Using pentaploids as tools for studying genomic composition and allele transmission patterns in octoploid Fragaria

Elizabeth Poulsen
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/thesis

Recommended Citation
https://scholars.unh.edu/thesis/585

This Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Master's Theses and Capstones by an authorized administrator of University of New Hampshire Scholars’ Repository. For more information, please contact nicole.hentz@unh.edu.
USING PENTAPLOIDS AS TOOLS FOR STUDYING GENOMIC COMPOSITION
AND ALLELE TRANSMISSION PATTERNS IN OCTOPLOID FRAGARIA

BY

ELIZABETH POULSEN
Bachelor of Science, University of New Hampshire, 2005

THESIS

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Master of Science
in
Genetics

September, 2010
This thesis has been examined and approved.

Thomas M. Davis, Professor of Biology and Plant Biology

Rebecca C. Sideman, Extension Associate Professor of Environmental Horticulture, Extension Associate Faculty of Plant Biology

J. Brent Loy, Professor of Biology and Environmental Horticulture and Plant Biology

August 16, 2010
Date
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Introduction to the Genus <em>Fragaria</em></td>
<td>1</td>
</tr>
<tr>
<td>Economics of Strawberries</td>
<td>2</td>
</tr>
<tr>
<td>Genetic Importance of <em>Fragaria</em></td>
<td>3</td>
</tr>
<tr>
<td><em>Fragaria</em> Ancestry</td>
<td>3</td>
</tr>
<tr>
<td><em>Fragaria vesca</em></td>
<td>3</td>
</tr>
<tr>
<td>History of <em>Fragaria xananassa</em></td>
<td>5</td>
</tr>
<tr>
<td><em>Fragaria</em> Octoploid Models</td>
<td>6</td>
</tr>
<tr>
<td>Disomic vs. Polysomic Inheritance</td>
<td>9</td>
</tr>
<tr>
<td>Explanation of Terms</td>
<td>9</td>
</tr>
<tr>
<td>Significance of Multivalents</td>
<td>11</td>
</tr>
<tr>
<td>CAPS, Isozyme, and other Genetic Evidence</td>
<td>12</td>
</tr>
<tr>
<td>Meiotic Pairing Evidence</td>
<td>14</td>
</tr>
<tr>
<td>Intergenic Regions in <em>Fragaria</em></td>
<td>14</td>
</tr>
<tr>
<td>Scope of Research</td>
<td>15</td>
</tr>
<tr>
<td>II. MATERIALS &amp; METHODS</td>
<td>19</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Breeding History and Pentaploid Production</td>
<td>19</td>
</tr>
<tr>
<td>Pentaploid Seed Sterilization and Planting</td>
<td>21</td>
</tr>
<tr>
<td>Confirmation of Ploidy</td>
<td>21</td>
</tr>
<tr>
<td>Chromosome Squash Slides</td>
<td>21</td>
</tr>
<tr>
<td>Flow-Sorting Outsourcing</td>
<td>22</td>
</tr>
<tr>
<td>Genomic DNA Isolation</td>
<td>23</td>
</tr>
<tr>
<td>Genomic PCR</td>
<td>24</td>
</tr>
<tr>
<td>Intergenic Region PCR</td>
<td>24</td>
</tr>
<tr>
<td>Expanded Region PCR</td>
<td>25</td>
</tr>
<tr>
<td>Gel Electrophoresis</td>
<td>26</td>
</tr>
<tr>
<td>LB Bacterial Medium Plates</td>
<td>27</td>
</tr>
<tr>
<td>Molecular Cloning and Colony PCR</td>
<td>27</td>
</tr>
<tr>
<td>Plasmid Isolation and PCR</td>
<td>29</td>
</tr>
<tr>
<td>DNA Sequencing</td>
<td>31</td>
</tr>
<tr>
<td>Fluorescent End Labeling by PCR</td>
<td>32</td>
</tr>
<tr>
<td>Enzyme Digestion of Fluorescent End-Labeled PCR Products</td>
<td>33</td>
</tr>
<tr>
<td>Genotyping of FEL-CAPS</td>
<td>35</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>36</td>
</tr>
<tr>
<td>Pentaploid Population</td>
<td>36</td>
</tr>
<tr>
<td>Chromosome Counting</td>
<td>36</td>
</tr>
<tr>
<td>Flow Cytometry</td>
<td>37</td>
</tr>
<tr>
<td>Electrophoresis of Genomic PCR Products</td>
<td>38</td>
</tr>
<tr>
<td>Intergenic Region PCR Products</td>
<td>38</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I am deeply grateful to my amazingly supportive family, without whom this thesis would not have been completed. My parents, Dianne and Val Poulsen have always encouraged me to pursue scholarly endeavors and have aided me both emotionally and financially throughout my college years at the University of New Hampshire. My brother, Christopher Kratt has been a superior role model for me and always inspires me to achieve my fullest potential. My best friend and boyfriend, DJ Ducharme has provided me with indispensable motivation at the most crucial times and has helped me achieve this goal in so many ways.

I am also very fortunate to have had the support and encouragement of so many knowledgeable and talented students, faculty, and staff at the University of New Hampshire. My advisor, Dr. Tom Davis, has constantly provided me with invaluable insight, guidance, and encouragement since I began working in his lab as an undergraduate. He is not only a fountain of wisdom but has also demonstrated immeasurable patience. I would also like to thank Melanie Shields, Lise Mahoney, Qian Zhang, Ben Orcheski, Bo Liu, Kelly Vining, Laura DiMeglio, and Kevin DeHaan of the Davis Laboratory for their technical guidance and generous assistance over the years. I am also greatly appreciative to my thesis committee members, Dr. Brent Loy and Dr. Becky Sideman for providing valuable insight to improve this thesis.
LIST OF TABLES

Table 1. Ploidy and ranges of *Fragaria* species. ............................................................... 2
Table 2. FEL-CAPS genotyping data................................................................................... 49

LIST OF FIGURES

Figure 1. Octoploid *Fragaria* evolution from diploid progenitors ................................. 7
Figure 2. Two published models for the genome composition of octoploid *Fragaria*;
AAA'A'BBB'B' and AAA'A'BBBB .................................................................................... 8
Figure 3. Gametes produced by AAA'A'BBB'B' and AAA'A'BBBB octoploids .......... 18
Figure 4. Genomic ancestry of an *F. virginiana* × *F. vesca* pentaploid population...... 20
Figure 5. PCR-amplified *gRGA1-Subtilase* intergenic region ....................................... 25
Figure 6. Expanded PCR-amplified *gRGA1-Subtilase* intergenic region ...................... 26
Figure 7. pCR®2.1-TOPO vector .................................................................................... 28
Figure 8. Equation for how many clones must be sequenced in a polyploid................. 31
Figure 9. The production of FEL-CAPS (Fluorescent End-Labeled Cleaved Amplified
Polymorphic Sequence) ................................................................................................. 35
Figure 10. Chromosomes in a pentaploid root tip cell ..................................................... 37
Figure 11. *gRGA-Subtilase* PCR products from LB48 and pentaploids ....................... 39
Figure 12. Expanded region *gRGA1-Subtilase* PCR products amplified with
AccuPrime™ Polymerase ............................................................................................... 40
Figure 13. Colony-PCR products .................................................................................... 41
Figure 14. Plasmid PCR Products ................................................................. 42
Figure 15. Schematic diagram of the gRGA1-Subtilase haplotype alignment ........... 43
Figure 16. PCR on haplotype-representative plasmids ........................................ 44
Figure 17. Diagram of expected gRGA1-Subtilase PCR product sizes in the pentaploids and family ................................................................. 45
Figure 18. FEL-CAPS products ....................................................................... 46
Figure 19. gRGA1-Subtilase haplotypes in LB48 and pentaploid population ........... 47
Figure 20. Genotyping chromatogram from LB48 ........................................... 50
Figure 21. Genotyping chromatogram from pentaploid 5XB ............................. 51
Figure 22. Genotyping chromatogram from pentaploid 5XD ............................. 52
ABSTRACT

USING PENTAPLOIDS AS TOOLS FOR STUDYING GENOMIC COMPOSITION AND ALLELE TRANSMISSION PATTERNS IN OCTOPLOID FRAGARIA

by

Elizabeth Poulsen

University of New Hampshire, September, 2010

A unique pentaploid population of Fragaria (strawberry) was developed to gain insight into octoploid genome composition and allele transmission patterns. This population was produced by crossing representatives of two divergent subspecies of octoploid F. virginiana and then by crossing an octoploid hybrid plant with diploid F. vesca. To enable examination of allele transmission patterns, the intergenic region of the gRGA1-Subtilase gene pair locus was amplified, and PCR products were cloned, sequenced, and compared to define the allele composition of the hybrid octoploid. The sequencing data revealed three distinct major haplotypes and additional subtypes. The pentaploids were genotyped utilizing the FEL-CAPS method. From these results, it was possible to detect disomic inheritance of an allele pair within one major haplotype. This result establishes the potential value of the pentaploid population as a resource for analysis of octoploid allele transmission patterns in strawberry.
CHAPTER I

INTRODUCTION

Introduction to the Genus *Fragaria*

The genus *Fragaria* (strawberry), which belongs to the family Rosaceae, comprises 23 species and is found all over Europe, Asia, North America, and along the western coast of South America (Hancock and Luby, 1993; Staudt, 2009). *Fragaria* is a member of the subclade in Rosoideae called Fragariinae (Lundberg et al., 2009). Other members of Fragariinae are the genera Alchemilla, Comarum, Dasiphora, Potaninia, Sibbaldia, Sibbaldianthe, Sibbaldiopsis, Chamaerhodos, and Drymocallis. The genera Potentilla, Rosa, and Rubus are the most closely related to Fragariinae. The ploidy levels in natural *Fragaria* species range from diploid (2x) to decaploid (10x) (see Table 1). The genome of *Fragaria* consists of seven chromosomes per set (Ichijima, 1926), and the chromosomes are very small. The C-value of the diploid *F. vesca* is approximately 206 Mbp, giving *Fragaria* species one of the smallest basic genomes of all major crop species (Folta and Davis, 2006). The morphology of the chromosomes is not consistent across all species (Iwatsubo and Naruhashi, 1991; 1989). Each chromosome set in *Fragaria* includes at least one chromosome with a secondary constriction containing a nucleolar organizing region. One representative of diploid *F. vesca* was reported to have three 45s rDNA loci among its seven chromosomes, and one 5s rDNA locus was co-localized with one of the 45s loci (Lim, 2004).
<table>
<thead>
<tr>
<th>Species</th>
<th>Ploidy</th>
<th>Range</th>
<th>Breeding System</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>F. vesca</em></td>
<td>2x</td>
<td>North America, Northern Asia,</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. viridis</em></td>
<td>2x</td>
<td>Europe, Western Asia</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. nilgerrensis</em></td>
<td>2x</td>
<td>Southern and Eastern Asia</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. daltoniana</em></td>
<td>2x</td>
<td>Eastern Himalayas</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. nubicola</em></td>
<td>2x</td>
<td>Eastern Himalayas</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. bucharica</em></td>
<td>2x</td>
<td>Western Himalayas</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. iinumae</em></td>
<td>2x</td>
<td>Japan, Southern Russia</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. mandshurica</em></td>
<td>2x</td>
<td>Northeastern Asia</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. nipponica</em></td>
<td>2x</td>
<td>Japan</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. hayatai</em></td>
<td>2x</td>
<td>Taiwan</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. pentaphylla</em></td>
<td>2x</td>
<td>Southwest China</td>
<td>Self-incompatible</td>
</tr>
<tr>
<td><em>F. xhifera</em></td>
<td>2x, 3x</td>
<td>Europe</td>
<td>Self-compatible</td>
</tr>
<tr>
<td><em>F. corymbosa</em></td>
<td>4x</td>
<td>Northern China</td>
<td>Dioecious</td>
</tr>
<tr>
<td><em>F. gracilis</em></td>
<td>4x</td>
<td>Northwest China</td>
<td>Dioecious</td>
</tr>
<tr>
<td><em>F. orientalis</em></td>
<td>4x</td>
<td>Northeastern Asia</td>
<td>Trioeious</td>
</tr>
<tr>
<td><em>F. moopinensis</em></td>
<td>4x</td>
<td>Southwestern China</td>
<td>Dioecious</td>
</tr>
<tr>
<td><em>F. tibetica</em></td>
<td>4x</td>
<td>Eastern Himalayas</td>
<td>Dioecious</td>
</tr>
<tr>
<td><em>F. xbringhurstii</em></td>
<td>5x, 6x</td>
<td>California</td>
<td>Dioecious</td>
</tr>
<tr>
<td><em>F. moschata</em></td>
<td>6x</td>
<td>Northern and Central Europe</td>
<td>Trioeious</td>
</tr>
<tr>
<td><em>F. chiloensis</em></td>
<td>8x</td>
<td>Western regions of the Americas,</td>
<td>Trioeious</td>
</tr>
<tr>
<td><em>F. virginiana</em></td>
<td>8x</td>
<td>North America</td>
<td>Trioeious</td>
</tr>
<tr>
<td><em>F. xananassa</em></td>
<td>8x</td>
<td>Pacific Northwest, North America</td>
<td>Trioeious</td>
</tr>
<tr>
<td><em>F. iturupensis</em></td>
<td>10x*</td>
<td>Japan</td>
<td>Trioeious</td>
</tr>
</tbody>
</table>

