted separately in *Citrus* and *Poncirus* for two distinct but related traits, nucellar embryony and polyembryony. Seeds collected from field trees were germinated, the number of seedlings per seed was counted as a measure of the frequency of polyembryonic seeds, and the germinated seedlings were genotyped with KASP SNP markers to determine whether they originated from a nucellar or a zygotic embryo. Using a high-density SNP array for citrus to genotype the seed source trees, a major dominant locus was identified for *Citrus* and a separate locus for *Poncirus*. This analysis identified a 101 kb region including a candidate gene CitRWP in *Citrus*, as previously reported. Significant regions identified by GWAS in *Poncirus* included many genes associated with stress response, oxidation-reduction, flower, seed, and embryo development, and six possible candidate genes. Additionally, this 2.2 Mb region identified by GWAS in *Poncirus* overlapped with the region previously identified by AFLP markers. These studies suggested that the genetic control of nucellar embryony in *Citrus* and *Poncirus* is indeed different.

Introduction

Citrus and its close relative, *Poncirus trifoliata*, are some of the most economically important perennial fruit crops in the world. Natural or artificial interspecific hybridization among citrus ancestral species [1, 2] during thousands years of cultivation lead to many commercially important citrus species having highly heterozygous genomes [3]. In addition, some citrus species are self-incompatible, and many commercial varieties are clonally-derived by mutations such as accumulated somatic mutations, sport mutations or nucellar variants [4, 5], further complicating genetic studies in *Citrus*. Citrus production has long relied on clonal vegetative propagation, due to high heterozygosity and long juvenility, to ensure uniform propagation and rapid fruiting of desired genotypes. Nearly all citrus is propagated by grafting a scion onto a seedling rootstock. Nearly all rootstock cultivars are genetically uniform due to selection for nucellar embryony. Nucellar embryony, or adventitious embryony, is a form of apomixis in which somatic embryos arise from the maternal nucellar tissue surrounding the embryo sac [6], therefore, while most apomixis is considered gametophytic [7, 8], nucellar embryony is a sporophytic apomixis. Because multiple nucellar embryos and sometimes the zygotic embryo develop in the same seed, nucellar embryony is associated with polyembryony and this term is sometimes (but incorrectly) used synonymously with nucellar embryony. Nucellar embryony is rare among plants, but common in some citrus species and mango (*Mangifera indica* L.) [9]. Citrus cultivars with nucellar embryony and polyembryony include sweet orange (*C. sinensis*), grapefruits (*C. paradisi*), lemon (*C. limon*), Satsuma (*C. unshiu*), Ponkan (*C. reticulata*) [10], some other mandarins (*C. reticulata*), and most trifoliolate orange cultivars (*P. trifoliata*) [11–13]. These polyembryonic cultivars can have 30 or more embryos developed within a single seed [10, 14, 15]. “Monoembryonic” citrus cultivars which lack nucellar embryony include clementine (*C. clementina*), Kinokuni mandarin (*C.*

kinokuni) [15], citron (*C. medica*), pummelo (*C. maxiuma*), *C. mangshanensis*, and Papeda [16]. There is a report of a monoembryonic cultivar arising from two polyembryonic cultivars [17].

A developmental study of nucellar embryony showed that fertilization events and subsequent endosperm development are required for nucellar embryos to mature and germinate [18], and that the induction of nucellar embryogenesis from the nucellar initial cells is coupled with the fertilization events [18, 19], leading to the development of both zygotic and nucellar embryos alongside each other within a seed. Therefore, one must consider several different scenarios when phenotyping seedlings with nucellar embryony. First, seeds can have one mature nucellar embryo; second, seeds can have one mature zygotic embryo; third, seeds can have multiple mature nucellar embryos; fourth, seeds can have one mature zygotic embryo and one mature nucellar embryo; fifth, seeds can have one mature and multiple mature nucellar embryos [20]; and lastly, seeds with zygotic twins [21–23]. Competition between sexual embryos and nucellar embryos has been reported [15, 18–20], in addition to varying proportions of seedlings of nucellar and zygotic origin [11–13] which is influenced both by genotype and environment [24, 25].

For rootstock propagation, nucellar embryony is often a desired trait, which allows the propagation of true-to-type genotypes. However, for both scion and rootstock breeding, nucellar embryony hinders the breeding effort as breeders obtain few to no hybrid seeds when genotypes with this trait are used as maternal parents. Therefore, citrus breeders must consider two separate strategies for breeding new varieties of rootstocks and scions. Elucidating genetic control for nucellar embryony has been an interest for many citrus breeders all over the world for many generations.

Apomixis in citrus was first observed by Leeuwenhoek in 1719 when he noticed orange seed having multiple embryos. Several hypotheses have been proposed during the last several decades to explain the characteristics and genetics underlying the inheritance of nucellar embryony.

The first reported hypothesis was a single dominant locus control [17], which was later questioned because this hypothesis could not explain nucellar embryony in both *Citrus* and *Poncirus* [26], leading to a two dominant loci control hypothesis [25, 27]. With the arrival of molecular markers, several studies suggested that the presence of nucellar embryony was under a single dominant locus while the proportion of nucellar embryony a given genotype could exhibit was under the control of multiple locus control [28], suggesting the possibility that the genetic control for *Citrus* and *P. trifoliata* was likely different. Subsequent analyses with sequencing technology allowed the identification of the *Citrus* polyembryony locus [10, 29–31]. After the release of *C. sinensis* [32] and *C. clementina* [33] reference genomes, the possible involvement in *Citrus* nucellar embryogenesis of stress related genes and miRNA were reported [34, 35].

The advent of citrus genomic studies allowed identification of the first candidate gene, CitRWP, responsible for nucellar embryony in *Citrus* [36] within the previously identified polyembryony locus [10]. The reported candidate gene CitRWP was highly expressed in the ovules of polyembryonic cultivars and a miniature inverted repeat transposable element (MITE) was found to be inserted in the promoter of CitRWP [36, 37]. These recent findings supported the standing hypothesis that one major dominant locus controls the induction of nucellar embryony in *Citrus*.

However, there hasn't been comparable progress in the investigation of the genetic control of nucellar embryony in *P. trifoliata* since Kapiro and Roose [38] reported their findings. Accessions belonging to *P. trifoliata* and its hybrids with *Citrus* are often used as rootstocks due to their ability to confer desirable traits to the grafted scion such as disease resistance, cold tolerance and dwarfing. Therefore, understanding the inheritance of nucellar embryony in *Poncirus* is important to plan and execute rootstock breeding programs.

The aim of this study was to understand the genetic control underlying nucellar embryony in both *Citrus* and *P. trifoliata* by conducting GWAS separately for *Citrus* and *P. trifoliata*, using

diverse citrus germplasm available from the University of California Riverside Citrus Variety Collection (CVC) and two segregating populations from crosses between monoembryonic cultivar ‘Fortune’ (*C. reticulata*) and polyembryonic cultivar ‘Fairchild’ (*C. reticulata*) and monoembryonic cultivar ‘Chandler’ (*C. maxima*) x polyembryonic *P. trifoliata*. Nucellar embryony phenotypic data was obtained by germinating seed samples from each accession, visually assessing each germinated seed as monoembryonic or polyembryonic, and genotyping each seedling with KASP single nucleotide polymorphism (SNP) markers (LGC. <https://www.biosearchtech.com/>) in order to correctly assign each seedling as of zygotic or nucellar origin. This study identified significant SNPs associated with a major dominant locus underlying nucellar embryony in *Citrus* and *P. trifoliata* as well as promising candidate genes.

Method

Sample preparation for array hybridization

Leaf samples (young flush) of the accessions of *C. reticulata*, *C. sinensis*, *C. aurantium*, *P. trifoliata*, their hybrids with *Citrus*, and two segregating populations F x F (‘Fortune’ x ‘Fairchild’) and Chand x Trifol (“Chandler” x *P. trifoliata* (‘Webber Fawcett’ or ‘Rubidoux’)) were collected in the field from the CVC or breeding program fields in (Riverside, CA) from 2014-2018. Collected leaf samples were placed on ice then returned to the laboratory and the leaf surface was quickly and gently wiped with chloroform prior to drying with silica packs in a sealed, labeled sample bag. Silica dried leaf samples were homogenized with FastPrep-24™ (MP Biomedicals, Santa Ana, CA) with ¼” ceramic spheres or ¼” stainless steel beads. High molecular weight genomic DNA was then isolated from homogenized dried leaf samples with DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). The quality of DNA was assessed by electrophoresis, 260/280 and 260/230 ratios with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) and genomic DNA

was quantified with a Qubit fluorometric system (Thermo Fisher Scientific, Wilmington, USA) according to manufacturer's instructions. Genomic DNA of concentration of 25 ng/μl in 50 μl and 30 ng/μl in 25 μl for Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array, respectively were prepared in Affymetrix™ bar-coded 96 well plates.

Fruit collection, seed extraction and treatment, and germination

Fruit samples (mature fruits) of the accessions from CVC, two segregating populations F x F were collected in the field from the CVC (Riverside, CA) from 2015-2019. A minimum of 10 fruits from north, south, east and west directions were collected from each tree. Collected fruit samples were then returned to the laboratory and stored in a 4 °C coldroom for at least 30 days. Seeds were then extracted and surface sterilized using the protocol described as follows prior to germination. The extracted seeds were immersed in 30 to 50% solution of Clariseb RL 2X to free seeds from pulp and were immersed in a 5% bleach solution for 10 minutes, followed by rinsing with warm water. The seeds were then subjected to the hot bath treatment for 10 minutes at 52 °C and immediately rinsed with the room temperature water. The seeds were then put back into the 5% bleach solution for 10 minutes, followed by rinsing with autoclaved water. Lastly, the seeds were immersed in a 1% solution of 8-hydroxyquinoline sulfate for 3 minutes prior to drying. Seeds were dried at room temperature for up to 24 hours then stored in a 4 °C coldroom. Seeds were germinated in sterilized Petri dishes in the growth chamber at 31 °C after carefully cutting the seed coat at the chalazal end to accelerate germination.

KASP marker design, sample preparation for KASP assay

SNP markers located on different chromosomes or on opposite ends of the same chromosome and heterozygous in the maternal parent genotype were selected based on genotype calls obtained

previously with Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array. KASP SNP markers (6 to 16) were designed by LGC Genomics from 50 bp of Clementine reference genome sequence flanking each SNP marker. The 2nd SNPs near the target SNPs were not specified in the design. Different sets of KASP SNP markers were designed to work separately in *Citrus* and *P. trifoliata* groups. Leaf samples (young flush) of the accessions in CVC, and individuals from the two segregating populations F x F and Chand x Trifol were collected in the field from the CVC (Riverside, CA) from 2017-2019. Collected leaf samples were then taken to the laboratory on ice and leaf disks were prepared using LGC-provided 6.0 mm biopsy punch and placed into a 96-well LGC sample plate. Plates were then stored at -80 °C prior to lyophilization. Germinated seeds were visually scored as monoembryonic or polyembryonic, root tip samples were clipped with a sterilized scalpel, collected into a 96-well LGC sample plate, and stored at -80 °C prior to lyophilization. Plates of lyophilized leaf and root tip samples were shipped to LGC Genomics for SNP marker analysis.

Phenotype evaluation and statistical analysis to estimate percent zygotic seedlings, percent polyembryony and between years and trees variation

For the *Citrus* GWAS panel, the percent zygotic seedling from the seeds of 71 accessions (Table 1) from the CVC and F x F progeny was estimated using KASP SNP markers. The estimated percent zygotic seedlings was calculated as observed number of zygotic seedlings (directly from marker data) divided by $(1-0.5^n)$ where n is the number of heterozygous markers in the seed parent. If the estimate was greater than 100%, the value was set to 100%. This equation assumes that zygotic seedlings originate by selfing or outcrossing to a genotype identical to the female parent. It likely overcorrects in accessions where many progeny result from outcrossing. In our dataset, n was greater than 2 for all 209 samples and the mean n was 5.6 (Supplemental Table 1). The

environmental effect on nucellar embryony and polyembryony was assessed by contingency chi-square for those accessions included in the GWAS panel and having both nucellar embryony and polyembryony data. A total of 15 accessions were collected over multiple collection years and a total of 8 accessions for which both trees of the same accession were collected (Table 2). For *P. trifoliata* 45 samples of *P. trifoliata* and 30 samples of *Poncirus* hybrids with *Citrus* were analyzed for polyembryony and for nucellar embryony with KASP markers using the same methods as for the citrus samples. However, for GWAS, polyembryony score was used as a phenotype in order for us to utilize the phenotypic data previously reported on Chand x Trifol individuals [38] in which multiple year data was compiled together. Germinated seeds of 88 accessions from CVC and Chand x Trifol segregating population were scored as monoembryonic or polyembryonic based on the number of root tips germinated from a single seed. Between-tree effects were assessed separately for polyembryony score by contingency chi-square test using 32 *Poncirus* accessions screened from CVC that had data for both trees of the same accession.

SNP genotyping

Hybridization data of Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array for accessions in the CVC and two segregating populations was analyzed with Affymetrix Axiom® Analysis Suite software v.4.0.3.3 by Affymetrix™ at Axiom® BestPractice workflow default setting (default DQC, QC, and % passing) to identify “Poly High Resolution” (PHR) probes. PHR probes are the most informative and accurate probe performance class on the Axiom® platform based on the genotype call cluster separation, cluster variance and cluster position. In addition to PHR probes, “NoMinorHomo” probes were included because these loci are useful for *Poncirus* where few accessions result from hybridization and therefore many loci have few or no accessions homozygous for the minor allele. The most stringent threshold QC97 was applied first, followed

by the lowered threshold of QC95 to allow inclusion of accessions belonging to distant taxa such as *Poncirus*. Genotyping data obtained from Axiom® Citrus Genotyping Array was further filtered with R package SNPolisher (SNPolisher, v2.0.3, Affymetrix™) using a lowered confidence threshold of 0.01.

