CHARACTERIZATION OF STRAWBERRY (*Fragaria ananassa*) BY GENOTYPING AND PHENOTYPING

By

ANJU BISWAS

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in the
Agriculture Graduate Program
of Delaware State University

DOVER, DELAWARE
May 2018

This thesis is approved by the following members of the Final Oral Review Committee:

Dr. Kalpalatha Melmaiee, Committee Chairperson, Department of Agriculture and Natural Resources, Delaware State University
Dr. Sathya Elavarthi, Committee Member, Department of Agriculture and Natural Resources, Delaware State University
Dr. Venu Kalavacharla, Committee Member, Department of Agriculture and Natural Resources, Delaware State University
Dr. Richard Braczewski, Committee Member, Department of Agriculture and Natural Resources, Delaware State University
Dr. Rose Ogutu, External Committee Member, Cooperative Extension, College of Agriculture and Related Sciences, Delaware State University

ACKNOWLEDGMENTS

I would like to acknowledge my advisor Dr. Kalpalatha Melmaiee for advising me throughout the entire project. Her guidance and support helped me to accomplish my research. My special thanks to Dr. Sathya Elavarthi who assisted me continuously from my admission until my graduation; provided his valuable guidance and stipend support to conduct my research. I am thankful to other committee members; Dr. Venu Kalavacharla and Dr. Richard Barczewski, for their mentorship and insightful comments. My sincere thanks to my external committee member Dr. Rose Ogutu who assisted me to start my project by providing strawberry plant samples from her high tunnel research. Thanks to Dr. Umesh Reddy and Ms. Venkata Lakshmi Abburi from West Virginia State University for their continuous help that enabled me to finish the significant part of my research. I would like to thank Dr. Cyril Broderick, Dr. Sigrid Smith for their constructive suggestions & assistance with my research analysis. Thanks to Dr. Karl Miletti-Gonzalez, Dr. Ayalew L. Osena and Mr. Michael Moore for their help in my project. I am grateful to all lab members from Dr. Venu Kalavacharla's lab especially Tara Kay Jones who supported me all the time. My heartiest thanks go to my lab members, especially Julian Jones who provided support and laughter throughout the time of my project. Besides, I would like to thank the McIntire Stennis Cooperative Forestry Research Programs and the College of Agriculture and Related Sciences (CARS) for providing me with the funding to conduct my research. Finally, I would like to thank my parents, my siblings, my nephews and nieces for their caring, love, motivation and support throughout the project. My mom, Dipti Rani Biswas gave me the extra energy to work harder and remained focus all the time; therefore I dedicated this thesis to her.

CHARACTERIZATION OF STRAWBERRY (*Fragaria ananassa*) BY GENOTYPING AND PHENOTYPING

Anju Biswas

Faculty Advisor: Dr. Kalpalatha Melmaiee

ABSTRACT

Strawberry is an important fruit crop for its nutritional value and is known to have a higher amount of antioxidants. Consuming strawberries can increase dietary antioxidants. Antioxidants in strawberry can act against cardiovascular diseases, cancer, and other health issues. Recently, use of molecular markers is becoming popular in strawberry breeding programs as well as in genotyping of existing varieties. Genotyping is the process of determining genetic differences among individual plants and can be readily applied to identify genetic diversity among the germplasm. Simple sequence repeat (SSR) markers present in every living organism. Majority of the SSR markers are used for genotyping includes cultivar identification, genetic diversity analysis, taxonomic analysis, marker-assisted breeding, and cloning. It is easier to identify a trait of interest in any genotype if known SSR markers are available. Cultivated strawberries have a wide range of SSR markers available which permits to detect polymorphism of closely related genotypes as the genome of strawberries has been sequenced. In this study, we genotyped thirty-three strawberry accessions utilizing thirty-five SSR markers and phenotyped by observing leaf total antioxidant (TA) content, leaf shape, leaf trichome density and petiole size. DNA extraction, PCR, and SSR analysis were done for genotyping. We identified two distinct genetic groups from those genotypes, one group is alpine (wild or woody) type strawberry and another group is the cultivated type. In this study, we identified 120 alleles with an average of 3.43

alleles per locus. The genetic polymorphism ranged from 0.1461 for the marker ARSFL_9 to 1.6635 for the marker FG1a/b. All thirty-three accessions were grouped into clusters based on the genetic diversity analysis and found that grouping them into three clusters was provided the maximum genetic diversity. The cluster three contained more diverse genotypes among those. Leaf total antioxidant content was measured in Trolox equivalent amounts by colorimetric assays. The highest antioxidant containing genotypes were Earliglow, Wendy, Elan hybrid, Clancy, and Record. Spectrophotometer and high-resolution camera were used for phenotyping. We classified all genotypes into five groups according to the leaf shape and three groups according to the leaf trichome density and leaf petiole size. From the association, the two genotypes named Clancy and Elan hybrid contained more total antioxidant content with high trichome density. Strawberry breeding programs can utilize information from this research during cultivar development and authentication of genotypes.