Table 1. **Ploidy and ranges of Fragaria species.** This table summarizes the species names, ploidy levels, geographical locations, and breeding systems of the members of the genus *Fragaria* (Hancock and Luby, 1993; Hancock, 1999; Folta and Davis, 2006; Staudt 2009). *New data indicate that recently collected representatives of this species are decaploid rather than octoploid (Hummer et al., 2009).

**Economics of Strawberries**

U.S. consumption, demand, acreage, and value of strawberries are increasing considerably (Folta and Davis, 2006). In 1970, the average yearly consumption of strawberries in the U.S. was just less than 2 lbs., and in 2004 the average was 6.8 lbs. per
person. The U.S. contributed over one-fourth of the total worldwide crop of over 4 million metric tons of strawberries in 2008 (FAOSTAT, 2009).

**Genetic Importance of Fragaria**

Because of *Fragaria*’s compact genome size and 80-100% sequence identity with housekeeping genes in other rosaceous species, it is a good genetic representative of the family Rosaceae. Studying *Fragaria* genetics is facilitated by the ease at which the plants can be grown in large numbers in greenhouses and by the ability to propagate the plants by either sexual or vegetative means (Folta and Davis, 2007).

**Fragaria Ancestry**

*Fragaria vesca*

The diploid *F. vesca* has been suggested as a likely ancestral species of the other diploids because of its wide geographic distribution and its ability to cross with nearly all of the other diploids in addition to itself (Hancock, 1999). Furthermore, *F. vesca* may be the most primitive extant *Fragaria* species (Longley, 1926), however (Matton et al., 1994) showed that self-compatibility is a derived state and self-incompatibility is the ancestral state. Also, the fact that *F. vesca* is self-compatible (rather than the fact that it may be the oldest of the species) may be the reason that it has a very wide range. *F. vesca* and *F. bucharica* (formerly *F. nubicola*) are among the diploids most closely related to the polyploids *F. orientalis, F. moschata, F. chiloensis,* and *F. virginiana* based on ITS and chloroplast sequence data (Potter et al., 2000). More recently, Rousseau-Gueutin et al. (2009) found that *F. inumae, F. vesca, F. bucharica,* and *F. mandshurica* were the diploids most closely related to the octoploids *F. chiloensis* and *F. virginiana* and the decaploid *F. iturupensis,* based on sequence data from the nuclear protein-
encoding genes DHAR (dehydro ascorbate reductase) and GBSSI-2 (granule-bound starch synthase, also known as "waxy").

Hancock (1999) also related cytological evidence that F. vesca has common ancestry with the commonly cultivated strawberry F. ×ananassa. Since the occurrence of unreduced gametes in Fragaria is somewhat common in natural populations, the polyploidization of Fragaria is thought to have involved unreduced gametes (Brinthurst and Senanayake, 1966; Hancock, 1999) which are functional in Fragaria as either eggs or pollen (Scott, 1951). Repeated fusions of unreduced gametes probably led to the creation of the species of Fragaria with higher ploidies, as was demonstrated by Ichijima (1926) when the crossing of two Fragaria diploids yielded a tetraploid progeny. Senanayake and Brinthurst (1967) dubbed F. vesca as having the genome type AA (although in Fedorova’s model [1946], F. vesca was CC). They suggested that an interspecific diploid hybrid such as AA’ may have arisen and undergone spontaneous somatic cell doubling to give rise to a tetraploid AAA’A’. Following this logic, an evolutionary story of polyploid Fragaria that does not involve unreduced gametes and instead depends on interspecific hybridization and subsequent somatic cell doubling from the starting point of two diploids, one being similar to modern F. vesca, is conceivable (Brinthurst and Gill, 1970). Cytological studies have prompted the claim that F. chiloensis, F. virginiana, and F. ×ananassa all have the same genomic constitution (Senanayake and Brinthurst, 1967; Brinthurst and Gill, 1970); however, this claim has been questioned by Davis et al. (2009)
History of *Fragaria xananassa*

The commonly cultivated strawberry, *F. xananassa*, is an octoploid with 2n=56 chromosomes, and was derived from the accidental cross-pollination of female *F. chiloensis* plants by hermaphroditic *F. virginiana* plants (both octoploids as well), perhaps in a botanical garden in Brest, France around 1750 (Hancock and Luby, 1993). Since then, crosses have been made using the original germplasm and also by crossing new germplasm from wild representatives of *F. chiloensis* and *F. virginiana*. Over the years since the first *F. xananassa* was created, many breeders have contributed to improving the germplasm. Most notably, cultivars that were winter-hardy and everbearing were produced by C. L. Powers and A. C. Hildreth (Hancock, 1999). Everbearing cultivars and cultivars with larger fruit were also produced by R. S. Bringhurst and V. Voth (Hancock and Luby, 1993). The germplasm used until the 1970’s for making crosses in order to explore the possibility of uncovering novel, beneficial genotypes in *F. xananassa* cultivars was said to be “of concern due to the possibility of detrimental inbreeding effects and a lack of diversity to face new environmental challenges” (Hancock and Luby, 1993). Since then, work has been done to characterize natural populations of *F. chiloensis*, and *F. virginiana*. Exploring new sources of germplasm has already proven beneficial, since these natural populations were found to have traits such as resistance to pests and disease, as well as higher photosynthetic rates, more efficient water usage, foliage hardiness, crown regrowth after cold exposure, drought tolerance, and day-neutral flowering (Hancock, 1999). Additionally, the natural populations had high variability in many other traits such as leaf, fruit, flower, and runner morphology, flavor, aroma, bloom period, and winter
hardiness. Of comparable importance was the diversity of environments these cultivars inhabited, varying in elevation, soil type, forest or field cohabitants, water availability, and climate. Many of the genotypes previously hidden in small, natural populations could be important to the future of the cultivated strawberry, if properly incorporated through “intensive breeding efforts” (Hancock and Luby, 1993).

**Fragaria Octoploid Models**

There are three published models for explaining the *Fragaria* octoploids’ (*F. virginiana*, *F. chiloensis*, and their hybrid *F. xananassa*) genomic composition: Model 1 - AABBBBCC (Fedorova, 1946 [Fedorova’s model was actually AAAABBC, but because it has been incorrectly printed in many papers it will be represented as it was cited by Senanayake and Bringhurst, 1967]); Model 2 - AAA’A’BBBB (Senanayake and Bringhurst, 1967); and Model 3 – AAA’A’BBB’B’ (Bringhurst, 1990). In these models, each letter designates one set of seven chromosomes. These models also imply that there are three (Models 1 and 2) or four (Model 3) ancestral diploid species (Model 1: AA, BB, CC; Model 2: AA, A’A’, BB; Model 3: AA, A’A’, BB, B’B’), each contributing two or four of the eight chromosome sets in the octoploid (see Figure 1). In the latter two models, the A and A’ genomes are relatively similar to each other, but are distinct from the correspondingly similar B and B’ genomes, and chromosomes from the A/A’ genomes are not expected to pair with chromosomes from the B/B’ genomes in meiosis. The second model, AAA’A’BBBB suggests that alleles on the A and A’ chromosomes are inherited in a disomic pattern, while alleles on the chromosomes in the four B genomes are inherited in a tetrasomic pattern (Model 2 is identical to Model 1 in terms of
inheritance patterns). The third model, AAA’A’BBB’B’ suggests only disomic inheritance (See Figure 2).

Figure 1. Octoploid Fragaria evolution from diploid progenitors. The left side shows the hypothetical evolution of a fully disomic octoploid Fragaria from four diploid progenitors (genomes represented by AA, A’A’, BB, B’B’), with two allotetraploid intermediates (genomes represented by AAA’A’ and BBB’B’). The right side shows the hypothetical evolution of a partially polysomic octoploid Fragaria from three diploid progenitors (genomes represented by AA, A’A’, BB), with one allotetraploid and one autotetraploid intermediate (genomes represented by AAA’A’ and BBBB, respectively).
Figure 2. Two published models for the genome composition of octoploid *Fragaria*; AAA'A'BBB'B' and AAA'A'BBBB. Part i shows the 56 chromosomes that would be present in an octoploid *Fragaria* nucleus, organized and color coded by subgenome type (A, A', B, or B'). Part ii draws attention to the subgenomes that differentiate the two models; two B and two B' subgenomes (left) versus four B subgenomes (right).
**Disomic vs. Polysomic Inheritance**

**Explanation of Terms**

The distinction between disomic and polysomic inheritance patterns is of great importance in studies concerning *F. ×ananassa* and other octoploid strawberry species because this distinction is the point of conflict between the two main genomic models. Other terms that commonly go hand-in-hand with the terms disomic and polysomic are allopolyploidy and autopolyploidy, respectively. Stebbins (1950) defined allopolyploidy as the condition of an organism having the chromosome combination of two or more distinct genomes and autopolyploidy as the condition of having either a doubled chromosome constitution or the chromosome combination of two or more genetically similar genomes. In autopolyploids, there are more than two homologous chromosomes, meaning chromosomes that are similar enough to pair with each other in meiosis. In allopolyploids, there are homoeologous chromosomes in addition to homologous chromosomes. Homoeologous chromosomes are identifiable as the same chromosome based on gene identities and synteny, but they are different enough from each other that they will not pair readily during meiosis. Instead, homologous chromosomes have pairing preference.

