UC Riverside

Title

Linkage Map Analysis of Segregation Distortion in Crosses Between Citrus and Poncirus

Author

Thomas, Zachary Mark

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

RIVERSIDE

Linkage Map Analysis of Segregation Distortion

in Crosses Between *Citrus* and *Poncirus*

A Thesis submitted in partial satisfaction

of the requirements for the degree of

Master of Science

in

Plant Biology

by

Zachary Mark Thomas

December 2023

Dissertation Committee:
     Dr. Mikeal L. Roose, Chairperson
     Dr. Timothy Close
     Dr. Zhenyu Jia

The Thesis of Zachary Mark Thomas is approved:

_____

_____

_____
                                                Committee Chairperson

University of California, Riverside

<u>**Dedication**</u>

I would like to dedicate this thesis to my mother Sheri Thomas who has always supported and helped me when I needed it and who supported my change to botany as a major because "I think you will be much happier". You could not have been more correct, thank you for everything.

I would also like to thank my professor Mikeal Roose for all the hard work and teaching he did to help me achieve this milestone, I hope to continue to work together. A large thank you to my committee Timothy Close and Zhenyu Jia for advice and support. Lastly a thank you to all the staff and students who have supported and assisted with my career over the past 10 years of working in the Roose lab especially to Claire Federici and Niel Stone who taught me so much and helped guide me.

Table of Contents	**Page Number**

List of Figures   **Page Number**

List of Tables   **Page Number**

Introduction

Citrus is one of the most important fruit crops worldwide, being grown in over 130 countries (Cuenca et al. 2018), with 158.5 million tons produced worldwide (citrus industry 2022) and 16.8 billion dollars in world trade (OCE 2020). The US citrus industry is valued at 3.4 billion dollars (USDA 2020). The family Rutaceae comprises three subtribes Triphasiinae, Balsamocitrinae, and Citrinae (Swingle and Reese 1967). Citrinae is the subtribe that contains most of the economically important genera, including *Citrus, Poncirus, Fortunella, Eremocitrus,* and *Microcitrus*. The last two have only recently gained interest due to studies showing tolerance and/or resistance in the species to Huanglongbing (HLB), also known as Citrus Greening (Ramadugu et al. 2016, Alves et al. 2021). *Citrus, Poncirus,* and *Fortunella* have been used extensively as scions, rootstocks, or both and in many hybrids for scion and rootstock breeding. The most commonly used classification system in the US was proposed by Swingle and Reese, containing 36 species in the genus *Citrus*. They were the first to include biochemical markers in their classification determination and the typical history and morphological characteristics. These classifications are still debated today as newer studies, such as the 2018 study by Wu et al., show that there were five progenitor species in the genus *Citrus* C. *medica*, C. *reticulata*, C. *maxima*, C. *micrantha*, and *Fortunella*, with most of the other species being complex hybrids of the five progenitor species. The new genetic information has prompted a new taxonomic system to show the complex admixture and reduce the number of genera in the subtribe (Mabberley 2004).

Citrus breeding has been done for hundreds of years with moderate success. Still, it has been "hampered by its' complex genetics and reproductive biology (apomixis, partial pollen and/or ovule sterility, cross- and self-incompatibility and high heterozygosity). In addition, citrus has a long juvenile period and usually takes several years for hybrids to set fruit." (Cuenca et al. 2018). Some species of citrus have nucellar embryony meaning that embryos of the species can form from nucellus tissue of the mother plant, making the embryo genetically identical to the mother barring mutations. These seeds can also be polyembryonic, meaning they contain multiple embryos from a single seed, which can be useful in creating plant replicates. Both of these traits can hamper a breeding program, as the primary goal in such a program is to create hybrids with beneficial traits. A high degree of nucellar embryony precludes the production of hybrids meaning that seeds collected may not be hybrid seeds even with controlled crossing. Nucellar embryony usually co-occurs with polyembryony, where multiple embryos almost always arise from nucellar tissue, with studies showing that even when a hybrid is present, it is often smaller than the nucellar embryo (Xiang 1988).

Most species within the genus Citrus are diploid (n=18), with a few examples of triploids and tetraploids (Krug 1948). These triploids and tetraploids, some created through natural mutation, have been propagated but don't survive well in the wild. Many more were created through manipulation, such as colchicine or fusion (Grosser and Gmitter 1990), in an effort to improve the cultivars with traits such as size increase or seedlessness. The citrus genome has nine chromosomes with an approximate size of 370

megabases. The complexity of citrus genetics has led to the slower development of genetic tools to assist breeding efforts.

The creation of markers associated with quantitative trait loci (QTLs) for genes is a step towards understanding Citrus genetics better and improve breeding results. Several types of markers have been developed over the years, and as new technologies have become available, they were applied to citrus as well. These include isozymes (Roose 1988; Torres et al. 1978, 1982); random amplified polymorphic DNA (RAPDs) (Luro et al. 1994); sequence characterized amplified regions (SCARs) (Nicolosi et al. 2000); restriction fragment length polymorphism (RFLPs) (Federici et al. 1998); simple sequence repeats (SSRs) (Barkley et al. 2006; Chen et al. 2008; Cuenca et al. 2011; Froelicher et al. 2008; Garcia-Lor et al. 2012, 2015; Kijas et al. 1997; Luro et al. 2001, 2008; Terol et al. 2007, 2008); intersimple sequence repeat (ISSRs) (Fang et al. 1997); amplified fragment length polymorphism (AFLPs) (Liang et al. 2007; Pang et al. 2007); cleaved amplified polymorphic sequences (CAPs) (Lotfy et al. 2003); insertion-deletion (Indel) and diversity arrays technology (DArT) (Curtolo et al. 2017); and single nucleotide polymorphism (SNPs) (Chen and Gmitter 2013; Cuenca et al. 2013b; Garcia-Lor et al. 2012, 2015; Ollitrault et al. 2012a, b). The first citrus maps were created using RFLP and isozyme markers, with subsequent maps utilizing different types of markers such as AFLP and RAPD (Durham et al. 1992, Jarrell et al. 1992, de Simone et al. 1998, Sankar and Moore 2001, de Oliveira et al. 2007, Gulsen et al. 2010). The first genome published in citrus was created by Xu et al. 2012 utilizing paired-end-tag sequencing of *Citrus sinensis.* A reference linkage map for citrus was created by comparing the *Citrus*

clementina genome and five parental genotypes using sequence-derived SNP, SSR, and Indel markers (Ollitrault et al. 2012a).

Simple sequence repeat markers are a type of microsatellite developed by Litt and Ludy in 1989 and Akkaya et al. (1992) First published work applying microsatellites in plants. These markers locate regions of the genome known as tandem repeats (TRs). The mutation rate in such regions is high, with tandem repeats being spread throughout the genome. Perfect repeat sections more extended than 20 base pairs were discovered to be highly polymorphic, with sections often having deletions or insertions of repeat elements (Vieira et al. 2016). Markers based on this variation are often co-dominant and multi-allelic, allowing for the confirmation of parentage when creating hybrids in citrus, as each parent will contribute one, an often unique, allele at each SSR locus. SSR markers were used in citrus to study genetic diversity and population structure in 2006 by Barkley et al. and have often been used to determine if a citrus tree was derived from hybridization or nucellar formation before use in the breeding program. SSR markers are also helpful in citrus for confirming the presence of each of the different chromosomes, as specific markers were designed to amplify only a region on each specific chromosome.

Single nucleotide polymorphisms (SNPs) are positions in a DNA sequence where a single base has been changed between the alleles. SNPs are the most common form of mutation found in DNA and have been used extensively in genetic studies to advance our knowledge of genetics. SNPs are used in plants to differentiate germplasm, track sections of DNA through lines, introgress important QTLs, select progeny, study diversity, identify cultivars, and develop linkage maps (Rafalski 2002). Specifically in breeding

programs, SNPs have been used to develop disease resistance, drought tolerance, consumer traits, processing traits, etc., in corn, wheat, barley, onion, melons, rice, mustard, tomato, and more (Morgil et al. 2020, Brenner et al. 2012, Scholten et al. 2016, Rana et al. 2019, Fridman 2004, Perpiñá 2016). SNPs are common within a genome, with a SNP occurring approximately every kilobase pair along the genome. SNP markers are generally biallelic, making the markers easier to use and analyze (Morgil et al. 2020). SNPs are located using in-vitro methods such as genome sequencing and restriction digestion and then markers can be designed around those regions (RFLPs and CAPS). There are also in-silico methods which use the DNA sequences or previously created genome libraries to find SNP sequences; the markers are then designed and tested. This is often called SNP mining (Morgil et al. 2020). These SNP regions can be used in citrus for marker-assisted breeding efforts allowing the tracking of important QTLs through generations. Those QTLs can also help inform gene editing techniques such as CRISPR or Agrobacterium.

Markers linked to various traits have been generated, such as dwarfing from Flying Dragon (Cheng and Roose 1995), fruit acidity (Feng et al. 1997), CTV resistance in *Poncirus* (Asins et al. 2004; Deng et al. 1997; Gmitter et al. 1996; Yang et al. 2003), anthocyanin in the pulp (Butelli et al. 2012, 2016), nucellar embryony (Wang et al. 2017), and Alternaria brown spot resistance (Dalkilic et al. 2005; Gulsen et al. 2010). These markers can accelerate citrus breeding efforts by allowing for screening populations at a seedling stage rather than phenotypic testing. Marker-based selection is

particularly valuable for traits like dwarfing where phenotyping is a long and cumbersome task that can be hindered by environmental factors.

The cost of genotyping large numbers of SNP markers and developing SNPs by sequencing many genotypes has decreased annually, allowing such technologies to be utilized in more plant genetic and breeding programs. SNP chip arrays are a technology used in genetic studies of citrus and other species. SNP arrays contain hundreds of thousands of unique sequences designed to bind to specific targets in the sample DNA. Laser confocal scanning is then used to read the fluorescence signal from each sequence location, with each SNP allele being represented at a different location. The signal measures the abundance of each queried SNP allele in the genome. The signal data can then be run through a program to determine which SNP alleles are present at each locus. In genetic studies, SNP markers are selected throughout the genome and can be used in linkage mapping to detect anomalies such as segregation distortion, loss of heterozygosity, and aneuploidy. Hiraoka and Roose developed an SNP array for use in citrus in 2020. They hybridized 924 citrus accessions from the Citrus Variety Collection at UC Riverside, creating a sizeable genomic database for comparing the accessions, determining admixture, phylogeny, introgression, and QTL analysis. SNP markers for traits of interest, such as the markers developed above, can be queried using other marker systems, such as Kompetitive allele-specific PCR **(**KASP) markers, for cheap and quick screening of many traits on various hybrids to identify appropriate parents and select progeny which have the genes associated with traits required for success as a scion or rootstock.

Citrus breeding is often separated into two categories: scion and rootstock. Scions are the top part of the tree grown for their fruit characteristics but also some disease tolerances, vigor, shape, and color of leaves. Rootstocks are the bottom part of the tree, selected to have disease resistance, good nutrient-acquiring ability, drought tolerance, salinity tolerance, and ease of replication. *Citrus* has been grown as grafted plants for hundreds of years, with the method's popularity increasing since the late 1800s (Castle 2010). Grafting is done because a rootstock affects the scion, including yield, fruit quality, health, dwarfing, precociousness, disease resistance, and flowering time (Wutscher 1970), traits for which a desired phenotype may not be present in a scion variety. The process can be done in several different ways, but all methods rely on aligning the cambium layers of the scion with the rootstock and allowing the plant to heal these together into a single tree. This allows the combination of scion traits such as fruit taste, yield, and tree vigor with rootstock traits of disease resistance, root vigor, and nutrient acquisition. The main rootstocks for *Citrus* include *Poncirus* and hybrids between the two species. *Poncirus* is a close relative of citrus. It is deciduous with a trifoliate leaf that has given it the common name trifoliate orange. *Poncirus* has many traits that have made it worthwhile as a rootstock, including phytophthora resistance, citrus nematode resistance, Citrus Tristeza Virus (CTV) resistance, cold tolerance, and huanglongbing (HLB) tolerance. The species tends to have many seeds, which contain nucellar embryos and are polyembryonic, meaning many copies of a plant can be created quickly and with little effort. Polyembryony however makes breeding difficult as most if not all of the seed produced from a cross will be clones of the parent and not hybrids. So

while the trait of polyembryony is desired in commercial nurseries for production it limits breeding efforts. *Poncirus* tends to be used as a male in many crosses, thereby avoiding polyembryonic seed by using a female parent which does not have polyembryony. The trifoliate leaf trait is dominant, a characteristic that allows leaf morphology to be used to identify successful crosses with citrus. However, as generations progress using citrus as a parent the use of leaf morphology is less effective to confirm that a seedling is a hybrid derived from trifoliate. Markers are often used to identify hybrids and thus have made crosses where few of the progeny will be hybrids more acceptable as the hybrids will quickly be separated from the nucellar seedlings at a young age, and resources aren't being wasted on nucellar progeny.

Poncirus was used as a rootstock sparingly until the 1940s when the graft-transmissible pathogen CTV began to spread throughout the *Citrus* growing regions of the world. This disease attacks the graft union between sweet orange and sour orange, causing quick decline and death in many cases. Alternatives to sour orange were necessary, and trifoliate orange was discovered to have resistance to CTV (Wallace 1978). Trifoliate orange started to become adopted by the US citrus industry due to the many valuable traits listed above and because sour orange, which was also resistant to many of the same pathogens and had other beneficial traits, could no longer be used as the trees would lose productivity and die. Issues with using trifoliate orange as a rootstock included slow growth, unsuitability in calcareous soils, reduced yield, and compatibility issues when grafted with some citrus scions. Breeding programs began to focus on combinations of *Citrus* and *Poncirus* that combined many of the beneficial traits

from both genera into one rootstock. One of the first successful hybrids was Carrizo, a cross of sweet orange x trifoliate. This hybrid is still one of the most used rootstocks in the US. Crosses between the two genera had a specific naming convention, with crosses sweet orange x trifoliate named citranges. In contrast, other hybrid names include citrumelo (grapefruit x trifoliate), citrandarin (mandarin x trifoliate), citradia (sour orange x trifoliate), etc. These hybrids combined many of the beneficial traits of both parents, reducing the negative effects of using pure trifoliate as rootstocks.

Crosses between *Citrus* and *Poncirus* have become the most popular type of rootstock used in the United States, with 80% of Florida propagations using these hybrids as rootstocks for 2020-2021(ccqc.org). The number is likely similar in many of the other *Citrus* growing regions in the US though the numbers of each cultivar grown is unavailable for the other states. However, these F1 hybrids still have limitations attributed to their trifoliate lineage, as they often have problems growing in calcareous soils and compatibility issues when grafted with certain scions. Backcrossing has been attempted many times by our breeding program and others. However, this has failed, with no commercial rootstocks being released from these backcross attempts despite several years and the use of varying parents. Both *Citrus* and *Poncirus* are diploid and are largely heterozygous in their genomes; this has likely allowed for many negative recessive alleles to arise within each species. These alleles lie dormant in the genome because the neutral or positive dominant allele prevents the negative trait from expressing. When outcrossed, the initial hybrids express a kind of hybrid vigor as the negative sections of the passed-along alleles are compensated for by the other species' allele. This includes negative

alleles expressed in the parents but masked in hybrids by the other species' alleles. However, the presence of negative recessive alleles in the F1 means that when attempting to back cross or hybridize with another *Citrus* x *Poncirus* hybrid, negative alleles are much more likely to express at a number of locations on the genome where these alleles can become homozygous. Consequently, when the F1s are hybridized with each other, those masked negative alleles often become homozygous in the offspring, leading to progenies that are weak or contain other detrimental phenotypes. This is a type of inbreeding depression. Inbreeding depression is often extreme in citrus and can prevent the plant from reaching maturity, with plants that do mature not performing as well as either of their parents and often worse than their grandparents. These issues may be exacerbated by segregation distortion within the population.

Segregation distortion occurs when the genotype frequency in observed progeny differs from Mendelian expectations. One of the mechanisms by which they are distorted is when specific alleles can be preferentially transmitted to progeny and can be either gametic or zygotic in origin. This type of distortion has been encountered in numerous species and populations in both animal and plant studies, including rice (Reflinur et al. 2014), maize (Tang et al. 2012), soybean (Zuo et al. 2019), and apple (Kenis and Keulemans 2005). The distorted regions can contain deleterious or beneficial genes; however, distortion also occurs around genes that prevent self-fertilization in some plant species. When the parental individuals are related, this can happen in crosses if a self-incompatibility (SI) region exists and the parents share alleles at the SI locus. Progeny homozygous for the shared allele will not occur, and the flanking region will show

segregation distortion. The segregation distortion that has been detected so far in both plants and animals is most often found in sexual crosses from divergent lines or species (Reflinur et al. 2014, Fu et al. 2020) so that a (*Poncirus* x *Citrus*) x *Citrus* should show more distortion than *Citrus* x *Citrus*. In these cases, it has been shown that segregation distortion mainly derives from meiotic failure though a few cases have been found to relate to the formation of the zygote (Lin and Ikehashi 1993, Li et al. 2011, Xu et al. 2013, Reflinur et al. 2014). The first type of meiotic failure is a male gametic failure that can be caused in several ways, leading to segregation distortion. After meiosis, the pollen cell could fail to form altogether due to a lethal gene on the single chromosome. In this case, the transmission of any gene near the lethal would be reduced and only present in offspring if there was a crossover event close to the lethal gene's location. The pollen genotype could also have reduced ability to compete with other pollen genotypes from the same plant. For example, if the pollen is smaller and has fewer nutrients to sustain its growth, or has genes which affect pollen tube growth, it could cause the progeny ratio to be distorted. However, the distortion may be less severe than in the previous case. The second type of meiotic failure comes similarly when the female gamete fails to form due to a lethal gene being present or, as above, when the egg cell is weaker and unable form a zygote that grows into a seed properly (it is not uncommon to find extremely tiny or flat seeds in citrus, this may sometimes be due to a lethal or deleterious gene).

Distortion may also arise after the two gametes fuse to form the zygote. In this case, a gene that is only active once the zygote forms, but which is lethal once activated can cause the seed to abort (Burbidge and James 1991) or, in citrus, may lead to that

11

particular embryo being too small and being crowded out by developing nucellar embryos. A common cause of distortion may also be the death of the seedling at a young age before the material is collected for genotyping, the genes may not be immediately lethal, but the low vigor of the plant could lead to differential survival which will be reflected in distortion around such loci.

Distortion in Citrus x Poncirus crosses has been detected using isozyme makers (Torres 1985, Ruiz and Asins 2002); however, the location and method of distortion were not ascertained. Other studies using isozymes or RFLP markers have detected segregation distortion in crosses between or within citrus species (Durham et al. 1992, Jarrell et al. 1992, Cai et al. 1994, Kijas et al. 1997). SNP marker genotyping has also been used more recently to create maps with more density and discover more specific regions where distortion events occur in populations (Ollitrault et al. 2021), leading to the discovery of an S gene related to self-incompatibility in mandarins, the mechanism being a failure of pollen tubes to complete its growth down the style. New methods are also being applied to study segregation distortion, such as single pollen grain whole genome amplification by Garavello et al. 2020. This method detects any pollen death-related distortion events and allow for distinguishing distortion related to pollen death from other sources of distortion by comparing the distortion in the pollen grains with the distortion in a related population.