Genome-wide association study

GWAS was conducted on estimated percent zygotic seedlings and polyembryony of selected accessions for *Citrus* and *Poncirus*. The maximum percent identities of included accessions were 92% and 99.4% for *Citrus* and *Poncirus*, respectively. Additionally, 19 selected accessions of *C. maxima* (non-admixed pummelo) and *C. reticulata* accessions from the CVC that are well documented as monoembryonic and known to produce only zygotic seedlings were added to the *Citrus* GWAS panel (Table 3). Totals of 97 accessions and 88 accessions composed the *Citrus* and *Poncirus* GWAS panels, respectively (Table 4). GWAS was conducted using the mixed linear model (MLM) [39] with the software TASSEL v5.2.31 (<http://www.maizogenetics.net/tassel>). In order to correct for population structure and relatedness of the accessions under study, GWAS was performed by accounting for principal component analysis (PCA) and a kinship coefficient matrix as implemented in TASSEL. The marker R^2 values were used to assess the SNP contribution to the total variation in the phenotype data, and Bonferroni correction at $\alpha = 0.05$ and False discovery rate (FDR) at 0.01 were applied to determine the genome wide significance threshold.

Results

PCA and phenotypic variation

PCA was carried out with Axiom® Citrus Genotyping Array genotyping data at 56,367 loci. The first three principal components (PCs), PC1, PC2 and PC3 explained 54.6% and 46.7% of the total

variation in the dataset for *Citrus* and *Poncirus* GWAS panels (Figure 1), respectively. Average, mean and standard deviation were determined for estimated percent zygotic seedlings for the *Citrus* GWAS panel and percent polyembryony for the *Poncirus* GWAS panel (Table 5). The frequency distributions were bimodal and skewed for the *Citrus* GWAS nucellar embryony panel (Supplemental Table 2, Supplemental Figure 1) and the *Poncirus* GWAS polyembryony panel, respectively (Supplemental Table 3, Supplemental Figure 2). Contingency chi-square test assessing between year and between tree effects on nucellar embryony and polyembryony for both *Citrus* and *Poncirus* GWAS panels were significant ($p < 0.001$). Of 15 samples assessed for between year comparisons, 13 were significant for nucellar embryony and 12 were significant for polyembryony, and of 8 samples assessed for between tree comparisons, all 8 were significant for nucellar embryony and 7 were significant for polyembryony (Table 2). When the analysis was restricted to between tree effects for the 32 *Poncirus* accessions, differences were not significant ($p = 0.995$)

KASP SNP estimated percent zygotic

The percent zygotic seedlings produced in open-pollinated seed was estimated with KASP SNP markers for 209 individuals in *Citrus*, *Poncirus* and *Poncirus* x *Citrus* hybrids that were also scored for visual scoring of polyembryony. The comparison between KASP estimated percent zygotic and visually assessed polyembryony score (100-polyembryo score) revealed a general trend but showed discrepancy as expected. Average, median, and standard deviation of KASP estimated percent zygotic is reported for individuals in blood orange, mandarin and mandarin hybrid, sour orange and sour orange hybrid, sweet orange, tangor, *Poncirus* and *Poncirus* hybrid and F x F segregating population (Supplemental Table 4)

Association mapping

GWAS was conducted with 56,367 loci for both *Citrus* and *P. trifoliata* and a major locus was detected in both *Citrus* and Poncirus panels (Figure 2, Figure 3). Effects and marker R^2 values of significant SNPs were calculated for each GWAS panel (Supplemental Data).

Citrus GWAS identified a major peak, AX-160087817, on chromosome 1 (25,483,078) on the *C. clementina* v1.0 reference genome (<http://www.phytozome.org/citrus/>) with log transformed *P* value of 8.1 (Table 6). The genome wide significant threshold was set to be 6.05 with Bonferroni correction. The peak region is 101,919 bp harboring 15 genes (<http://www.phytozome.net>) (Table 7). Although not significant, *Citrus* GWAS also identified (FDR 0.01) potential QTLs in the 7.6 Mb region on chromosome 3 (21,082,175-28,662,001) and 130 kb region on chromosome 4 (22,548,638-22,678,775) (Supplemental data)

Poncirus GWAS identified a major peak, at SNP AX-160970742, on chromosome 3 (49,634,176) on the *C. clementina* reference genome with log transformed *P* value of 7.93 (Table 8). The significance threshold was set to 4.62 with FDR. The estimated candidate region of 2.2 Mb on chromosome 3 harbored 324 genes (Supplemental Data) (Phytozome, JGI annotation). A significant peak was also detected (FDR 0.05) on chromosome 9 (10169754).

Discussion

Nucellar embryony is an important trait for both scion and rootstock breeding and played a significant role in the evolution of the citrus genomes [40]. Understanding the underlying genetics controlling nucellar embryony in *Citrus* and *P. trifoliata* can significantly improve citrus breeding strategies. Previous studies conducted using various different segregating populations have identified a major dominant locus and several quantitative trait loci (QTL) for nucellar embryony in *Citrus* [10, 28–30, 41]. With the availability of reference genomes, the first candidate gene for

nucellar embryony in *C. reticulata* was reported [36], using polyembryony score as a phenotype. However, there has been no study reporting a candidate gene(s) for nucellar embryony in *P. trifoliata* since Kapiro and Roose reported AFLP markers closely linked to a major gene that is heterozygous in some *Poncirus* cultivars and for which the dominant allele is apparently essential for nucellar embryony in trifoliolate orange and its hybrids with monoembryonic *Citrus* [38]. In order to further validate previous findings and place the *Poncirus* analysis into the context of a citrus reference genome, GWAS was performed using KASP SNP-genotyped nucellar embryony score in *Citrus* and visually scored polyembryony in *Poncirus* as phenotypes.

GWAS conducted for *Citrus* and *Poncirus* each identified a major peak apparently corresponding to a candidate major dominant locus underlying nucellar embryony (Figure 2, Figure 3). *Citrus* GWAS identified a 101 kb region on chromosome 1, in the polyembryonic locus [10, 30] harboring 15 candidate genes (Table 7). Synteny (<https://www.citrusgenomedb.org/>) between *C. clementina* (JGI) and *C. maxima* [36] reference genomes revealed that candidate genes Ciclev10008257m.g (Cg4g018940.1), Ciclev10008710m.g (Cg4g018960.1), Ciclev10010497m (Cg4g018970.1), Ciclev10009915m.g (Cg4g018890.1), Ciclev10010315m.g (Cg4g018980.1), Ciclev10010497m.g (Cg4g018970.1), Ciclev10010511m.g (Cg4g018930.1), and Ciclev10010725m.g (Cg4g018900.1), were also identified in the 80 kb candidate region responsible for *Citrus* polyembryony [36]. Ciclev10010497m.g, encoding RWP-RK domain-containing protein (Cg4g01897 encoding CitRWP) [36, 37], is located 101 kb away from the peak SNP AX-160087817. Ciclev10010497m.g, encoding RWP-RK domain-containing protein, corresponding to reported candidate gene Cg4g01897 encoding CitRWP [36], was located 1041 bp downstream of peak SNP marker AX-160087817 on chromosome 1 (25,483,078).

Poncirus GWAS identified a major peak spanning a candidate 2.2 Mb region on chromosome 3. In contrast to the result obtained from *Citrus* GWAS, the resolution of *Poncirus*

GWAS seems to have suffered from limited genetic variability within *Poncirus*, consequently among its hybrids with *Citrus* (Figure 1b) and low sample size, which are both known to affect the statistical power of the association study [42, 43]. Variation in *Citrus* is not relevant for *Poncirus* GWAS as previous study found no QTL affecting polyembryony in Chandler [38]. There are 324 genes included in the identified candidate 2.2 Mb region which include many genes associated with stress response [34, 36, 44], oxidation reduction [45], flower, seed, and embryo development (Supplemental Data), similar to previous reports on the *Citrus* Polyembryony region [36, 46]. Nucellar embryogenesis is expected to follow similar developmental patterns and gene expression patterns as somatic embryogenesis and therefore of sexual embryogenesis [36], and the processes involve gaining of embryogenic competence, and induction of somatic embryogenesis.

Poncirus GWAS identified several candidate nucellar related genes. Ciclev10023521m.g, encoding digalactosyl diacylglycerol (DGDG) synthase, is involved in synthesis of one of the main chloroplast lipids in higher plants and was reported to be one of the three targets of a novel miRNA miRN23-5p (Ciclev10023521m.g) which was found to have a higher accumulation in polyembryonic ovules [34]. Ciclev10018991m.g, encoding transketolase, is reported to be correlated with the active proliferation of somatic embryos in *Vitis vinifera* [47]. Ciclev10019056m.g, encoding leucine-rich repeat protein kinase family protein (LRR-RLK), is a large and complex gene family in plants whose members are mainly involved in development and stress response. Somatic embryogenesis receptor-like kinase 1 (*SERK1*) is a LRR-RLK that is involved in the acquisition of embryogenic competence [48, 49]. Ciclev10019083m.g, encodes a transcription factor *SCARECROW-LIKE (SCL)* which is a member of GRAS family TF. *LEAFY COTYLEDON 2 (LEC2)* genes are important embryonic regulators and are involved in the induction of somatic embryogenesis [50], and reported to activate several regulatory genes such as *SERK1* and GRAS family protein [51]. Expression analysis in Valencia orange revealed up-

regulation of *SCL* during somatic embryogenesis and reported *SCL* as a target of miRNA171c [52]. Ciclev10020532m.g encodes a transcription factor *SQUAMOSA* promoter binding protein-like 9 (*SPL9*) which is involved in the transition from vegetative to reproductive phase [53]. The members of *SPL* are reported to be regulated by miRNA156 [54]. A transcriptomic study revealed that the interaction of miR156 and *SPL9* was relevant to embryogenic capability and somatic embryogenesis in *Citrus* [55, 56]. Ciclev10022030m.g encodes *GLUTATHIONE S-TRANSFERASE (CLASS ZETA) 2 (GSTZ2)*. GSTs belong to diverse multigene families and their possible involvement in somatic embryogenesis has been reported [57, 58]. Additionally, a study in *Vitis vinifera* L. suggested the role of GST genes to attain embryogenic competence [59].

AX-160928087 (chromosome 9 10169754) was associated with a second peak detected ($R^2=0.17346$) in the *Poncirus* GWAS. Although this locus is not considered significant, previous studies indicated evidence of a minor QTL affecting the proportion of polyembryonic seeds on chromosome [38]. Percent identity and phylogeny analysis showed that there are two major *P. trifoliata* groups that are genetically distinguishable with ISSR and other markers [60] and also with SNP array data. The “large-flower” group includes 17 accessions with concordance of at least 99.1% and the “small-flower” group includes 14 accessions with at least 99.4 % concordance. We limited our sample to 7 accessions including representatives of each major group and other accessions with less than 99.4% concordance. Monoembryonic trifoliolate CRC 3888 and Small leaf CRC 4017, both monoembryonic trifoliolate accessions, were added as a third group based on the differences in the phenotype. In addition to the little genetic diversity within accessions of *P. trifoliata*, a majority of trifoliolate accessions in the CVC are polyembryonic possibly due to selection for inclusion in the collection because of their potential value as rootstock varieties. Many of these accessions were collected from citrus groves in California and their uniformity may reflect a genetic bottleneck caused by importation to California of a limited sample of *Poncirus* from Asia and

subsequent propagation of nucellar seedlings. Genome-wide association studies (GWAS) provide us with a powerful tool to statistically associate the genomic loci to the trait of interest. However, the association and the prediction of the underlying genetics relies on genetic polymorphism and phenotypic differences between the individuals under study [61]. Therefore, the lack of power in *Poncirus* GWAS is likely due, in part, to the small sample size in which GWAS was conducted, which resulted from the limited genetic variability within this group in CVC, and the phenotypic resolution, which can also greatly influence the power of detection in GWAS [62, 63]. The number of SNP markers polymorphic in *Poncirus* and included on the array is only a small fraction of the polymorphic markers present in *Poncirus* so a high-resolution GWAS based on whole genome sequencing would likely be more informative about genetic control of nucellar embryony in *Poncirus*.

This study identified a 101 kb region overlapping previously identified polyembryony locus in *Citrus* using KASP marker estimated percent nucellar embryony as a phenotype and high-density SNP array data from 90 individuals. The survey comparing two types of phenotypic data, KASP SNP marker estimated nucellar embryony and visually scored polyembryony, of accessions in CVC and mapping populations highlighted the complexity of processes affecting polyembryony and nucellar embryony. High correlation between these traits was only observed for those accessions that exhibit near 100% polyembryony or monoembryony (Supplemental Table 1). Screening of a large number of accessions over multiple years and individual trees of some accessions also revealed significant environmental influences on expression of nucellar embryony in both *Citrus* and *Poncirus* (Supplemental Table 1 Table 2). Such environmental influences will further obscure genetic variation affecting the percentage of nucellar vs zygotic embryos. Interestingly, when the segregation pattern at loci responsible for nucellar embryony and polyembryony in *Poncirus* hybrids with *Citrus* were examined, there were accessions that showed

the phenotype polyembryony that carry the *Citrus* polyembryony allele on chromosome 1, while lacking the *Poncirus* allele at candidate loci on chromosome 3. This observation suggests that the presence of *Citrus* nucellar embryony is sufficient for the induction of nucellar embryony in these genotypes.