Stebbins stresses that the distinction between autopolyploids and allopolyploids is strictly the origin of their genomes, and not the subsequent pairing behavior of the chromosomes. However, the criterion by which the subgenomes that make up the polyploid constitution are classified as similar or distinct is simply whether they pair together in meiosis or not (Stebbins, 1950). Sybenga (1999) suggested that chains of proteins extend out from chromosomes during meiosis, and upon meeting with a
homologous chain, the chromosomes are pulled together to synapse. This leads to the idea that homoeologous chromosomes may be able to pair in a hybrid, but if a homologous chromosome is present, the homologues will pair preferentially. In the case of a hypothetical AAA'ABBB'B octoploid, chromosomes from an A genome may theoretically be able to pair with chromosomes from an A' genome, but with another set of A-genome chromosomes present, the homologues will pair rather than the homoeologues. This idea lent itself to the first revision of the octoploid Fragaria genome model from AABBBBCC to AAA'A'BBBB which was based on the pairing behavior of chromosomes in pentaploid and hexaploid hybrids of octoploids crossed with F. vesca 'colchiploids' (a diploid that has been treated with a metaphase arresting agent, causing it to become an extreme autopolyploid). Chromosomes that no longer had a homologue were found to be pairing with what was previously a homoeologue, prompting the claim that Fedorova's AA and CC genomes were closely related, and that the CC genome should be relabeled as A'A' to reflect this (Senanayake and Bringhurst, 1967).

Allopolyploidy by definition leads to disomic inheritance because in the hypothetical allotetraploid AABB, the homologous chromosomes of the two A genomes will pair with and disjoin from each other, as will those of the two B genomes, causing all gametes to have the genomic constitution of AB. Autopolyploidy usually leads to polysomic inheritance because in the hypothetical autotetraploid AAAA, there is no preferential or exclusive bivalent pairing, and therefore homologues can pair with and disjoin from each other in various possible combinations. At a given locus, if the alleles in the hypothetical autotetraploid are numbered as A_1A_2A_3A_4, the resulting gametes could have any combination of two alleles: A_1A_2, A_1A_3, A_1A_4, A_2A_3, A_3A_4, A_2A_4 (assuming
only bivalent pairing; the effects of multivalent formation are discussed in the following paragraph). In contrast, in an allotetraploid, if the alleles of a locus are numbered as A_1A_2B_1B_2, the resulting gametes could have only four combinations of two alleles: A_1B_1, A_1B_2, A_2B_1, A_2B_2 (see Figure 3).

**Significance of Multivalents**

In many cases, autopolyploidy allows for the formation of multivalents in meiosis, where a multivalent is a group of more than two synapsed chromosomes. The presence of multivalents at meiosis is directly associated to the condition of autopolyploidy and therefore polysomic inheritance. Although some authors seem to equate polysomic inheritance with the necessity of multivalent formation (Pairon and Jacquemart, 2005), the absence of multivalents at meiosis does not necessarily reject the presence of autopolyploidy or polysomic inheritance. Even in an autotetraploid that was created by the somatic doubling of its chromosomes, mechanisms may exist which give preference to bivalent pairing or restrict multivalent formation so that only bivalents are visible in meiosis (Sybenga, 1996; Lerceteau-Kohler et al., 2003). Therefore, cytological studies of polyploids cannot conclude the mode of inheritance based solely on the finding of exclusive bivalent pairing. The above examples of auto- and allotetraploidy do not take into account the other source of imbalanced chromosome transmission that is caused by multivalent formation: aneuploidy. In an autotetraploid, if three chromosomes synapse into a multivalent, the remaining lone chromosome will line up on the metaphase plate alone, likely resulting in aneuploid gametes (Sorrells, 1992).

In the case of octoploids, the distinction between disomic and polysomic inheritance gains complexity. In model 2, AAA’A’BBB’B’, the implications are fairly
straightforward: fully disomic inheritance in all chromosomes, at all loci; exclusively bivalent pairing; and the species is an allo-octoploid. In model 1, AAA’A’BBBB, the implications are that half of the chromosomes have disomic inheritance and half have polysomic inheritance; at least half of the chromosomes exhibit only bivalent pairing; and the species is an auto-allo-octoploid. Fjellstrom et al. (2001) found evidence of both tetrasomic inheritance and preferential pairing in the segmental allopolyploid Lotus corniculatus. Welch (1962) reported that different chromosomes in a polyploid can vary in their tendencies to form multivalents, and even different loci on a chromosome can vary in the frequency at which they undergo double reduction. Wu et al. (2001) described a mathematical method for determining the preferential pairing factor (related to disomic inheritance) and the frequency of double reductions (related to multivalent formation and polysomic inheritance), which only together can help to characterize the gene segregation patterns in polyploids.

**CAPS, Isozyme, and other Genetic Evidence**

Many researchers have used genetic methods to investigate the question of disomic vs. polysomic inheritance in Fragaria. However, genetic approaches to transmission genetics questions face unique difficulties in polyploid systems. One difficulty is high heterozygosity (never or rarely observing homozygosity across eight or more alleles) which may prevent genetic markers from “revealing co-dominance” (Kunihisa et al., 2003). Sorrels (1992) remarked that RFLP (restriction fragment length polymorphism) mapping in polyploids is complicated by “a large number of segregating genotypes, co-migration of fragments, poorly characterized genome constitution and/or chromosome pairing behavior, and complicated genotype characterization due to multiple
fragments”. Lerceteau-Kohler et al. (2003) used AFLP (amplified fragment length polymorphism) mapping on F. ×ananassa and found mixed disomic and polysomic inheritance. They concluded that F. ×ananassa is neither fully polyploidized nor fully diploidized and supported a model such as AAA’A’BBBB. Their evidence was based upon the ratio of markers linked in coupling phase to those linked in repulsion phase that did not fit the expected ratios for fully disomic inheritance (based on the calculation tool developed by Wu et al., 1992). However, the alternative hypothesis of segregation distortion was not considered (Folta and Davis, 2006). Kunihisa et al. (2003 and 2005) studied the segregation of CAPS (cleaved amplified polymorphic sequence) markers in F. ×ananassa and found them to be transmitted in a manner that is consistent with disomic Mendelian inheritance. Arulsekar et al. (1981) analyzed the inheritance of PGI (phosphoglucoisomerase) and LAP (leucine amino peptidase) isozymes in F. ×ananassa, and asserted that the genome was highly diploidized. However, Gastony (1991) showed that in polyploid homosporous fern gene silencing can make polyploids look like diploids in terms of isozyme identity, so the F. ×ananassa genome appearing highly diploidized in terms of isozymes may or may not mean that it’s highly diploidized in terms of meiotic pairing.

Haymes et al. (1997) noted that the resistance gene Rpf1 (which confers resistance to red stele root rot in Fragaria) and seven linked RAPD (randomly amplified polymorphic DNA) markers co-segregated in a manner indicative of disomic inheritance. Ashley et al. (2003) observed microsatellite loci segregating in a disomic manner in both the wild and cultivated Fragaria octoploids, and concluded that the cultivated strawberry genome is highly diploidized. Rousseau-Gueutin et al. (2008) performed comparative
genetic mapping between octoploid and diploid *Fragaria* and found high degrees of colinearity and macrosynteny between the genomes. They noted that there may be some residual polysomic inheritance, but their data suggested that the inheritance in *F. ×ananassa* is mainly disomic.

**Meiotic Pairing Evidence**

Several studies reported that only bivalents were observed in cultivars of *F. ×ananassa* (Ichijima, 1926; Longley, 1926; Powers, 1944), suggesting disomic inheritance and supporting the AAA’A’BBB’B’ genome composition model for the octoploids. However, as mentioned before, while the presence of multivalents indicates polysomic inheritance, the absence of them does not necessarily indicate disomic inheritance. Alternately, Mok and Evans (1971) observed that more than half of the pollen mother cells of nine cultivars of *F. ×ananassa* had chromosomes associating as multivalents. They observed that an average of six chromosomes per cell, and a range of zero to sixteen chromosomes per cell, were associated as multivalents.

**Intergenic Regions in Fragaria**

Many of the intergenic regions in *Fragaria* are small (Davis et al., 2008), in the range of 1 -5 kb in length, which is beneficial for the prospect of studying the divergence of intergenic regions. Different varieties of *Fragaria* are expected to have similar coding regions, but the less highly-conserved intergenic regions have had opportunity to accumulate sequence divergence, which is the basis of the present study. When an intergenic region of interest contains several different polymorphic sites, gene-pair haplotypes can be defined by working out the complete set of character states that encompasses all of the polymorphic sites in each of the different sequences of an
intergenic region (Davis et al., 2008). The particular intergenic region of interest here is that between the genes gRGA1 and Subtilase. RGA stands for resistance gene analog, which means its sequence is similar to disease resistance genes in other dicot species. Resistance genes generally confer resistance to bacterial, fungal, or viral plant pathogens (Collins et al., 1998). The Subtilase gene encodes a serine protease of the family Subtilisin, and is involved in prepro-protein processing (Janzik et al., 2000). In F. vesca ssp. americana var. Pawtuckaway the intergenic region of gRGA1-Subtilase is short (578 bp; Shields and Davis, unpublished) and therefore ideal for PCR amplification and sequencing. This makes the gRGA1-Subtilase intergenic region an ideal gene-pair marker sequence.

Gene-pair markers are a good tool for marker-assisted selection in plants that have small genomes (Davis et al., 2008). The idea of using intergenic regions to study polymorphisms is essentially the same as CAPS (cleaved amplified polymorphic sequences) except that the focus is on intergenic regions rather than introns, which are typically of shorter sequence. However, to amplify an intron, sequence information about only one gene is required, while for gene-pair markers sequence information about two adjacent genes is required for primer construction (Davis et al., 2008). The methods that follow employ the use of fluorophores to attach to the end of restriction enzyme-cleaved PCR products, and are hence referred to as FEL-CAPS (fluorescent end-labeled cleaved amplified polymorphic sequence).

Scope of Research

The objectives of this project were to: 1. Cross a potentially heterozygous octoploid Fragaria with a diploid Fragaria to create a segregating pentaploid population;
2. Sequence all alleles of the gene pair locus gRGA1-Subtilase in the octoploid Fragaria;
3. Genotype the members of the pentaploid population utilizing FEL-CAPS; 4. Discern whether the segregation of the octoploid alleles fits either of the published models.

Recently, a novel system of determining the composition of the octoploid strawberry genome became available in the form of a population of F₁ hybrid pentaploid plants (Davis et al., 2009). These plants were created by crossbreeding an F. vesca diploid, as female, with an F. virginiana (wild strawberry) octoploid hybrid. The genome composition of F. vesca can be represented as AA. Therefore, the egg cell donated by this parent would have one copy of the A genome in it. The genome composition of the pollen from the octoploid parent depends on which (if either) of the proposed genome models is correct: An AAA’A’BBBB parent would yield AA’BB pollen, but an AAA’A’BBB’B’ parent would yield AA’BB’ pollen. The resulting pentaploid progeny would have either AAA’BB or AAA’BB’ genomic compositions, with the fifth-listed genome’s identity of either B or B’ providing the evidence to accept one proposed model of octoploid genome composition and reject the other. The identities of these genomes could be determined by identifying the presence of genome-specific alleles that had been previously identified as such via comparisons to reference diploid species by other researchers (Zhang and Davis, unpublished).