Linkage mapping is the primary method for discovering segregation distortion in a population. During the prophase of meiosis, the chromatids pair up, and crossover events can occur, allowing sections of the chromatids to exchange with each other. The

chromatids then separate into individual gametes over the course of meiosis. Linkage

mapping takes advantage of these crossover events to determine the distance between any

two locations on the genome as measured by the number of crossovers. When markers

are on the same physical chromosome they may show linkage if recombination events

between them are rare. Genes far apart on the same chromosome are likely to have

multiple recombination events between them which results in a recombination frequency

estimate of 50%, the same as observed for genes on non-homologous chromosomes.

Using recombination estimates for many pairs of genes, a map can be created showing

the distance between two markers as a function of the number of crossover events

detected between those two markers in a particular hybrid population. Closer markers

will have few to no crossover events, while distant markers will have more crossover

events detected in the progeny. The number of progeny or markers used in such mapping

can increase the accuracy and resolution of the map and enable locating specific markers

and, when combined with phenotype data on progeny, determining the effects and

position of trait genes of interest. Several different programs are used to create a map

from marker data. Map distance is calculated using recombination with this equation

(Map distance = 100* recombination frequency= 100*(number of recombinant

progeny/total number of progeny)) (Sturtevant 1913). By comparing the recombination

frequency between all marker pairs, the order of the genes in terms of distance from each

other can be calculated.

This study aims to discover genome regions distorted in crosses involving *Citrus*

x *Poncirus* hybrids. This distortion may help to determine which genes will be

13

challenging to introgress into *Citrus*. These types of analyses should allow breeders to determine what parents to use, which parent to use as male and which as female, and how many progenies may be necessary to achieve success in combining important traits into a single tree, be it rootstock or scion. Many rootstock traits, such as tolerance to disease, are likely quantitative traits. If regions are more challenging to introgress, then using different parents or discovering other loci that contribute to these traits should be the focus of breeding efforts rather than searching for progeny with traits in a highly distorted region where the positive allele is rarely transmitted. However, these distortion can also favor the introgression of desired alleles making the parent ideal for hybridizing as the trait would be much more likely to present in the offspring. Combining the information from studies of different populations and crosses within *Citrus* will allow breeders to make better choices and succeed tremendously, especially when combined with marker-assisted breeding.

Mandarin orange (*Citrus reticulata*) is one of the three ancestral species of citrus. They are often combined commercially with other similar-looking "species", such as clementines (*Citrus clementina*), satsumas (*Citrus unshiu*), and tangerines (*Citrus tangerina*). However, these groups are all considered mandarin oranges under the Swingle taxonomic structure. China is the largest producer of mandarins worldwide, with over 20 million metric tons produced per year (Citrus Industry 2021), with the US being a major importer of mandarins despite production growth. The number of mandarins grown in California has risen since the early 2000s (Geisseler 2016). This is likely due to the release of seedless mandarin varieties, such as Tango, which allow for cheaper

14

production. Prior to these releases, the trees required netting to exclude pollinators which resulted in seedless fruit. Mandarins have grown in popularity, acreage, and tons sold in the US. This can be attributed to marketing varieties such as seedless W. Murcott sold by Mulholland farms as Delite® (Citrusvariety.ucr.edu) and Sun Pacific's seedless fruit sold as Cuties®. These marketing names and others have become popular in California and other parts of the country for their look, taste, and lack of seed.

With the growing popularity combined with the industry's desire for more novel types of mandarins to meet consumer desire for new flavors and colors, work on the genetics of mandarins has risen in importance within our breeding program. Some varieties of mandarins are self-incompatible, which further complicates breeding efforts. Determining regions of distortion will allow for the study of those regions to determine what effect this has on the scion traits. This study shows some of the results from one cross between two mandarin types and the distortion that results and will need to be considered when breeding mandarins in the future.

Procedure

Development of mapping population. A population created from two parents, *Citrus grandis* cv. Chandler the male parent, and *Citrus grandis* cv. Tahitian pummelo *x Poncirus trifoliata* cv. Flying Dragon, (Tahitian pummelo x Flying Dragon (TPxFD)) the female parent. Chandler was chosen as the male parent because, in previous populations studied by the Roose lab (unpublished), Chandler showed very little to no segregation distortion in markers transmitted through pollen. The female TPxFD was chosen for two

reasons, being a pummelo x trifoliate hybrid, the resulting population would reflect what

can be expected from backcrossing a female with *Citrus* x *Poncirus* parentage, and the

selected TPxFD hybrid is also monembryonic, meaning that the seeds contain only one

embryo, and that embryo is zygotic in nature. Using a monembryonic parent allows for

the production of many hybrid seeds quickly and without having to screen the seedlings

for nucellar plants. Approximately 20 unopened mature flowers of Chandler were

collected into a labeled paper bag and immediately transported to the lab. The flowers

were opened, and the anthers were removed and placed on wax paper and left underneath

an incandescent light overnight to dry the anthers and cause dehiscence. The dehisced

anthers were then placed inside a labeled tube which can be stored at 4℃ for up to a

week. The tubes of pollen were stored on ice while transported and used in the field.

Unopened mature flowers on the TPxFD tree were located, and the petals were forced

open and, in many cases, removed. The flowers were then emasculated, and a paintbrush

of Chandler pollen was applied to the flower's stigma. Then a Uline organza bag was

placed over the flower or group of flowers to prevent pollination from other sources.

Labels with the cross information were attached to the branch above the bag. The flowers

were then left to develop into fruit, and the fruit was allowed to ripen before the fruit was

harvested for seed extraction, keeping the labels with the bags. In some cases, during the

summer, as the fruit developed, the bags were not large enough, so the original organza

bag was removed, and a larger mesh bag was placed over the fruit and tied in place. The

reciprocal cross was also attempted using TPxFD as a male, but this cross failed, perhaps

due to self-incompatibility genes shared by the pummelo parents.

Ripe fruits were collected in December 2019, then the fruit was cut shallowly along the equator, and the fruit was then twisted to separate the halves without damaging any of the seeds. The seeds were then washed with soap and water to clean and placed in ClariSEB RL (a pectinase) for 30 minutes to remove the leftover pulp and reduce stickiness. The seeds were then placed in a 10% commercial bleach solution for 10 minutes to sterilize the surface. The seeds were rinsed in DI water and then placed in a 52°C-water bath for 10 minutes to heat treat the seeds removing some internal pathogens such as citrus stubborn (*Spiroplasma citri*) and Phytophthora. The seeds were then rinsed in cool DI water and treated with a 10% commercial bleach solution for 10 minutes. Then the seeds were rinsed and placed in a 1% 8-hydroxyquinoline sulfate solution for 3 minutes to prevent mold when packaged (Bridges 1966, Castle 1981). The seeds were dried overnight and stored in a plastic Ziploc bag at 4°C until planted.

A total of 327 seeds were collected. These seeds were left unsorted to prevent bias in the population chosen later. The seeds were planted into 98-well container trays in January 2020. The plants were then grown for 11 months in a greenhouse. Ninety-five plants were chosen randomly, and eight leaf discs per sample were collected and placed in a 96-deep well plate, the last well containing eight leaf discs from the female parent TPxFD. Dr. Hiraoka included the male parent Chandler in previous years' DNA extractions. After collecting, the samples were kept on ice and transferred to a -80 freezer to completely freeze them before processing. The samples were then placed in a lyophilizer until completely dry. The eight leaf discs which are estimated to be the same size, together comprising approximately 100mg of tissue. This was based on the weight

of fresh tissue from 3 samples, though samples may have varied slightly due to variations in leaf thickness.

DNA extraction and quantification. DNA extraction was performed in December 2020 using the Magattract 96 DNA Plant Core Kit from Qiagen (Qiagen.com). The eight leaf discs were ground using 10-2.5mm zirconia beads placed in each well with the lyophilized leaf samples. The plate was placed in a Mini-Bead Beater 96+ from Biospec for 3 minutes. The protocol from Qiagen was modified using 500µl of lysis buffer instead of 300µl to compensate for the oil that citrus leaves contain. The solution was then bead-beat for another 5 minutes and then centrifuged. The supernatant was transferred to a new plate and combined with the magnetic beads in ethanol. This causes the DNA to bind to the beads. The samples were then washed several times and treated with RNase A to degrade RNA and remove non-DNA contaminants from the sample. Elution was done using 200ul of 10mM Tris, .5mM EDTA (TE) buffer and stored at -20°C until needed for PCR or shipping.

DNA was quantified with Thermo Scientific NanoDrop 2000c spectrophotometer (Thermofisher). Since the spectrophotometer is not specific to DNA, a sub-sample of 5 extracts was quantified using the Invitrogen Qubit 2.0 fluorometer, the kit used to stain the DNA was the iQuant dsDNA HS assay kit following standard protocol included with the kit (Thermofisher) to ensure that there was enough DNA in some of the samples with lowers ng/µl reads. The amount of DNA detected by the fluorometer did not correlate well with the amount of total nucleic acids present when the spectrophotometer was used; however, even with the lower samples, there was sufficient DNA present for all

downstream processes, so it was determined that the rest of the samples could be used without further analysis.

SSR marker testing. SSR markers were used to verify that each of the chromosomes was present and that the DNA in the select samples was citrus DNA (Barkley 2006). Each sample was diluted to 5ng/µl before being used in the PCR reaction. The PCR was done using a reaction buffer consisting of 2.5µl 10mg/ml BSA, .2 µl 5U/µl Platinum taq, 2.5µl 10x taq buffer, .5µl .2mM dNTPs, 1µl 2mM MgCl2, and 1µl .2uM SSR primers. The 5µl of DNA template was added, and the total volume was adjusted to 25µl with water. The sample was then run in a Genomyx Cycle LR thermocycler with the program set for 5 minutes, first denaturing at 94℃, denaturing 94℃, annealing 55℃, extension 72℃, for forty cycles, with a final extension at 72℃. The samples were stored in a -20℃ freezer. 5µl of each sample was run on a 2% agarose sodium borate gel using sodium borate as the buffer in a HE99X Max Horizontal Agarose Electrophoresis Unit. The agarose was mixed with sodium borate solution and heated using a microwave in 30-second intervals until the agarose completely melted. The mixture was then left to cool, then Gel Red (Biotium) was added to the gel before being poured in a cast and left to solidify fully. This included time in a refrigerator to ensure the gel was completely solidified. 8µl of the sample with 2µl of 5x loading dye were loaded into the wells with a Fisher BioReagents 100bp low-scale DNA ladder (Fisher Scientific). The gel was run at 100 volts for 1.5 hours using a Biorad Powerpac 300 power supply. The Biorad Molecular Imager Gel Doc XR+ was set to detect weak bands and allowed to automatically activate the ultraviolet

light causing fluorescence in the Gel Red bound to the DNA in the lanes. The SSR

markers showed that citrus DNA was indeed present in the samples (Figure 1).
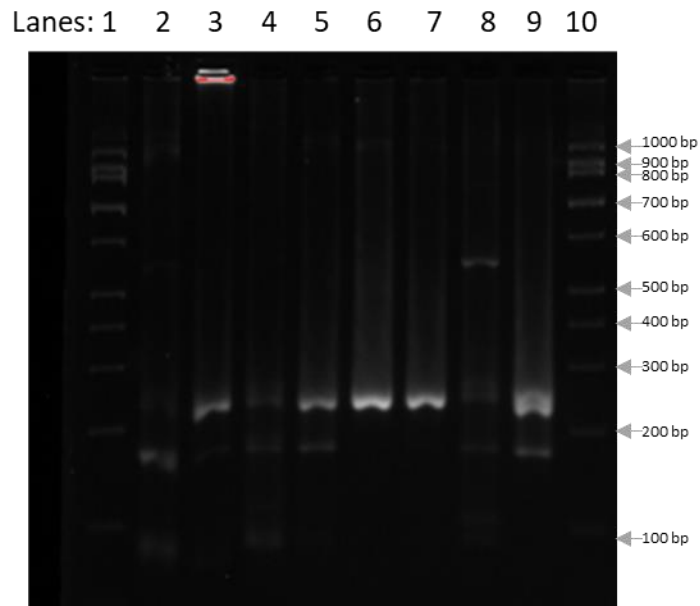
Figure 1: Agarose gel with random subset of (TP x FD) x Chandler
samples in wells 2-8, well 9 contains the parent TP x FD. Lanes 1
and 10 are 100 bp ladder.

SNP Array genotyping. For each sample, 40 µl was transferred to a new plate without

dilution to ensure enough DNA was present for SNP array analysis. As stated, we did not

quantify every sample using the more specific method. The samples were then shipped to

Affymetrix for analysis on the Axiom™ Citrus56 Array (Affymetrix), which includes

57,933 autosomal and 500 chloroplast SNPs. This chip was designed using the

Clementine genome sequence as a base for the design of the probesets used in the chip.

These probesets were designed from SNPs detected across the entire genome to allow for

genomic comparisons in other citrus and relative species. The quality control for the

DQC (a measure of discrimination between positive signal and background signal) was set to >.82. These markers are intended to be present in all citrus and relatives. They are used to ensure that the data we obtain is accurate. Then call rate was set to 97% to eliminate samples with a relatively high no-call rate associated with a higher error rate. To the (TP x FD) x Chandler population were added 49 files containing Affymetrix data collected from various citrus in previous years. This was done to improve the calls at loci where the population did not contain all three possible SNP genotypes.

SNPs were analyzed using the Axiom Analysis Suite software, which analyzes the fluorescence levels in the plates and determines the allele calls for the samples involved. The data type provided by each marker is then determined and based on the alleles detected. The type of markers used in the downstream analysis were the PolyHighResolution marker types, as these marker types show at least two occurrences of the minor allele allowing for the determination of distortion events (AppliedBiosystems). A few markers were checked using the cluster plot view to determine if they had the correct distribution and if there was visible distortion in the plots (Figure 2).
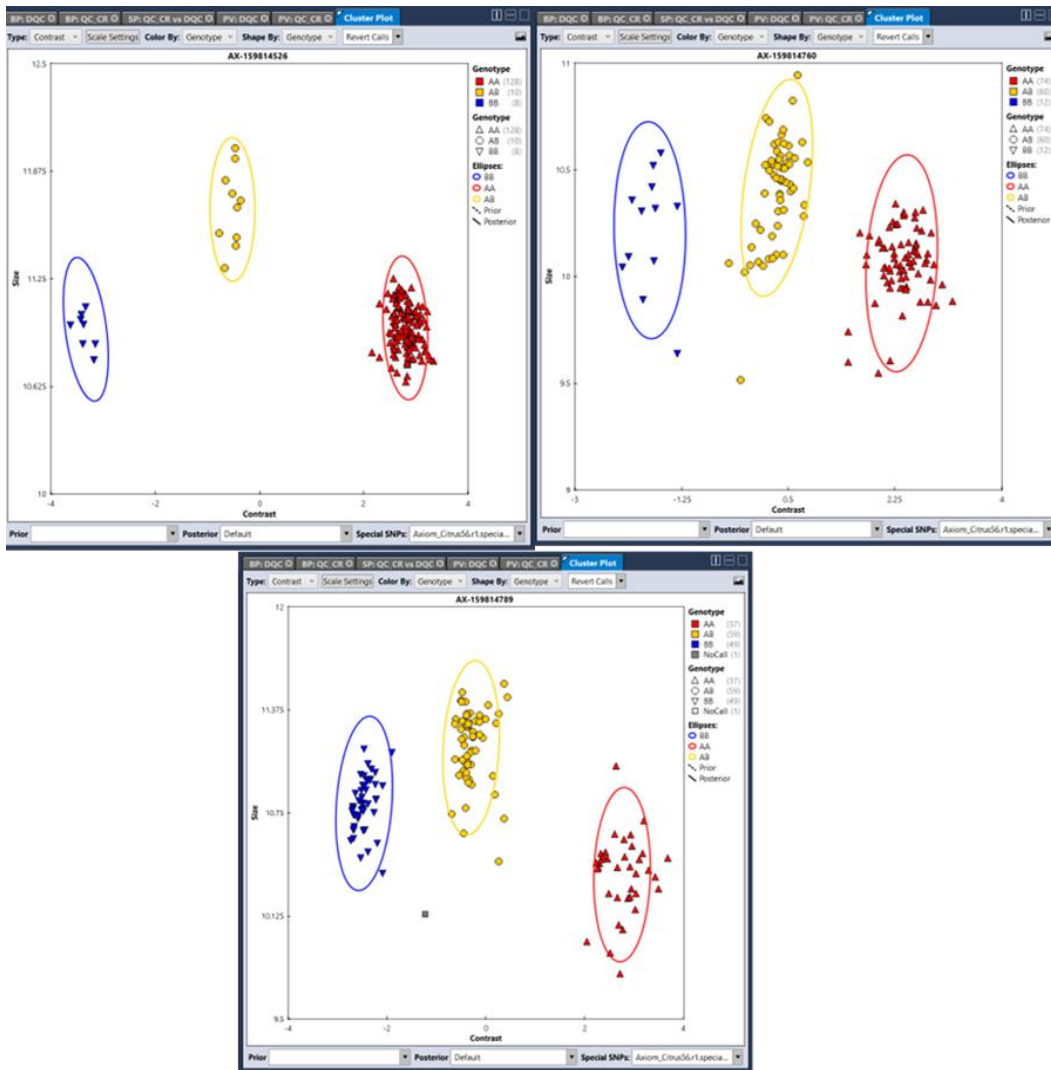
Figure 2: Cluster plots of three markers from the Axiom Suite analysis. Marker AX-159814526 (upper left) shows high segregation distortion with many more AA calls than AB or BB (1:2:1 expected ratio). Marker AX-159814760 upper right and AX-159814789 lower plots show markers with low or no distortion.

The SNP calls needed to be formatted before they could be used by the JoinMap software. This transformation process was conducted using Microsoft Excel. The SNP calls were copied from the Axiom output, and then the Chandler parent calls were added from a previous analysis of samples by Dr. Hiraoka. The parents of TPxFD were also added solely to compare with TPxFD. The homozygous markers in TPxFD and Chandler

were removed from the data set. The heterozygous markers were then distributed into three different sheets. Those markers are heterozygous in Chandler, heterozygous in TPxFD, and heterozygous in both. These sheets were then transformed as follows: those heterozygous only in Chandler were transformed to nn:np, those heterozygous only in TPxFD were transformed to lm:ll, and those heterozygous in both were transformed to hk:hk. In JoinMap, the markers were named according to the location on the Clementine 1.0 genome, for example, 1_176583 corresponds to a marker on chromosome 1 with an SNP at position 176583. The transformed data was then input into JoinMap.