Additionally, *Poncirus* and *Citrus* reportedly diverged 9 million years ago [64], suggesting possible limitations in conducting fine scale genomic studies in *Poncirus* using *Citrus* reference genomes. Availability of a *Poncirus* genome will significantly improve our understanding of the genetics and genomics underlying unique traits associated with *Poncirus*. Association studies showed that the high-density SNP array data combined with KASP SNP marker estimated nucellar embryony provided sufficient power and resolution to identify a candidate region underlying nucellar embryony in *Citrus* and *Poncirus*, and revealed that the gene controlling nucellar embryony in *Citrus* and *Poncirus* are likely to be different. Further study will be necessary to study and fine-map candidate genes in *Poncirus* and to identify minor affecting loci controlling the proportion of nucellar embryony in polyembryonic accessions in both *Citrus* and *Poncirus*, in addition to understanding the possible relationships between *Citrus* nucellar embryony locus and *Poncirus* polyembryony locus.

References cited

1. Nicolosi E. Origin and taxonomy. Citrus genetics, breeding and biotechnology. CAB International, Cambridge, MA. 2007:19–43. doi:10.1079/9780851990194.0019.
2. Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor Appl Genet. 2000;100:1155–66. doi:10.1007/s001220051419.
3. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet. 2013;45:487–94. doi:10.1038/ng.2586.
4. Frost HB. Nucellar embryony and juvenile characters in clonal varieties of *Citrus*. J Hered. 1938;29:423–92. doi:10.1093/oxfordjournals.jhered.a104431.
5. Soost RK, Cameron JW. Citrus. Advances in fruit breeding. Purdue Univ. Press, West Lafayette, Ind. 1975;:507–40.
6. Kobayashi S, Ikeda I, Nakatani M. Studies on nucellar embryogenesis in *Citrus*. J Jpn Soc Hortic Sci. 1979;48:179–85. doi:10.2503/jjshs.48.179.
7. Bicknell RA, Koltunow AM. Understanding apomixis: recent advances and remaining conundrums. Plant Cell. 2004;16 Suppl:S228–45. doi:10.1105/tpc.017921.
8. Spillane C, Steimer A, Grossniklaus U. Apomixis in agriculture: the quest for clonal seeds. Sex Plant Reprod. 2001;14:179–87. doi:10.1007/s00497-001-0117-1.
9. Mukherjee SK. Origin of mango (*Mangifera indica*). Economic Botany. 1972;26:260–4. doi:10.1007/bf02861039.
10. Nakano M, Shimada T, Endo T, Fujii H, Nesumi H, Kita M, et al. Characterization of genomic sequence showing strong association with polyembryony among diverse Citrus species and cultivars, and its synteny with *Vitis* and *Populus*. Plant Sci. 2012;183:131–42. doi:10.1016/j.plantsci.2011.08.002.
11. Moore GA, Castle WS. Morphological and Isozymic Analysis of open-pollinated citrus rootstock populations. J Hered. 1988;79:59–63. doi:10.1093/oxfordjournals.jhered.a110448.
12. Ruiz C, Paz Breto M, Asíns MJ. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. Euphytica. 2000;112:89–94. doi:10.1023/A:1003992719598.
13. Xiang C, Roose ML. Frequency and characteristics of nucellar and zygotic seedlings in 12 citrus rootstocks. Sci Hortic. 1988;37:47–59. doi:10.1016/0304-4238(88)90150-1.
14. Koltunow AM. Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. Plant Cell. 1993;5:1425–37. doi:10.1105/tpc.5.10.1425.

15. Ueno, I. Embryo number of various varieties of *Citrus* and its relatives. Bull Hort Res Sta Japan, ser B. 1967;7:11–22. <https://ci.nii.ac.jp/naid/10024030075/>. Accessed 19 Apr 2020.
16. Zhang S, Liang M, Wang N, Xu Q, Deng X, Chai L. Reproduction in woody perennial *Citrus*: an update on nucellar embryony and self-incompatibility. Plant Reprod. 2018;31:43–57. doi:10.1007/s00497-018-0327-4.
17. Parlevliet, E J. Evidence on the inheritance of nucellar embryony in *Citrus*. Proc Am Soc Hortic Sci. 1959;74:252–60. <https://ci.nii.ac.jp/naid/10020830116/>. Accessed 20 Apr 2020.
18. Esen A, Soost RK. Adventive embryogenesis in citrus and its relation to pollination and fertilization. American Journal of Botany. 1977;64:607–14. doi:10.1002/j.1537-2197.1977.tb11899.x.
19. Koltunow AM, Soltys K, Nito N, McClure S. Anther, ovule, seed, and nucellar embryo development in *Citrus sinensis* cv. Valencia. Canadian Journal of Botany. 1995;73:1567–82. doi:10.1139/b95-170.
20. Wakana A, Uemoto S. Adventive embryogenesis in *Citrus* (Rutaceae) II. Postfertilization development. American Journal of Botany. 1988;75:1033–47. doi:10.1002/j.1537-2197.1988.tb08810.x.
21. Aleza P, Juárez J, Ollitrault P, Navarro L. Polyembryony in non-apomictic citrus genotypes. Annals of Botany. 2010;106:533–45. doi:10.1093/aob/mcq148.
22. Cameron JW, Garber MJ. Identical-twin hybrids of *Citrus* x *Poncirus* from strictly sexual seed parents. Am J Bot. 1968;55:199–205. <https://bsapubs.onlinelibrary.wiley.com/doi/abs/10.1002/j.1537-2197.1968.tb06962.x>.
23. Frost HB. Polyembryony, heterozygosis and chimeras in *Citrus*. Hilgardia. 1926;1:365–402. <http://hilgardia.ucanr.edu/Abstract/?a=hilg.v01n16p365>.
24. Cameron JW, Frost HB. Genetics, breeding, and nucellar embryony. The Citrus Industry (eds W. Reuther, HJ Webber & LD Batchelor), vol. II. 1968.
25. Cameron JW, Soost, RK. Sexual and nucellar embryony in F1 hybrids and advanced crosses of *Citrus* with *Poncirus*. 1979. <https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCALAGROLINEINRA8010051248>.
26. Iwamasa, M. Inheritance of nucellar embryony in citrus. Bull Hort Res Stn. 1967;7:1–10. <https://ci.nii.ac.jp/naid/10006289656/>. Accessed 20 Apr 2020.
27. Hong Q-B, Xiang S-Q, Chen K-L, Chen L-G. Two complementary dominant genes controlling apomixis in genus *Citrus* and *Poncirus*. Yi chuan xue bao= Acta genetica Sinica. 2001;28:1062–7. <https://europepmc.org/abstract/med/11725641>.
28. Raga V, Bernet GP, Carbonell EA, Asins MJ. Segregation and linkage analyses in two complex populations derived from the citrus rootstock Cleopatra mandarin. Inheritance of seed reproductive traits. Tree Genet Genomes. 2012;8:1061–71. doi:10.1007/s11295-012-0486-7.

29. Nakano M, Shimizu T, Fujii H, Shimada T, Endo T, Nesumi H, et al. Marker enrichment and construction of haplotype-specific BAC contigs for the polyembryony genomic region in *Citrus*. *Breed Sci.* 2008;58:375–83. https://www.jstage.jst.go.jp/article/jsbbs/58/4/58_4_375/_article/-char/ja/.
30. Nakano M, Kigoshi K, Shimizu T, Endo T, Shimada T, Fujii H, et al. Characterization of genes associated with polyembryony and in vitro somatic embryogenesis in *Citrus*. *Tree Genet Genomes.* 2013;9:795–803. doi:10.1007/s11295-013-0598-8.
31. Kumar V, Malik SK, Pal D, Srinivasan R, Bhat SR. Comparative transcriptome analysis of ovules reveals stress related genes associated with nucellar polyembryony in citrus. *Tree Genet Genomes.* 2014;10:449–64. doi:10.1007/s11295-013-0690-0.
32. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al. The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet.* 2013;45:59–66. doi:10.1038/ng.2472.
33. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat Biotechnol.* 2014;32:656–62. doi:10.1038/nbt.2906.
34. Long J-M, Liu Z, Wu X-M, Fang Y-N, Jia H-H, Xie Z-Z, et al. Genome-scale mRNA and small RNA transcriptomic insights into initiation of citrus apomixis. *J Exp Bot.* 2016;67:5743–56. doi:10.1093/jxb/erw338.
35. Wu XM, Kou SJ, Liu YL, Fang YN, Xu Q. Genomewide analysis of small RNA s in nonembryogenic and embryogenic tissues of citrus: micro RNA-and si RNA-mediated transcript cleavage involved in somatic embryogenesis. *Plant Biotechnol.* 2015. <https://onlinelibrary.wiley.com/doi/abs/10.1111/pbi.12317>.
36. Wang X, Xu Y, Zhang S, Cao L, Huang Y, Cheng J, et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat Genet.* 2017;49:765–72. doi:10.1038/ng.3839.
37. Shimada T, Endo T, Fujii H, Nakano M, Sugiyama A, Daido G, et al. MITE insertion-dependent expression of CitRKD1 with a RWP-RK domain regulates somatic embryogenesis in citrus nucellar tissues. *BMC Plant Biol.* 2018;18:166. doi:10.1186/s12870-018-1369-3.
38. Kepiro JL, Roose ML. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*. *Tree Genet Genomes.* 2010;6:1–11. doi:10.1007/s11295-009-0223-z.
39. Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. *Nat Genet.* 2010;42:355–60. doi:10.1038/ng.546.
40. Oueslati A, Salhi-Hannachi A, Luro F, Vignes H, Mournet P, Ollitrault P. Genotyping by sequencing reveals the interspecific *C. maxima* / *C. reticulata* admixture along the genomes of modern citrus varieties of mandarins, tangors, tangelos, orangelos and grapefruits. *PLoS One.* 2017;12:e0185618. doi:10.1371/journal.pone.0185618.

41. García R, Asíns MJ, Forner J, Carbonell EA. Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. *Theor Appl Genet.* 1999;99:511–8. doi:10.1007/s001220051264.
42. Teo YY. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Curr Opin Lipidol.* 2008;19:133–43. doi:10.1097/MOL.0b013e3282f5dd77.
43. Wang M, Xu S. Statistical power in genome-wide association studies and quantitative trait locus mapping. *Heredity* . 2019;123:287–306. doi:10.1038/s41437-019-0205-3.
44. Pasternak TP, Prinsen E, Ayaydin F, Miskolczi P, Potters G, Asard H, et al. The Role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of alfalfa. *Plant Physiol.* 2002;129:1807–19. doi:10.1104/pp.000810.
45. Zavattieri MA, Frederico AM, Lima M, Sabino R, Arnholdt-Schmitt B. Induction of somatic embryogenesis as an example of stress-related plant reactions. *Electron J Biotechnol.* 2010;13:12–3. https://scielo.conicyt.cl/scielo.php?pid=S0717-34582010000100012&script=sci_arttext&tlng=en.
46. Schmidt A, Schmid MW, Klostermeier UC, Qi W, Guthörl D, Sailer C, et al. Apomictic and sexual germline development differ with respect to cell cycle, transcriptional, hormonal and epigenetic regulation. *PLoS Genet.* 2014;10:e1004476. doi:10.1371/journal.pgen.1004476.
47. Marsoni M, Bracale M, Espen L, Prinsi B, Negri AS, Vannini C. Proteomic analysis of somatic embryogenesis in *Vitis vinifera*. *Plant Cell Rep.* 2008;27:347–56. doi:10.1007/s00299-007-0438-0.
48. Hecht V, Vielle-Calzada J-P, Hartog MV, Schmidt EDL, Boutilier K, Grossniklaus U, et al. The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiol.* 2001;127:803–16. <http://www.plantphysiol.org/content/127/3/803.short>.
49. Steiner N, Santa-Catarina C, Guerra MP, Cutri L, Dornelas MC, Floh EIS. A gymnosperm homolog of SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE-1 (SERK1) is expressed during somatic embryogenesis. *Plant Cell Tissue Organ Cult.* 2012;109:41–50. doi:10.1007/s11240-011-0071-z.
50. Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, et al. Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci U S A.* 2006;103:3468–73. doi:10.1073/pnas.0511331103.
51. Guo F, Liu C, Xia H, Bi Y, Zhao C, Zhao S, et al. Induced expression of AtLEC1 and AtLEC2 differentially promotes somatic embryogenesis in transgenic tobacco plants. *PLoS One.* 2013;8:e71714. doi:10.1371/journal.pone.0071714.
52. Wu X-M, Kou S-J, Liu Y-L, Fang Y-N, Xu Q, Guo W-W. Genomewide analysis of small RNAs in nonembryogenic and embryogenic tissues of citrus: microRNA-and siRNA-mediated transcript cleavage involved in somatic embryogenesis. *Plant Biotechnology Journal.* 2015;13:383–94. doi:10.1111/pbi.12317.