After sequencing and cataloguing the gRGA1-Subtilase alleles of the octoploid parent of a hybrid pentaploid F. virginiana population, these alleles and their segregation pattern needed to be identified in each pentaploid plant. The FEL-CAPS method employed for genotyping was expected to identify which alleles were present in each pentaploid, assuming that the pollen donor (F. virginiana) had eight alleles with
sufficient polymorphism for this method. To test whether the pattern of inheritance from the octoploid parent is fully disomic or disomic and tetrasomic, the B (and B') alleles, once identified in the octoploid parent, could be tallied up in each pentaploid progeny plant. Gametes from an octoploid with fully disomic inheritance would have four possible allele combinations for any given locus. Gametes from an octoploid with two pairs of disomically inherited genomes and one set of tetrasomically inherited genomes would have six possible allele combinations for any given locus (See Figure 3).
Figure 3. Gametes produced by AAA'A'BBB'B' and AAA'A'BBBB octoploids. Part i shows the 56 chromosomes that would be present in an octoploid *Fragaria* nucleus, organized and color coded by subgenome type (A, A', B, or B'). Part ii draws attention to the subgenomes that differentiate the two models; two B and two B' subgenomes versus four B subgenomes. The B (and B') genomes are numbered from 1 to 4 to differentiate these four subgenomes. Part iii illustrates the four versus six combinations of alleles (numbered 1,2,3,4) at a locus on chromosome 1 of the B/B' genomes that would be found in gametes produced by, respectively, octoploids of AAA'A'BBB'B' versus AAA'A'BBBB constitutions.
CHAPTER II

MATERIALS & METHODS

Breeding History and Pentaploid Production

A population of octoploid F1 hybrids from the cross *F. virginiana* ssp. *virginiana* (known as “L1”, collected in Lincoln, NH) x *F. virginiana* ssp. *glauc*a (known as “BC6”, collected in British Columbia) had been previously cultivated (T.M. Davis, unpublished). This F1 population was named “LB” after the two parent plants, and was expected to be highly heterozygous because of the evolutionary divergence implied by the subspecies distinction and the geographical distance between the locations in which the parental plants had been collected. One of the F1 plants, LB48, was used to create pentaploids by crossing it as male with inbred *F. vesca* diploid varieties “Yellow Wonder” (a.k.a. YW,) or “Hawaii” a.k.a. H) (see Figure 4).
Figure 4. Genomic ancestry of an *F. virginiana* *F. vesca* pentaploid population. This illustration diagrams the predicted chromosomal constitution of the pentaploid population and its parents, incorporating the fully disomic octoploid model (AAA’A’BBB’B’) and the partially polysomic octoploid model (AAA’A’BBBB). The different colors of the chromosomes represent the different genomes they belong to.
To perform the cross, the petals of young flower buds of the diploid *F. vesca* were removed and the buds were emasculated. Flowers with mature anthers were removed from LB48 and rubbed on the “naked” *F. vesca* (YW or H) buds. These buds were tagged and the *F. vesca* plants were separated from other *Fragaria* plants to avoid possible outcrossing.

**Pentaploid Seed Sterilization and Planting**

Any fruit that grew from the pollinated *F. vesca* var. Yellow Wonder or Hawaii buds were collected when ripe, and the seeds were removed, air-dried and surface-sterilized. Seeds were washed in a 1.5 mL tube with tap water and a few drops of Tween-20. The tubes were shaken and briefly vortexed. The soapy water was pipetted off of the seeds, and a 15% bleach solution in sterile deionized water was added. The tubes were agitated constantly, but gently for 10 minutes at room temperature. The bleach solution was replaced with sterile deionized water, and then the water was changed 2-3 more times, with brief shaking each time. The seeds were mixed again for 20 minutes in sterile deionized water, with another water change after 10 minutes. The seeds were then spread onto the surface of sterilized pots of soil. The pots were watered with sterile deionized water and then sprinkled with sphagnum moss. The pots were kept warm (about 25°C) and under constant fluorescent light. Plants were grown and maintained in the University of New Hampshire greenhouse facilities.

**Confirmation of Ploidy**

**Chromosome Squash Slides**

Runners from the pentaploid plants were pinned down on potting soil to stimulate rooting, and the distal 1 cm root tips were harvested after 7-10 days. Root tips were
pretreated in ice water for 12h and then fixed in 3 parts 95% ethanol: 1 part glacial acetic acid for at least 24 hours. Fixed root tips were transferred to 70% ethanol within a week of fixation. For making a slide, root tips were first treated in 1N HCl at 60°C for 2-5 minutes and then transferred to ice-cold water for at least 1 minute. A root tip was placed in a drop of aceto-orcein or aceto-carmine stain on a microscope slide, then the proximal portion was cut off, leaving the distal 1 mm root tip to soak in the stain for at least 2 minutes. A cover slip was added over the root tip and stain and the slide was briefly heated over an alcohol lamp. The cover slip was tapped with a dissecting needle or eraser end of a pencil until the tissue spread out. Then, the tissue was squashed between the cover slip and slide by pressing down on the cover slip for a few seconds. The slides were viewed on a compound microscope.

**Flow-Sorting Outsourcing**

Due to difficulties in getting accurate counts because of the small size of the *Fragaria* chromosomes, a second approach was taken to confirm the ploidy of the putative pentaploids. Tissue samples consisting of 2-3 unexpanded leaf bundles from each plant were packaged in a small resealable plastic bag containing a dampened paper towel and sent to Plant Cytometry Services (Schijndel, The Netherlands; http://www.plantcytometry.nl/inhoud/taalkeuze.htm) for flow-sorting. In addition to sampling 41 pentaploids, samples from several other *Fragaria* species of various known ploidies were submitted to calibrate the data: diploids *F. vesca* (varieties Hawaii, Yellow Wonder, BS2X, Pawtuckaway) and *F. viridis*; a tetraploid *F. vesca* (accession BS4X); a reportedly pentaploid hybrid of *F. vesca* and *F. chiloensis* (accession CFRA 373); a hexaploid *F. moschata* (accession CFRA 376); octoploids *F. virginiana* ssp. glauca
(accession BC6), *F. virginiana* ssp. *virginiana* (accession L1), *F. virginiana* hybrid (accession LB48), *F. iinumae* (accession J7), and *F. chiloensis* (accession 743.001); and a decaploid (CFRA 374) that was derived from CFRA 373 (described above).

**Genomic DNA Isolation**

Genomic DNA was isolated from the following: the parental octoploid *F. virginiana* plants, “L1” and “BC6”; hybrid LB48; *F. vesca* vars. Yellow Wonder and Hawaii, and each member of the pentaploid population, named in serial “5XA”, “5XB” and so on. Several young, unexpanded leaves were collected from each of the plants and were stored at 4 °C in sterile microfuge tubes until isolation. DNA was isolated from each plant separately. The leaves were placed into a mortar and enough liquid nitrogen was added to cover the leaves. A pestle was used to gently grind the leaves into a powder before the nitrogen boiled away. Without allowing the powder to thaw, 1 ml grinding buffer (2% CTAB buffer [0.1 M Tris-HCl, pH 8; 20 mM Na-EDTA; 1.4 M NaCl; 2% (w/v) CTAB (Hexadecyltrimethylammonium bromide)] with the addition of 4 µl β-mercaptoethanol) was added to the mortar while continuing to grind the plant material. When the material became a slurry, it was transferred to a new sterile microfuge tube and incubated at 60 °C for one hour. The tube was allowed to cool at room temperature for 10 minutes and then 24:1 chloroform: octanol was added to fill the tube (500-800 µl). The tube was vortexed until the contents were uniformly mixed and then centrifuged at 14,000 x g for 5 minutes to separate the phases. The upper, aqueous phase (500-800 µl) was transferred to a new sterile microfuge tube and to it was added 95% ethanol to fill the tube. The tube was held on ice or in the freezer until a milky precipitate appeared (5-10 minutes) and the tube was gently inverted several times before being centrifuged at
14,000 x g for 5 minutes. The supernatant was decanted and 1 ml cold 70% ethanol was added. The tube was held on ice for 10 minutes and centrifuged again at 14,000 x g for 5 minutes. The supernatant was decanted again and the residual ethanol was removed by pipette. The DNA pellets were dried for 2-5 minutes in a vacuum centrifuge. Then, the pellet was resuspended in 50 µl TE (10 mM Tris-HCl, pH 8.0; 1 mM Na-EDTA) at 4 °C overnight. The following day, 50 µl RNase solution (10 µg/ml RNase in sterile deionized water) was added to the tube and mixed in by gently pipetting the liquid up and down 10 times with a pipette tip that had an enlarged opening. The tube was incubated at 37 °C for one hour, and the DNA in solution was stored at 4 °C thereafter.

**Genomic PCR**

**Intergenic Region PCR**

The genomic DNA from L1, BC6, LB48, “Yellow Wonder”, and “Hawaii” was amplified via PCR in order to subsequently locate polymorphisms in a specific intergenic region known as “gRGA1-Subtilase” (the region between the gRGA1 gene and the Subtilase gene; see Figure 5). Each PCR volume contained 1x EconoTaq® (Lucigen, Middleton, WI) PCR buffer containing Mg++, 1.25 units EconoTaq DNA Polymerase, 0.8 mM dNTP's (0.2 mM each of dATP, dGTP, dCTP, and dTTP), 0.4 µM forward primer 19M24F (5’ AAGGCGTTGATCGTTTTCTGTTG), 0.4 µM reverse primer 19M24R (5’ CCCGAGTTTGGAGGCTTTGTTGC), 10 ng genomic DNA, and sterile deionized water to a final volume of 25 µL. PCR was carried out in a Mastercycler® ep Gradient (Eppendorf, Hamburg, Germany) thermocycler. PCR was initiated with a 10 minute denaturation at 94 °C which was followed by forty cycles of denaturation at 94 °C for 30 seconds, annealing at 60 °C for 30 seconds, and elongation at 72 °C for one
minute, and then a final elongation step at 72 °C for 10 minutes. The PCR products from LB48 were sequenced, methods follow.

![Figure 5. PCR-amplified gRGA1-Subtilase intergenic region.](image)

This diagram illustrates the adjacent gRGA1 and Subtilase genes. The gray region between them is the gRGA1-Subtilase intergenic region. The black bar represents the location of the PCR-amplified region.

**Expanded Region PCR**

New primers were designed using fosmid sequence data from *F. vesca* var. Pawtuckaway. The new primers were designed (with the aid of Lasergene® PrimerSelect software © 2010 DNASTAR, Inc., Madison, WI) to anneal further outside of the intergenic region of gRGA1-Subtilase so that the forward primer region sits at the end of the coding region of gRGA1 and the reverse primer region sits at the beginning of the Subtilase gene, based on the “Pawtuckaway” sequence. The new primers were seated outside the original primer annealing sites in order to amplify and subsequently sequence those original primer sites; and deeper into the coding regions (see Figure 6) in order to maximize primer region sequence identity between the sequence data used (*F. vesca* var. Pawtuckaway) and the intended primer target sequence (*F. vesca* vars. Yellow Wonder and Hawaii). The expected product length, based on the *F. vesca* var. Pawtuckaway fosmid sequence, was 1,992 bp. A new protocol was used which employed AccuPrime™ Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA). In addition to greatly improving primer-template hybridization, this Taq DNA polymerase includes an enzyme that possesses a 3’-5’ exonuclease activity and can therefore “proofread” the primers so
they anneal to the target more accurately. Each PCR volume contained 1x AccuPrime™ PCR Buffer II (600 mM Tris-SO₄ [pH 8.9], 180 mM [NH₄]₂SO₄, 20 mM MgSO₄, 2 mM dGTP, 2 mM dATP, 2 mM dTTP, 2 mM dCTP, thermostable AccuPrime™ protein, 10% glycerol), 0.2 µM forward primer 19M24-20862F (5’ CAGTTAACTGGGCGGCTGAC), 0.2 µM reverse primer 19M24-22853 (5’ CATTGGATCATGGAGCTGGACACA), 0.2 µL (1.0 unit) AccuPrime™ Taq DNA Polymerase High Fidelity, and 50 ng genomic template DNA. PCR was carried out in an Eppendorf thermocycler. PCR was initiated with a 2 minute denaturation at 94 °C which was followed by forty cycles of denaturation at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds, and elongation at 68 °C for two minutes, and then a final elongation step at 68 °C for 5 minutes.