Linkage mapping. The SNPs were mapped using JoinMap software (version 5.0, Van Ooijen 2018). The software was set to analyze a cross-pollinated progeny type to allow for the multiple segregation types from the parents of the cross and for the program to calculate linkage phases. The system sorts the data using the specified segregation type; in this case, we have lmxll, nnxnp, and hkxhk. The data was then split into three datasets to be mapped separately, the lm:ll loci used to create a map of TPxFD, nn:np loci mapping Chandler, and a combination of nn:np with hk:hk loci mapping. This was to separate the two parental genotypes to simplify the output, and the hk:hk were included in determining their usefulness in helping to align areas of the chromosomes that may otherwise have problems and to create a more complete and accurate map by using as many markers as possible. An attempt was made to combine hk:hk and lm:ll loci on a single map. However, this caused the 3rd and 8th chromosome markers to group together rather than separately. JoinMap then determined the phase of each locus, estimated the recombination frequencies between the markers, and output a map using the maximum

likelihood method. Maximum likelihood is a method by which the map distance of the markers can be determined by calculating the likelihood that any two markers are close to each other. JoinMap uses an Expectation-maximization algorithm in which the unobserved data/latent variable is predicted by applying an expectation equation followed by a maximization step. This process is repeated over many samples, allowing for the prediction of the unobserved data using the log-likelihood over all the possible values. The latent variable can then be applied to the original data to generate a more accurate map (kyazma.nl).

JoinMap also notes in a tab where there are possible problems with the map and assigns them a number indicating the level of stress created on the map by each marker. By sorting manually through each problematic marker and determining if the marker is in error, those markers can be removed from the mapping, and the map can be reconstructed from the remaining markers. Markers may have issues if the genotype call of an individual output by Axiom Suite is incorrect. These loci cause "stress" to the map, discrepancies between the calculated map location and the expected genotype of specific progeny given this map. The locus itself may also not segregate in a Mendelian fashion, such as having more than two alleles (rare for SNPs) that were not estimated in this study. JoinMap also calculates the segregation distortion shown by each marker and presents a chi-square value for the goodness of fit to the expected ratio and a significance test.

The map was then copied into an Excel sheet so that gaps could be calculated, markers could be compared against their annotated locations, markers that mapped to

different chromosomes from their annotation could be noted, and segregation distortion could be visualized. Each of the three maps were derived from the heterozygous markers of one of the parents, so the parental source of segregation distortion could be determined.

The map location, marker location, and distortion could then be displayed using Mapchart (Voorrips 2002). This program creates chromosome maps from the linkage analysis from JoinMap. The chromosome maps display a large amount of information from JoinMap in a format that can be analyzed and displayed more efficiently. These maps plot the markers along the chromosome, providing the distance as a visual rather than a number so visualization of gaps and other issues with the map can be seen.

The population used to study mandarin distortion was Fortune x Fairchild. These hybrids share a parent of Clementine according to the cross records but the male parent of Fairchild was recorded as Orlando and that of Fortune as Dancy mandarin (Citrusvariety.ucr.edu). However, genetic research by Barry et al. 2015 showed that Fortune's male parent was also likely Orlando rather than Dancy. The DNA was collected by Dr. Federici and others in the Roose lab, extracted, and sent to Affymetrix for analysis. Dr. Roose then organized the calls for the loci into an Excel spreadsheet and sent the spreadsheet to me for mapping. I used JoinMap to assemble the maps, including the hk:hk loci, to improve the map's overall structure and allow for detecting distortion related to the hk:hk loci.

Results and Discussion

(Tahitian pummelo x Flying Dragon) x Chandler maps. The almost 58,000 markers that are included on the chip after being sorted several ways and filtered for polymorphic markers resulted in 1747 nnxnp markers that are heterozygous in the male parent Chandler, 12508 lmxll markers that are heterozygous in the female parent Tahitian Pummelo x Flying Dragon, and 1695 hkxhk markers that are heterozygous in both parents. These loci were mapped using the software JoinMap. These markers spanned the nine chromosomes. Several markers were eliminated on each of the chromosomes due to stress effects. The markers were designed using only the *Citrus clementina* genome as a template for sequence flanking each targeted SNP because the genome for *Poncirus* had yet to be published when the array was designed. The probes on the array may match imperfectly with *Poncirus* or other taxa distant from *C. clementina*. This creates many instances where markers may fail or detect SNPs other than the intended target causing false calls. In addition, if the signal is weak from the chip analysis, then the call may also be incorrect. In some regions, many markers map to the same location on a chromosome. This is due to the lack of recombination between the markers indicating those markers are likely in close proximity on the genome or located in a region with few recombination events. This resulted in a lower number of unique loci across all nine chromosomes compared with the total number of heterozygous loci. Increased population size should result in a higher proportion of markers with unique map locations. The nn:np map includes 804 unique loci, the lm:ll map includes1989 unique loci, and the map of nn:np combined with hk:hk includes 1751 unique loci. Figures 3, 4, and 5 compare the

annotated genome locations in C. *clementina* to the mapped locations. Chromosome 3 seems to have an inverted and displaced region at about 32 Mb in all of the three maps for (TP x FD) x Chandler, and this is also true of the maps in Fortune x Fairchild, so it likely indicates that there is either an assembly issue in the original clementine genome or that the variety of clementine mapped has an inversion and an intrachromosomal translocation in this region. The gaps seen in chromosomes 2 and 9 in the np maps are regions in Chandler with low heterozygosity (Wu et al. 2014), likely the array used did not contain any markers in these regions due to the lack of heterozygosity.
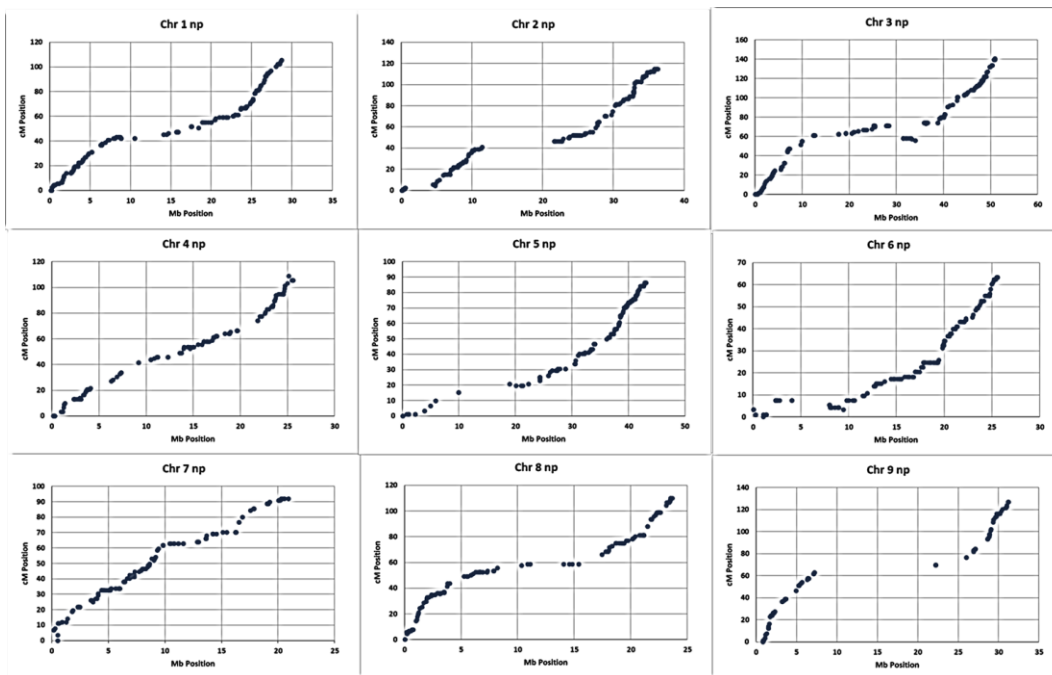
Figure 3: Comparison of mapped positions with annotated positions for nn:np loci in the cross (TP x FD) x Chandler.
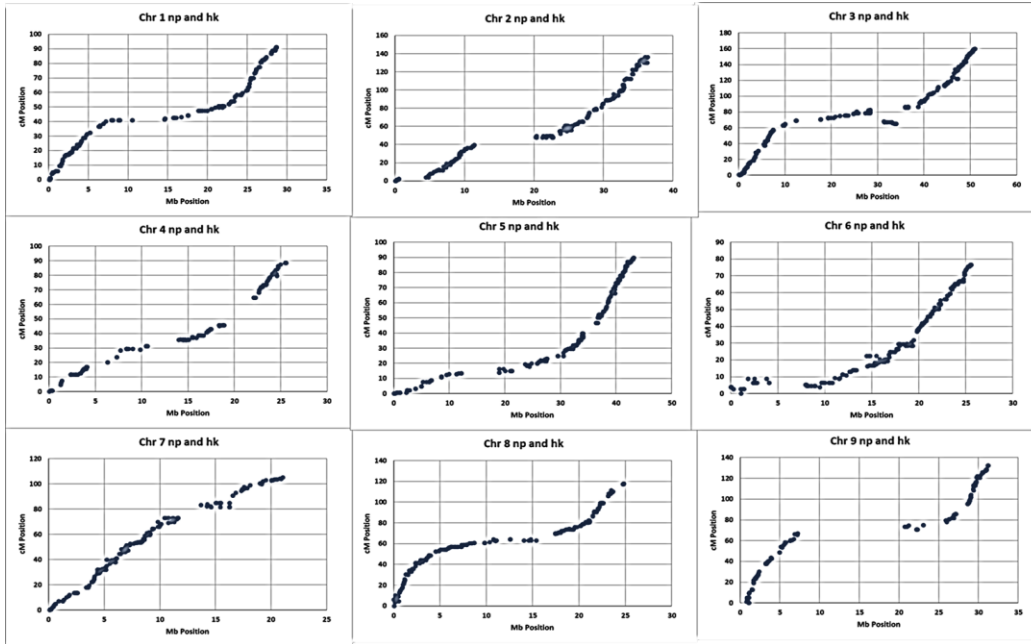
Figure 4: Comparison of mapped positions with annotated positions for nn:np and hk:hk loci combined in the cross (TP x FD) x Chandler.
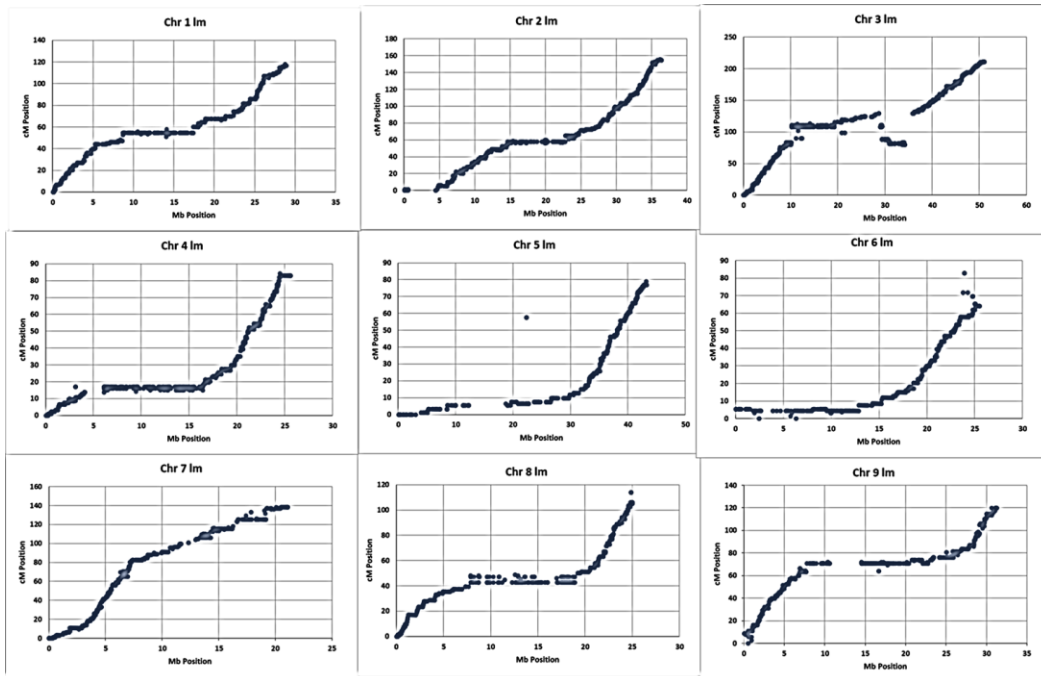
Figure 5: Comparison of mapped positions with annotated positions for lm:ll loci in the cross (TP x FD) x Chandler.

Main Chrom	Initial # Loci (all)	# unique loci	Total # Mapped	Elim. loci	Map Len (cM)	Start (1st marker)	End (last marker)	Gaps >1 Mb	P2 rec max	P2 rec > 4	# P2 NNstress in cM >2	Gt Prob Ind Means Max	# Gt Prob Ind > 0.01	# Gt Prob Loci >0.1	Segreg Distortion : # loci ** or more
1	210	105	208	2	105.5	1_176583	1_28706015	5	3	0	3	0.033	1	0	41
2	236	112	235	1	114.5	2_72223	2_36320367	2	9	1	4	0.147	2	0	0
3	231	106	231	0	140.5	3_91203	3_50965125	6	4	0	4	0.058	5	0	4
4	147	81	146	1	108.7	4_195253	4_25149789	6	4	0	3	0.054	12	0	3
5	161	81	159	2	86.3	5_125942	5_43075379	7	2	0	1	0.055	2	0	1
6	176	77	173	3	63.4	6_1097322	6_25576696	3	2	0	0	0.027	3	9	1
7	147	87	147	0	91.8	7_523761	7_20923918	3	9	1	5	0.218	3	0	0
8	177	93	174	3	109.7	8_50763	8_23676056	4	4	0	0	0.058	7	8	0
9	97	61	97	0	126.9	9_814402	9_31247698	4	4	0	1	0.125	13	0	0

Table 1: Summary of JoinMap output and statistics for nn:np loci in the cross (TP x FD) x Chandler. Note ** denotes a p-value of .05. Details on table located in supplementary material.

Chandler map. Table 1 summarizes the output from JoinMap for the nn:np loci. This map is focused on the heterozygous loci from the male parent Chandler. Few loci were removed for poor fit, and the chromosome lengths are reasonable for pummelo. There is good coverage of the chromosomes and reasonable numbers of loci present in each chromosome. The gaps are generally small in most of these maps and may be due to regions that are homozygous in Chandler; Wu et al. 2014 showed low levels of heterozygosity (around or below 1%) for Chandler in both chromosomes 2 and 9, particularly in those regions which were not able to be mapped. Since the mined SNPs on the Axiom Array chip were selected based on sequence variation in a range of 41 citrus accessions including Chandler, it is possible those regions have no heterozygous SNPs on the array in the gap regions simply by chance sampling. The majority of the segregation distortion was detected on chromosome 1, with 41 unique markers showing significant distortion; this is 39% of the total unique markers for chromosome 1.

Segregation distortion in Chandler Pummelo. The most significant distortion for the Chandler map was a peak around 23 cM, with a few markers showing a distortion ratio of about 61:33 nn:np (Figure 6). The significant segregation distortion spanned approximately 40 cM though some non-significant distortion can be seen even as far out as 60 cM. It is likely that a gene which distorted the rest of the chromosome is located around 23 cM. The distortion decreases the more genetically distant any marker is from the gene causing the distortion. Chromosomes 3, 4, 5, and 6 had a few non-significantly distorted markers. It is possible that those are distorted due to calling errors or are very minorly distorted and have little effect on nearby loci.

Figure 6: Distortion of chromosome 1 for nn:np loci in the cross (TP x FD) x Chandler with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84.

Main Chrom	Initial # Loci (all)	# unique loci	Total # Mapped	Elim. loci	Map Len (cM)	Start (1st marker)	End (last marker)	Gaps >1 Mb	P2 rec max	P2 rec > 4	# P2 NNstress in cM >2	Gt Prob Ind Means Max	# Gt Prob Ind > 0.01	# Gt Prob Loci >0.1	Segreg Distortion : # loci ** or more
1	404	173	299	105	91.0	1_100968	1_28706015	5	3	0	0	0.023	13	1	61
2	441	204	358	83	136.1	2_72223	2_36341036	3	19.6	3	20	0.125	20	2	3
3	527	272	444	83	159.7	3_243250	3_50964047	6	8	9	33	0.027	1	2	4
4	291	111	168	123	88.5	4_57951	4_25576395	6	5	1	0	0.036	29	2	4
5	478	190	413	65	89.6	5_119526	5_43189791	5	5	1	16	0.044	9	0	4
6	350	160	288	62	76.5	6_1097322	6_25574837	6	10.4	1	17	0.13	10	5	0
7	318	173	268	50	105.1	7_114911	7_21038711	5	19.5	1	10	0.229	11	0	9
8	343	197	317	26	117.7	8_50763	8_24818296	6	10	1	25	0.143	14	0	0
9	199	131	178	21	132.2	9_1073527	9_31247698	5	18.1	2	24	0.418	19	1	0

Table 2: Summary of JoinMap output and statistics for nn:np and hk:hk loci in the cross (TP x FD) x Chandler. Note ** denotes a p-value of .05. Details on table located in supplementary material.

Chandler hk and np map. Table 2 summarizes the output from JoinMap for the

nn:np and hk:hk loci. This Chandler map was created to improve the previous map by

increasing the number of markers, allowing for better recombination calculations. This

map with more total loci is likely more accurate concerning chromosome size and fills in

some previous gaps. However, the significant gaps in chromosomes 2 and 9 reinforce the

idea that the sections are likely homozygous in Chandler. More loci were removed for

poor fit, mostly the hk:hk loci. These loci are primarily errors in calls from the Axiom

software resulting in heterozygous calls where none exists or only the female parent is

heterozygous.