Keywords: Strawberry **(***Fragaria ananassa)*, marker-assisted breeding, SSR, phenotyping, genotyping, antioxidants, leaf traits

Table of contents

List of tables

List of figures

LIST OF ABBREVIATIONS

AATI: Advanced analytical technologies, inc.

AMOVA: Analysis of molecular variance

AFLP: Amplified fragment length polymorphismh

Bp: Base pair

CTAB: Cetyltrimethylammonium bromide

DNA: Deoxyribonucleic acid

EST: Expressed sequence tag

EDTA: Ethylenediaminetetraacetic acid

ISSR: Inter simple sequence repeats

IQR: Inter quartile range

MAB: Marker assisted breeding

MAS: Marker-assisted selection

NJ: Neighbor joining

PCR: Polymerase chain reaction

PCA: Principle component analysis

RAPD: Random amplification of polymorphic DNA

RCF: Relative centrifuge force

RFLP: Restriction fragment length polymorphism

SSR: Simple sequence repeats

STRs: Short tandem repeat

SNP: Single nucleotide polymorphism

SSLPs: Simple sequence length polymorphism

SD: Standard deviation

SORC: Smyrna outreach & research centre

mM: Millimolar

µl: Micro liter

CHAPTER I: INTRODUCTION

Strawberry is a promising fruit crop in most of the temperate regions of the world. It is one of the most enjoyable small fruit with high nutritional value. The cultivated strawberry, *Fragaria×ananassa* belongs to Rosaceae family, subfamily Rosoidaceae, home to some of the other berry crop species like raspberries and blackberries. There are four naturally occurring ploidy levels in strawberry such as diploid, tetraploid, hexaploid and octoploid (Hancock et al., 1993; DiMeglio et al., 2014). Another rare decaploidy was reported by some authors (Folta & Kole, 2011).The base chromosome number of strawberry is 7.

Fragaria vesca is one of the representative diploid species (2n=2x=14). It is also known as the alpine or woodland type and is an ancestor of octaploid species (Folta & Kole, 2011). It was cultivated for a century in Europe (Darrow, G. M. 1966.). Due to its ubiquitous nature, small size, simple genetic makeup and phenotypic variability, it is an attractive species for strawberry research. The genome size (240Mb) is almost similar to *Arabidopsis* (Folta & Davis, 2006). Despite its small fruit size, it can be of interest to breeders for its flavor and aroma. The existence of natural hexaploid (2n=6X=42) is insufficient. There was only one hexaploid strawberry that has been cultivated in Europe on a limited scale. However, the fruits were considered the biggest until the introduction of octaploids (Folta & Kole, 2011). Like other essential crops such as cotton and wheat, the history of strawberry is conveyed about hybridization and polyploidization (DiMeglio et al., 2014). Two American octaploid (2n=8X=56) species *F. virginiana* and *F. chiloensis* are the ancestors of modern cultivated strawberry. The octaploid strawberry is originated from intercrossing between two species and were found in North America, South America and in Europe in the mid-1700s (Hancock et al., 1994). The hybrid fruit of these two species is different from parents by their distinctive and desirable traits. Octoploid strawberry is

cultivated in 60 different countries and it is becoming a favorite fruit in some tropical and sub-tropical areas. It is grown as a perennial fruit crop in temperate regions and as an annual in other parts of the world.

The high amount of antioxidant and attractive taste are some of the positive traits of strawberry. So, one can increase their dietary antioxidants by consuming strawberry (Cook & Samman, 1996). Research shows that strawberry acts against cardiovascular diseases, cancer and other health issues (Capocasa et al., 2008). Numerous studies have demonstrated that phytochemical content and similar antioxidant activity of fruits and vegetables are correlated with protection against chronic and other degenerative diseases (Record et al., 2017). These antioxidants scavenge free radicals including superoxide radicals, hydrogen peroxide, hydroxyl radicals, and singlet oxygen which are responsible for chronic diseases (Wang & Jiao, 2000). Phytochemicals that exhibits antioxidant activity include phenolic compounds such as flavonoids, anthocyanin, tocopherols, carotenoids as well as ascorbic acid (Meyers et al., 2003). These natural antioxidants scavenge free radicals, inhibit reactive species and prevent damage of cellular lipid, protein, and nucleic acid (Heinonen et al., 1998).