Figure 6. Expanded PCR-amplified gRGA1-Subtilase intergenic region. This diagram illustrates the adjacent gRGA1 and Subtilase genes. The gray region between them is the gRGA1-Subtilase intergenic region. The black bar represents the location of the original PCR-amplified region, while the yellow bar represents the location of the expanded PCR-amplified region.

**Gel Electrophoresis**

Gel electrophoresis was carried out typically using 2% (w/v) standard agarose in 1x TBE (0.089 M Tris base, 0.09 M boric acid, 2 mM Na-EDTA) with 1x TBE as a running buffer. To each 10-20 µl sample to be electrophoresed was typically added 1-2 µl of loading dye (0.25% [w/v] bromophenol blue, 0.25% [w/v] xylene cyanol, in sterile 30% glycerol). Gels were electrophoresed at approximately 5 volts per centimeter between electrodes. After electrophoresis, the gels were stained for 20 minutes in 10
µg/ml ethidium bromide and then destained for 20 minutes in water. The gels were viewed on a FOTO/Analyst® Luminary/FX® gel imaging system (FOTODYNE, inc., Hartland, WI), and the photos were captured with FOTO/Analyst® PC Image version 10.10 © 2008 FOTODYNE®, inc.

**LB Bacterial Medium Plates**

LB plates were made by making a solution of 1% bacto-tryptone, 0.5% yeast extract, 1% NaCl, and 1.5% agar, pH 7.0. The solution was autoclaved at 221 °C for 23 minutes, and then cooled to 55°C before adding Ampicillin to a final volume of 25 µg/mL. The mixture was poured into sterile plastic plates and cooled to solidify. If X-GAL was to be added, the plates were warmed to 37°C and then 40 µL of 40 mg/mL X-GAL was added and spread evenly with a sterilized glass wand. The X-GAL on the plates was allowed to dry before *E. coli* cells were added to the plates.

**Molecular Cloning and Colony PCR**

*gRGA1-Subtilase* PCR products from LB48 (amplified using primers 19M24F & R and EconoTaq Taq DNA polymerase) were cloned using the TOPO TA Cloning kit from Invitrogen Life technologies. If PCR products were not fresh (less than 24 hours old), they were readenylated by adding 1 unit of Taq polymerase to a 25 µL PCR reaction volume and incubating the tubes at 72°C for 10 minutes, and then holding on ice. For each PCR product to be cloned, the following cloning reaction mixture was made using a TOPO® TA cloning Kit (Invitrogen): 4 µL fresh or readenylated PCR product, 1 µL salt solution, 0.25 µL pCR® 2.1 -TOPO vector (see Figure 7), and 0.75 µL sterile deionized water. The volume was mixed gently and centrifuged briefly before being incubated at room temperature for 30 minutes. A tube of One Shot® TOP10 Chemically Competent
E. coli (Invitrogen) was taken out of the -70 °C freezer and thawed on ice for 8 minutes. Once thawed, 2 µL of the cloning reaction mixture was added to the E. coli, and the cells were stirred gently with the pipette tip. The cells were then heat shocked for 30 seconds at 42 °C, placed on ice for 1 minute, and 250 µL SOC medium was added. The tubes were shaken horizontally at 190-210 rpm for 1 hour at 37 °C before being spread on prewarmed (37 °C) LB plates containing 25 µg/mL ampicillin and coated with 1.6 g X-GAL. The plates were incubated at 37 °C overnight, and then a number of white colonies were selected from each plate and transferred onto a gridded LB plate without X-GAL. The gridded plates were incubated at 37 °C overnight, and then stored at 4 °C until the clones were analyzed.

**Figure 7.** pCR® 2.1-TOPO vector. This figure diagrams the sequence layout of the plasmid vector that was used to clone the gRGA1-Subtilase PCR product from LB48.
Colonies were analyzed by PCR to make sure that the plasmids in the cells contained the intended PCR product inserts. A sterile toothpick was used to transfer a small amount of cells from a colony onto a gridded plate into a tube containing 1X MasterTaq® (Eppendorf) Taq buffer (50 mM KCl, 10 mM Tris-HCl [pH 8.3 at 25 °C], 1.5 mM Mg[OAc]₂), 1X TaqMaster PCR Enhancer (Eppendorf), 1 unit Taq polymerase, 0.8 mM dNTP's and 0.4 µM each of two primers, specific to the regions of the vector flanking the insert site (M13F 5' GTAAAACGACGGCCAGT and M13R 5' CAGGAAACAGCTATGAC), and sterile deionized water, in a final volume of 20 µL. PCR was initiated by denaturation at 94 °C for 10 minutes, followed by forty cycles of denaturation at 94 °C for 30 seconds, annealing at 60 °C for 30 seconds, and elongation at 72 °C for 1 minute, and finished by a final elongation at 72 °C for 10 minutes.

**Plasmid Isolation and PCR**

Samples of colonies that were to have their inserts sequenced were transferred by sterile pipette tips into tubes containing 2 mL LB broth and 100 µg Ampicillin. The tubes were incubated at 37 °C overnight on an orbital shaker set to 200 rpm. Plasmid DNA was isolated according to the directions from Promega’s Wizard® Plus SV Minipreps DNA Purification System (Madison, WI). The overnight culture was centrifuged at 10,000 x g for 5 minutes and the supernatant was decanted. 250 µl of Cell Resuspension Solution (50 mM Tris-HCl [pH 7.5], 10 mM EDTA, 100 µg/ml RNase A) was added, and the tubes were inverted several times to thoroughly resuspend the cells. 250 µl of Cell Lysis Solution (0.2 M NaOH, 1% SDS) was added and the tubes were inverted four times to mix. The tubes were incubated at room temperature for approximately five minutes or until partial clearing of the lysate was observed. 10 µl of Alkaline Protease Solution was
added and the tubes were inverted four times to mix before incubating at room
temperature for five minutes. After cell lysis, 350 µl of Neutralization Solution (4.09 M
guanidne hydrochloride, 0.759 M potassium acetate, 2.12 M glacial acetic acid; pH 4.2)
was added and the tubes were inverted four times before centrifuging the tubes for 10
minutes at 14,000 x g. The cleared lysate was decanted into a spin column. The
supernatant was centrifuged at 14,000 x g for one minute and the flowthrough was
discarded. The spin column was moved to a clean collection tube. Column Wash
Solution (750 µl) was added to the column, and then the spin column was centrifuged for
one minute at 14,000 x g. The flowthrough was discarded and 250 µl of Column Wash
Solution was added and the spin column was centrifuged for two minutes at 14,000 x g.
The spin column was transferred to a new, sterile microcentrifuge tube. The plasmid
DNA was eluted by adding 100 µl of nuclease-free water to the spin column. The
nuclide-free water was allowed to sit in the spin column for five minutes before
centrifuging for one minute at 14,000 x g. The flowthrough containing the plasmid was
stored in a -20 °C freezer.

The plasmid DNA was checked for content by PCR. A sterile pipette tip was
dipped in the tube containing the plasmid DNA, and then stirred into a tube containing 1x
MasterTaq® (Eppendorf) Taq buffer (50 mM KCl, 10 mM Tris-HCl [pH 8.3 at 25 °C],
1.5 mM Mg[OAc]₂), 1x TaqMaster PCR Enhancer (Eppendorf), 1 unit Taq polymerase,
0.8 mM dNTP's and 0.4 µM each of two primers, specific to the regions of the vector
flanking the insert site (M13F and M13R, sequences given previously), and sterile
deionized water, in a final volume of 25 µL. PCR was initiated by denaturation at 94 °C
for 2 minutes, followed by forty cycles of denaturation at 94°C for 30 seconds, annealing
at 60°C for 30 seconds, and elongation at 72°C for 1 minute, and finished by a final elongation at 72°C for 10 minutes.

**DNA Sequencing**

The plasmid DNA isolated from approximately 76 of the LB48 *gRGA1-Subtilase* clones was sequenced. In an octoploid having maximal heterozygosity, we can expect a maximum of eight alleles for any given locus, meaning that a large number of clones must be sequenced to ensure that each of the eight alleles is randomly chosen to be sequenced at least once. A mathematical formula was devised by Linyuan Li (Li and Davis, unpublished) and employed to calculate that 38 was the minimum number of clones that must be sequenced to ensure 95% confidence that all eight alleles of LB48 were sequenced (see Figure 8).

\[
\text{Prob}(d, n) = 1 - \binom{d}{1} \left( \frac{d-1}{d} \right)^n + \binom{d}{2} \left( \frac{d-2}{d} \right)^n - \binom{d}{3} \left( \frac{d-3}{d} \right)^n + \cdots + (-1)^j \binom{d}{j} \left( \frac{d-j}{d} \right)^n + \cdots + (-1)^{d-1} \binom{d}{d-1} \left( \frac{1}{d} \right)^n,
\]

where

\[
\binom{d}{j} = \frac{d!}{j! \times (d-j)!}, \quad d! = d \times (d-1) \times (d-2) \times \cdots \times 3 \times 2 \times 1.
\]

**Figure 8. Equation for how many clones must be sequenced in a polyploid.** This mathematical equation was developed by Linyuan Li, University of New Hampshire, Department of Mathematics and Statistics. It was employed to calculate how many clones, *n*, must be randomly sequenced to ensure that a maximum of *d* alleles at one locus are sequenced at least once, and the corresponding probability. *J* ranges from 1 to *d*-1, and specifies the sequential placement number of the respective term in the equation.
Each sequencing reaction was prepared by mixing approximately 150-300 ng of plasmid DNA with 0.2 µL of the 20 µM primer M13F (5' GTAAAACGACGGCCAGT). To this was added 1 µl sequencing mix, 3 µl sequencing buffer (BigDye® Terminator Cycle Sequencing Kits [v1.1 and v3.1], Applied Biosystems, Life Technologies, Carlsbad, CA), and sterile deionized water to a final volume of 10 µl. PCR was initiated by denaturation at 96°C for 1 minute, followed by 25 cycles of denaturation at 96°C for 10 seconds, annealing at 50°C for 5 seconds, and elongation at 60°C for 4 minutes, and ending with refrigeration at 4°C until ready to analyze. After clean up (according to kit directions), the DNA samples were then resolved by capillary electrophoresis on an ABI 3130 DNA Analyzer. The data was analyzed using FinchTV software, version 1.4.0, © Geospiza Inc. 2004-2006.

A simple Chi-square goodness of fit test was performed on the sequencing data to determine whether the haplotypes found were present as sequenced clones in a 1:1:2 ratio as would be expected in an AAA'A'BBBB or AABBBBCC octoploid. For 44 sequenced clones, the expected numbers of subgenome-representative haplotypes would be 11:11:22.