Segregation distortion with hk loci included. The map distances and size changed

for nearly every chromosome when the hk:hk loci were included as compared to the

nn:np loci maps alone, and this means that the distorted region seen in the nn:np map

(Figure 6) is closer to 30 cM in the combined map (Figure 7), and the percentage of

unique loci significantly distorted is 35.1%. However, segregation distortion detected at

hk loci could occur in the TPxFD parent, so the values are not comparable. There also

seems to be a separate event around 50 cM in which the distortion is mainly seen in the

hk:hk loci (figure 10). It is possible these regions do overlap and that the loci are slightly

misplaced due to an issue with the mapping of those loci as the loci that are heterozygous

in both parents can often cause issues with mapping as we don't know which parent is

contributing which alleles. This type of mapping error likely contributed to the spiked

pattern in the hk:hk distortion map. There is likely a low distortion region and a relatively

high distortion region close to each other. With the poor alignment from the hk:hk, those

regions are mixed rather than distinct.
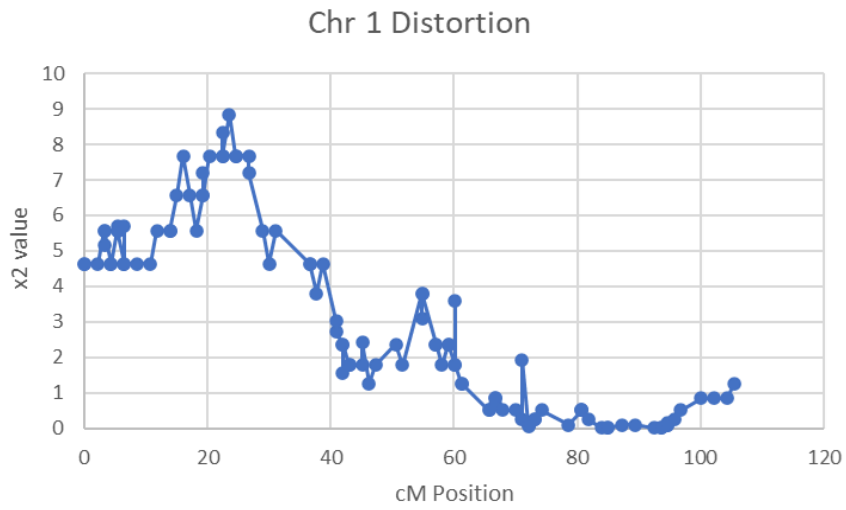

Figure 7: Distortion of chromosome 1 for nn:np and hk:hk loci in the cross (TP x FD) x
Chandler with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84
for the np markers and 6 in the hk markers top left map plots the hk and np markers
together, top right is just the markers from np, and bottom is just the markers from hk.

Main Chrom	Initial # Loci (all)	# unique loci	Total # Mapped	Elim. loci	Map Len (cM)	Start (1st marker)	End (last marker)	Gaps >1 Mb	P2 rec max	P2 rec > 4 in cM >2	# P2 NNstress in cM >2	Gt Prob Ind Means Max	# Gt Prob Ind > 0.01	# Gt Prob Loci >0.1	Segreg Distortion : # loci ** or more
1	1341	192	1326	15	117.7	1_67231	1_28735824	0	5	2	15	0.051	15	0	2
2	1526	235	1469	57	155.8	2_73445	2_36199144	1	6	3	2	0.041	2	0	8
3	2337	353	2276	61	211.1	3_55100	3_51018918	4	5	4	0	0.114	1	0	0
4	1419	201	1396	23	84.1	4_72678	4_24522146	1	2	1	0	0.02	1	0	0
5	1348	138	1335	13	79.0	5_79513	5_43196677	2	3	0	0	0.028	4	0	18
6	1169	131	1125	44	82.9	6_2451391	6_23915796	2	2	1	1	0.081	2	1	1
7	1152	217	1125	27	138.0	7_27358	7_21083554	2	5	2	11	0.122	9	0	36
8	1031	170	1005	26	114.0	8_29665	8_24864336	3	3	1	0	0.092	3	1	6
9	1150	197	1118	32	119.8	9_520569	9_31250717	2	4	1	2	0.066	5	0	0

Table 3: Summary of JoinMap output and statistics for lm:ll loci in the cross (TP x FD) x Chandler. Note ** denotes a p-value of .05. Details on table located in supplementary material.

Tahitian pummelo x Flying Dragon map. Table 3 summarizes the output from

JoinMap for the lm:ll loci. This map is focused on the heterozygous loci from the female

parent (TP x FD). Many loci were mapped for each chromosome, and the loci eliminated

for poor fit were relatively low. The chromosome lengths are reasonable. The gaps are

generally small, with larger ones on chromosomes 5 and 9. Chromosome 5 had the gap

region mapping to chromosome 6, which may indicate that the region isn't a gap and

belongs on the other chromosome. The same is true for chromosome 9 with the loci

mapping to chromosome 8.

Segregation Distortion in Tahitian pummelo x Flying Dragon map. Segregation

distortion in TPxFD was observed on three chromosomes, 2, 5, and 7, with the distortion

being much less severe than seen with the nn:np maps (Table 3). Chromosome 2 has

3.4% of the unique loci significantly distorted. However, a more significant number

(16.7%) are minorly distorted with a p-value of .1, chromosome 5 has 13% of the unique

loci significantly distorted with a minor distortion of 31.9%, and chromosome 7 has

16.6% of the unique loci significantly distorted with a minor distortion of 32.3%.

The distortion ratios on these chromosomes are close to about 37:55, each with a

distortion pattern. Chromosome 2 shows distortion at 0 cM, which is primarily

insignificant, with just one locus at 1 cM above the significance threshold, indicating the

region is likely, not distorted. Instead, the error is due to a small sample size and minor

ratio changes causing significant changes in chi-square values. The second distorted

region starts at 80 cM and continues until 140 cM, peaking at 115 cM. Chromosome 5

shows distortion at 0 cM, which continues until about 60 cM. There are several peaks and

3 cM, 10 cM, 23 cM, and 35 cM. It is unclear if each of these peaks marks a locus that alters segregation, as likely in the peaks seen in chromosome 1 of the nn:np map, or if this is a region with one major gene causing the distortion to occur. Other possible explanations are that the map shows multiple peaks due to incorrect locus positions on the map, or that we were not able to map the region where the major distortion event occurs though this is unlikely given the high numbers of loci unless the region is absent in the Clementine genome that the markers were derived from. Chromosome 7 shows minor distortion (rarely above the significance threshold of 4.5) throughout the entire chromosome, with peaks on both ends (Figure 8).
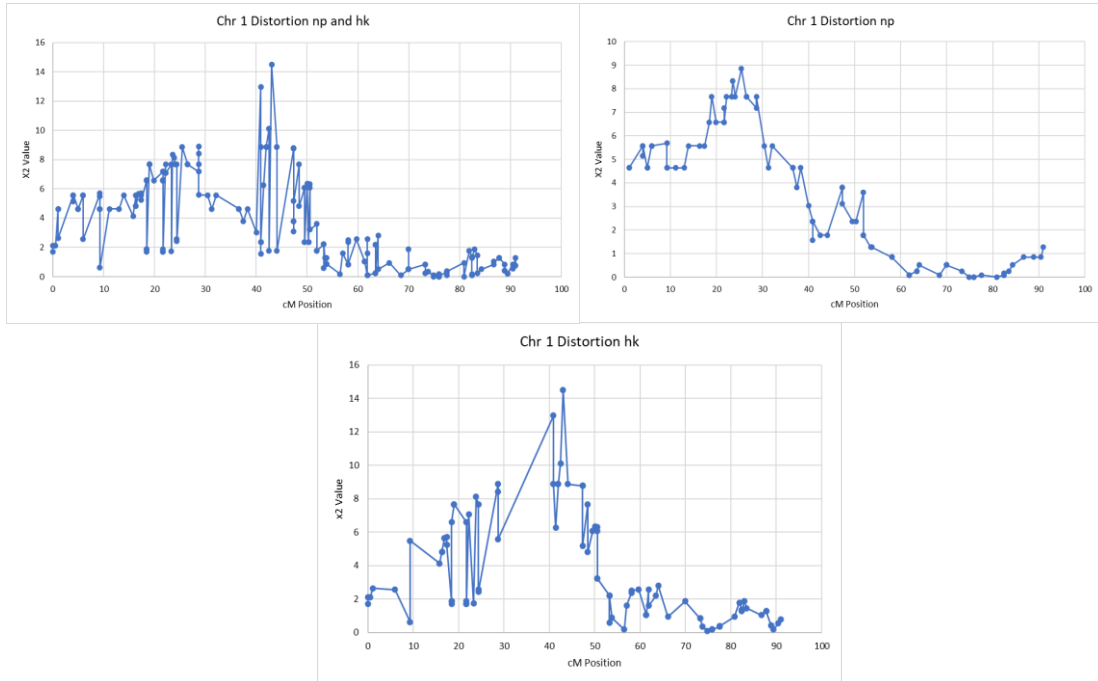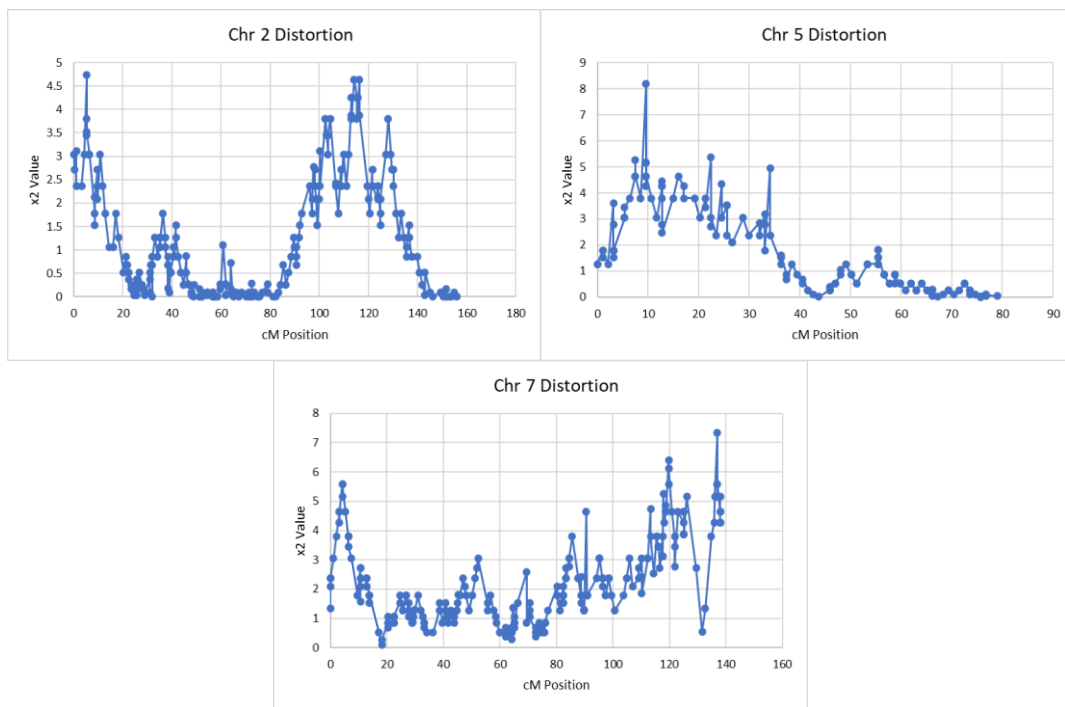

Figure 8: Distortion of chromosome 2, 5, and 7 for lm:ll loci in the cross (TP x FD) x Chandler with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84.

The maps also showed several locations where markers mapped to a different chromosome from what was initially annotated when the marker was designed. It is unclear if this is an artifact of the original assembly of the clementine genome (Wu et al. 2014), which used a double haploid as the primary source of genetic material and was assembled from Sanger sequence reads, which may have resulted in errors when the scaffolds were assembled. It is also possible that these sections mark differences between the clementine genome and the genomes of either pummelo or trifoliate orange.

Figures 9, 10, and 11 compare the locations annotated for each locus and the chromosomes they mapped to in the three maps made using JoinMap. The most consistent patterns show that a region of chromosome 2 maps to chromosome 4, regions of both chromosomes 5 and 4 map to chromosome 7, and regions of chromosomes 3 and 9 map to chromosome 8. The percentage of markers that mapped to chromosomes other than that expected from Clementine genome sequence 1.0 was 2.8%, 2.7%, and 2.3%. The most extensive set of markers was on the lm:ll map and was annotated on chromosome 5 but mapped to chromosome 7. This is consistent with other work comparing trifoliate maps to the clementine genome by Huang et al. 2018.

Summary of Loci by Group for (Tahitian Pumello x Flying Dragon) x Chandler with 95 progeny									
Group	1=Chr3	2=Chr1	3=Chr5	4=Chr4	5=Chr7	6=Chr2	7=Chr6	8=Chr8	9=Chr9
Chr 1		299							
Chr 2	1				19	385			
Chr 3	443							7	
Chr 4				144	3		1		
Chr 5			413		16				
Chr 6					5		282		
Chr 7					249				
Chr 8							5	294	
Chr 9								16	178
total	444	299	413	168	268	385	288	317	178

Figure 9: Chart showing the loci which mapped to different chromosomes when compared to annotated regions in nn:np loci for cross (TP x FD) x Chandler.

Summary of Loci by Group for (Tahitian Pumello x Flying Dragon) x Chandler with 95 progeny									
Group	1=Chr6	2=Chr7	3=Chr4	4=Chr9	5=Chr2	6=Chr3	7=Chr1	8=Chr8	9=Chr5
Chr 1							175		
Chr 2		15			235	2			
Chr 3						228		4	
Chr 4	1	3	131						
Chr 5		5							159
Chr 6	172								
Chr 7		139							
Chr 8								159	
Chr 9				97				11	
total	173	147	146	97	235	230	175	174	159

Figure 10: Chart showing the loci which mapped to different chromosomes when compared to annotated regions in nn:np and hk:hk loci combined for cross (TP x FD) x Chandler.

Summary of Loci by Group for (Tahitian Pumello x Flying Dragon) x Chandler with 95 progeny									
Group	1=Chr3	2=Chr9	3=Chr2	4=Chr5	5=Chr4	6=Chr7	7=Chr1	8=Chr6	9=Chr8
Chr 1			1				1325		1
Chr 2		2	1466	1	58				
Chr 3	2272	11			1				42
Chr 4	2				1335	19			
Chr 5			1	1332		123			
Chr 6	2	1			2			1114	
Chr 7		1				982	1		
Chr 8				2				10	873
Chr 9		1103	1			1		1	89
total	2276	1118	1469	1335	1396	1125	1326	1125	1005

Figure 11: Chart showing the loci which mapped to different chromosomes when compared to annotated regions in lm:ll loci for cross (TP x FD) x Chandler.

Fortune x Fairchild linkage maps. The Fortune x Fairchild population analysis

resulted in two maps, the first containing loci for nn:np and hk:hk, the second for lm:ll

and the same set of hk:hk. These maps were constructed using the same procedure as for

the (TP x FD) x Chandler maps. Figures 12 and 13 compare the annotated positions to the

mapped positions. Chromosome 8 mapped poorly on both maps due to many of the loci

erroneously mapping to chromosome 3; this likely indicates a region of similarity

between chromosome 8 and 3 as the same issue occurred in the (TP x FD) x Chandler

population. Chromosome 8 has also been shown to have very low heterozygosity in both

Fortune and Fairchild (Roose lab unpublished) leading to the erroneous mapping and

gaps. To help correct for this mapping error the loci which grouped with chromosome 3

were manually added to the chromosome 8 map. A 50cM gap was included to estimate

the distance between the 2 separately mapped regions. The np map also had very few loci

for chromosome 7, which explains the poor mapping, as most loci are hk:hk loci.

Chromosome 3 still shows the misplaced segment in all maps of the 3rd chromosome. The

np+hk map shows chromosome 9 as having many misplaced loci; these loci are likely

differences between the two cultivars and may indicate where a translocation may be

present or may be an error in the mapping of this region. It is also possible that since the

hk loci segregate in both parents, one parent has more recombination resulting in the

displacement of the loci relative to np.
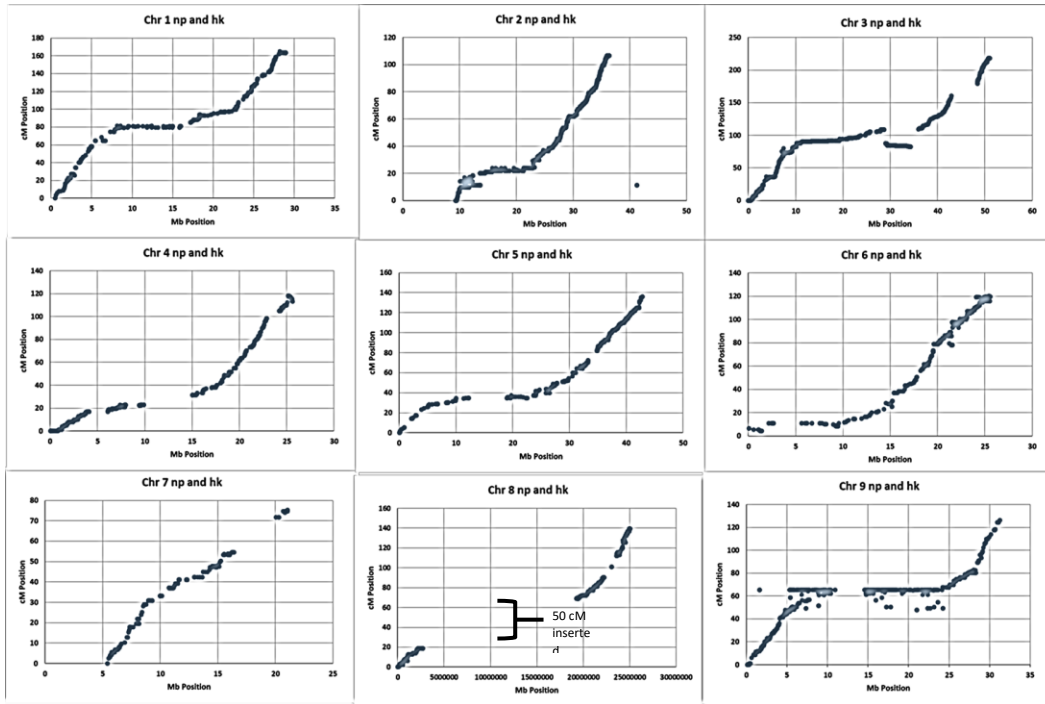
Figure 12: Comparison of mapped positions with annotated positions for the nn:np and hk:hk loci combined for the cross Fortune x Fairchild. Note reduced scale for
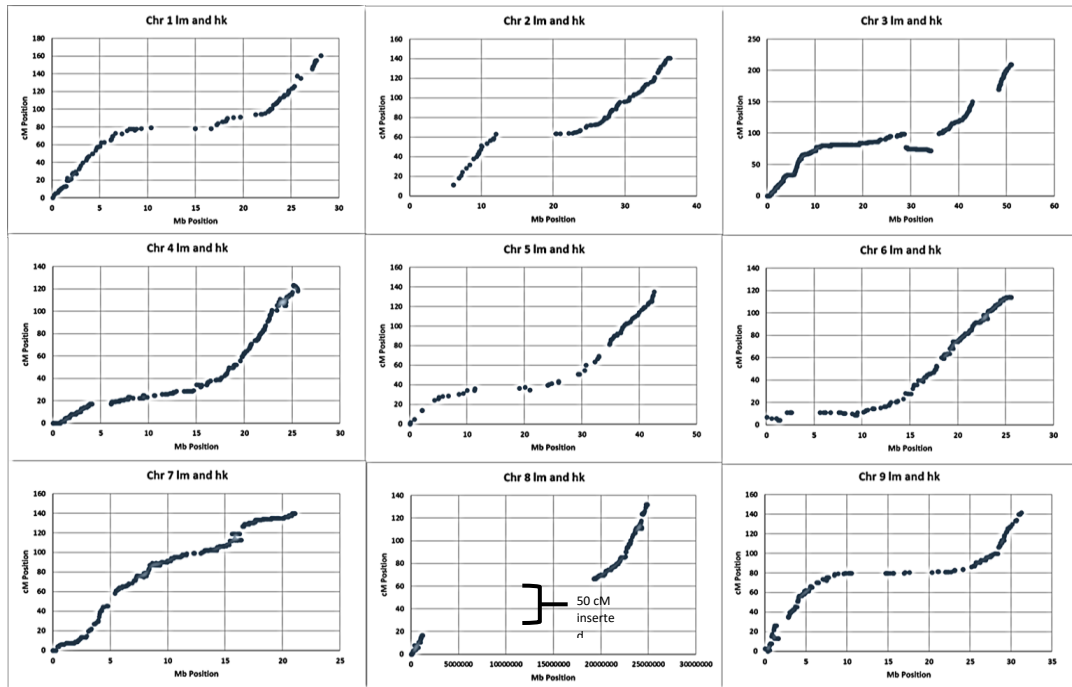
Figure 13: Comparison of mapped positions with annotated positions for the lm:ll and hk:hk loci combined for the cross Fortune x Fairchild.