53. Wang J-W. Chapter 18 - The multifaceted roles of miR156-targeted SPL transcription factors in plant developmental transitions. In: Gonzalez DH, editor. *Plant Transcription Factors*. Boston: Academic Press; 2016. p. 281–93. doi:10.1016/B978-0-12-800854-6.00018-X.
54. Wang H, Wang H. The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Mol Plant*. 2015;8:677–88. doi:10.1016/j.molp.2015.01.008.
55. Long J-M, Liu C-Y, Feng M-Q, Liu Y, Wu X-M, Guo W-W. miR156-SPL modules regulate induction of somatic embryogenesis in citrus callus. *Journal of Experimental Botany*. 2018;69:2979–93. doi:10.1093/jxb/ery132.
56. Wu X-M, Liu M-Y, Ge X-X, Xu Q, Guo W-W. Stage and tissue-specific modulation of ten conserved miRNAs and their targets during somatic embryogenesis of Valencia sweet orange. *Planta*. 2011;233:495–505. doi:10.1007/s00425-010-1312-9.
57. Galland R, Randoux B, Vasseur J, Hilbert JL. A glutathione S-transferase cDNA identified by mRNA differential display is upregulated during somatic embryogenesis in *Cichorium*. *Biochim Biophys Acta*. 2001;1522:212–6. doi:10.1016/s0167-4781(01)00314-1.
58. Kitamiya E, Suzuki S, Sano T, Nagata T. Isolation of two genes that were induced upon the initiation of somatic embryogenesis on carrot hypocotyls by high concentrations of 2,4-D. *Plant Cell Rep*. 2000;19:551–7. doi:10.1007/s002990050772.
59. Malabadi RB, Chalannavar RK, Meti NT, Mulgund GS, Nataraja K, Vijaya Kumar S, et al. Detection of Glutathione S-Transferase gene (GST2, GST3) during induction of somatic embryogenesis in grape (*Vitis vinifera* L.). 1. 2013;4. <http://updatepublishing.com/journal/index.php/rib/article/view/2349>. Accessed 3 May 2020.
60. Fang DQ, Roose ML, Krueger RR, Federici CT. Fingerprinting trifoliate orange germ plasm accessions with isozymes, RFLPs, and inter-simple sequence repeat markers. *Theor Appl Genet*. 1997;95:211–9. doi:10.1007/s001220050550.
61. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*. 2013;9:29. doi:10.1186/1746-4811-9-29.
62. van der Sluis S, Verhage M, Posthuma D, Dolan CV. Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS One*. 2010;5:e13929. doi:10.1371/journal.pone.0013929.
63. Crowell S, Korniliev P, Falcão A, Ismail A, Gregorio G, Mezey J, et al. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat Commun*. 2016;7:10527. doi:10.1038/ncomms10527.
64. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of *Citrus*. *Nature*. 2018;554:311–6. doi:10.1038/nature25447.

Table 1. 71 accessions in *Citrus* GWAS panel that were screened for nucellar embryony
71 accessions including 23 *Citrus* accessions in the CVC and 48 individuals from F x F segregating
population in *Citrus* GWAS panel were screened for nucellar embryony with KASP SNP.

Sample Name	CRC	Type
12C_10_01	NA	FxF
12C_10_04	NA	FxF
12C_10_07	NA	FxF
12C_10_12	NA	FxF
12C_10_18	NA	FxF
12C_10_19	NA	FxF
12C_11_01	NA	FxF
12C_11_12	NA	FxF
12C_11_14	NA	FxF
12C_11_15	NA	FxF
12C_11_19	NA	FxF
12C_11_21	NA	FxF
12C_12_06	NA	FxF
12C_12_07	NA	FxF
12C_12_17	NA	FxF
12C_12_21	NA	FxF
12C_15_05	NA	FxF
12C_15_10	NA	FxF
12C_15_20	NA	FxF
12C_15_23	NA	FxF
12C_15_24	NA	FxF
12C_15_25	NA	FxF
12C_16_09	NA	FxF
12C_16_16	NA	FxF
12C_16_18	NA	FxF
12C_16_19	NA	FxF
12C_16_20	NA	FxF
12C_16_21	NA	FxF
12C_16_23	NA	FxF
12C_17_15	NA	FxF
12C_17_17	NA	FxF
12C_17_18	NA	FxF
12C_17_19	NA	FxF
12C_17_22	NA	FxF
12C_17_24	NA	FxF
12C_17_25	NA	FxF

12C_18_01	NA	FxF
12C_18_08	NA	FxF
12C_18_11	NA	FxF
12C_18_15	NA	FxF
12C_19_01	NA	FxF
12C_19_10	NA	FxF
12C_19_12	NA	FxF
12C_19_15	NA	FxF
12C_19_16	NA	FxF
12C_19_20	NA	FxF
12C_19_21	NA	FxF
12C_19_23	NA	FxF
C. reticulata	3812	Mandarin Hybrid
Fairchild	3559	Mandarin Hybrid
C. amblycarpa	2485	Mandarin Hybrid
Konejime	3611	Sour Orange_Hyb
King	3845	Mandarin Hybrid
Fremont	3558	Mandarin Hybrid
Fortune	3560	Mandarin Hybrid
Lee	3851	Mandarin Hybrid
Scarlet Emperor	3226	Mandarin Hybrid
Soh Niamtra	3260	Mandarin Hybrid
Dweet	3018	Tangor
Tien Chieh	2590	Mandarin Hybrid
Sun Chu Sha	4003	Mandarin
C. leiocarpa	3147	Mandarin Hybrid
C. sunki	3143	Mandarin Hybrid
Som Keowan	3752	Mandarin Hybrid
Cutter	3030	Valencia Orange
Delfino	4114	Blood_Orange
Goutoucheng	4004	Sour_Orange
East Highlands	2369	Sweet Orange
Zhuluan	3981	SourOrange
Koster	3958	Tangor
Frost Owari	3178	Satsuma

Table 2 Contingency Chi-square test assessing environmental effects on nucellar embryony and polyembryony in GWAS panel

Variation	Nucellar embryony	Polyembryony
between years (N=15)	1.2785E-194	1.0353E-100
between trees (N=8)	3.3082E-102	2.4434E-62

Table 3. 90 accessions in Citrus GWAS panel

Sample Name	CRC	Type
12C_10_01	NA	FxF
12C_10_04	NA	FxF
12C_10_07	NA	FxF
12C_10_12	NA	FxF
12C_10_18	NA	FxF
12C_10_19	NA	FxF
12C_11_01	NA	FxF
12C_11_12	NA	FxF
12C_11_14	NA	FxF
12C_11_15	NA	FxF
12C_11_19	NA	FxF
12C_11_21	NA	FxF
12C_12_06	NA	FxF
12C_12_07	NA	FxF
12C_12_17	NA	FxF
12C_12_21	NA	FxF
12C_15_05	NA	FxF
12C_15_10	NA	FxF
12C_15_20	NA	FxF
12C_15_23	NA	FxF
12C_15_24	NA	FxF
12C_15_25	NA	FxF
12C_16_09	NA	FxF
12C_16_16	NA	FxF
12C_16_18	NA	FxF
12C_16_19	NA	FxF
12C_16_20	NA	FxF
12C_16_21	NA	FxF
12C_16_23	NA	FxF
12C_17_15	NA	FxF
12C_17_17	NA	FxF
12C_17_18	NA	FxF
12C_17_19	NA	FxF
12C_17_22	NA	FxF
12C_17_24	NA	FxF
12C_17_25	NA	FxF
12C_18_01	NA	FxF

12C_18_08	NA	FxF
12C_18_11	NA	FxF
12C_18_15	NA	FxF
12C_19_01	NA	FxF
12C_19_10	NA	FxF
12C_19_12	NA	FxF
12C_19_15	NA	FxF
12C_19_16	NA	FxF
12C_19_20	NA	FxF
12C_19_21	NA	FxF
12C_19_23	NA	FxF
African	2346	Pummelo
Algerian	279	Clementine
C. amblycarpa	2485	Mandarin Hybrid
C. leiocarpa	3147	Mandarin Hybrid
C. reticulata	3812	Mandarin Hybrid
C. sunki	3143	Mandarin Hybrid
Chinese	1224	Pummelo
Clement	2603	Tangelo
Cutter	3030	Valencia Orange
Deep Red	2347	Pummelo
Delfino	4114	Blood Orange
Dweet	3018	Tangor
EastHighlands	2369	Sweet Orange
Ellendale	3960	Tangor
Encore	3569	Mandarin Hybrid
Fairchild	3559	Mandarin Hybrid
Fortune	3560	Mandarin Hybrid
Fremont	3558	Mandarin Hybrid
Frost	3178	Satsuma
Goutoucheng	4004	Sour Orange
King	3845	Mandarin Hybrid
Konejime	3611	Sour Orange_Hyb
Koster	3958	Tangor
Lee	3851	Mandarin Hybrid
Mandarinette	3405	Mandarin Hybrid
Pink	2246	Pummelo
Reinking	3805	Pummelo
Robinson	3850	Mandarin Hybrid
Scarlet Emperor	3226	Mandarin Hybrid

Seedless Kishu	3906	Mandarin Hybrid
Siamese Acidless	2240	Pummelo
Soh Niamtra	3260	Mandarin Hybrid
Som Keowan	3752	Mandarin Hybrid
Sun Chu Sha	4003	Mandarin
Tahitian	3806	Pummelo
Temple	2598	Tangor
Thong Dee	3927	Pummelo
Tien Chieh	2590	Mandarin Hybrid
Ugli	4171	Tangelo
Webber Shaddock	4233	Pummelo
Wilking	3020	Mandarin Hybrid
Zhuluan	3981	Sour_Orange

Table 4. 88 accessions in Poncirus GWAS panel

Sample Name	CRC	Type
#27	3938	Trifoliolate
Argentina	3206	Trifoliolate
C32	3911	Trifoliolate Hybrid
C35	3912	Trifoliolate Hybrid
Chandxtrif_6B_12-16	NA	Chand x Trifol
Chandxtrif_6B_12-18	NA	Chand x Trifol
Chandxtrif_6B_12-20	NA	Chand x Trifol
Chandxtrif_6B_12-21	NA	Chand x Trifol
Chandxtrif_6B_12-22	NA	Chand x Trifol
Chandxtrif_6B_13-01	NA	Chand x Trifol
Chandxtrif_6B_13-03	NA	Chand x Trifol
Chandxtrif_6B_13-04	NA	Chand x Trifol
Chandxtrif_6B_13-05	NA	Chand x Trifol
Chandxtrif_6B_13-06	NA	Chand x Trifol
Chandxtrif_6B_13-08	NA	Chand x Trifol
Chandxtrif_6B_13-09	NA	Chand x Trifol
Chandxtrif_6B_13-13	NA	Chand x Trifol
Chandxtrif_6B_13-14	NA	Chand x Trifol
Chandxtrif_6B_13-15	NA	Chand x Trifol
Chandxtrif_6B_13-16	NA	Chand x Trifol
Chandxtrif_6B_13-17	NA	Chand x Trifol
Chandxtrif_6B_13-18	NA	Chand x Trifol
Chandxtrif_6B_13-20	NA	Chand x Trifol
Chandxtrif_6B_13-21	NA	Chand x Trifol
Chandxtrif_6B_14-01	NA	Chand x Trifol
Chandxtrif_6B_14-02	NA	Chand x Trifol
Chandxtrif_6B_14-03	NA	Chand x Trifol
Chandxtrif_6B_14-05	NA	Chand x Trifol
Chandxtrif_6B_14-07	NA	Chand x Trifol
Chandxtrif_6B_14-08	NA	Chand x Trifol
Chandxtrif_6B_14-10	NA	Chand x Trifol
Chandxtrif_6B_14-12	NA	Chand x Trifol
Chandxtrif_6B_14-13	NA	Chand x Trifol
Chandxtrif_6B_14-14	NA	Chand x Trifol
Chandxtrif_6B_14-15	NA	Chand x Trifol
Chandxtrif_6B_14-16	NA	Chand x Trifol
Chandxtrif_6B_14-17	NA	Chand x Trifol

Chandxtrif_6B_14-18	NA	Chand x Trifol
Chandxtrif_6B_14-20	NA	Chand x Trifol
Chandxtrif_6B_14-22	NA	Chand x Trifol
Chandxtrif_6B_15-01	NA	Chand x Trifol
Chandxtrif_6B_15-02	NA	Chand x Trifol
Chandxtrif_6B_15-03	NA	Chand x Trifol
Chandxtrif_6B_15-08	NA	Chand x Trifol
Chandxtrif_6B_15-09	NA	Chand x Trifol
Chandxtrif_6B_15-10	NA	Chand x Trifol
Chandxtrif_6B_15-15	NA	Chand x Trifol
Chandxtrif_6B_15-17	NA	Chand x Trifol
Chandxtrif_6B_15-18	NA	Chand x Trifol
Chandxtrif_6B_16-01	NA	Chand x Trifol
Chandxtrif_6B_16-02	NA	Chand x Trifol
Chandxtrif_6B_16-03	NA	Chand x Trifol
Chandxtrif_6B_16-06	NA	Chand x Trifol
Chandxtrif_6B_16-09	NA	Chand x Trifol
Chandxtrif_6B_16-12	NA	Chand x Trifol
Chandxtrif_6B_16-14	NA	Chand x Trifol
Chandxtrif_6B_16-15	NA	Chand x Trifol
Chandxtrif_6B_16-17	NA	Chand x Trifol
Chandxtrif_6B_16-18	NA	Chand x Trifol
Chandxtrif_6B_16-19	NA	Chand x Trifol
Chandxtrif_6B_16-20	NA	Chand x Trifol
Chandxtrif_6B-16-10	NA	Chand x Trifol
Chandxtrif_6B-16-21	NA	Chand x Trifol
Chandxtrif_6B-16-22	NA	Chand x Trifol
Chandxtrif_6E-32-01	NA	Chand x Trifol
Chandxtrif_6E-32-02	NA	Chand x Trifol
Citrandarin	2618	Trifoliate Hybrid
Citradia	1436	Trifoliate Hybrid
Citrangle #1416	2866	Trifoliate Hybrid
Citrangor	1447	Trifoliate Hybrid
Citremon	1449	Trifoliate Hybrid
Cunningham	271	Trifoliate Hybrid
Flying Dragon	3330A	Trifoliate
Glen Citrangedin	3573	Trifoliate Hybrid
Monoembryonic	3888	Trifoliate
Morton	2748	Trifoliate Hybrid
S281	3552	Trifoliate Hybrid

S302	3415	Trifoliate Hybrid
Sacaton	3337	Trifoliate Hybrid
Savage	275	Trifoliate Hybrid
Simmons	3549	Trifoliate
Small Leaf	4017	Trifoliate
Spanish Swo x P. trifoliata	3336	Trifoliate Hybrid
Swingle	3771	Trifoliate Hybrid
Swingle	3767	Trifoliate Hybrid
Trifeola	3954	Trifoliate Hybrid
Uvalde	2865	Trifoliate Hybrid
X639	3957	Trifoliate Hybrid