The production and demand of strawberry are increasing continuously in recent decades due to its health benefits and the use of it in both food and cosmetic industry. The United States is the highest strawberry producing country in the world. The total production of strawberries in the United States was approximately 1.57 million tons in 2016 (https://www.statista.com/statistics/193288/us-total-strawberry-production-since-2000/). Other dominant strawberry producing countries are Spain, Korea, Poland, Mexico, Turkey, Egypt (Wu et al., 2012). The strawberry industry is mainly located in the southern and coastal areas of California in the United States (Geisseler and Horwath, 2014). According to Wu et al., (2012),

California, Florida, and Oregon are the three top strawberry producing states, others are North Carolina, Washington, New York, Michigan, Pennsylvania, Wisconsin, and Ohio.

To keep pace with the increasing demand for strawberry, genetic improvement programs, production, and post-harvest practices are adopting many new technologies. One plant differs from another plant because of differences in their genetic make-up (DNA). Exploitation of these genetic differences is one of the promising tools for strawberry research and industry (Whitaker, 2011). Molecular markers are one of the promising tools for detection and exploitation of DNA (deoxyribonucleic acid) polymorphism. Development and the use of DNA markers are one of the most critical tools for strawberry research, and has been addressed by several authors beginning with Hokanson and Mass (2001). These markers are used for analysis of genetic diversity, cultivar identification, polymorphism identification, and genetic mapping.

Polymorphism identification is becoming popular in many crops including strawberries and the discovery of simple sequence repeats (SSRs) marker makes it even more popular. SSR is also known as microsatellite, short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs). SSRs are the smallest class of simple repeated DNA sequences. Some authors (Armour et al., 1999) reported that SSR is 2-8bp repeats while others (Goldstein & Pollock, 1997) reported 1-6bp repeats. The main reasons for wide use of SSR markers is that they are highly polymorphic even between closely related species (Govan et al., 2008), requires less amount of DNA, easily exchangeable between laboratories and highly transferable between populations, can be easily automated for high throughput screening (Gupta et al., 1999). SSR markers are excellent markers for fluorescent dye techniques, multiplexing, and high throughput analysis as well as for the studies of gene mapping and population genetics (Goldstein & Schlotterer, 1999).

According to Whitaker et al., (2011), marker-trait associations are used for disease resistance, metallothionein-like protein, antioxidant identification, and sex expression. Marker-assisted breeding (MAB) of strawberry is becoming popular in private sector (Folta & Kole, 2011) as private breeding program realized the power of marker-assisted technology however it has been limited in public sector. They have been using molecular biology technologies for screening plant lines as well as to describe genome of strawberry and it is one of the best tools for cultivar identification (Brunings et al., 2010).

MAB is cost-effective compared to conventional breeding methods. For instance, the disease resistant trait is very expensive or very difficult to screen or apply selection phenotypically compared to marker-assisted selection (MAS) as disease occurrence depends on correct environmental conditions, aggressive pathogen and susceptible host (Luby & Shaw, 2001). It may not be possible to create this conducive condition all the time. MAS is cost-effective compared to plastic culture or other systems. For example, when we consider the cost of plastic culture, maintenance and field preparation is expensive compared to MAS. So, financially MAS can be more attractive (Whitaker, 2011).

According to Luby and Saw (2001), MAS will be more attractive and economical if the heritability is below 0.2 because selection will be based on the presence of that DNA fragment will be more convenient. Flavor is such a kind of trait with low heritability. It is affected by multiple chemical components like sugars, volatile compounds, organic acids, etc. It is quite expensive to isolate volatiles to evaluate genotypes by using standard methods, but the use of markers for selection can be cost-effective and is easy (Ulrich et al., 2007). Thirty three strawberry genotypes were utilized from different USA vendors representing different continents. The collected accessions represent both cultivated and wild genotypes. These are a

new collection of strawberry germplasm at Delaware State University. These genotypes may have genetic polymorphism, differ in their total antioxidant content and also differ in different leaf traits. Genetic polymorphism was identified utilizing 35 SSR markers, measured total antioxidant content, observed leaf traits such as leaf shape, petiole sizes and trichome density. Desirable traits are associated with genetic constituents. Molecular markers (SSR markers) assisted in finding out those traits among thirty-three genotypes.

Research objectives are as follows.

1. To find out SSR markers which show polymorphism and identify polymorphism among the thirty-three genotypes
2. To observe the total antioxidant content and other leaf traits among all strawberry genotypes under the study.

MAB speeds up the varietal development process. Strawberry industry and public programs are in need of characterized markers for many desirable traits. The above-characterized makers will be helpful in the varietal developmental process of strawberry or closely related species like blueberry and raspberry.