Table 4 summarizes the nn:np and hk:hk maps. This map shows significant distortion in chromosomes 5, 6, and 7, with other chromosomes having low distortion levels in a few loci. Chromosome 5 has a distortion in 16.8% of the unique loci, chromosome 6 has a distortion of 12.9%, and chromosome 7 has a distortion in 21.6%.

Main Chrom	Initial # Loci (all)	# unique loci	Total # Mapped	Elim. loci	Map Len (cM)	Start (1st marker)	End (last marker)	Gaps >1 Mb	P2 rec max	P2 rec > 4	# P2 NNstress in cM >2	Gt Prob Ind Means Max	# Gt Prob Ind > 0.01	# Gt Prob Loci >0.1	Segreg Distortion : # loci ** or more
1	473	175	447	26	164.7	1_528900	1_28223198	3	5.2	3	17	0.046	19	1	16
2	2248	370	2238	10	106.4	2_9337069	2_36364386	1	5	1	11	0.014	2	4	14
3	1956	389	1954	2	218.0	3_13771	3_51024788	4	5	2	13	0.013	1	6	16
4	1037	246	1031	6	117.9	4_57652	4_25192228	4	4	0	31	0.026	1	0	3
5	1466	279	1452	14	136.1	5_72022	5_42848457	8	5	1	17	0.017	1	2	36
6	1155	215	1155	0	120.3	6_1226345	6_25409477	5	5	1	18	0.031	9	3	36
7	245	86	230	15	75.3	7_5410613	7_21066040	3	2	0	0	0.031	31	3	53
8	933	153	932	1	139.0	8_19873	8_2707267	1	3	0	1	0.006	0	0	17
9	1138	191	1089	49	126.1	9_138077	9_31268589	3	5	1	4	0.046	34	5	2

Table 4: Summary of JoinMap output and statistics for nn:np and hk:hk loci in the cross Fortune x Fairchild. Note ** denotes a p-value of .05. Details on table located in supplementary material.

Distortion rates vary based on the chromosome. The highest distortion rate on Chromosome 5 is about 30:42:12, and all loci in this region that show distortion are hk:hk loci (Figure 14), so the maps for this region are the same for both of the maps made (Figures 12 and 13). The significant peak is around 100 cM though there are also non-significant peaks around 5 cM and 30 cM. The distorted regions appear to be from 70 cM to 130 cM.
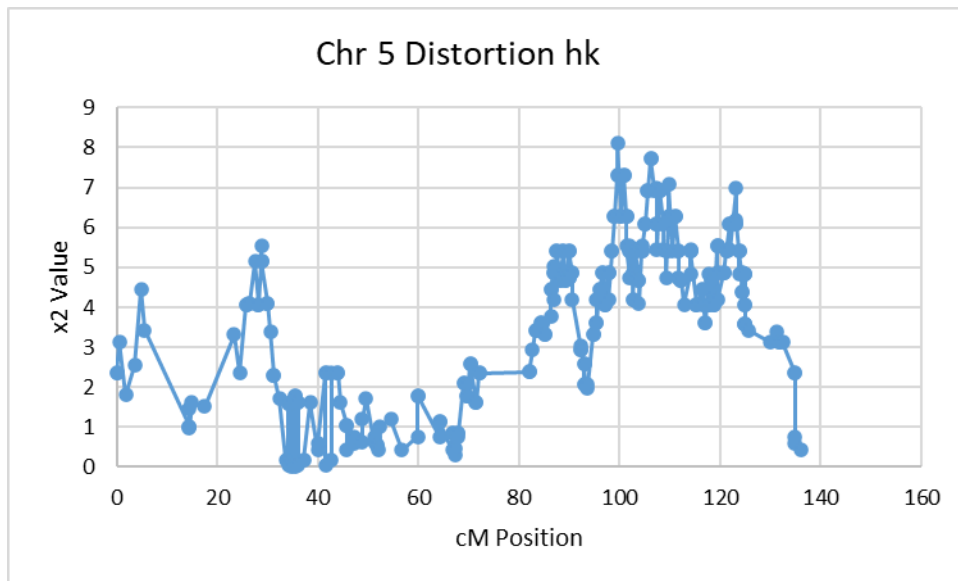
Figure 14: Distortion of chromosome 5 hk:hk loci in the cross Fortune x Fairchild with the chi square value calculated by JoinMap. A p-value of .05 is about 6.

The highest distortion ratio on Chromosome 6 is 26:57 for the nn:np loci and 39:28:17 for the hk:hk loci; this region shows significantly more distortion than any of the previous maps. The distortion patterns also seem to vary in this chromosome when comparing the two types of loci with the nn:np having two peaks at about 79 cM and 95 cM with the region of distortion affecting about 65 cM to the end of the chromosome at

120 cM. The hk:hk loci have peaks around 70 cM and 85 cM - 95 cM; the span is similar

to the nn:np loci (Figure 15). It is unclear if the two peaks are separate or just a single

distortion event with a few progenies differing, creating multiple peaks.

Figure 15: Distortion of chromosome 6 for nn:np and hk:hk loci in the cross Fortune x Fairchild with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84 for the np markers and 6 in the hk markers. Top left map plots the hk and np markers together, top right is just the markers from np, and bottom is just the markers from hk.

The highest distortion rate on Chromosome 7 is 21:63 for the nn:np loci and

38:35:10 for the hk:hk loci. The distortion positions for this chromosome are much better

aligned and have a peak around the 20 cM position. The range of the distorted region is

from the beginning on the chromosome 0 cM to about 45 cM. Including the hk:hk loci

allow for the complete chromosome analysis; the lower number of markers in this region

may have affected the analysis (Figure 16). There are no np markers in the first 20 cM of

chromosome 7 because both parents are identically heterozygous in this region, likely

due to their identical parentage.

Figure 16: Distortion of chromosome 7 for nn:np and hk:hk loci in the cross Fortune x Fairchild with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84 for the np markers and 6 in the hk markers. Top left map plots the hk and np markers together, top right is just the markers from np, and bottom is just the markers from hk.

Table 5 summarizes the lm:ll and hk:hk loci. This map shows significant

distortion in chromosomes 5, 6, and 7, with other chromosomes having low distortion

levels in a few loci. Chromosome 5 has a distortion in 13.8% of the unique loci,

chromosome 6 has a distortion of 60.2%, and chromosome 7 has a distortion in 40.1%.

Main Chrom	Initial # Loci (all)	# unique loci	Total # Mapped	Elim. loci	Map Len (cM)	Start Mb (1st marker)	End Mb (last marker)	Gaps >1 Mb	P2 rec max	P2 rec > 4	# P2 NNstress in cM >2	Gt Prob Ind Means Max	# Gt Prob Ind > 0.01	# Gt Prob Loci >0.1	Segreg Distortion : # loci ** or more
1	803	237	799	4	160.62	1_110163	1_28114866	2	5	2	19	0.046	21	6	11
2	1943	362	1937	6	140.572	2_360086	2_36364386	3	5.5	3	11	0.014	3	1	10
3	2228	409	2226	2	209.243	3_13771	3_51024788	6	4	4	5	0.015	8	9	16
4	1218	250	1189	29	122.747	4_57652	4_25192228	3	5.5	4	16	0.031	2	4	3
5	1384	254	1359	25	134.714	5_72022	5_42631559	10	5	2	21	0.039	2	0	36
6	1144	221	1119	25	114.326	6_1226345	6_25409477	3	4.5	1	20	0.031	14	5	133
7	1275	187	1271	4	139.651	7_24721	7_21086362	0	4.7	3	6	0.033	34	7	75
8	674	158	659	15	131.645	8_19873	8_24891150	0	3	0	8	0.017	42	3	3
9	536	168	504	32	141.566	9_352587	9_31268589	7	5.8	6	20	0.037	22	3	13

Table 5: Summary of JoinMap output and statistics for lm:ll and hk:hk loci in the cross Fortune x Fairchild. Note ** denotes a p-value of .05. Details on table located in supplementary material.

Chromosome 5 only contains the same distortion in the hk:hk loci as previously reported in the np+hk map; since they are identical, the data will not be repeated here. Chromosome 6 is also identical for the hk:hk loci; however, the lm:ll loci's highest distortion rate is 54:29. The peak is about 65 cM. The range for the distortion is 45 cM to about 95 cM (Figure 17).
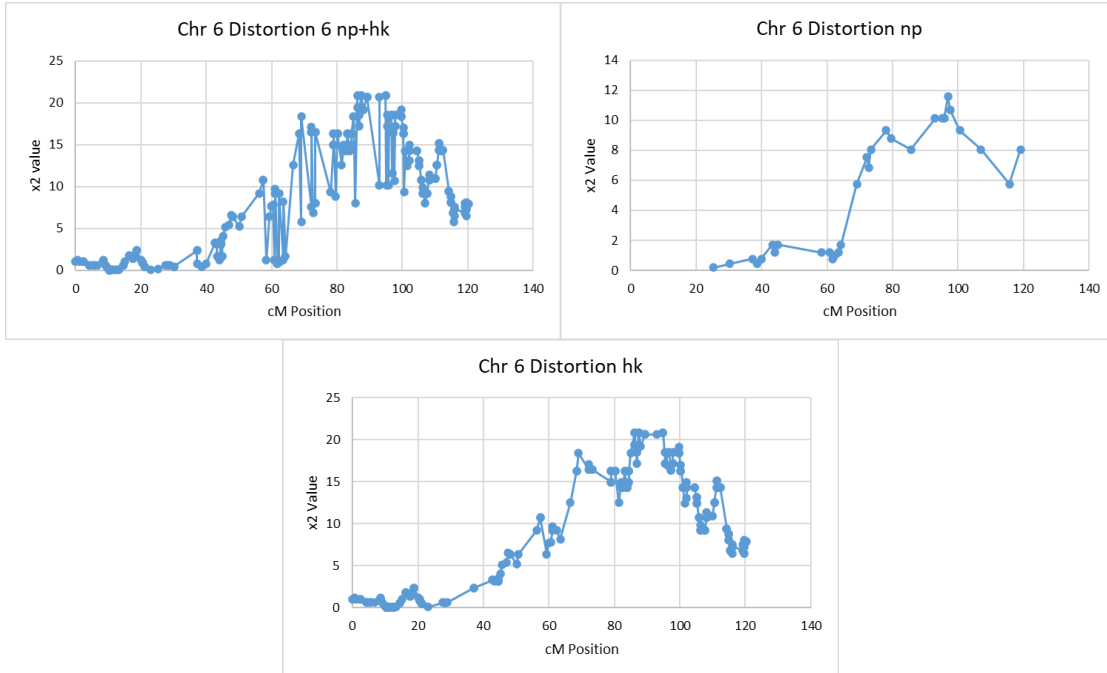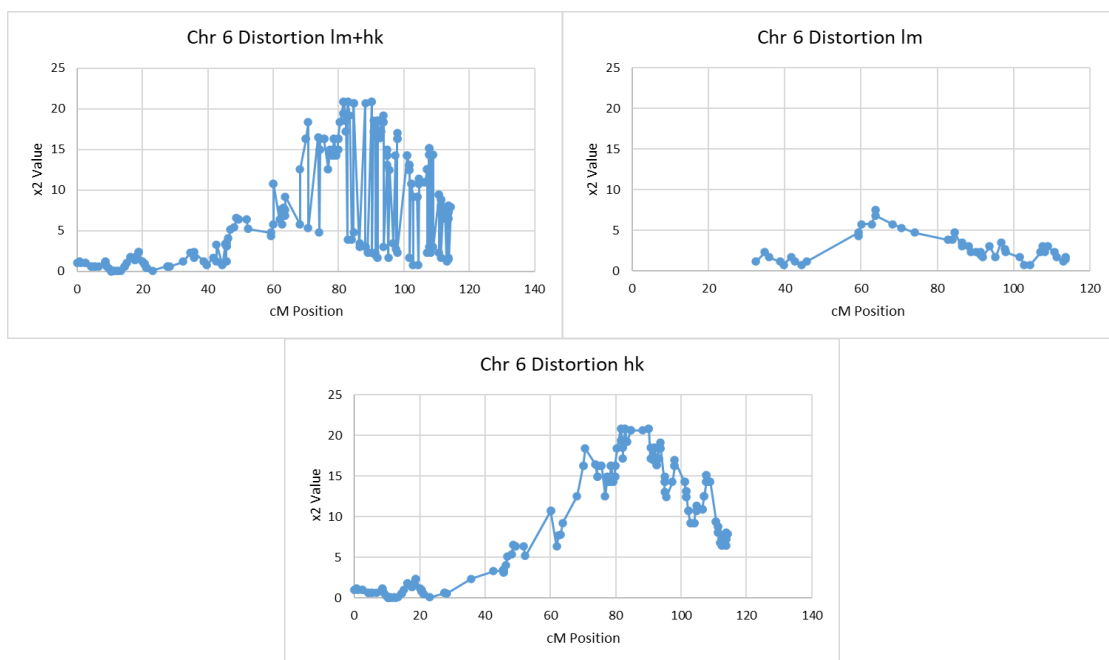

Figure 17: Distortion of chromosome 6 for lm:ll and hk:hk loci in the cross Fortune x Fairchild with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84 for the np markers and 6 in the hk markers. Top left map plots the hk and lm markers together, top right is just the markers from lm, and bottom is just the markers from hk.

The highest distortion rate on Chromosome 7 is 54:30 for the lm:ll loci and 38:36:10 for the hk:hk loci. The lm:ll loci have three peaks at 5 cM, 90 cM, and 130 cM. The hk:hk has peaks at 60 cM and 75 cM. The lm:ll range for distortion is 0 cM to 30 cM and from 45 cM to the end of the chromosome at 140 cM. This chromosome differs

significantly from this region's np+hk map, possibly indicating different sources of segregation or greater overall segregation than a single map (Figures 16 and 18).
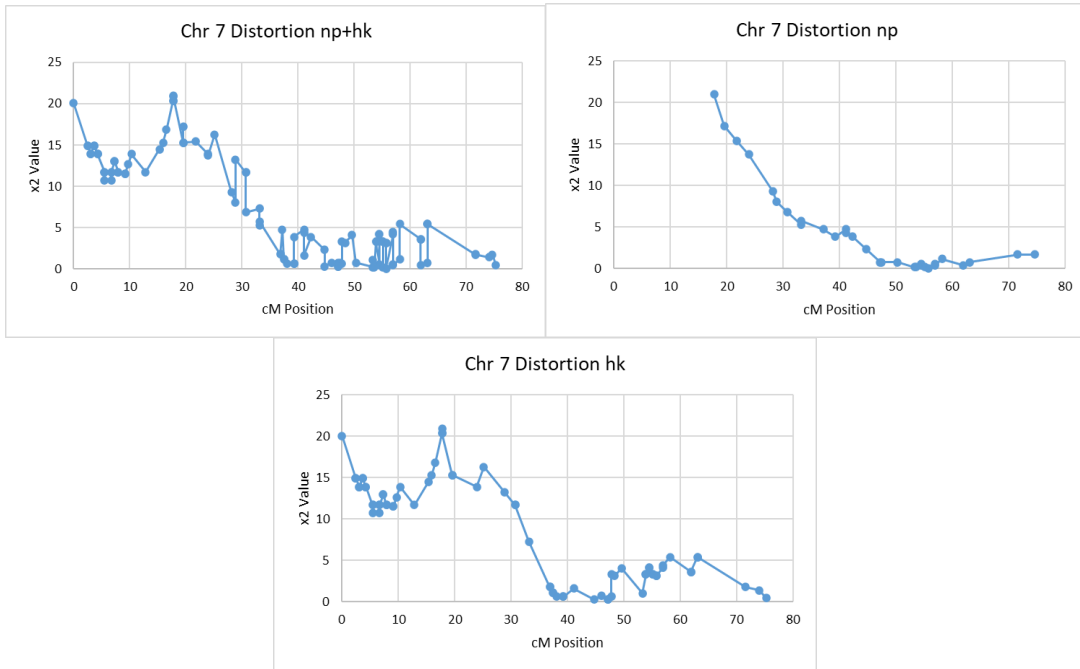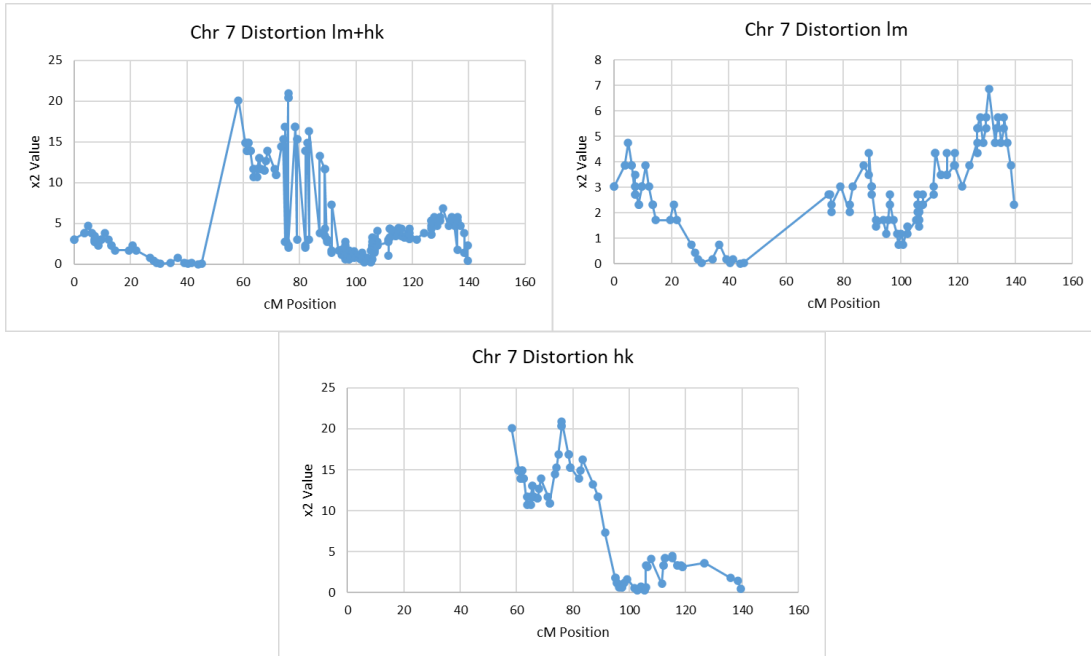

Figure 18: Distortion of chromosome 7 for lm:ll and hk:hk loci in the cross Fortune x Fairchild with the chi square value calculated by JoinMap. A p-value of .05 is about 3.84 for the np markers and 6 in the hk markers. Top left map plots the hk and lm markers together, top right is just the markers from lm, and bottom is just the markers from hk.