**Fluorescent End Labeling by PCR**

The genomic DNA isolated from the pentaploid plants was prepared for genotyping by PCR with modified primers. The primer 19M24F was modified to have a FAM fluorophore at the 5’ end, while the primer 19M24R was given a HEX fluorophore at the 5’ end. These modifications to the primers were designed to dually label each PCR product (see Figure 9). Each PCR volume contained 1x EconoTaq PCR buffer (containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM MgCl₂, and 0.1% Triton X-
100), 1.25 units EconoTaq Taq DNA polymerase, 0.8 mM dNTP’s (0.2 mM each of dATP, dGTP, dCTP, and dTTP), 0.4 µM forward primer RS(HEX)F (5’ /5HEX/ AAGGCCTGATCGTTTTCTGTTGC), 0.4 µM reverse primer RS(FAM)R (5’ /56FAM/ CCCGAGTTTGGAGGCTTGGTGT), 20 ng genomic DNA, and sterile deionized water to a final volume of 25 µL. PCR was carried out in an Eppendorf Thermocycler. PCR was initiated with a 10 minute denaturation at 94 °C which was followed by forty cycles of denaturation at 94 °C for 30 seconds, annealing at 60 °C for 30 seconds, and elongation at 72 °C for one minute, and then a final elongation step at 72 °C for 10 minutes. Ten microliters of each unpurified PCR product was removed for gel electrophoresis. The remaining PCR product volumes were purified by ethanol precipitation. To each remaining 15 µL reaction volume was added 200 µL 100% ethanol and the tubes were inverted several times, and then chilled for 30 minutes at -70 °C. The tubes were centrifuged in the cold (4 °C) at 13,000 x g for 15 minutes, the supernatant was decanted, 200 µL 70% ethanol was added, and the tubes were centrifuged again under the same conditions for 5 minutes. The final supernatant was decanted and the pellets were dried in a vacuum centrifuge or air dried in a laminar flow hood. The pellets were resuspended in 15 µL sterile deionized water overnight at 4 °C before the subsequent enzyme digestion.

**Enzyme Digestion of Fluorescent End-Labeled PCR Products**

*grGA1-Subtilase* sequence data from previous work on L1 and BC6 (Zhang and Davis, unpublished) was used as a basis to choose a restriction endonuclease with a confirmed cut site in every *grGA1-Subtilase* PCR product from the pentaploids. All of the PCR products from the hybrid LB48 or from the four *F. virginiana* chromosome sets
in the pentaploid population were expected to have been included in the LB48 sequencing data pool because of their direct inheritance from L1 and BC6. The subsequent sequencing of the gRGA1-Subtilase intergenic region in LB48 was consistent with the L1 and BC6 sequencing data. The enzyme ApaI (cut sequence 5' GGGCC^C) was chosen because it was expected to cut each PCR product only once and roughly in the middle. The sequencing data suggested that the PCR products would be approximately 675-750 bp and that the products of the enzyme digestion would be 250-475 bp. Because the PCR products used in this segment were dually labeled (with a fluorophore at each end), each enzyme digestion product was expected to be conjugated to either a HEX fluorophore if it was the forward primer end of the original labeled PCR product, or a FAM fluorophore if it was the reverse primer end of the original labeled PCR product (see Figure 9). Each enzyme digestion volume contained 1x NEBuffer 4 (New England BioLabs, Ipswich, MA), 100 µg/mL BSA, 10 units ApaI (New England BioLabs), 10 µL labeled purified PCR product (described above) and sterile deionized water to a final volume of 20 µL. The reactions were incubated at 25 °C for 60 minutes and then the enzyme was deactivated by heating to 60 °C for 10 minutes.
Figure 9. The production of FEL-CAPS (Fluorescent End-Labeled Cleaved Amplified Polymorphic Sequence). The red line represents primer sequences, while the green and blue asterisks attached to the primers represent FAM and HEX molecules, respectively. The intergenic region is amplified by PCR utilizing a FAM-labeled forward primer and a HEX-labeled reverse primer, yielding a dually labeled PCR product. This product is then cleaved into two or more fragments by enzyme digestion. Different sequences will optimally yield FAM and HEX-labeled fragments of differential sizes. FEL-CAPS are easily measurable by genotyping; however any middle fragments are unlabeled and therefore not measurable by traditional genotyping.

**Genotyping of FEL-CAPS**

The labeled enzyme digestion products were genotyped on an ABI3130 genetic analyzer using the GS500 size standard. The data was analyzed using Peak Scanner™ Software Version: 1.0, © 2006 Applied Biosystems.
CHAPTER III

RESULTS

Pentaploid Population

Chromosome Counting

In total, approximately 65 putative pentaploid progeny plants grew from the seeds of the *F. vesca* (var. Yellow Wonder or Hawaii) x LB48 crosses. A small number of those were quickly identified as maternal self-crosses based on their morphology, and several died before runnering so no clonal duplicates could be made. Several of the remaining 58 putative pentaploids were subjected to chromosome counting by the root tip squash method, and approximately 35 chromosomes were counted in each root tip cell, although the low degree of chromosome spreading made it difficult to get an indisputably accurate count (see Figure 10).
Figure 10. Chromosomes in a pentaploid root tip cell. This photograph shows a cell with approximately 35 metaphase-condensed chromosomes. This photograph was taken on an Olympia C-5050 digital camera attached to an Olympus BX40 compound microscope under the 100x oil immersion objective.

Flow Cytometry

In order to verify the ploidy level in the putative pentaploids, young leaf samples were subjected to C-value testing by flow cytometry. The pentaploid samples were accompanied by samples of several other Fragaria species of various known ploidies (2X, 4X, 6X, and 8X) in order to calibrate the data. In total, samples of 41 of the putative pentaploids were tested, and all 41 were shown to have 5X DNA content. Additionally, the F. virginiana octoploids L1, BC6, and LB48 were all verified to have 8X DNA.
content, while the *F. vesca* diploid vars. Yellow Wonder and Hawaii were verified to have 2X DNA content.

**Electrophoresis of Genomic PCR Products**

**Intergenic Region PCR Products**

PCR with the original *gRGAI-Subtilase* primers amplified products that appeared as three gel bands in the pentaploid population's pollen donor, LB48, and in the members of the pentaploid population. All of the PCR products fell between the 650 bp and 850 bp markers; however, the products were so close in size that they were difficult to clearly differentiate on agarose gels (See Figure 11). Genomic DNA samples from *F. vesca* vars. Yellow Wonder and Hawaii showed no PCR product from this protocol after repeated attempts, which means that one or both of the 19M24 primers was not homologous enough to be able to anneal to the *F. vesca* genomic DNA.
Figure 11. *gRGA-Subtilase* PCR products from LB48 and pentaploids. This is a photo of an electrophoretic agarose gel under UV light after ethidium bromide staining. Lane 1 has PCR products from LB48 amplified with unlabeled *gRGA1-Subtilase* primers; lanes 2, 4, 5, 6 have PCR products from individual pentaploids; lane 3 is a 1Kb+ molecular weight marker; and lane 7 is a PCR negative control (no template). All of the PCR products fell between the 650 bp and 850 bp markers. Some aberrant and unclear bands can be seen outside of the expected size range, which are probably the result of heteroduplexing or template jumping.

**Expanded Region PCR Products**

PCR's using the expanded region primers did not yield products when used to amplify "Yellow Wonder" DNA when standard *Taq* DNA polymerase was used, however the PCR reactions using AccuPrime™ *Taq* DNA Polymerase High Fidelity (along with the expanded region primers) yielded 1 or 2 products of the expected size of just under 2 kb. In *F. vesca* var. Pawtuckaway, gel electrophoresis showed one band, while the *F. vesca* var. Yellow Wonder showed one major band and one or more much
fainter bands, depending on the particular tube of YW genomic DNA (see Figure 12). Attempted sequencing of the YW expanded PCR products provided incomplete information that did not explain why the original 19M24 primers had failed to generate a product from YW.

Figure 12. Expanded region gRGA1-Subtilase PCR products amplified with AccuPrime™ Polymerase. This is a photo of an electrophoretic gel under UV light after ethidium bromide staining. Lanes 1 and 16 are 1Kb+ molecular weight markers; lane 3 is LB48; lanes 5 and 6 are pentaploid products; lanes 6 and 13 are F. vesca var. Pawtuckaway products; lane 7 is products from L1; lane 8 is products from BC6; lanes 9-12 are products from various samples of F. vesca var. Yellow Wonder genomic DNA.

Cloning and Electrophoresis of Colony PCR Products

A total of 72 clones of the LB48 gRGA1-Subtilase intergenic region were found to contain the anticipated product size, approximately 675-750 bp in length (see Figure 13).
Figure 13. Colony-PCR products. This is a photo of an electrophoretic gel under UV light after ethidium bromide staining. Lanes 10 and 31 are the 1Kb+ molecular weight marker. All other lanes are colony PCR products, amplified with gRGA1-Subtilase primers. Each colony has one PCR product in the expected size range (between the 650 bp and 850 bp bands of the marker).

Colony PCR and Electrophoresis of Plasmid PCR Products

The plasmid DNA isolated from the LB48 gRGA1-Subtilase colonies yielded DNA that was clean and high enough quality to go through PCR. The initial plasmid PCR contained far too much plasmid DNA and gave significant amounts of aberrant PCR
products (See Figure 14), but later PCR's on haplotype-representative plasmid DNA showed one band per plasmid, in the expected size range of 650-850 bp (See Figure 16).

![Image of electrophoretic gel](image)

**Figure 14. Plasmid PCR Products.** This is a photo of an electrophoretic gel under UV light after ethidium bromide staining. Lanes 11 and 31 are the 1Kb+ molecular weight marker. All other lanes are PCR products of plasmid DNA, amplified by M13 (vector) primers. In the reaction that generated these products, too much plasmid DNA was added to the mixture, causing aberrant PCR products to be amplified in many cases. However, most lanes have a prominent band in the expected size range (between the 650 bp and 850 bp markers).

The sequencing of the *gRGA1-Subtilase* alleles in the octoploid parent (LB48) revealed three distinct haplotypes, which is consistent with previous data from the
sequencing of the same locus in the F1 parents L1 and BC6 (Zhang and Davis, unpublished). The three haplotypes were color-coded as Blue, Orange, and Green. The orange haplotype was found to comprise two easily distinguishable alleles that differ by a 29 base pair indel (See Figure 15). The smaller of the two alleles is approximately 704 bp and the larger is approximately 733 bp in length. The green haplotype sequences are approximately 744 bp in length, and the blue haplotype sequences are approximately 681 bp in length (See Figure 16 and Figure 17).

**Figure 15. Schematic diagram of the gRGA1-Subtilase haplotype alignment.** There are three distinct haplotypes (blue, orange, and green) and one additional subdivision within the orange haplotype. Where the colored bars are aligned vertically, the sequence is highly conserved. Breaks in the continuity of the colored bars represent indels. The *ApaI* cut site is indicated by the red line, highlighting that this cut site is in a block of sequence that is conserved over the three haplotypes.
Figure 16. PCR on haplotype-representative plasmids. This is a photo of an electrophoretic gel under UV light after ethidium bromide staining. Lane 6 is a 1Kb+ molecular weight marker; lanes 1-2 are PCR products from blue haplotype plasmids (inserted with a 681 bp fragment); lanes 3-4 are products from green haplotype plasmids (inserted with a 744 bp fragment); lanes 5 and 7 are products from the smaller of the orange haplotype plasmids (inserted with a 704 bp fragment); and lanes 8-9 are products from the larger of the orange haplotype plasmids (inserted with a 733 bp fragment).
Figure 17. Diagram of expected gRGA1-Subtilase PCR product sizes in the pentaploids and family. The orange haplotype comprises two alleles that are easily distinguishable because of a 29 bp indel.