Table 5. Average, mean and standard deviation for estimated percent zygotic seedlings and percent polyembryony for GWAS panels

Average, mean and standard deviation were determined for estimated percent zygotic seedlings for the *Citrus* GWAS panel and percent polyembryony for the *Poncirus* GWAS panel

Trait	GWAS population	Average	Median	Standard Deviation
Nucellar embryony	Citrus (n=90)	41.6	0.85	45.4
Polyembryony	<i>P. trifoliata</i> (n=88)	21.9	2.38	28.6

Table 6. Peak SNPs for Citrus NE identified by GWAS LML analysis with threshold corrected with Bonferroni

Peak SNP	Chromosome	Position	-LOG10(P)	Marker R ²
AX-160087817	1	25483078	8.399909651	0.58829
AX-160826823	1	25483447	7.658981172	0.52453
AX-161046513	1	25485429	7.658981172	0.52453
AX-160167504	1	25498118	7.42251969	0.50474
AX-160220206	1	25396199	7.403578857	0.55997

Table 7. *Citrus* nucellar embryony candidate genes

locusName	Pfam	Panther	KOG	KEG G/ec	KO	GO	Best-hit- arabi-name	arabi- symbol	arabi-defline
Ciclev10007 465m.g	PF00097,PF0009 8,PF08783	PTHR15439,PTHR1 5439:SF0	KOG 0314		K10 624	GO:0046872,GO:0008270,GO:00 03676,GO:0005634	AT5G47430.1		
Ciclev10007 516m.g	PF00924	PTHR31618,PTHR31618:SF3				GO:0055085,GO:0016020	AT5G12080 .3	ATMSL10, MSL10	DWNN domain, a CCHC-type zinc finger mechanosensitive channel of small conductance- like 10
Ciclev10007 870m.g	PF01535,PF1304 1	PTHR24015,PTHR24015:SF707					AT1G05670.1		Pentatricopeptide repeat (PPR-like) superfamily protein
Ciclev10008257m.g		PTHR33541,PTHR33541:SF2					AT3G13980.1		
Ciclev10008 710m.g	PF00795	PTHR23088,PTHR2 3088:SF20	KOG 0806	3.5.1. 3	K13 566	GO:0016810,GO:0006807	AT5G12040.1		Nitrilase/cyanide hydratase and apolipoprotein N- acyltransferase family protein
Ciclev10008 847m.g	PF01564	PTHR11558,PTHR1 1558:SF30	KOG 1562	2.5.1. 79	K18 787		AT5G19530 .1	ACL5	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Ciclev10009 272m.g	PF00628,PF1216 5	PTHR12321,PTHR1 2321:SF61	KOG 1632			GO:0005515,GO:0042393,GO:00 06355	AT5G20510 .1	AL5	alfin-like 5
Ciclev10009 286m.g	PF00249	PTHR10641,PTHR10641:SF459			K09 422		AT1G22640 .1	ATMYB3, MYB3	myb domain protein 3
Ciclev10009 451m.g	PF00226						AT5G21430.1		Chaperone DnaJ-domain superfamily protein
Ciclev10009 915m.g	PF06596	PTHR34455,PTHR34455:SF1				GO:0016020,GO:0015979,GO:00 09523	AT2G06520 .1	PSBX	photosystem II subunit X
Ciclev10010315m.g		PTHR33984,PTHR33984:SF4					AT4G12070.1		
Ciclev10010 430m.g	PF00189	PTHR11760,PTHR11760:SF22			K02 985		AT5G35530.1		Ribosomal protein S3 family protein
Ciclev10010 497m.g	PF02042	PTHR32002					AT4G35590.1		RWP-RK domain-containing protein
Ciclev10010 511m.g	PF07714	PTHR23257,PTHR2 3257:SF504	KOG 0192	2.7.12 .1		GO:0006468,GO:0004672	AT5G49470.2		PAS domain-containing protein tyrosine kinase family protein
Ciclev10010 725m.g	PF00924	PTHR31618,PTHR31618:SF3				GO:0055085,GO:0016020	AT5G12080 .1	ATMSL10, MSL10	mechanosensitive channel of small conductance- like 10

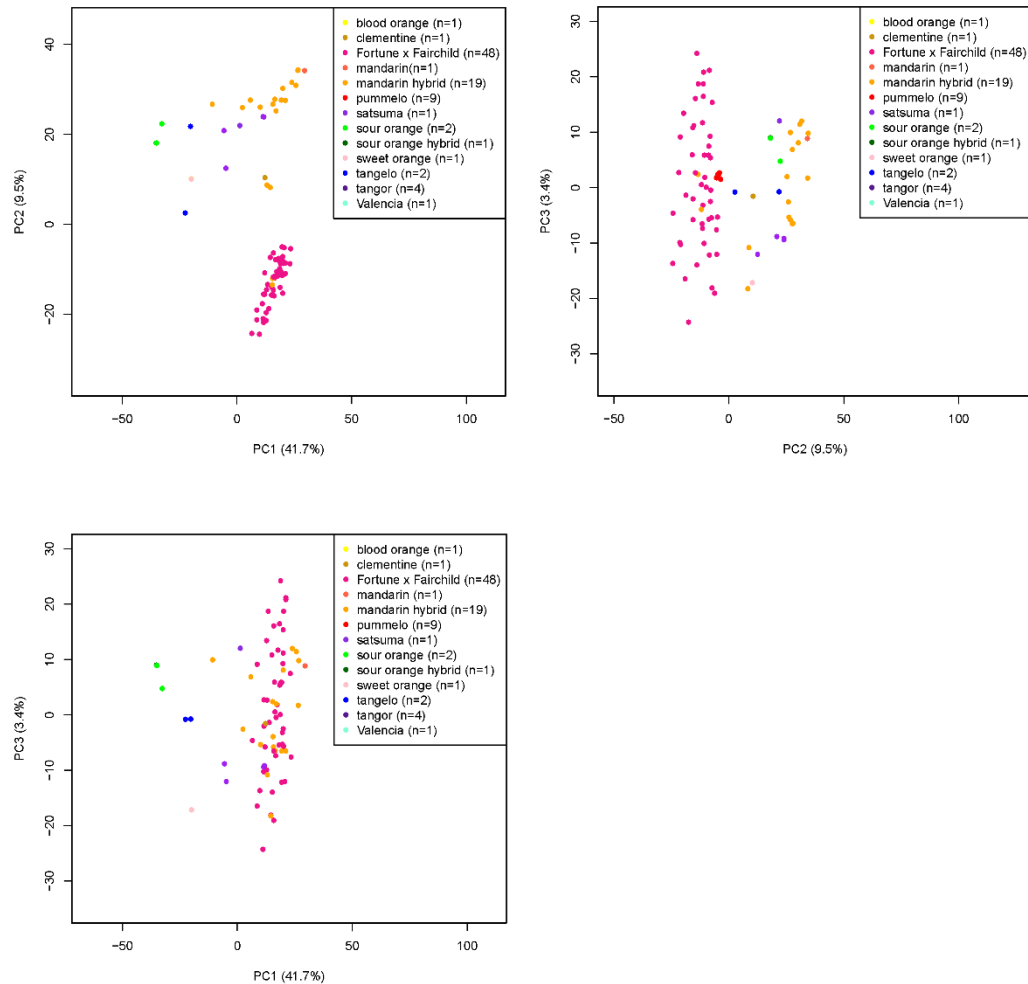
Table 8. Peak SNPs for Citrus NE QTL identified by GWAS LML analysis with threshold corrected with FDR 0.01

Peak SNP	Chromosome	Position	-LOG10(P)	Marker R ²
AX-160970742	3	49634176	7.931035	0.25857
AX-160513213	3	50511411	7.819043	0.23974
AX-160672623	3	50686141	7.563996	0.25176
AX-160251964	3	49731144	7.260048	0.26822
AX-160503210	3	49765164	7.09489	0.19889
AX-160138631	3	49765712	6.993192	0.19628
AX-160825202	3	49875794	6.915781	0.19659
AX-160688684	3	49510130	6.528488	0.22889
AX-161079647	3	50656481	6.500038	0.24085
AX-161057100	3	50962064	6.402053	0.18083
AX-159957648	3	50911960	6.32879	0.20588
AX-160693900	3	49572330	6.308936	0.20144
AX-160416987	3	49516128	6.071548	0.17199
AX-160950521	3	49578081	6.071548	0.17199
AX-161096688	3	50069242	5.994691	0.16992
AX-159968820	3	49266936	5.821483	0.18835
AX-159914171	3	50806724	5.324497	0.17802
AX-160108021	3	49077764	5.205115	0.14815
AX-160423982	3	50966758	5.155597	0.1884
AX-161084663	3	49365327	5.135507	0.14781
AX-160459126	3	50843122	4.790271	0.17075
AX-160538030	3	48768749	4.625398	0.19508

Figure 1. PCA for GWAS panels

PCA was carried out with Axiom® Citrus Genotyping Array genotyping data at 56,367 loci. The first three principal components (PCs), PC1, PC2 and PC3 explained 54.6% and 46.7% of the total variation in the dataset for (a) *Citrus* and (b) *Poncirus* GWAS panel

(a)



(b)

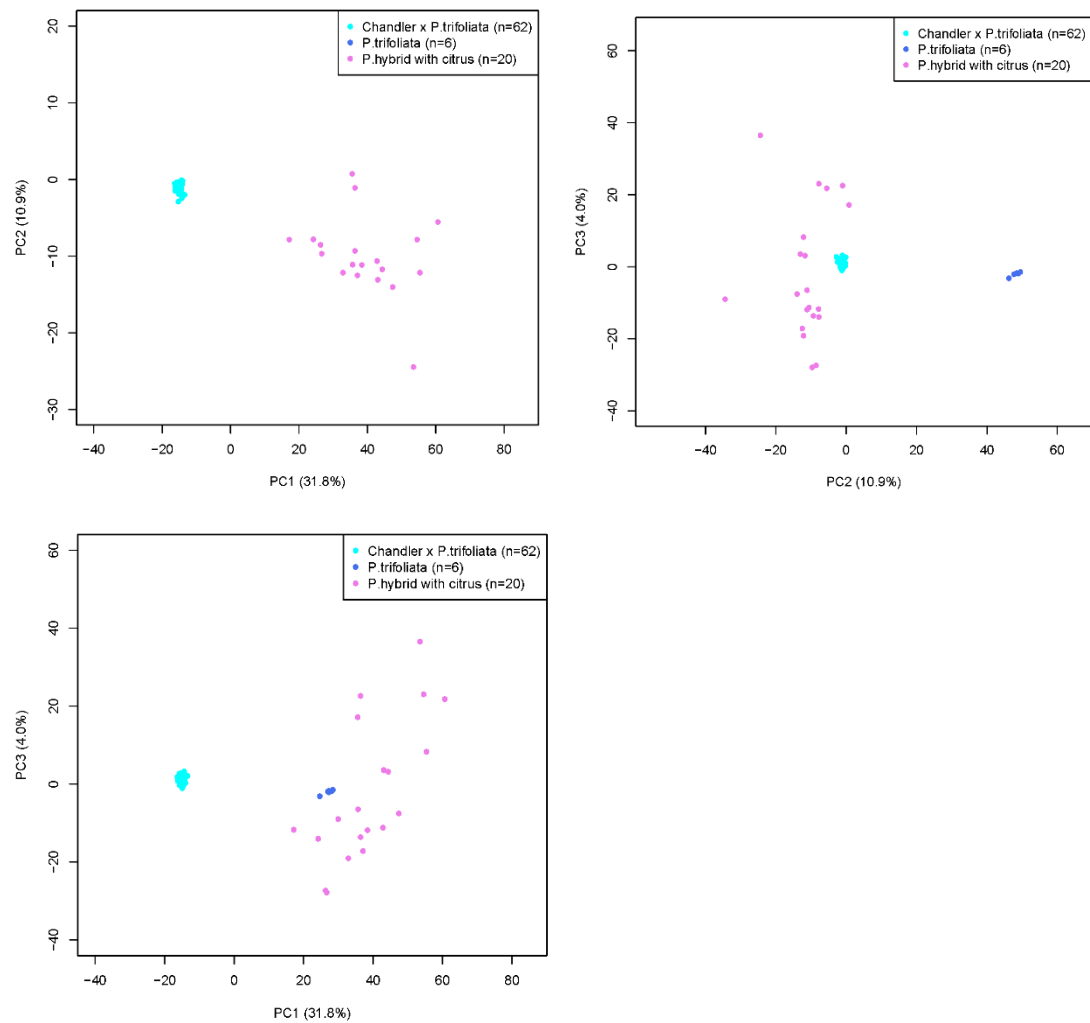


Figure 2. Manhattan plot of *Citrus* GWAS for nucellar embryony
 GWAS for KASP SNP marker estimated percent nucellar embryony in *Citrus*. Negative $\text{Log}_{10}p$ values are plotted against genomic position on each of the 9 chromosomes. Dashed line indicates Bonferroni-corrected genome-wide significance threshold