Figures 19 and 20 compare the locations annotated for each locus and the chromosomes they mapped to in the two maps made using JoinMap. The most consistent patterns show that a region of chromosome 8 maps to chromosome 6, and regions of chromosomes 5 and 4 maps to chromosome 7. The regions of chromosome 7 are consistent with the previous three maps. The percentage of markers that mapped to chromosomes other than that expected from Clementine genome sequence 1.0 was .5% and 2.4%.

| Summery of np and hk markers for 84 progeny of Fortune x Fairchild | | | | | | | | | |
Group	1=Chr3	2=Chr2	3=Chr6	4=Chr8	5=Chr5	6=Chr9	7=Chr4	8=Chr1	9=Chr7
Chr 1								447	
Chr 2		2237							
Chr 3	1954	1							
Chr 4							1031		18
Chr 5					1452				21
Chr 6			1150						
Chr 7									191
Chr 8			5	397					
Chr 9						1089			
total	1954	2238	1155	397	1452	1089	1031	447	230

Figure 19: Chart showing the loci which mapped to different chromosomes when compared to the annotated regions in nn:np and hk:hk loci combined for the cross Fortune x Fairchild.

| Summery of lm and hk markers for 84 progeny of Fortune x Fairchild | | | | | | | | | |
Group	1=Chr3	2=Chr8	3=Chr2	4=Chr5	5=Chr1	6=Chr9	7=Chr7	8=Chr4	9=Chr6
Chr 1					799				
Chr 2			1937				1	6	
Chr 3	2226					2			
Chr 4							37	1183	
Chr 5				1359			215		
Chr 6									1114
Chr 7							1018		
Chr 8		634							5
Chr 9						502			
total	2226	634	1937	1359	799	504	1271	1189	1119

Figure 20: Chart showing the loci which mapped to different chromosomes when compared to the annotated regions in lm:ll and hk:hk loci combined for the cross Fortune x Fairchild.

The (TPxFD) x Chandler maps show a remarkably small segregation distortion for a cross between two genera. Different cross directions would allow for better analysis of the overall distortion in these two genera as Chandler x (TPxFD) may have had different distortion due to male gametic failure. Chromosome 1 of Chandler shows the most significant segregation distortion and is the only chromosome distorted in the nn:np maps. The low level of segregation distortion in TPxFD indicates that using a *Citrus x Poncirus* hybrid as the female parent does not always lead to much segregation distortion.

This is consistent with other studies, which find more segregation distortion in the male parent probably due to pollen failure, but this was not tested in TPxFD. The distortion on Chromosome 1 of Chandler is consistent with an S locus proposed for citrus. However, in a previous study by Ollitrault et al. 2021, they discovered the gene on chromosome 7 for mandarin-type citrus. Perhaps pummelo has an S locus located on chromosome 1. The three less distorted chromosomes in the lm:ll are not likely to indicate a complete failure of a zygote or gamete formation, as that would likely have produced a much larger distortion. These regions may be more associated with growth and vigor; if a severe negative vigor characteristic was present, the plant would likely have died before being collected or would never have germinated.

The distorted loci were checked against the original parents, and in most cases, the distortion favors homozygosity, with a few loci breaking this pattern. The alleles also tended to prefer the allele from the Chandler parent when it could be determined which parent contributed the allele, but two examples also showed a preference for the Flying Dragon allele. Hence, the distortions are not specific to any one conformation or ancestry.

The populations in this study were collected differently, which may have led to distortions outside of the significant reasons above. The (TP x FD) x Chandler population was collected from seeds that germinated exceptionally well and were sampled for DNA isolation not long after initial planting. This allowed the population to include even weaker genotypes that may not have survived outside the greenhouse. By contrast, the Fortune x Fairchild population was collected as young field trees, increasing the chance that trees died before being grafted and moved to the field. Fortune x Fairchild population

is partially inbred as the parents are full siblings. This likely would have contributed to the inbreeding depression in the greenhouse and reduced the number of plants that made it to the field or survived once in the field. This may have resulted in distortions from Fortune x Fairchild related to loci related to survival and vigor. The (TP x FD) x Chandler will have fewer issues related to the collection due to the young age of the plants sampled. However, some seeds didn't germinate, and some seedlings never grew beyond a couple of centimeters, making them impossible to sample with the others as they quickly died.

The Fortune x Fairchild population's results show distortion on the identical chromosomes in the lm+hk and np+hk maps indicating those regions are affected regardless of the parent. However, this may also be due to the close relationship between these two cultivars, as Fortune and Fairchild share parents. The distortion on the beginning of chromosome 7 is likely the same S-locus identified by Ollitrault et al. 2021; however, the second region in the lm+hk map is not likely to be explained by the same locus. It could be a second gene with a similar mechanism or something unrelated to self-incompatibility.

Studying segregation distortion in pollen from the intended parent should allow for the prediction of distortion in the offspring, as the male parent is likely the main contributor to segregation distortion. This is also supported by there being little to no distortion in the lm:ll markers. Thus, the segregation likely stems from the male gamete rather than the female gamete or zygotic formation.

Citations:

Aleza P, Juárez J, Hernández M *et al. (*2009). 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 9:110 https://doi.org/10.1186/1471-2229-9-110.

Alves MN, Lopes SA, Raiol-Junior LL, Wulff NA, Girardi EA, Ollitrault P and Peña L. (2021). Resistance to 'Candidatus Liberibacter asiaticus,' the Huanglongbing associated bacterium, in sexually and/or graft-compatible Citrus relatives. Front. Plant Sci. 11:617664. doi: 10.3389/fpls.2020.617664.

Asins MJ, Bernet GP, Ruiz C, et al. (2004). QTL analysis of Citrus Tristeza Virus-citradia interaction. Theor App Genet 108 603–611.

Axiom™ Genotyping Solution Data Analysis. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/axiom_genotyping_solution_analysis_guide.pdf.

Barkley NA, Roose ML, Krueger RR, Federici CT. (2006). Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). Theor Appl Genet 112 1519–1531. https://doi.org/10.1007/s00122-006-0255-9.

Barry G, Gmitter FG, Chen C, Roose ML, Federici CT, and McCollum GT (2015). Investigating the parentage of 'Orri' and 'Fortune' mandarin hybrids, Proc. XII th Intl. Citrus Congress, Acta Hort. 1065, ISHS.

Biotium. https://biotium.com/product/gelred-nucleic-acid-gel-stain/.

Brenner EA, Blanco M, Gardner C, et al. (2012). Genotypic and phenotypic characterization of isogenic doubled haploid exotic introgression lines in maize. Mol Breeding 30 1001–1016. https://doi.org/10.1007/s11032-011-9684-5.

Bridges GD & Youtsey CO (1966). Improved disease control through hot water treatment of citrus seed Proc. Fla. State Hort. Soc. 79 114-115.

Burbidge AH, James AH (1991). Journal of Heredity, Volume 82, Issue 4, July 1991, Pages 319–328, https://doi.org/10.1093/oxfordjournals.jhered.a111092.

Butelli E, Licciardello C, Zhang Y et al. (2012). Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24:1242–1255.

Cai Q, Guy CL & Moore GA (1994). Extension of the linkage map in Citrus using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theoret. Appl. Genetics 89 606–614. https://doi.org/10.1007/BF00222455.

Castle WS (1981). A review of Citrus seed biology and its relationship to nursery practices Proc. Intl. Soc. Citricult. 1 113-119.

Castle W S (2010). A career perspective on citrus rootstocks, their development, and commercialization, HortScience, 45(1) 11-15.

CD Genomics. The Principles and Workflow of SNP Microarray (2022). https://www.cd-genomics.com/the-principles-and-workflow-of-snp-microarray.html.

Chen C, Lyon MT, O'Malley D et al. (2008). Origin and frequency of 2n gametes in Citrus sinensis x Poncirus trifoliata and their reciprocal crosses. Plant Sci 174 1–8.

Chen C, Gmitter FG (2013). Mining of haplotype-based expressed sequence tag single nucleotide polymorphisms in citrus. BMC Genom 14 746.

Cheng FS, Roose ML (1995) Origin and inheritance of dwarfing by the citrus rootstock Poncirus trifoliate `Flying Dragon'. J Am Soc Hort Sci 120 286–291.

Citrus Industry. Tangerine and Mandarin Global Production to Increase. Citrus Industry (2021).

Citrus Industry. Global citrus statistics released (2022).
https://citrusindustry.net/2022/11/30/global-citrus-statistics-released/

Citrus Variety Collection. https://citrusvariety.ucr.edu/.

Cuenca J, Froelicher Y, Aleza P et al. (2011). Multilocus half-tetrad analysis and centromere mapping in citrus: evidence of SDR mechanism for 2n megagametophyte production and partial chiasma interference in mandarin cv. 'Fortune'. Hered 107 462–470.

Cuenca J, Aleza P, Vicent A et al. (2013a). Genetically based location from triploid populations and gene ontology of a 3.3-Mb genome region linked to Alternaria brown spot resistance in citrus reveal clusters of resistance genes. PLoS ONE 76755.

Curtolo M, Cristofani-Yali M, Gazaffi R et al. (2017). QTL mapping for fruit quality in Citrus using DArTseq markers. BMC Genom 18 289.

Cuties. Sunpacific.com. https://sunpacific.com/our-produce/clementines-mandarins.

Dalkilic Z, Timmer LW, Gmitter FG (2005). Linkage of an Alternaria disease resistance gene in mandarin hybrids with RAPD fragments. J Am Soc Hort Sci 130 191–195.

de Oliveira AC, Bastianel M, Cristofani-Yaly M, do Amaral AM, Machado MA (2007) Development of genetic maps of the citrus varieties 'Murcott' tangor and 'Péra' sweet orange by using fluorescent AFLP markers. J Appl Genet 48 219–231.

de Simone M, Russo MP, Puleo G, Marsan PA, Lorenzoni C, Marocco A, Reforgiato-Recupero G (1998) Construction of genetic maps for Citrus aurantium and C. latipes based on AFLP, RAPD and RFLP markers. Fruits 53 383–390.

Deng Z, Huang S, Xiao S, Gmitter FG (1997). Development and characterization of SCAR markers linked to the Citrus Tristeza Virus resistance gene from Poncirus trifoliata. Genome 40 697–704.

Durham RE, Liou PC, Gmitter FG et al. (1992). Linkage of restriction fragment length polymorphisms and isozymes in Citrus. Theoret. Appl. Genetics 84 39–48. https://doi.org/10.1007/BF00223979.

Fang D, Roose ML, Krueger RR, Federici CT (1997). Fingerprinting trifoliate orange germ plasm accessions with isozymes, RFLPs, and inter-simple sequence repeat markers. Theor Appl Genet 95 211–219.

Federici CT, Fang DQ, Scora RW, Roose ML (1998). Phylogenetic relationships within the genus Citrus (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. Theor Appl Genet 94 812–822.

Fridman E, Carrari F, Liu Y, Fernie AR, Zamir D (2004). Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science. https://www.science.org/doi/full/10.1126/science.1101666.

Froelicher Y, Dambier D, Bassene JB et al. (2008). Characterization of microsatellite markers in mandarin orange (Citrus reticulata Blanco). Mol EcolRes 8 119–122.

Fu Q, Meng X, Luan S et al. (2020). Segregation distortion: high genetic load suggested by a Chinese shrimp family under high-intensity selection. Sci Rep 10 21820. https://doi.org/10.1038/s41598-020-78389-w.

Garavello M, Cuenca J, Dreissig S, Fuchs J, Navarro L, Houben A, Aleza P (2020). Analysis of crossover events and allele segregation distortion in interspecific citrus hybrids by single pollen genotyping. Frontiers in Plant Science. https://www.frontiersin.org/articles/10.3389/fpls.2020.00615.

Garcia-Lor A, Curk F, Snoussi-Trifa H et al. (2012). 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 111 1–19.

Garcia-Lor A, Luro F, Ollitrault P, Navarro L (2015). Genetic diversity and population structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers. Tree Genet Genom 11 123.

Garnsey SM, Miyakawa T (1988). Tristeza. Compendium of Citrus Diseases.. The American Phytopathological Society 48-49.

Geisseler D and Horwath WR (2016.) Citrus Production in California. CDFA.

Gmitter FG, Xiao SY, Huang S et al. (1996). A localized linkage map of the Citrus tristeza virus resistance gene region. Theor Appl Genet 92 688–695.

Grosser JW, Gmitter FG Jr. (1990). Protoplast fusion and citrus improvement. Plant Breeding Reviews Vol.8 339-374.

Gulsen O, Uzun A, Canan I et al. (2010). A new citrus linkage map based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. Euphytica 173 265–277.

Hayden MJ, Sharp PJ (2021). Targeted development of informative microsatellite (SSR) markers. Nucleic Acids Res. doi: 10.1093/nar/29.8.e44. PMID: 11292858; PMCID: PMC31328.

Hiraoka, Y (2020). Application of High-Density SNP Genotyping Array in Citrus Germplasm Characterization and Genetic Dissection of Traits. UC Riverside Electronic Thesis and Dissertations. https://escholarship.org/uc/item/2498q3xj.

Huang M, Roose ML, Yu Q, Du D, Yu Y, Zhang Y, Deng Z, Stover E, Gmitter FG. (2018). Construction of high-density genetic maps and detection of QTSLs associated with huanglongbing tolerance in citrus. Frontiers in Plant Science. https://www.frontiersin.org/articles/10.3389/fpls.2018.01694. 10.3389/fpls.2018.01694.

Inglese, P., Sortino, G (2023) Citrus History, Taxonomy, Breeding, and Fruit Quality. Oxford Research Encyclopedia of Environmental Science. https://oxfordre.com/environmentalscience/view/10.1093/acrefore/9780199938 9414.001.0001/acrefore-9780199389414-e-221.

Jarrell DC, Roose ML, Traugh SN et al. (1992). A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. Theoret. Appl. Genetics 84 49–56. https://doi.org/10.1007/BF00223980.

Kijas JMH, Thomas MR, Fowler JCS et al. (1997). Integration of trinucleotide microsatellites into a linkage map of Citrus. Theor Appl Genet 94 701–706. https://doi.org/10.1007/s001220050468.

Krug CA (1943). Chromosome numbers in the subfamily Aurantioideae with special reference to the genus Citrus. Bot.Gaz. 104 602—611.

Li X, Wang X, Wei Y et al. Prevalence of segregation distortion in diploid alfalfa and its implications for genetics and breeding applications. (2011). Theor Appl Genet 123 667–679. https://doi.org/10.1007/s00122-011-1617-5.

Liang G, Xiong G, Guo Q et al. (2007). AFLP analysis and the taxonomy of Citrus. Acta Hort 760 137–142.

Lin, SY, Ikehashi, H (1993). A gamete abortion locus detected by segregation distortion of isozyme locus Est-9 in wide crosses of rice (Oryza sativa L.). Euphytica 67 35–40. https://doi.org/10.1007/BF00022722.

Lobo I, Shaw K (2008). Thomas Hunt Morgan, genetic recombination, and gene mapping. NatureEducation 1(1):205. https://www.nature.com/scitable/topicpage/thomas-hunt-morgan-genetic-recombination-and-gene-496/.

Lotfy S, Luro F, Carreel F et al. (2003). Application of cleaved amplified polymorphic sequence method for analysis of cytoplasmic genome among Aurantioideae intergeneric somatic hybrids. J Am Soc Hort Sci 128 225–230.

Luro F, Loireux M, Laigret F et al. (1994). Genetic mapping of an intergeneric citrus hybrid using molecular markers. Fruits 49 404–408.

Luro F, Rist D, Ollitrault P (2001). Evaluation of genetic relationships in Citrus genus by means of sequence tagged microsatellites. Acta Hort 546 237–242.

Luro F, Costantino G, Terol J et al. (2008). Transferability of the EST-SSRs developed on Nules clementine (Citrus *clementina* Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping. BMC Genomics 9 287.

Lodish H, Berk A, Matsudaira P, Kaiser CA, Kreiger M, Scott MP, Zipursky L, & Darnell J (2004). Molecular Cell Biology 5th 396. San Francisco: W. H. Freeman.

Mabberley DJ (2004). Citrus (Rutaceae): a review of recent advances in etymology, systematics and medical applications. Blumea, 49 481-498. https://doi.org/10.3767/000651904X484432.

Morgil H, Gercek YC and Isil Tulum I. (2020). Single nucleotide polymorphisms (SNPs) in plant genetics and breeding. The Recent Topics in Genetic Polymorphisms, IntechOpen. Crossref, doi:10.5772/intechopen.91886.

NASS. https://www.nass.usda.gov/Publications/Todays_Reports/reports/cfrt0820.pdf.

Nicolosi E, Deng ZN, Gentile A et al. (2000). Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor Appl Genet 100 1155–1166.

OEC. Citrus (2020). https://oec.world/en/profile/hs/citrus.

Ollitrault P, Terol J, Chen C et al. (2012a). A reference genetic map of C. clementina Hort. ex Tan.; citrus evolution inferences from comparative mapping. BMC Genom 13 593.

Ollitrault P, Terol J, Garcia-Lor A et al. (2012b). SNP mining in C. clementina BAC end sequences; transferability in the Citrus genus (Rutaceae), phylogenetic inferences and perspectives for genetic mapping. BMC Genom 13 13.

Ollitrault P, Ahmed D, Costantino G, Evrard JC, Cardi C, Mournet P, Perdereau A, Froelicher Y (2021). Segregation distortion for male parents in high density genetic maps from reciprocal crosses between two self-incompatible cultivars confirms a gametophytic system for self-incompatibility in citrus. Agriculture 11 379. https://doi.org/10.3390/agriculture11050379.

Pang XM, Hu CG, Deng XX (2007). Phylogenetic relationships within Citrus and its related genera as inferred from AFLP markers. Genet Res Crop Evol 54 429–436.

Perpiñá, G, Esteras C, Gibon Y et al. (2016). A new genomic library of melon introgression lines in a cantaloupe genetic background for dissecting desirable agronomical traits. BMC Plant Biol 16 154. https://doi.org/10.1186/s12870-016-0842-0.