The random sequencing of the alleles in LB48 yielded 25 blue sequences, 12 orange sequences, and 7 green sequences. With 44 total sequences, there is 97.7% confidence that if there were as many as eight alleles, they were all sequenced at least once (Figure 8). Several additional non-duplicate sequences were obtained that contained sequence apparently from two different haplotypes. In addition to being present only once each in the LB48 sequence pool, their sequences were not represented in the sequence pool from the parents L1 and BC6. These were deemed PCR chimeras, probably caused by template switching during the initial PCR, and dismissed from the study.

**Electrophoresis of FEL PCR Products and FEL-CAPS**

The products of the FEL-PCR on the pentaploids appeared to be roughly the same size as the unlabeled products, based on agarose gel electrophoresis. However, these products were visible under UV light without staining the gel with ethidium bromide, verifying that they were attached to a fluorophore (See Figure 18, which shows the gel after ethidium bromide staining in order to see the molecular weight marker). *ApaI* digestions were carried out on all of the pentaploid PCR products because the total
sequence length differences between the haplotypes were not expected to be enough to
distinguish which alleles were present in each pentaploid. Agarose gel electrophoresis of
the *ApaI*-cut *gRGA1-Subtilase* PCR products was unable to cleanly resolve the relatively
small sequence length differences between some of the products.

Figure 18. **FEL-CAPS products.** This is a photo of an electrophoretic gel after
ethidium bromide staining and under UV light. Lane 7 is a 1Kb+ molecular weight
marker; lane 1 is products of *gRGA1-Subtilase* PCR on LB48 genomic DNA that
were digested with *ApaI*; lanes 2-6 are products of *gRGA1-Subtilase* PCR on pentaploid
genomic DNA that were digested with *ApaI*. 
**Genotyping of FEL-CAPS**

The *gRGA1-Subtilase* PCR products that had been cut by *ApaI* matched the size expectations based on the LB48 sequence data (See Figure 19 and ). After the DNA from each of the pentaploids was digested with *ApaI*, the size difference between the two orange alleles was easily detected by genotyping. Each of the 42 pentaploid plants received either the 704 bp orange allele or the 733 bp orange allele, along with at least one green and one blue allele (multiple blue or green alleles were not detectable by *ApaI* digestion).

<table>
<thead>
<tr>
<th>gRGA1-Subtilase Haplotypes in LB48 and Pentaploid Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green Fluorescent Tag</td>
</tr>
<tr>
<td>287 bp</td>
</tr>
<tr>
<td>262 bp</td>
</tr>
<tr>
<td>262 bp</td>
</tr>
<tr>
<td>284 bp</td>
</tr>
</tbody>
</table>

**Figure 19.** *gRGA1-Subtilase* haplotypes in LB48 and pentaploid population. The ends of each PCR product are shown with the fluorescent tags they would have after labeled-primer PCR. The *ApaI* restriction site is present in all of the haplotypes.

The genotyping data revealed that all of the pentaploids had fragments that approximately matched the sizes of the expected fragments. However, in practice, the fragment lengths obtained from the genotyping data were skewed from the predicted numbers by four or fewer bases (see Table 2). Some pentaploid plants were found to have an additional HEX-labeled fragment that did not correspond exactly to the LB48 haplotype model (see ). Twelve pentaploid plants had a 396 or 398 bp fragment, while twelve had a 456 bp fragment in addition to the four expected HEX-labeled fragments.
shown above (see Figure 19). Those fragments that were 396 or 398 bp in length were classified as belonging to the blue haplotype because of their similarity in length to the expected blue fragment (394 bp), while those fragments that were 456 bp in length were classified as belonging to the green haplotype because of their similarity in length to the expected green fragment (460 bp). The basis for the observed small variations in size within the blue and green allele classes was not determined.
Table 2. Genotyping Fragment Sizes After Digestion By *ApaI*

<table>
<thead>
<tr>
<th>Plant</th>
<th>Orange Haplotype</th>
<th>Green Haplotype</th>
<th>Blue Haplotype</th>
<th>Orange Haplotype (O-subtype)</th>
<th>Green Haplotype</th>
<th>Blue Haplotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fragments With a FAM Fluorescent Tag</td>
<td>Fragments With a HEX Fluorescent Tag</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXA</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXB</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXD</td>
<td>265 285 288 442</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXE</td>
<td>265 285 288 442</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXF</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXG</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXH</td>
<td>265 285 288 442</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXI</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXJ</td>
<td>265 285 288 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXL</td>
<td>265 285 288</td>
<td>456 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXN</td>
<td>265 285 288</td>
<td>462 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXO</td>
<td>265 285 288</td>
<td>470 461 393 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXQ</td>
<td>265 285 288</td>
<td>470 456 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXR</td>
<td>265 285 288</td>
<td>470 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXU</td>
<td>265 285 288 442</td>
<td>461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXY</td>
<td>265 285 287 442</td>
<td>461 393 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXZ</td>
<td>265 285 287 442</td>
<td>462 393 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAA</td>
<td>265 285 288</td>
<td>470 462 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAB</td>
<td>265 285 288</td>
<td>470 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAC</td>
<td>265 285 287 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAD</td>
<td>265 285 288</td>
<td>456 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAF</td>
<td>265 285 288 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAG</td>
<td>265 285 287 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAH</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAJ</td>
<td>265 285 287 442</td>
<td>456 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAK</td>
<td>265 285 288 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAL</td>
<td>265 285 288</td>
<td>470 462 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAM</td>
<td>265 285 288 442</td>
<td>462 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAN</td>
<td>265 285 288 442</td>
<td>461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAO</td>
<td>265 285 288</td>
<td>470 462 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAP</td>
<td>265 285 288</td>
<td>470 461 393 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAQ</td>
<td>265 285 287 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAR</td>
<td>265 285 288</td>
<td>470 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAS</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAT</td>
<td>265 285 288 442</td>
<td>462 393 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAU</td>
<td>265 285 288</td>
<td>470 456 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAV</td>
<td>265 285 288</td>
<td>470 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAW</td>
<td>265 285 288 442</td>
<td>456 460 393 398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAX</td>
<td>265 285 288 442</td>
<td>461 396</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAY</td>
<td>265 285 288</td>
<td>470 456 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXAZ</td>
<td>265 285 288 442</td>
<td>460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SXBA</td>
<td>265 285 288</td>
<td>470 460 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LB48</td>
<td>265 285 288 442</td>
<td>470 461 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. FEL-CAPS genotyping data. All plants had a labeled fragment of the expected size for the FAM end of the orange, green, and blue haplotypes. All of the plants had a labeled fragment of the expected size for the HEX end of the green and blue haplotypes. 21 pentaploids had a labeled fragment of the expected size of the HEX end of the larger orange haplotype, and 21 had a labeled fragment of the expected size of the HEX end of the smaller orange haplotype.
The octoploid LB48 displays both expected orange alleles (Figure 20), whereas, in the pentaploid population, the two alleles of the orange haplotype were found to be segregating in a disomic pattern, where each pentaploid had one and only one of the two orange alleles. Of the 42 pentaploids genotyped, 21 had the 704 bp allele and 21 had the 733 bp allele (Figure 21 and Figure 22). This outcome exactly matches the expectation for a 1:1 ratio.

**Figure 20. Genotyping chromatogram from LB48.** Both of the two distinct orange haplotype fragments are present, showing that this F1 plant has both orange alleles. Additionally, this plant has at least one blue and one green allele.
Figure 21. Genotyping chromatogram from pentaploid 5XB. The only orange HEX-labeled orange fragment is the 471 bp, showing that the plant only received longer of the two orange alleles.
Figure 22. Genotyping chromatogram from pentaploid 5XD. The only orange HEX-labeled fragment is the 442 bp, showing that the plant only received shorter of the two orange alleles.
CHAPTER IV

DISCUSSION

As anticipated, the data obtained in this study illuminated the *Fragaria* octoploid genome composition in three ways. First, three different allele types were distinguished. This result contrasts to those of previous phylogenetic studies, in which only two different allele types were defined at each of three different loci in the examined octoploids. Second, the three allele types occurred in relative frequencies that fit a ratio of 2:2:4 (p-value = 0.38). This ratio is consistent with two of the published genome composition models (AABBBBCC and AAA’A’BBBB) but not the other (AAA’A’BBBB’). Third, some segregation data was obtained from FEL-CAPS on the pentaploids, providing partial insight into the issue of disomic versus polysomic inheritance. Each of these aspects of the study will be explored.

**Significance of Haplotype Variation**

Because the sequencing data from LB48 showed three haplotypes of *gRGA1-Subtilase* alleles, a three genome model such as AAA’A’BBBB is supported, but a two genome model is not. The disomic segregation of one of the haplotypes in the pentaploid population means that, if the model is accurate, the orange haplotype must represent the alleles from either the A or A' genome because the B genome alleles are expected to segregate in a tetrasomic pattern. We hypothesize that the blue haplotype may represent the B genomes because, of the 44 PCR product clones that were randomly sequenced
from LB48, over half (57%) were of the blue haplotype. Lundberg and Davis (unpublished data) examined the *gRGA1-Subtilase* intergenic region in a wide sampling of octoploid *Fragaria* taxa and through phylogenetic analysis identified three clades corresponding to the three haplotypes described in this study, and also found that sequences of the blue haplotype were almost twice as numerous as either the orange or the green haplotype sequences.

In 2009, Rousseau-Gueutin et al. examined two nuclear protein-coding genes in several species of *Fragaria* with the purpose of clarifying phylogenetic relationships. The two genes studied, as mentioned earlier, are *DHAR* and *GBSSI-2*. The two octoploids examined in that study were *F. chiloensis* and *F. virginiana*. *F. iturupensis* (a decaploid, previously thought to be octoploid) sequences were also included, as well as sequences from other *Fragaria* diploids, tetraploids, and a hexaploid. Three clades were distinguished from these data by phylogenetic analysis; however, all of the octoploid and decaploid alleles for both of these genes only fell into two of these clades. DiMeglio and Davis (unpublished) had similar results when the alcohol dehydrogenase (*ADH*) gene was examined in various species of *Fragaria*. Three clades were distinguished from the sequence data and, again, the octoploid sequences occurred in only two of the clades. The octoploid sequences falling into two clades in these studies does not fit well into the published octoploid models which suggest either three (AAA'A'BBBB or AABBCCBCC) or four (AAA'A'BBBB'B') subgenome types. It seems that the genes *GBSSI-2, DHAR*, and *ADH* are simply not polymorphic enough to resolve more than two allele types and therefore these studies do not support nor negate any of the published subgenome models. Additionally, none of the octoploid or decaploid accessions were maximally
heterozygous (yielding 8 alleles in an octoploid or 10 alleles in a decaploid), which would be highly desirable in a gene inheritance study.

**Haplotype Ratios**

In the case of this study, the *gRGA1-Subtilase* intergenic region proved to be more polymorphic than the loci examined in previous studies, such that three haplotypes were identifiable in octoploids rather than the previous maximum of two. This on its own supports a three-genome model of octoploid genome composition, and coupled with the observed haplotype ratio of 1:1:2, it matches the implied subgenome ratio of 2:2:4 of the three-genome models. The subgenome ratios draw attention to the blue haplotype as the possible domain of polysomic inheritance, but not enough sequence polymorphism was present to allow allelic resolution using the FEL-CAPS method.