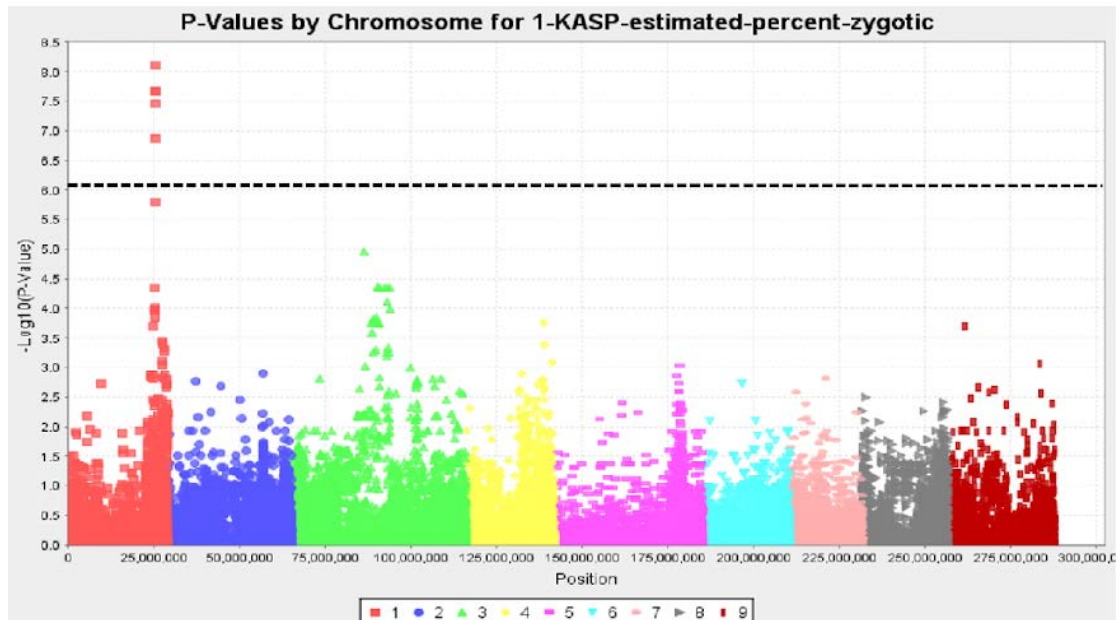
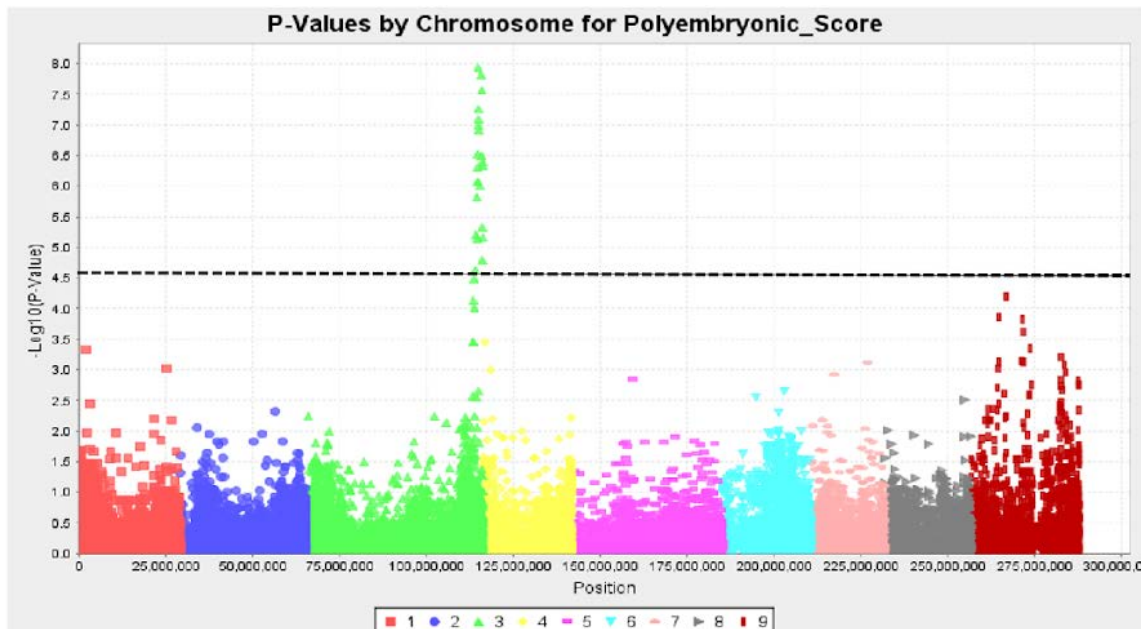


Figure 3. Manhattan plot of *Poncirus* GWAS for polyembryony
 GWAS for visually scored polyembryony in *Poncirus*. Negative Log₁₀p values are plotted against genomic position on each of the 9 chromosomes. Dashed line indicates FDR (0.01)-corrected genome-wide significance threshold



Supplemental Table 1. 209 individuals in CVC and F x F segregating population screened for nucellar embryony and polyembryony

Accession	CRC	Type	Estimated % zygotic	100- Polyembry ony	Informative KASP SNP Count	Number of seedlings
#22	4172	Trifoliate	23.3	84.1	6	48
#26	3939	Trifoliate	0.8	2.8	6	47
#26	3939	Trifoliate	0.0	37.5	6	48
#27	3938	Trifoliate	21.2	54.0	6	48
12C-10-01	FxF	FxF	3.3	59.7	6	31
12C-10-01	FxF	FxF	26.7	38.0	11	45
12C-10-04	FxF	FxF	100.0	95.9	3	47
12C-10-07	FxF	FxF	0.0	89.3	8	47
12C-10-09	FxF	FxF	100.0	100.0	5	48
12C-10-12	FxF	FxF	23.8	22.1	6	47
12C-10-13	FxF	FxF	98.3	100.0	8	47
12C-10-14	FxF	FxF	0.0	22.7	5	47
12C-10-15	FxF	FxF	100.0	100.0	3	47
12C-10-16	FxF	FxF	54.9	100.0	5	47
12C-10-17	FxF	FxF	13.0	3.2	6	47
12C-10-18	FxF	FxF	0.0	66.7	2	47
12C-10-19	FxF	FxF	100.0	100.0	7	47
12C-11-01	FxF	FxF	100.0	98.6	7	47
12C-11-03	FxF	FxF	38.1	79.2	6	48
12C-11-04	FxF	FxF	98.9	100.0	4	55
12C-11-04	FxF	FxF	98.6	100.0	7	47
12C-11-05	FxF	FxF	99.3	100.0	8	94
12C-11-06	FxF	FxF	100.0	100.0	7	47
12C-11-09	FxF	FxF	2.9	100.0	8	35
12C-11-11	FxF	FxF	100.0	100.0	9	47
12C-11-12	FxF	FxF	5.6	10.1	2	48
12C-11-14	FxF	FxF	4.5	13.3	4	47
12C-11-15	FxF	FxF	33.3	48.7	2	24
12C-11-15	FxF	FxF	8.7	17.5	6	47
12C-11-17	FxF	FxF	2.2	30.8	4	48
12C-11-19	FxF	FxF	9.7	31.7	3	47
12C-11-20	FxF	FxF	100.0	100.0	6	47
12C-11-21	FxF	FxF	100.0	80.0	5	47
12C-12-04	FxF	FxF	24.2	18.2	5	47
12C-12-06	FxF	FxF	100.0	100.0	6	8

12C-12-07	FxF	FxF	0.0	46.2	6	9
12C-12-09	FxF	FxF	4.8	40.0	3	48
12C-12-09	FxF	FxF	100.0	100.0	6	47
12C-12-17	FxF	FxF	100.0	100.0	7	47
12C-12-21	FxF	FxF	6.6	41.1	5	47
12C-15-01	FxF	FxF	2.2	47.5	5	47
12C-15-02	FxF	FxF	100.0	91.5	7	47
12C-15-04	FxF	FxF	100.0	100.0	7	34
12C-15-05	FxF	FxF	0.0	17.7	8	47
12C-15-10	FxF	FxF	4.5	16.0	4	47
12C-15-12	FxF	FxF	100.0	100.0	5	10
12C-15-13	FxF	FxF	61.4	56.0	6	4
12C-15-13	FxF	FxF	40.3	13.8	11	71
12C-15-15	FxF	FxF	6.8	54.5	4	47
12C-15-15	FxF	FxF	10.0	18.4	9	40
12C-15-20	FxF	FxF	100.0	100.0	5	47
12C-15-23	FxF	FxF	4.3	31.9	6	47
12C-15-24	FxF	FxF	100.0	100.0	8	47
12C-15-25	FxF	FxF	100.0	97.3	8	47
12C-16-01	FxF	FxF	100.0	100.0	4	47
12C-16-05	FxF	FxF	2.8	78.0	2	48
12C-16-09	FxF	FxF	8.3	59.6	2	48
12C-16-13	FxF	FxF	100.0	100.0	4	48
12C-16-16	FxF	FxF	38.1	50.0	3	6
12C-16-16	FxF	FxF	23.5	11.8	8	47
12C-16-18	FxF	FxF	100.0	100.0	4	48
12C-16-19	FxF	FxF	100.0	100.0	8	47
12C-16-20	FxF	FxF	100.0	98.0	4	8
12C-16-21	FxF	FxF	100.0	100.0	6	47
12C-16-23	FxF	FxF	2.2	36.8	5	48
12C-16-23	FxF	FxF	2.1	19.7	10	47
12C-16-24	FxF	FxF	0.0	10.8	8	47
12C-17-09	FxF	FxF	100.0	98.2	3	88
12C-17-11	FxF	FxF	100.0	100.0	3	48
12C-17-15	FxF	FxF	100.0	100.0	2	47
12C-17-17	FxF	FxF	100.0	100.0	3	48
12C-17-18	FxF	FxF	100.0	100.0	6	47
12C-17-19	FxF	FxF	100.0	100.0	6	47
12C-17-22	FxF	FxF	100.0	100.0	5	47
12C-17-24	FxF	FxF	100.0	100.0	8	47

12C-17-25	FxF	FxF	99.4	100.0	6	47
12C-18-01	FxF	FxF	0.0	67.3	4	47
12C-18-04	FxF	FxF	100.0	100.0	3	47
12C-18-05	FxF	FxF	35.9	83.1	3	35
12C-18-07	FxF	FxF	29.6	41.3	6	48
12C-18-08	FxF	FxF	0.0	25.0	3	6
12C-18-08	FxF	FxF	4.3	5.9	6	47
12C-18-11	FxF	FxF	20.7	51.6	5	50
12C-18-11	FxF	FxF	5.3	9.1	8	39
12C-18-15	FxF	FxF	100.0	100.0	3	48
12C-18-16	FxF	FxF	9.4	35.8	5	88
12C-18-17	FxF	FxF	100.0	100.0	5	47
12C-18-18	FxF	FxF	97.8	100.0	4	48
12C-18-24	FxF	FxF	100.0	100.0	6	47
12C-19-01	FxF	FxF	100.0	100.0	8	47
12C-19-07	FxF	FxF	0.0	49.3	6	48
12C-19-12	FxF	FxF	98.4	100.0	5	86
12C-19-15	FxF	FxF	3.2	38.7	3	36
12C-19-16	FxF	FxF	34.0	26.7	2	47
12C-19-18	FxF	FxF	100.0	100.0	2	48
12C-19-19	FxF	FxF	31.8	52.9	3	47
12C-19-20	FxF	FxF	99.9	100.0	4	47
12C-19-21	FxF	FxF	21.2	68.3	6	48
12C-19-23	FxF	FxF	98.9	100.0	5	48
Argentina	3206	Trifoliate	13.9	88.3	6	88
Barnes	2554	Trifoliate	40.2	82.9	6	43
Benecke	3338	Trifoliate	9.2	82.8	6	44
Benoit	3547	Trifoliate	16.5	68.6	6	86
C.	2485	Mandarin	0.0	84.0	5	48
<i>amblyscarpa</i>		Hybrid				
<i>C. leiocarpa</i>	3147	Mandarin	0.0	26.2	5	47
		Hybrid				
<i>C. leiocarpa</i>	3147	Mandarin	3.4	64.6	2	39
		Hybrid				
<i>C. reticulata</i>	3812	Mandarin	6.6	95.9	5	47
		Hybrid				
<i>C. reticulata</i>	3812	Mandarin	0.0	81.8	3	7
		Hybrid				
<i>C. sunki</i>	3143	Mandarin	31.2	46.7	2	47
		Hybrid				
<i>C. sunki</i>	3143	Mandarin	54.5	54.1	2	36
		Hybrid				
C32	3911	Trifoliate	24.2	42.0	6	42
		Hybrid				

C35	3912	Trifoliolate Hybrid	21.4	62.5	7	94
Carrizo	2863	Trifoliolate Hybrid	0.0	33.3	5	41
Carrizo	2863	Trifoliolate Hybrid	0.0	45.4	5	46
Citradia	1436	Trifoliolate Hybrid	74.4	80.0	2	43
Citrandarin	2618	Trifoliolate Hybrid	100.0	100.0	4	47
Citrangle #1416	2866	Trifoliolate Hybrid	0.0	48.9	7	29
Citrangle #1416	2866	Trifoliolate Hybrid	0.0	48.9	7	23
Citrangor	1447	Trifoliolate Hybrid	33.3	79.2	4	48
Citremon	1449	Trifoliolate Hybrid	21.5	90.4	5	48
Citremon	1449	Trifoliolate Hybrid	29.1	83.8	5	39
Cunningham	271	Trifoliolate Hybrid	0.0	0.0	6	33
Cutter	3030	Valencia orange	11.0	7.7	5	47
Delfino	4114	Blood orange	11.5	10.7	12	26
Delfino	4114	Blood Orange	6.5	8.8	5	48
Dweet	3018	Tangor	37.3	79.7	5	47
EastHighlands	2369	Sweet orange	10.5	7.5	12	57
English Dwarf	3876	Trifoliolate	2.3	74.7	6	44
English Large	3548	Trifoliolate	30.7	79.6	6	43
Fairchild	3559	Mandarin Hybrid	0.0	40.5	5	40
Fairchild	3559	Mandarin Hybrid	0.0	42.7	5	48
Fairhope	3351	Trifoliolate	2.3	62.5	6	44
Fairhope	3351	Trifoliolate	8.5	62.5	6	48
Flying Dragon B	3330	Trifoliolate	33.9	73.6	6	48
Flying Dragon A	3330	Trifoliolate	46.6	86.6	6	48
Fortune	3560	Mandarin Hybrid	100.0	100.0	5	48
Fremont	3558	Mandarin Hybrid	2.2	2.7	5	47
Frost	3178	Mandarin Hybrid	0.0	0.0	9	19
Frost	3484	Trifoliolate	23.9	85.4	5	95