Rafalski A (2002). Applications of single nucleotide polymorphisms in crop genetics. Curr O pin Plant Biol. 94-100. Doi: 10.1016/s1369-5266(02)00240-6. PMID: 11856602.
Ramadugu C, Keremane ML, Halbert SE, Duan YP, Roose ML, Stover E, and Lee RF (2016). Long-term field evaluation reveals huanglongbing resistance in Citrus relatives, Plant Dis. 100 1858-1869.

Rana K, Atri C, Akhatar J, Kaur R, Goyal A, Singh Mohini P, Kumar N, Sharma A, Sandhu PS, Kaur G, Barbetti MJ, Banga SS (2019). Detection of first marker trait associations for resistance against Sclerotinia sclerotiorum in Brassica juncea–Erucastrum cardaminoides introgression lines. Frontiers in Plant Science. https://www.frontiersin.org/articles/10.3389/fpls.2019.01015.

Reece PC. (1967). "The botany of Citrus and its wild relatives." The Citrus Industry 1 190-430.

Reflinur, Kim B, Jang SM, Chu SH, Bordiya Y, Akter MB, Lee J, Chin JH, Koh HJ (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. Rice 7 3. doi: 10.1186/s12284-014-0003-8. Epub 2014 Aug 7. PMID: L6055992; PMCID: PMC4884001.

Roose ML (1988). Isozymes and DNA restriction fragment length polymorphisms in citrus breeding and systematics. Proc Int Soc Citricult 1 155–165.

Ruiz C and Asins MJ (2002). Comparison between Poncirus and Citrus genetic linkage maps. Theor Appl Genet 106:826–836. DOI 10.1007/s00122-002-1095-x.

Sankar AA, Moore GA (2001). Evaluation of inter-simple sequence repeat analysis from mapping in Citrus and extension of the genetic linkage map. Theor Appl Genet 102 206–214.

Scholten OE, van Kaauwen MP, Shahin A et al. (2016). SNP-markers in Allium species to facilitate introgression breeding in onion. BMC Plant Biol 16 187 https://doi.org/10.1186/s12870-016-0879-0.

Sturtevant AH (1913). The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. Journal of Experimental Zoology 14 43-59.

Tang Z, Wang X, Zhang M et al. (2013). The maternal cytoplasmic environment may be involved in the viability selection of gametes and zygotes. Heredity 110 331–337. https://doi.org/10.1038/hdy.2012.89.

Taylor DR, Ingvarsson PK (2003). Common features of segregation distortion in plants and animals. Genetica. 27-35. doi: 10.1023/a:1022308414864. PMID: 12656570.

Terol J, Conesa A, Colmenero JM et al. (2007). Analysis of 13000 unique Citrus clusters associated with fruit quality, production and salinity tolerance. BMC Genom 8 31.

Terol J, Naranjo MA, Ollitrault P, Talon M (2008). Development of genomic resources for Citrus *clementina*: characterization of three deep-coverage BAC libraries and analysis of 46,000 BAC end sequences. BMC Genom 9 423.

Torres AM, Soost RK, Diedenhofen U (1978). Leaf isozymes as genetic markers in Citrus. Am J Bot 65:869–881.

Torres AM, Mau-Lastovicka T, Williams TE, Soost RK (1985). Segregation distortion and linkage of Citrus and Poncirus isozyme genes. J Heredity 76 289-294.
USDA (2020).Citrus fruits 2020 summery. United States Department of Agriculture.

Van Ooijen JW (2018). JoinMap ® 5, software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma B.V., Wageningen, Netherlands.

Vieira ML, Santini L, Diniz AL, Munhoz Cde F (2016). Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol. doi: 10.1590/1678-4685-GMB-2016-0027. Epub 2016 Aug 4. PMID: 27561112; PMCID: PMC5004837.

Voorrips RE (2002). MapChart: software for the graphical presentation of linkage maps and QTLs, Journal of Heredity 93 77–78. https://doi.org/10.1093/jhered/93.1.77.

Wallace JM (1978). Virus and virus-like diseases. In: Reuther W, EC Calavan, and GE Carman (eds.). The Citrus Industry. 4 67-184. Univ. Calif. Div. Agr. Sci., Berkeley.

Wang L, He F, Huang Y, He J, Yang S, Zeng J, Deng C, Jiang X, Fang Y, Wen S, Xu R, Yu H, Yang X, Zhong G, Chen C, Yan X, Zhou C, Zhang H, Xie Z, Larkin RM, Deng X, XuQ (2018). Genome of wild mandarin and domestication history of mandarin. Molecular Plant 11 1024-1037 https://doi.org/10.1016/j.molp.2018.06.001. (https://www.sciencedirect.com/science/article/pii/S1674205218301874).

Wang X, Xu Y, Zhang S et al. (2017). Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. Nat Genet 49 765–772.

WiCell (2022). Single nucleotide polymorphism (SNP). https://www.wicell.org/home/characterization/cytogenetics/snp-microarray/single-nucleotide-polymorphism-snp-mircroarray-.cmsx.

Wu G, Terol J, Ibanez V et al. (2018). Genomics of the origin and evolution of Citrus. Nature 554 311–316. https://doi.org/10.1038/nature25447.

Wu GA, Prochnik S, Jenkins J et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotech 32 656–662.

Wutscher HK (1979). Citrus rootstocks. Horticultural Reviews 1 237-269.
Wu G, Prochnik S, Jenkins J et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat Biotechnol* 32 656–662. https://doi.org/10.1038/nbt.2906.

Xu X, Li L, Dong X, Jin W, Melchinger AE, Chen S (2013). Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J Exp Bot. doi: 10.1093/jxb/ers393.

Xiang C, Roose ML (1988). Frequency and characteristics of nucellar identified zygotic seedlings in 12 Citrus rootstocks. *Scientia Hortic.* 37 47-59.

Yang ZN, Ye XR, Molina J et al. (2003). Sequence analysis of a 282-kilobase region surrounding the citrus tristeza virus resistance gene (CTV) locus in Poncirus trifoliata L. Raf Plant Phys 131 482–492.

Zuo, JF, Niu Y, Cheng P et al. (2019). Effect of marker segregation distortion on high density linkage map construction and QTL mapping in Soybean (Glycine max L.). Heredity 123 579–592. https://doi.org/10.1038/s41437-019-0238-7.

Supplementary:

S/n	Nr	Locus	Segregation	Phase	Position	nn	np	X2	Df	Signif.	Classification
1	224	1_243159	<nnxnp>	{-0}	0	37	58	4.64	1	**	[nn:np]
2	391	1_176583	<nnxnp>	{-1}	0	58	37	4.64	1	**	[nn:np]
3	343	1_333321	<nnxnp>	{-0}	2.151	37	58	4.64	1	**	[nn:np]
4	349	1_419310	<nnxnp>	{-1}	3.215	59	36	5.57	1	**	[nn:np]
6	1023	1_424856	<nnxnp>	{-1}	3.215	58	36	5.15	1	**	[nn:np]
7	1513	1_412679	<nnxnp>	{-0}	3.215	36	59	5.57	1	**	[nn:np]
8	1161	1_705038	<nnxnp>	{-1}	4.279	58	37	4.64	1	**	[nn:np]
10	547	1_698886	<nnxnp>	{-0}	4.279	37	58	4.64	1	**	[nn:np]
12	651	1_101766:	<nnxnp>	{-1}	5.342	59	36	5.57	1	**	[nn:np]
16	898	1_110431	<nnxnp>	{-1}	5.342	58	35	5.69	1	**	[nn:np]
17	409	1_926186	<nnxnp>	{-0}	5.342	36	59	5.57	1	**	[nn:np]
20	647	1_150450:	<nnxnp>	{-0}	6.406	37	58	4.64	1	**	[nn:np]
21	901	1_153468:	<nnxnp>	{-1}	6.406	58	35	5.69	1	**	[nn:np]
22	1041	1_139966'	<nnxnp>	{-1}	6.406	58	37	4.64	1	**	[nn:np]
24	1223	1_165628:	<nnxnp>	{-0}	8.557	37	58	4.64	1	**	[nn:np]
27	498	1_174118(	<nnxnp>	{-0}	10.708	37	58	4.64	1	**	[nn:np]
30	732	1_179991!	<nnxnp>	{-1}	11.772	59	36	5.57	1	**	[nn:np]
31	1322	1_263201:	<nnxnp>	{-0}	13.923	36	59	5.57	1	**	[nn:np]
32	1190	1_210526!	<nnxnp>	{-1}	13.923	59	36	5.57	1	**	[nn:np]
34	97	1_278604!	<nnxnp>	{-0}	14.987	35	60	6.58	1	**	[nn:np]
38	152	1_291161:	<nnxnp>	{-1}	16.051	61	34	7.67	1	***	[nn:np]
40	1699	1_299585:	<nnxnp>	{-1}	17.114	60	35	6.58	1	**	[nn:np]
41	1191	1_301098!	<nnxnp>	{-1}	18.178	59	36	5.57	1	**	[nn:np]
42	195	1_312432(	<nnxnp>	{-1}	19.242	60	35	6.58	1	**	[nn:np]
47	929	1_329272:	<nnxnp>	{-1}	19.242	60	34	7.19	1	***	[nn:np]
48	463	1_353474	<nnxnp>	{-0}	19.242	35	60	6.58	1	**	[nn:np]
49	214	1_345433:	<nnxnp>	{-1}	20.306	61	34	7.67	1	***	[nn:np]
50	339	1_381821!	<nnxnp>	{-0}	22.457	34	61	7.67	1	***	[nn:np]
52	94	1_370932!	<nnxnp>	{-1}	22.457	61	33	8.34	1	****	[nn:np]
53	235	1_376157:	<nnxnp>	{-1}	22.457	61	34	7.67	1	***	[nn:np]
58	221	1_405994!	<nnxnp>	{-1}	23.521	62	33	8.85	1	****	[nn:np]
60	563	1_415361!	<nnxnp>	{-0}	24.585	34	61	7.67	1	***	[nn:np]
62	1365	1_412708:	<nnxnp>	{-1}	24.585	61	34	7.67	1	***	[nn:np]
65	568	1_446951!	<nnxnp>	{-1}	26.736	61	34	7.67	1	***	[nn:np]
66	950	1_434910:	<nnxnp>	{-0}	26.736	34	60	7.19	1	***	[nn:np]
67	1071	1_475289!	<nnxnp>	{-1}	28.886	59	36	5.57	1	**	[nn:np]
68	1424	1_487001:	<nnxnp>	{-1}	29.95	58	37	4.64	1	**	[nn:np]
69	151	1_523463:	<nnxnp>	{-1}	31.014	59	36	5.57	1	**	[nn:np]
73	171	1_633523!	<nnxnp>	{-1}	36.575	58	37	4.64	1	**	[nn:np]
77	703	1_644476:	<nnxnp>	{-0}	36.575	37	58	4.64	1	**	[nn:np]
78	604	1_648366:	<nnxnp>	{-1}	37.639	57	38	3.8	1	*	[nn:np]
79	448	1_692079	<nnxnp>	{-0}	38.703	37	58	4.64	1	**	[nn:np]
81	1084	1_722272!	<nnxnp>	{-0}	40.854	39	56	3.04	1	*	[nn:np]
82	933	1_740326!	<nnxnp>	{-0}	40.854	39	55	2.72	1	*	[nn:np]
83	1262	1_887012:	<nnxnp>	{-1}	41.918	55	40	2.37	1		[nn:np]

S1: JoinMap output of Chromosome 1 for nn:np loci of the cross (TP x FD) x Chandler complete

range.

S/n	Nr	Locus	Segregation	Phase	Position	hh	hk	kk	nn	np	X2	Df	Signif.	Classification
1	265	1_128690	<hkxhk>	{11}	0	28	47	19	0	0	1.72	2		[hh:hk:kk]
2	1249	1_100968	<hkxhk>	{00}	0	19	47	29	0	0	2.12	2		[hh:hk:kk]
4	979	1_165943	<hkxhk>	{01}	0.532	29	47	19	0	0	2.12	2		[hh:hk:kk]
6	1919	1_243159	<nnxnp>	{-0}	1.064	0	0	0	37	58	4.64	1	**	[nn:np]
7	2086	1_176583	<nnxnp>	{-1}	1.064	0	0	0	58	37	4.64	1	**	[nn:np]
5	1461	1_195387	<hkxhk>	{01}	1.064	30	46	19	0	0	2.64	2		[hh:hk:kk]
8	3208	1_412679	<nnxnp>	{-0}	4.026	0	0	0	36	59	5.57	1	**	[nn:np]
9	2718	1_424856	<nnxnp>	{-1}	4.026	0	0	0	58	36	5.15	1	**	[nn:np]
10	2044	1_419310	<nnxnp>	{-1}	4.026	0	0	0	59	36	5.57	1	**	[nn:np]
12	2242	1_698886	<nnxnp>	{-0}	4.992	0	0	0	37	58	4.64	1	**	[nn:np]
14	2856	1_705038	<nnxnp>	{-1}	4.992	0	0	0	58	37	4.64	1	**	[nn:np]
16	2353	1_924680	<nnxnp>	{-0}	5.958	0	0	0	36	59	5.57	1	**	[nn:np]
18	2346	1_1017662	<nnxnp>	{-1}	5.958	0	0	0	59	36	5.57	1	**	[nn:np]
22	841	1_916185	<hkxhk>	{11}	5.958	29	48	18	0	0	2.56	2		[hh:hk:kk]
23	2596	1_1534681	<nnxnp>	{-1}	9.271	0	0	0	58	35	5.69	1	**	[nn:np]
24	2342	1_1504501	<nnxnp>	{-0}	9.271	0	0	0	37	58	4.64	1	**	[nn:np]
28	2736	1_1399667	<nnxnp>	{-1}	9.271	0	0	0	58	37	4.64	1	**	[nn:np]
27	897	1_1368686	<hkxhk>	{01}	9.271	31	49	15	0	0	5.48	2	*	[hh:hk:kk]
25	498	1_1521895	<hkxhk>	{00}	9.271	22	46	27	0	0	0.62	2		[hh:hk:kk]
30	2918	1_1656283	<nnxnp>	{-0}	11.142	0	0	0	37	58	4.64	1	**	[nn:np]
33	2428	1_1773619	<nnxnp>	{-0}	13.013	0	0	0	37	58	4.64	1	**	[nn:np]
35	2427	1_1799915	<nnxnp>	{-1}	13.938	0	0	0	59	36	5.57	1	**	[nn:np]
36	327	1_2036390	<hkxhk>	{01}	15.809	31	47	17	0	0	4.14	2		[hh:hk:kk]
38	2885	1_2105269	<nnxnp>	{-1}	16.341	0	0	0	59	36	5.57	1	**	[nn:np]
40	1513	1_2096271	<hkxhk>	{01}	16.341	32	46	17	0	0	4.83	2	*	[hh:hk:kk]
41	516	1_2103979	<hkxhk>	{10}	16.341	17	46	32	0	0	4.83	2	*	[hh:hk:kk]
43	341	1_2469791	<hkxhk>	{01}	16.873	33	45	17	0	0	5.65	2	*	[hh:hk:kk]
55	3017	1_2632012	<nnxnp>	{-0}	17.405	0	0	0	36	59	5.57	1	**	[nn:np]
49	77	1_2568529	<hkxhk>	{01}	17.405	33	42	18	0	0	5.71	2	*	[hh:hk:kk]
50	629	1_2529413	<hkxhk>	{10}	17.405	18	44	33	0	0	5.25	2	*	[hh:hk:kk]
56	1792	1_2786049	<nnxnp>	{-0}	18.469	0	0	0	35	60	6.58	1	**	[nn:np]
64	1253	1_2807190	<hkxhk>	{10}	18.469	17	44	34	0	0	6.6	2	**	[hh:hk:kk]
60	1418	1_2674759	<hkxhk>	{00}	18.469	18	51	26	0	0	1.86	2		[hh:hk:kk]
62	257	1_2677759	<hkxhk>	{11}	18.469	25	51	18	0	0	1.72	2		[hh:hk:kk]
63	1233	1_2768185	<hkxhk>	{11}	18.469	26	51	18	0	0	1.86	2		[hh:hk:kk]
65	1847	1_2911618	<nnxnp>	{-1}	19.001	0	0	0	61	34	7.67	1	***	[nn:np]
67	439	1_2930006	<hkxhk>	{01}	19.001	35	43	17	0	0	7.67	2	**	[hh:hk:kk]
69	557	1_2964061	<hkxhk>	{10}	19.001	17	43	35	0	0	7.67	2	**	[hh:hk:kk]
73	3394	1_2995851	<nnxnp>	{-1}	19.889	0	0	0	60	35	6.58	1	**	[nn:np]
75	2624	1_3292723	<nnxnp>	{-1}	21.684	0	0	0	60	34	7.19	1	***	[nn:np]
74	2158	1_3534744	<nnxnp>	{-0}	21.684	0	0	0	35	60	6.58	1	**	[nn:np]
76	1890	1_3124320	<nnxnp>	{-1}	21.684	0	0	0	60	35	6.58	1	**	[nn:np]
81	1673	1_3205369	<hkxhk>	{10}	21.684	17	44	34	0	0	6.6	2	**	[hh:hk:kk]
82	261	1_3224353	<hkxhk>	{00}	21.684	18	51	25	0	0	1.72	2		[hh:hk:kk]
83	159	1_3218377	<hkxhk>	{11}	21.684	25	51	18	0	0	1.72	2		[hh:hk:kk]
84	1353	1_3282414	<hkxhk>	{11}	21.684	26	51	18	0	0	1.86	2		[hh:hk:kk]
85	1909	1_3454333	<nnxnp>	{-1}	22.216	0	0	0	61	34	7.67	1	***	[nn:np]
86	1484	1_3443788	<hkxhk>	{10}	22.216	16	45	34	0	0	7.08	2	**	[hh:hk:kk]
87	2034	1_3818219	<nnxnp>	{-0}	23.291	0	0	0	34	61	7.67	1	***	[nn:np]
89	107	1_3695977	<hkxhk>	{00}	23.291	18	50	26	0	0	1.74	2		[hh:hk:kk]
90	1789	1_3709325	<nnxnp>	{-1}	23.557	0	0	0	61	33	8.34	1	****	[nn:np]
91	1413	1_3757955	<hkxhk>	{10}	23.823	16	44	35	0	0	8.12	2	**	[hh:hk:kk]
92	1930	1_3761573	<nnxnp>	{-1}	24.089	0	0	0	61	34	7.67	1	***	[nn:np]
97	438	1_3948041	<hkxhk>	{10}	24.355	17	43	35	0	0	7.67	2	**	[hh:hk:kk]
98	547	1_3827122	<hkxhk>	{00}	24.355	17	52	26	0	0	2.56	2		[hh:hk:kk]
99	172	1_3816710	<hkxhk>	{11}	24.355	25	52	17	0	0	2.43	2		[hh:hk:kk]
100	1916	1_4059949	<nnxnp>	{-1}	25.434	0	0	0	62	33	8.85	1	****	[nn:np]