From the random sequencing of the *gRGA1-Subtilase* intergenic region, approximately 57% of the total sequences obtained were blue haplotype sequences. Along with the fact that only three distinct haplotypes were found in LB48, this indirectly supports a mixed disomic and polysomic inheritance model such as AAA’A’BBBB (Senanayake & Bringhurst, 1967) or AABBBBCC (Fedorova, 1946), rather than a fully disomic model such as AAA’A’BBB’B’ (Bringhurst, 1990). The mixed disomic and polysomic models each require there to be a maximum of three genome types in the octoploid, while the fully disomic model requires a maximum of four genome types. Also, the mixed models require that two pairs (or 2/8 and 2/8) of the genomes have more similarity within the pair than between the pairs, and are disomically inherited; and that two pairs (or 4/8) of the genomes have more similarity among the foursome than between the foursome and any of the other genomes, and are tetrasomically inherited. So, with
approximately half of all randomly sequenced clones representing the blue haplotype, it can be assumed that the blue haplotype’s genomes make up 50% of the genomes in the octoploid, which supports a mixed disomic and polysomic inheritance model. However, since multiple blue type alleles were not differentiable, the genotyping data does neither support nor oppose any model based on the blue haplotype.

**Segregation Data from FEL-CAPS**

The orange sequences appeared to be segregating in a disomic pattern in the pentaploid population, based on the fact that each pentaploid had one kind of orange allele and that approximately one half of the pentaploids received the smaller allele and half received the larger allele. Disomic segregation is based on the genes and genomes in question segregating as if they were in a diploid organism. Necessarily, each chromosome pairs with only one homolog in meiosis (in a disomic situation) and the resulting gametes will receive one or the other homologous chromosome in a 1:1 ratio. Contrastingly, polysomic inheritance involves a non-diploidized manner of meiotic pairing. Each chromosome has more than one homolog, and therefore has more than one potential meiotic pairing partner. The chromosomes may pair with zero, one, or more than one other homolog during meiosis, causing the resulting gametes to have differing genotypic ratios. Some plants may have mechanisms that favor bivalent formation over multivalent formation, which would disallow aneuploid gametes. In this case, the only difference between the gametes formed by a polysomic system vs. the gametes formed by a disomic system would be the number of allele combinations possible. In a disomic system, two homologs must pair with each other and are subsequently separated into different gametes. In a polysomic system with a preference for bivalents, each
chromosome has more than once choice for a pairing and segregation, such that each different meiotic event can produce different gametes, based on which specific homologs pair together. Figure 3 is an illustration of this idea, using the premise of four genomes that are inherited by either a polysomic system or two separate disomic systems. As can be seen, the same four genomes can produce two more unique gamete types with a polysomic system than with two disomic systems, even when pairing is restricted to bivalents in the polysomic system.

The identification of two mutually exclusive orange alleles is indicative of disomic inheritance. This is because polysomic inheritance requires the existence of more than two homologs. However, if the gRGA1-Subtilase intergenic region only had two different orange alleles, but yet these two alleles were present on four genomes and therefore part of a polysomic inheritance system, the genotypic data would have revealed 50% of the pentaploids to have one or the other orange allele, and 50% to have both orange alleles because of random bivalent formation among the four genomes. Since this was not the case, we can conclude that the orange haplotype and its two segregating alleles represent a subgenome that has a disomic pattern of inheritance. In Senanayake & Bringhurst’s genome model (AAA’A’BBBB, 1967), the subgenome with the orange haplotype could be represented by either of the subgenomes with disomic inheritance, A or A’. In Bringhurst’s model (AAA’A’BBB’B’, 1990) the orange haplotype could belong to any of the four subgenomes, since they all have disomic inheritance in this model. In Fedorova’s model (AABBBBCC, 1946), the orange haplotype could belong to either the A or the C subgenome.
Implications

Completely understanding the genomic inheritance patterns of the cultivated strawberry will first put an end to the deliberation about which (if any) of the three published genome models is correct. Since evidence has been found supporting both a fully disomic model and a partially polysomic model, a much more comprehensive investigation must be done on the inheritance patterns in octoploid Fragaria. The hybrid pentaploid plants created in this project would be an excellent starting point for such an investigation. Sufficiently polymorphic loci would need to be discovered on each arm of each of the seven Fragaria chromosomes, using the parent LB48 as the basis upon which to judge the polymorphic value of each locus. Gene pair loci are very well suited for this purpose. Such gene pair loci would need to be amplified and 40-50 of the copies sequenced. Haplotypes could be described from the sequencing data, which would show whether there are three or four distinct haplotypes in a 1:1:2 ratio at every locus or whether the numbers and/or ratios of haplotypes differ across chromosomes or even across loci on the same chromosome. Then, whether by FEL-CAPS genotyping or by direct sequencing, the meiotic segregation pattern of the octoploid chromosomes would be easily deducible by examining the pentaploids for allele content.

After a thorough investigation has shown direct evidence for the meiotic segregation pattern of LB48, the chromosomal transmission pattern for F. virginiana could be inferred to be similar since LB48 is a hybrid of two F. virginiana subspecies. However, there is no reason to expect that all Fragaria octoploids will conform to a common model of subgenome composition. Earlier investigations into the genome content, meiotic pairing behavior, or transmission patterns of the Fragaria octoploids
may have appeared to be conflicting because each octoploid species may be different. An understanding of *Fragaria* octoploid subgenome composition will: help guide future efforts to sequence an octoploid strawberry genome; and enable breeders to employ appropriate models of inheritance in planning crosses and population sizes needed to recover the desired genotypes.

**Limitations of FEL-CAPS**

The original idea of this study was to identify and sequence all of the *gRGA1-Subtilase* alleles in the octoploid LB48, and then recover those alleles in groups of four by FEL-CAPS genotyping of each member of the pentaploid progeny. If only four genotypes were represented among the pentaploid progeny, the octoploid parent (LB48) must be fully disomic at this locus. However, if six genotypes were represented among these progeny, the octoploid parent must be partially polysomic at this locus. Unfortunately, the polymorphism at this locus was insufficient, as only four distinguishable alleles were identified in the octoploid parent LB48, so the full genetic complement of each pentaploid plant could not be elucidated. Additional work could be done on this particular locus to try to detect other subtypes, including resequencing all of the seventy-six LB48 plasmids from the tail end (using the M13 Reverse primer) to more precisely identify the expected maximum of eight alleles for this locus in LB48. It is possible that there is sufficient polymorphism yet at this locus to achieve the original goals of the study, however, even the smallest of replication errors in the PCR that created the cloned fragments or in any of the subsequent replications would continue to muddle the data.
The failure of the *gRGA1-Subtilase* alleles in “Yellow Wonder” to be amplified by the original primers is an example of how different the sequences can be in even in closely related species. The primer recognition sites of the original primers (19M24F & R) are just inside the coding regions of the *gRGA1* and *Subtilase* genes and so it was expected that the primers would work well in any *Fragaria* plant. Since this was not the case, it is also possible that the sequences could be so different between haplotypes that one or more haplotypes of LB48 may not have been amplified at all. However, the three haplotypes that were observed in this study are the same as were found in a previous study of *gRGA1-Subtilase* alleles in BC6 and L1, which are the parent plants of LB48. It is possible, though, that since the same primers were used in that study, the same additional sequences were not amplified in either study. This problem could potentially be solved by resequencing the *gRGA1-Subtilase* intergenic region, starting with the expanded region primers and employing the AccuPrime™ Taq DNA Polymerase High Fidelity. The primer recognition sites are farther into the coding region of the genes, so it is possible that there would be more sequence conservation in that area. This approach, however would more than double the fragment sizes for sequencing, so it would take more than one sequencing run (with current standard sequencing technology) to sequence each fragment.

The initial plan to differentiate and conveniently genotype eight alleles in the pentaploid population could not be fully realized for several reasons. One obstacle was the inherent limitation of the FEL-CAPS approach to genotyping. If a sequence polymorphism didn't result in an altered restriction fragment pattern, it could not be detected at all with this genotyping method. For example, at position 754 of the *gRGA1*
Subtilase region, the blue haplotype sequences have either a T or C. This sequence polymorphism did not result in an altered restriction site and therefore could not have been genotyped using FEL-CAPS. However, regardless of the limitations of the genotyping method, ultimately it was the lack of incontrovertible DNA sequence polymorphism that precluded resolution of eight alleles at the gRGA1-Subtilase locus.

**Alternate Methods**

To ensure the success of a future study, it would be necessary to identify a more polymorphic site and employ a genotyping method that will allow maximum resolution of allelic diversity. It would be beneficial to have a genotyping method that is better able to distinguish between eight alleles without having to rely on changes at restriction sites or on restriction fragment sizes. One such a method involves having sequences that differ only by SNP's. PCR products, rather than clones, can be directly sequenced and the sequencing chromatograms can be examined for multiple peaks, which is indicative of a SNP. Then, each progeny plant's PCR products can be sequenced and examined to see which nucleotide state is present. This genotyping by sequencing method could only be used for small areas of the gRGA1-Subtilase intergenic region because just one indel will offset the frame and cause the sequencing chromatogram to be unreadable. Additionally, in order for this protocol to yield valuable results, the SNP's would need to be informative in some way. For example, at the end of the gRGA1-Subtilase region, there is a small stretch of nucleotides that contains a SNP and no indels. The blue haplotype sequences have either a T or C at position 754 of the sequence alignment. This may represent a genuine subtype of the blue haplotype, however, the other two haplotypes also have C at this position. Any genotyping method that was able to show
which pentaploids had T and which had C at this position would only be able to detect the presence vs. absence of an allele with a T at this position. In an octoploid this is not incredibly useful, as a plant with one allele that has T at this position would appear the same as a plant with four alleles that have a T at this position. Therefore, this particular method would work only if there is a character state unique to one allele only. Another method for genotyping multiple alleles is third generation sequencing, which involves anchoring single molecules of DNA polymerase in individual wells and reading the sequence of fluorescently labeled dNTP's with a camera as each single molecule of DNA is being polymerized (http://www.pacificbiosciences.com). This technology would allow very quick and easy sequencing of all alleles in polyploids by sequencing, rather than genotyping by fragment labeling or other genotyping methods. However, if each and every pentaploid was fully sequenced in this region, it is unclear whether there would be enough resolution to differentiate the sequences even then.

**Future Directions**

In this study, the plant LB48 was chosen as a potentially maximally heterozygous octoploid representative of *Fragaria* octoploids, but the study was limited by insufficient polymorphism to differentiate 8 alleles, and further limited by the genotyping method which could not exploit all of the polymorphisms that did exist. However, the haplotypes, their ratios, and the segregation data that was gained was evidence to support a partially disomic and partially polysomic inheritance model such as AAA'A'BBBB or AABBBBBC. This same kind of study could be done directly on the cultivated strains of *F. xananassa* to see if all of the *Fragaria* octoploids have the same genetic constitution. Additionally, these same methods should be used on additional intergenic regions in
Fragaria, preferably with loci on other chromosomes, since each particular locus will only represent one chromosome out of seven in each of the genomes, for a total of eight chromosomes out of the 56 that octoploid Fragaria has. Ideally, at least one locus from each of Fragaria's seven chromosomes should be tested in an inheritance study such as this before assumptions are made about the behavior of any genome as a whole. To be even more comprehensive, it might be considered to look at one locus near each end of every chromosome to rule out or investigate the possibility of translocation-like synapsis patterns.
REFERENCES


Bennett, M.D., Leitch, I.J., Price, H.J., Johnston, J.S. 2003. Comparisons with Caenorhabditis (approximately 175 Mb) and Drosophila (approximately 175 Mb) using flow Cytometry show genome size in Arabidopsis to be approximately 157 MB and thus approximately 25% larger than the Arabidopsis genome initiative estimate of approximately 125 Mb. Annals of Botany 91: 547-557.


Zhang, L. and Davis, T.M. Unpublished. *gRGA1-Subtilase* and *ADH* sequencing data.