GlenCitrang edin	3573	Trifoliolate Hybrid	1.9	78.0	6	53
Goutouchen g	4004	Sour Orange	17.5	35.0	8	46
HallgpftXtri fol	3821	Trifoliolate Hybrid	100.0	28.7	8	16
Hiryu	3795	Trifoliolate	48.6	91.7	6	92
Hiryu	3882	Trifoliolate	38.9	78.7	6	47
Hongjian	4124	Sweet orange	4.5	0.0	5	46
Indian River	3859	Sweet orange	2.2	28.8	5	47
Indian River	3859	Sweet orange	4.3	26.1	5	48
Jacobson	3411	Trifoliolate	9.2	85.2	6	44
King	3845	Mandarin Hybrid	74.3	97.1	3	40
Koethen	1106	Sweet orange	0.0	30.0	5	40
Konejime	3611	Sour orange Hybrid	10.3	90.9	5	40
Konejime	3611	Sour orange Hybrid	0.0	52.4	5	8
Koster	3958	Tangor	100.0	98.4	3	48
Kryder 16 -6	3210	Trifoliolate	30.1	81.9	6	44
Kryder 5- 5	3586	Trifoliolate	18.5	84.4	6	44
Kryder 55- 1	3486	Trifoliolate	8.5	95.7	6	48
Kryder 55 -5	3215	Trifoliolate	13.9	90.1	6	44
Kryder 60 -2	3213	Trifoliolate	14.9	87.4	6	48
Kryder 8 -5	3218	Trifoliolate	12.7	68.2	6	48
Kryder Medium	3212	Trifoliolate	23.1	75.3	6	44
Lee	3851	Mandarin Hybrid	100.0	100.0	8	48
MalteseOval	274	Sweet orange	37.3	84.3	5	47
Monoembry onic	3888	Trifoliolate	92.3	100.0	5	145
Morton	2748	Trifoliolate Hybrid	0.0	2.8	9	48
Olivelands	3863	Sweet orange	0.0	9.1	5	48
P. trifoliata Sdlg	4009	Trifoliolate	6.9	50.0	6	44
P. trifoliata Sdlg	4007	Trifoliolate	2.1	86.6	6	48
P. trifoliata Sdlg	4006	Trifoliolate	24.2	94.4	6	42

P. trifoliata Sdlg	4008	Trifoliata	6.6	94.4	6	46
Pineapple	3858	Sweet orange	48.0	54.8	5	48
Rich 12-2	3209	Trifoliata	13.0	84.0	6	47
Rich 16-6	3485	Trifoliata	22.2	91.6	6	96
Rich 22-2	3211	Trifoliata	40.0	95.6	6	41
Rich 7-5	3587	Trifoliata	23.3	89.4	6	48
Rubidoux 123	NGR 5134	Trifoliata Hybrid	2.1	89.4	6	48
Ruby Blood orange	3827	Blood Orange	0.0	23.3	5	47
S281	3552	Trifoliata Hybrid	0.0	33.3	5	47
S302	3415	Trifoliata Hybrid	25.4	23.0	6	8
S302	3415	Trifoliata Hybrid	7.7	23.0	8	39
Sacaton	3417	Trifoliata Hybrid	27.3	23.0	4	43
Sacaton	3337	Trifoliata Hybrid	100.0	66.2	9	46
Sacaton	3414	Trifoliata Hybrid	56.4	97.3	9	48
Sacaton	3337	Trifoliata Hybrid	100.0	28.7	4	47
Sathgudi	3083	Sweet orange	6.5	40.3	5	48
Scarlet Emperor	3226	Mandarin Hybrid	12.9	17.3	5	16
Simmons	3549	Trifoliata	41.7	81.9	4	46
Small Leaf	4017	Trifoliata	95.4	100.0	5	92
Soh	3269	Mandarin	0.0	2.7	7	47
Niamtra		Hybrid				
Som	3752	Mandarin	0.0	96.7	6	47
Keowan		Hybrid				
Sun Chu Sha	4003	Mandarin	15.5	7.9	7	52
Swingle	3767	Trifoliata Hybrid	6.4	50.0	9	47
Swingle	3771	Trifoliata Hybrid	10.4	66.4	9	48
Swingle	3771	Trifoliata Hybrid	15.2	69.0	9	46
Taylor	3571	Trifoliata	15.5	93.1	6	92
Texas	2861	Trifoliata	38.7	74.6	6	42
Tien Chieh	2376	Mandarin Hybrid	24.6	37.5	4	26
Towne F	3572	Trifoliata	23.8	90.1	6	47
Trifeola	3954	Trifoliata Hybrid	10.9	89.9	4	39

Troyer	1459	Trifoliate Hybrid	0.0	20.5	5	47
USDA	1498	Trifoliate	11.5	78.5	6	44
Uvalde	2865	Trifoliate Hybrid	0.0	69.4	9	47
Uvalde	2865	Trifoliate Hybrid	0.0	87.5	9	41
Webber Fawcett #22	2552	Trifoliate	6.9	88.9	6	44
Webber Fawcett #22	2552	Trifoliate	16.9	67.1	6	48
X639	3957	Trifoliate Hybrid	4.4	72.9	7	46
Xuegan	3984	Sweet orange	0.0	16.9	5	48
Yamaguchi	3412	Trifoliate	13.9	94.7	6	47
Yamaguchi	3412	Trifoliate	23.8	94.2	6	47
Zhuluan	3981	Sour Orange	25.8	43.8	5	48

Supplemental Table 2. KASP SNP estimated % nucellar embryony for 90 accessions in *Citrus* GWAS panel

Sample Name	CRC	Type	NE (100-KASP SNP estimated % zygotie)
12C_10_01	NA	FxF	73.3
12C_10_04	NA	FxF	0
12C_10_07	NA	FxF	100
12C_10_12	NA	FxF	76.2
12C_10_18	NA	FxF	100
12C_10_19	NA	FxF	0
12C_11_01	NA	FxF	0
12C_11_12	NA	FxF	94.4
12C_11_14	NA	FxF	95.5
12C_11_15	NA	FxF	91.35
12C_11_19	NA	FxF	90.3
12C_11_21	NA	FxF	0
12C_12_06	NA	FxF	0
12C_12_07	NA	FxF	100
12C_12_17	NA	FxF	0
12C_12_21	NA	FxF	93.4
12C_15_05	NA	FxF	100
12C_15_10	NA	FxF	95.5
12C_15_20	NA	FxF	0
12C_15_23	NA	FxF	95.7
12C_15_24	NA	FxF	0
12C_15_25	NA	FxF	0
12C_16_09	NA	FxF	91.7
12C_16_16	NA	FxF	76.5
12C_16_18	NA	FxF	0
12C_16_19	NA	FxF	0
12C_16_20	NA	FxF	0
12C_16_21	NA	FxF	0
12C_16_23	NA	FxF	97.87
12C_17_15	NA	FxF	0
12C_17_17	NA	FxF	0
12C_17_18	NA	FxF	0
12C_17_19	NA	FxF	0
12C_17_22	NA	FxF	0
12C_17_24	NA	FxF	0
12C_17_25	NA	FxF	0.6
12C_18_01	NA	FxF	100

12C_18_08	NA	FxF	95.7
12C_18_11	NA	FxF	94.7
12C_18_15	NA	FxF	0
12C_19_01	NA	FxF	0
12C_19_10	NA	FxF	13.9
12C_19_12	NA	FxF	1.6
12C_19_15	NA	FxF	96.83
12C_19_16	NA	FxF	66
12C_19_20	NA	FxF	0.1
12C_19_21	NA	FxF	78.8
12C_19_23	NA	FxF	1.1
African	2346	Pummelo	0
Algerian	279	Clementine	0
C. amblyscarpa	2485	Mandarin Hybrid	100
C. leiocarpa	3147	Mandarin Hybrid	100
C. reticulata	3812	Mandarin Hybrid	93.4
C. sunki	3143	Mandarin Hybrid	68.8
Chinese	1224	Pummelo	0
Clement	2603	Tangelo	0
Cutter	3030	Valencia	89
Deep Red	2347	Pummelo	0
Delfino	4114	Blood Orange	88.5
Dweet	3018	Tangor	62.7
EastHighlands	2369	Sweet orange	89.5
Ellendale	3960	Tangor	0
Encore	3569	Mandarin Hybrid	0
Fairchild	3559	Mandarin Hybrid	100
Fortune	3560	Mandarin Hybrid	0
Fremont	3558	Mandarin Hybrid	97.8
Frost	3178	Satsuma	100
Goutoucheng	4004	Sour Orange	82.5
King	3845	Mandarin Hybrid	25.7
Konejime	3611	Sour Orange_Hyb	100
Koster	3958	Tangor	0
Lee	2851	Mandarin Hybrid	0
Mandarinette	3405	Mandarin Hybrid	0
Pink	2246	Pummelo	0
Reinking	3805	Pummelo	0
Robinson	3850	Mandarin Hybrid	0
Scarlet Emperor	3326	Mandarin Hybrid	87.1

Seedless kishu	3906	Mandarin Hybrid	0
Siamese Acidless	2240	Pummelo	0
Soh Niamtra	3260	Mandarin Hybrid	100
Som Keowan	3752	Mandarin Hybrid	100
Sun Chu Sha	4003	Mandarin	84.5
Tahitian	3806	Pummelo	0
Temple	2598	Tangor	0
Thong Dee	3927	Pummelo	0
Tien Cieh	2590	Mandarin Hybrid	75.4
Ugli	4171	Tangelo	0
Webber Shaddock	4233	Pummelo	0
Wilking	3020	Mandarin Hybrid	0
Zhuluan	3981	Sour Orange	74.2

Supplemental Table 3. Percent polyembryony for 88 accessions in Poncirus GWAS panel

Sample Name	CRC	Type	Polyembryony Score
#27	3938	Trifoliolate	46.0
Argentina	3206	Trifoliolate	11.7
C32	3911	Trifoliolate Hybrid	58.0
C35	3912	Trifoliolate Hybrid	37.5
Chandxtrif_6B_12-16	NA	Chand x Trifol	73.1
Chandxtrif_6B_12-18	NA	Chand x Trifol	0.0
Chandxtrif_6B_12-20	NA	Chand x Trifol	0.0
Chandxtrif_6B_12-21	NA	Chand x Trifol	46.4
Chandxtrif_6B_12-22	NA	Chand x Trifol	2.3
Chandxtrif_6B_13-01	NA	Chand x Trifol	0.0
Chandxtrif_6B_13-03	NA	Chand x Trifol	0.5
Chandxtrif_6B_13-04	NA	Chand x Trifol	0.9
Chandxtrif_6B_13-05	NA	Chand x Trifol	0.9
Chandxtrif_6B_13-06	NA	Chand x Trifol	69.1
Chandxtrif_6B_13-08	NA	Chand x Trifol	0.0
Chandxtrif_6B_13-09	NA	Chand x Trifol	75.0
Chandxtrif_6B_13-13	NA	Chand x Trifol	71.1
Chandxtrif_6B_13-14	NA	Chand x Trifol	0.9
Chandxtrif_6B_13-15	NA	Chand x Trifol	0.0
Chandxtrif_6B_13-16	NA	Chand x Trifol	0.4
Chandxtrif_6B_13-17	NA	Chand x Trifol	0.0
Chandxtrif_6B_13-18	NA	Chand x Trifol	0.4
Chandxtrif_6B_13-20	NA	Chand x Trifol	1.0
Chandxtrif_6B_13-21	NA	Chand x Trifol	67.2
Chandxtrif_6B_14-01	NA	Chand x Trifol	0.4
Chandxtrif_6B_14-02	NA	Chand x Trifol	2.8
Chandxtrif_6B_14-03	NA	Chand x Trifol	8.5
Chandxtrif_6B_14-05	NA	Chand x Trifol	2.4
Chandxtrif_6B_14-07	NA	Chand x Trifol	57.1
Chandxtrif_6B_14-08	NA	Chand x Trifol	3.6
Chandxtrif_6B_14-10	NA	Chand x Trifol	61.3
Chandxtrif_6B_14-12	NA	Chand x Trifol	0.0
Chandxtrif_6B_14-13	NA	Chand x Trifol	1.1
Chandxtrif_6B_14-14	NA	Chand x Trifol	0.5
Chandxtrif_6B_14-15	NA	Chand x Trifol	0.5
Chandxtrif_6B_14-16	NA	Chand x Trifol	0.3
Chandxtrif_6B_14-17	NA	Chand x Trifol	56.8