| 102	3060	1_4127082	<nnxnp>	{-1}	26.513	0	0	0	61	34	7.67	1	***	[nn:np]
105	2258	1_4153618	<nnxnp>	{-0}	26.513	0	0	0	34	61	7.67	1	***	[nn:np]
108	2645	1_4349101	<nnxnp>	{-0}	28.694	0	0	0	34	60	7.19	1	***	[nn:np]
111	2263	1_4469515	<nnxnp>	{-1}	28.694	0	0	0	61	34	7.67	1	***	[nn:np]
107	175	1_4566813	<hkxhk>	{10}	28.694	20	38	36	0	0	8.89	2	**	[hh:hk:kk]
109	1472	1_4366641	<hkxhk>	{01}	28.694	36	39	20	0	0	8.43	2	**	[hh:hk:kk]
110	1070	1_4494063	<hkxhk>	{11}	28.694	25	56	14	0	0	5.59	2	*	[hh:hk:kk]
112	2766	1_4752895	<nnxnp>	{-1}	30.422	0	0	0	59	36	5.57	1	**	[nn:np]
113	3119	1_4870013	<nnxnp>	{-1}	31.277	0	0	0	58	37	4.64	1	**	[nn:np]
114	1846	1_5234633	<nnxnp>	{-1}	32.132	0	0	0	59	36	5.57	1	**	[nn:np]
118	1866	1_6335239	<nnxnp>	{-1}	36.602	0	0	0	58	37	4.64	1	**	[nn:np]
122	2398	1_6444768	<nnxnp>	{-0}	36.602	0	0	0	37	58	4.64	1	**	[nn:np]
123	2299	1_6483668	<nnxnp>	{-1}	37.457	0	0	0	57	38	3.8	1	*	[nn:np]
124	2143	1_6920794	<nnxnp>	{-0}	38.312	0	0	0	37	58	4.64	1	**	[nn:np]
126	2779	1_7222729	<nnxnp>	{-0}	40.04	0	0	0	39	56	3.04	1	*	[nn:np]
132	307	1_8856293	<hkxhk>	{01}	40.896	38	33	24	0	0	12.98	2	****	[hh:hk:kk]
137	917	1_8982157	<hkxhk>	{00}	40.896	16	62	17	0	0	8.87	2	**	[hh:hk:kk]
127	2957	1_8870121	<nnxnp>	{-1}	40.896	0	0	0	55	40	2.37	1		[nn:np]
131	2577	1_10553126	<nnxnp>	{-0}	40.896	0	0	0	40	52	1.57	1		[nn:np]
134	2889	1_8820437	<nnxnp>	{-0}	40.896	0	0	0	40	55	2.37	1		[nn:np]
139	36	1_14594975	<hkxhk>	{00}	41.428	17	58	17	0	0	6.26	2	**	[hh:hk:kk]
140	1675	1_14662281	<hkxhk>	{00}	41.96	17	62	16	0	0	8.87	2	**	[hh:hk:kk]
141	495	1_14652960	<hkxhk>	{11}	41.96	16	62	17	0	0	8.87	2	**	[hh:hk:kk]
142	1631	1_16104507	<hkxhk>	{11}	42.492	16	63	16	0	0	10.12	2	***	[hh:hk:kk]
143	2348	1_15847100	<nnxnp>	{-0}	42.492	0	0	0	41	54	1.78	1		[nn:np]
146	802	1_16720701	<hkxhk>	{10}	43.024	26	31	38	0	0	14.49	2	*****	[hh:hk:kk]
147	848	1_17553348	<hkxhk>	{00}	44.087	16	62	17	0	0	8.87	2	**	[hh:hk:kk]
148	2074	1_17574263	<nnxnp>	{-0}	44.087	0	0	0	41	54	1.78	1		[nn:np]
158	364	1_19348357	<hkxhk>	{01}	47.35	36	37	22	0	0	8.77	2	**	[hh:hk:kk]
159	3	1_19984097	<hkxhk>	{00}	47.35	16	56	13	0	0	8.79	2	**	[hh:hk:kk]
152	2863	1_18898058	<nnxnp>	{-1}	47.35	0	0	0	57	38	3.8	1	*	[nn:np]
153	1990	1_19127461	<nnxnp>	{-0}	47.35	0	0	0	38	57	3.8	1	*	[nn:np]
157	2603	1_19331985	<nnxnp>	{-0}	47.35	0	0	0	38	55	3.11	1	*	[nn:np]
160	926	1_19352877	<hkxhk>	{11}	47.35	21	58	16	0	0	5.17	2	*	[hh:hk:kk]
161	1485	1_20475467	<hkxhk>	{10}	48.425	23	37	35	0	0	7.67	2	**	[hh:hk:kk]
162	459	1_20519205	<hkxhk>	{00}	48.425	17	58	20	0	0	4.83	2	*	[hh:hk:kk]
167	1626	1_20964434	<hkxhk>	{01}	49.5	34	39	22	0	0	6.07	2	**	[hh:hk:kk]
164	1832	1_21007580	<nnxnp>	{-0}	49.5	0	0	0	40	55	2.37	1		[nn:np]
168	966	1_21357684	<hkxhk>	{10}	50.032	23	38	34	0	0	6.35	2	**	[hh:hk:kk]
169	2057	1_21474679	<nnxnp>	{-1}	50.298	0	0	0	55	40	2.37	1		[nn:np]
173	174	1_21641420	<hkxhk>	{10}	50.564	21	39	34	0	0	6.32	2	**	[hh:hk:kk]
179	1047	1_21879286	<hkxhk>	{01}	50.564	34	39	22	0	0	6.07	2	**	[hh:hk:kk]
174	446	1_22009994	<hkxhk>	{00}	50.564	18	56	21	0	0	3.23	2		[hh:hk:kk]
180	1530	1_21410901	<hkxhk>	{11}	50.564	21	56	18	0	0	3.23	2		[hh:hk:kk]
181	1702	1_22723834	<nnxnp>	{-1}	51.912	0	0	0	54	36	3.6	1	*	[nn:np]
182	3118	1_22693926	<nnxnp>	{-1}	51.912	0	0	0	54	41	1.78	1		[nn:np]
184	3111	1_22754098	<nnxnp>	{-0}	51.912	0	0	0	41	54	1.78	1		[nn:np]
185	1605	1_22358380	<hkxhk>	{01}	53.259	30	44	21	0	0	2.22	2		[hh:hk:kk]
186	856	1_23017599	<hkxhk>	{11}	53.259	23	51	21	0	0	0.6	2		[hh:hk:kk]
187	1014	1_22959548	<hkxhk>	{10}	53.259	21	44	30	0	0	2.22	2		[hh:hk:kk]
188	3299	1_22988598	<nnxnp>	{-1}	53.525	0	0	0	53	42	1.27	1		[nn:np]
189	596	1_23241657	<hkxhk>	{00}	53.791	21	52	22	0	0	0.87	2		[hh:hk:kk]
190	1813	1_23363519	<nnxnp>	{-0}	53.791	0	0	0	42	53	1.27	1		[nn:np]
194	1360	1_23228712	<hkxhk>	{11}	53.791	22	52	21	0	0	0.87	2		[hh:hk:kk]
195	1601	1_23454677	<hkxhk>	{00}	56.498	24	49	22	0	0	0.18	2		[hh:hk:kk]
196	1682	1_23488016	<hkxhk>	{01}	57.029	29	45	21	0	0	1.61	2		[hh:hk:kk]

S2: JoinMap output of Chromosome 1 for nn:np and hk:hk loci combined of the cross (TP x FD) x

Chandler complete range.

S3: Mapchart Maps of Chromosomes for the cross (TP x FD) x

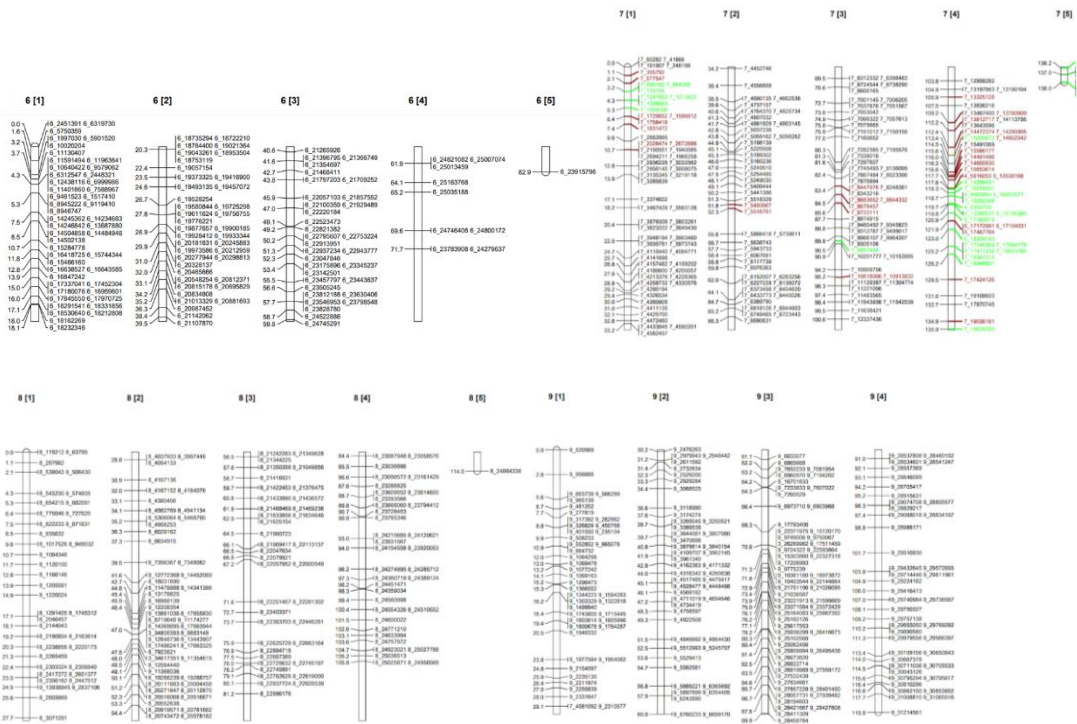Chandler nn:np and hk:hk loci.
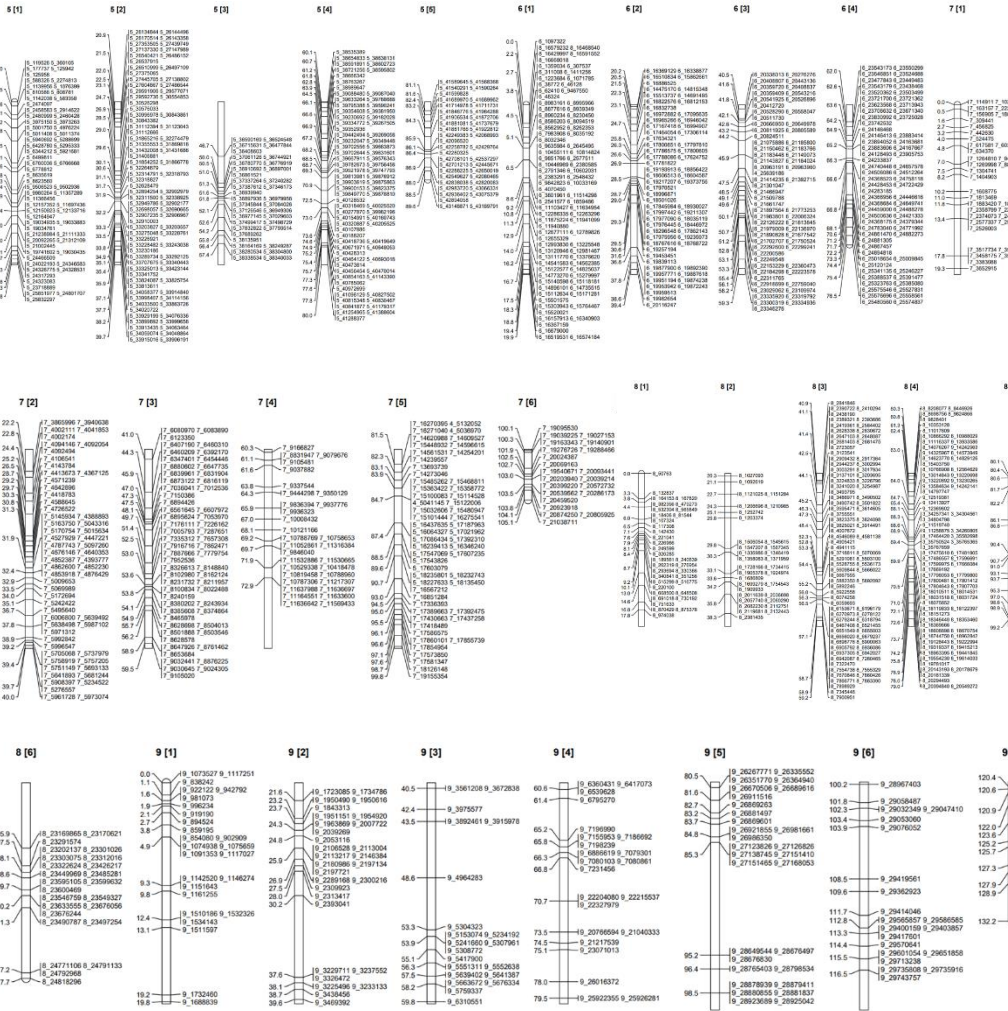
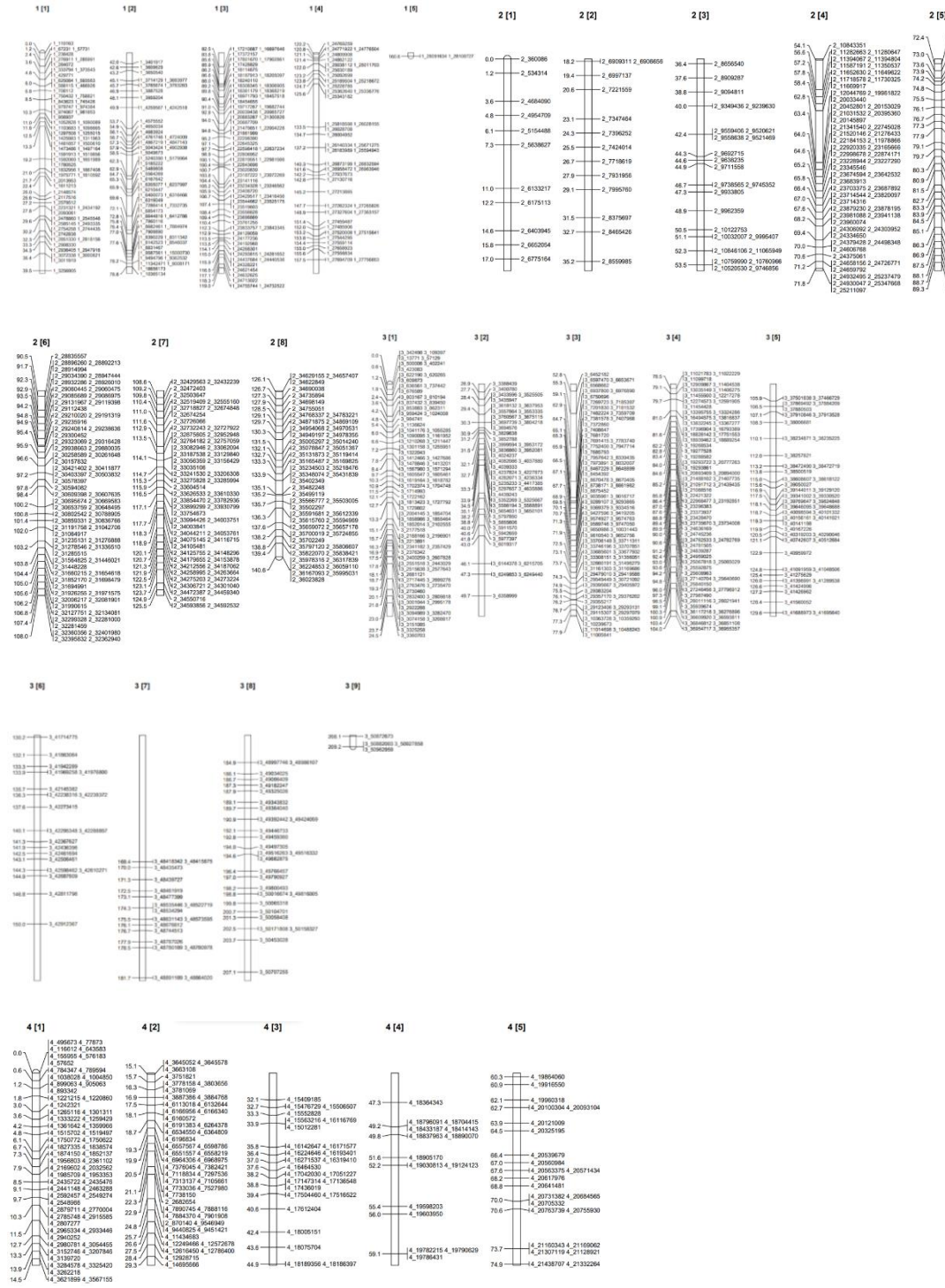S4: Mapchart Maps of Chromosomes for the cross (TP x FD) x Chandler nn:np loci.

S5: Mapchart Maps of Chromosomes for the cross (TP x FD) x Chandler lm:ll loci.

S6: Mapchart Maps of Chromosomes for the cross Fortune x Fairchild lm:ll and hk:hk loci
combined.

Supplemental Explanations for Tables:

P2 (or P1 in other tables) Rec max: The largest number of expected crossovers observed in gametes from the male parent Chandler in tables 1 and 2 (table 3 refers to the female parent TP x FD, table 4 male parent Fortune, and table 5 female parent Fairchild). A larger value may indicate some bad SNP calls.

P2 (or P1 in other tables) rec >4: The number of progeny having more than 4 crossovers in gametes from Parent.

 # P2 (or P1 in other tables) NNstress in cM>2: Number of P2 loci with nearest neighbor stress values >2 cM. Stress for a locus will be high if it is placed in the wrong map position.

Gt. Prob. Ind Means Max: For each unique locus and individual, the program calculates the probability of each genotype call given the map and genotypes of the neighboring unique loci and records these as the negative of the log10 value, so a probability of .01 would have a value of 2. These are then averaged over all loci for each individual and chromosome. An individual with many genotype calls not consistent with the map (many error calls or not a member of the population) will have a large value (> .02).

#Gt. Prob. Ind >.01: This shows, for each chromosome, the number of individuals with genotype probability averaged over all unique loci that are greater than .01 (that is, not a near perfect fit to the map).

#Gt. Prob. Loci >.1: Shows, for each chromosome, the number of loci with genotype probabilities averaged over individuals that are greater than 0.1, a value consistent with about 3% of loci not fitting well.

S7: Mapchart Maps of Chromosomes for the cross Fortune x Fairchild nn:np and hk:hk loci combined.