Chandxtrif_6B_14-18	NA	Chand x Trifol	0.5
Chandxtrif_6B_14-20	NA	Chand x Trifol	0.7
Chandxtrif_6B_14-22	NA	Chand x Trifol	0.4
Chandxtrif_6B_15-01	NA	Chand x Trifol	0.0
Chandxtrif_6B_15-02	NA	Chand x Trifol	66.5
Chandxtrif_6B_15-03	NA	Chand x Trifol	0.3
Chandxtrif_6B_15-08	NA	Chand x Trifol	55.6
Chandxtrif_6B_15-09	NA	Chand x Trifol	0.0
Chandxtrif_6B_15-10	NA	Chand x Trifol	1.9
Chandxtrif_6B_15-15	NA	Chand x Trifol	3.5
Chandxtrif_6B_15-17	NA	Chand x Trifol	41.7
Chandxtrif_6B_15-18	NA	Chand x Trifol	0.2
Chandxtrif_6B_16-01	NA	Chand x Trifol	0.6
Chandxtrif_6B_16-02	NA	Chand x Trifol	1.2
Chandxtrif_6B_16-03	NA	Chand x Trifol	0.4
Chandxtrif_6B_16-06	NA	Chand x Trifol	27.3
Chandxtrif_6B_16-09	NA	Chand x Trifol	0.0
Chandxtrif_6B_16-12	NA	Chand x Trifol	0.0
Chandxtrif_6B_16-14	NA	Chand x Trifol	0.4
Chandxtrif_6B_16-15	NA	Chand x Trifol	55.3
Chandxtrif_6B_16-17	NA	Chand x Trifol	1.8
Chandxtrif_6B_16-18	NA	Chand x Trifol	0.0
Chandxtrif_6B_16-19	NA	Chand x Trifol	1.1
Chandxtrif_6B_16-20	NA	Chand x Trifol	4.7
Chandxtrif_6B-16-10	NA	Chand x Trifol	0.0
Chandxtrif_6B-16-21	NA	Chand x Trifol	0.0
Chandxtrif_6B-16-22	NA	Chand x Trifol	48.6
Chandxtrif_6E-32-01	NA	Chand x Trifol	0.0
Chandxtrif_6E-32-02	NA	Chand x Trifol	1.1
Citradarin	2618	Trifoliate Hybrid	0.0
Citradia	1436	Trifoliate Hybrid	20.0
Citrang #1416	2866	Trifoliate Hybrid	51.1
Citrangor	1447	Trifoliate Hybrid	20.8
Citremon	1449	Trifoliate Hybrid	13.3
Cunningham	271	Trifoliate Hybrid	100.0
Flying Dragon	3330A	Trifoliate	13.4
Glen Citrangedin	3573	Trifoliate Hybrid	22.0
Monoembryonic	3888	Trifoliate	0.0
Morton	2748	Trifoliate Hybrid	97.2
S281	3552	Trifoliate Hybrid	66.7

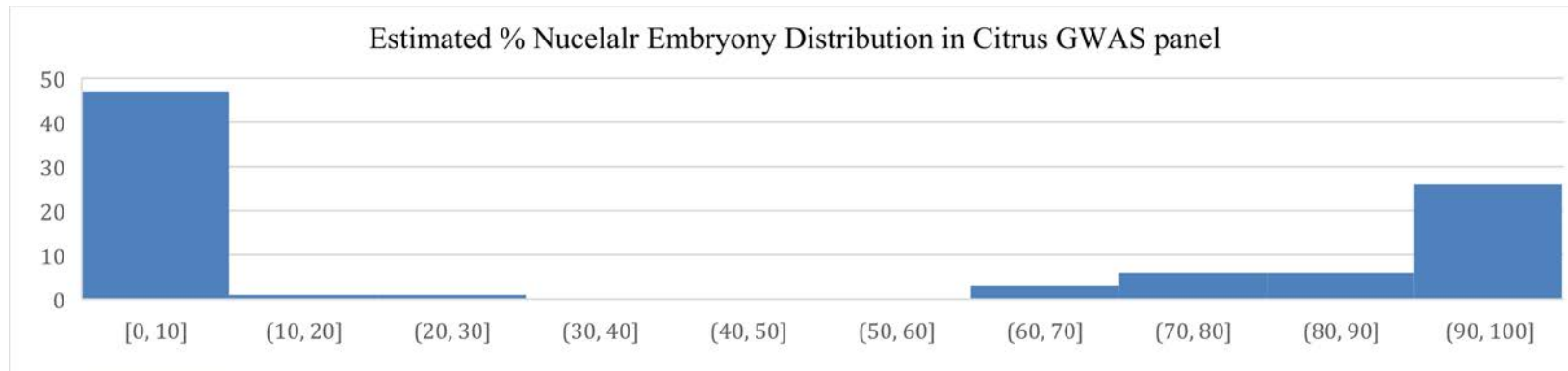
S302	3415	Trifoliate Hybrid	77.0
Sacaton	3337	Trifoliate Hybrid	33.8
Savage	275	Trifoliate Hybrid	67.9
Simmons	3549	Trifoliate	18.1
Small Leaf	4017	Trifoliate	0.0
Spanish Swo x P. trifoliata	3336	Trifoliate Hybrid	51.7
Swingle	3771	Trifoliate Hybrid	31.0
Swingle	3767	Trifoliate Hybrid	43.2
Trifeola	3954	Trifoliate Hybrid	30.2
Uvalde	2865	Trifoliate Hybrid	88.0
X639	3957	Trifoliate Hybrid	27.1

Supplemental Table 4. Average, median, and standard deviation of KASP estimated percent zygotie by groups.

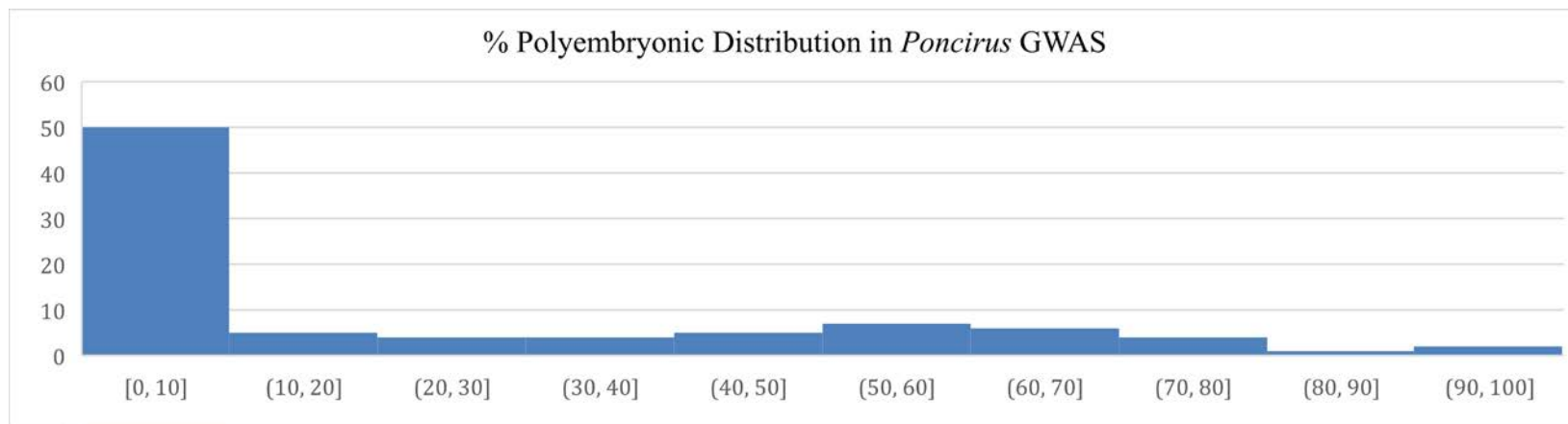
Average, median, and standard deviation of KASP estimated percent zygotie by groups was calculated for 210 individuals screened for both nucellar embryony and polyembryony

Type	N	Average	Median	Standard Deviation
Blood Orange	3	5.98	6.45	5.76
Mandarin/Mandarin hybrid	19	22.54	3.4	34.56
<i>Poncirus</i>	45	22.5	16.9	20.26
<i>Poncirus</i> hybrids	30	29.21	10.65	45.08
Sour Orange/sour orange hybrid	4	13.4	13.9	10.95
Sweet orange	11	11.3	4.5	16.15
Tangor	2	72.15	72.15	49.29
FxF	95	58.4	61.4	46.91

Supplemental Figure 1. Frequency distribution of nucellar embryony in *Citrus* GWAS panel
Frequency distribution of nucellar embryony score is plotted for 90 accessions in *Citrus* GWAS panel



Supplemental Figure 2. Frequency distribution of polyembryony in *Poncirus* GWAS panel
Frequency distribution of polyembryony score is plotted for 88 accessions in *Poncirus* GWAS panel



GENERAL CONCLUSION

High density Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array were developed with more than 1.4M and 58K autosomal, mostly gene-focused SNPs. The arrays also include 1002 and 500 chloroplast SNPs and both autosomal and chloroplast SNPs represented on the arrays are well distributed over the citrus genome. In general, high concordance of the genotyping calls between the sequence data and Axiom® citrus array data was observed, except in more distant taxa where concordance of 97.3 to 99.2% was observed. Sequence read alignment for variant discovery in the citrus variant discovery panel and probe sequence design for Axiom® citrus arrays used the *C. clementina* v1.0 reference genome. The lower concordance observed in *Poncirus* and some more distant taxa was presumably the result of reduced probe hybridization due to a higher frequency of additional, Clementine-divergent SNPs and indels in regions covered by SNP probes. However, possible genotyping error within the range observed in the study is expected to have little to no effect on the genome-wide analyses employed in the study based on near identity of analyses with and without simulated genotyping errors. Phylogeny inferred with Axiom® citrus arrays showed that these arrays provide sufficient marker density and information content to reveal similar relationships as analyses with whole genome sequence data. This study demonstrates that the Axiom® arrays for *Citrus* and some related genera provide a tool that is very powerful, useful and increases throughput for analyzing diversity of complex citrus genomes. They also provide us a robust, standardized set of markers that we can use repeatedly for mapping traits in various different populations and for characterization of germplasm collections.

A systematic diversity study using 924 CVC accessions was conducted with high density Axiom® citrus array data. The diversity study integrated multiple complementary analyses that elucidate the complex relationships among different accessions in CVC. Inclusion of many accessions of *Citrus* and related genera that have never been analyzed with a high-density marker

system revealed new information on their genetic profiles. This study identified additional candidate type-1 (non-introgressed) and type-2 (*C. sunki* or Cleopatra-like) mandarins with minimal introgression from pummelo. The CVC includes many non-admixed accessions of the ancestral taxa including 39 pummelos, 26 citrons, 8 papedas, and only 5 mandarins. Phylogenetic trees of these non-admixed accessions show the most closely related groups are *Fortunella* and *Microcitrus*, and that citron has little population structure. In addition, our observations based on multiple complementary analyses strongly suggest polyphyly of the papeda group and *Citrus hongheensis*, currently classified as papeda, is a sister species to pummelo. A total of 382 accessions were identified as genetically unique (not derived by mutation from other accessions), while 542 accessions were considered identical at the 51,296 SNP loci studied with percent identity ranging from 97.7% in orange group to 99.8% in clementine and considered either duplicates or differing only by mutation. A phylogeny network analysis of these 382 diverse accessions shows that relationships are affected by extensive hybridization, a conclusion also consistent with admixture analysis. For accessions showing admixture from two ancestral species a more detailed analysis of the ancestry of each 0.5 Mb chromosome segment was conducted and further documents complex hybridization in the history of citrus. This extensive inventory of the genetic profiles of the accessions in the CVC will assist researchers in efficient germplasm conservation and utilization in breeding. This information could be used to update the sometimes questionable passport records of many accessions added to the CVC during its more than 100 years of history.

To determine genetic control of nucellar embryony and better characterize variation in the percent nucellar vs zygotic seedlings produced, a total of 114 CVC accessions and 95 hybrids between a monoembryonic mandarin and a polyembryonic mandarin were analyzed for both % nucellar seedlings using KASP SNP marker analysis and % polyembryony scored from germinated seeds. This showed that these traits are only loosely correlated. Comparison of seed samples

collected from replicate trees and in different years revealed significant environmental influences on expression of nucellar embryony in both *Citrus* and *Poncirus*. This survey data could be useful for breeders in identifying parents for future hybridization. GWAS was also conducted using high-density Axiom® citrus array data, and identified a 101 kb region on chromosome 1 overlapping a previously identified polyembryony locus in *Citrus* and potential minor loci on chromosome 3 and 4 using KASP marker estimated % nucellar embryony as a phenotype. GWAS in *Poncirus* using polyembryony score as a phenotype identified a 2.2 Mb region on chromosome 3 overlapping the previously identified polyembryony locus in *Poncirus*. No other significant QTLs were detected possibly because the study lacked power due to a small population size and a lower number of segregating markers in *Poncirus* in comparison to *Citrus*. Additionally, the use of polyembryony score as a phenotype could have contributed to the results as a large discrepancy between KASP estimated % nucellar seedlings and visually assessed % polyembryonic seeds was observed in some samples studied.

Poncirus and *Citrus* reportedly diverged 9 million years ago, therefore, it is possible that the lack of a *Poncirus* reference genome limits the accuracy of a fine scale genomic study in *Poncirus* using *Citrus* reference genomes. Availability of a *Poncirus* genome will significantly improve our understanding of genetics and genomics underlying unique traits associated with *Poncirus* in the future. The association studies showed that the high-density SNP array data combined with KASP SNP marker estimated nucellar embryony phenotypic data provided sufficient power and resolution to identify a candidate region underlying nucellar embryony in *Citrus* and *Poncirus*, and revealed that the genetics controlling for nucellar embryony in *Citrus* and *Poncirus* are likely different. Further study will be necessary to fine-map candidate genes in *Poncirus* and to identify minor affecting loci controlling the proportion of nucellar embryony in polyembryonic accessions in both *Citrus* and *Poncirus*. Additionally, further investigation into the

possible relationships or interaction between the *Citrus* nucellar embryony locus and the *Poncirus* polyembryony locus should be valuable.

High-density Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array have promising utility in genetic and genomic analyses to dissect important traits such as fruit quality, nucellar embryony, and disease resistance, given the high number of robust SNPs represented on the arrays and the observations from this study. High density Affymetrix™ Axiom® Citrus HD Genotyping Array and Axiom® Citrus Genotyping Array are now publicly available, and currently used to investigate loss of heterozygosity (LOH), phylogeny inference, characterization of germplasm and trait mapping, providing citrus breeders and geneticists with a valuable new tool for genomic analysis.