CHAPTER II: REVIEW OF LITERATURE

An improved understanding of the vital role of fruits in the diet in maintaining the human health has led to increasing of berry crop production (Debnath et al., 2012). Humans suffer from chronic diseases like cancer and heart disease at some point in time. There is an association between fruits consumption and reduced incidences of cancer and mortality were observed (Wang et al., 1996). In other studies, a negative association between intake of total fresh fruits and ischemic heart disease was reported (Armstrong et al., 1975; Acheson & Williams, 1983). Regular fruit consumption reduces blood pressure (Sacks & Kass, 1988). The positive health benefits from berries can be attributed mostly to the presence of antioxidants in berries (Bors & Saran, 1987). Changes in the dietary pattern can help to reduce these diseases by 4% (Joshipura et al., 2001). Berry crops contain high level of Vitamin C, antioxidant and anti-inflammatory activities. Strawberry contains a high amount of antioxidants and studies shown that they act against chronic diseases (Felgines et al., 2003), act as antitumor, antiulcer agents (Debnath et al., 2012). Flavonoids, anthocyanin are some phenolic compounds with antioxidant properties. The amount of flavonoids present in strawberry have been directly or indirectly associated with antioxidant capacity. Total flavonoid content varies from cultivar to cultivar, and it may affect the total antioxidant capacity and overall protective benefits for health (Meyers et al., 2003).

The development of molecular markers for the detection of DNA polymorphism is one of the significant building blocks in molecular biology. Molecular markers have been significantly used in the molecular breeding of horticultural crops to create a new source of genetic variation. Recently, molecular markers have been used to protect plant breeders' right and to maintain the genetic integrity of strawberry. Plant molecular markers are popular and advantageous over phenotypic characteristics (Govan et al., 2008). Molecular markers have a pivotal role to study

genetic variability and diversity. Analysis of genetic diversity can help to classify accessions and also help to identify a subset of core accessions (Debnath et al., 2012). Analysis of genetic diversity also help to find out genetic variability among genotypes (Cox et al., 1986), identify diverse parental genotype to create maximum genetic variability in progenies (Barrett et al., 1998) and introgressing desirable genes from diverse genetic resources (Thompson, Nelson, & Vodkin, 1998).

MAB is used for breeding including cultivar identification, genetic diversity analysis or taxonomic analysis. There are several types of markers have been using for this purpose. The molecular markers are one of them. Molecular markers refer to the fragment of DNA which is associated with a certain location of a genome. The molecular markers can be grouped into hybridization-based and PCR-based markers (Semagn et al., 2006). They vary in principle, methodologies derived from and their applications. The most commonly used molecular markers are; RFLP (restriction fragment length polymorphism), RAPD (random amplification of polymorphic DNA), AFLP (amplified fragment length polymorphism), ISSR (inter-simple sequence repeat), SSR (simple sequence repeat) or microsatellites, EST (expressed sequence tag), and SNP (single nucleotide polymorphism).

PCR based molecular markers utilize PCR technique. No living organism is needed here. This technique is used to amplify short, well-defined part of DNA strand from a single gene or portion of a gene. PCR based molecular technique was invented by Kary Mullis in 1983 (Semagn et al., 2006). The advantages of PCR based markers compared to a hybridization-based technique are that they require only a small amount of DNA, and can be usable even in small labs regarding equipment, facilities, and cost (Wolfe & Liston, 1998). This technique does not require any prior

sequence knowledge for many applications (RAPD, AFLP etc.) except for developing SSR markers.

PCR (polymerase chain reaction) based techniques are grouped into two types (Semagn et al., 2006) depending upon how the primers were derived whether as arbitrary or semi-arbitrary primer-based PCR techniques or site targeted PCR techniques. No prior sequence information is needed in arbitrary or semi-arbitrary primer-based PCR techniques for generating primers such techniques includes RAPD (Random Amplification of Polymorphic DNA), AFLP (amplified fragment length polymorphism), and ISSRs (inter-simple sequence repeat) etc. Known DNA sequence is used for generating primers in site-targeted PCR techniques such techniques include EST (expressed sequenced tag), CAPS (cleaved amplified polymorphic sequences), SSR (simple sequence repeat), and SCAR (sequence characterized amplified region) etc.

For a period of time RFLP was the most commonly used hybridization based molecular marker. It was first used in 1975 for identifying DNA sequence polymorphism (Sambrook et al., 1975). This technique is based on restriction enzymes and it helps to find out the pattern differences between DNA fragment sizes in any organism. Some of the limiting factors of RFLP is it requires a high amount of quality DNA (Roy et al., 1992), the level of polymorphism is low, time-consuming, laborious and expensive compared to other techniques (Semagn et al., 2006).

The RAPD technique is based on the use of single arbitrary oligonucleotide primer to amplify template DNA without previous knowledge of the sequence of the targeted DNA. The main limitations of RAPD are reproducibility, dominant inheritance and homology (Adams & Demeke, 1993).

AFLP combines the power of RFLP with flexible PCR based technique such as ligation of primer recognition sequences with the restriction enzyme sequences (Lynch & Walsh, 1998). AFLP analysis is more efficient, reproducible, reliable and representable than RFLP (Mueller et al., 1996). One of the main limitations of AFLP technique is that it requires many steps to produce the result, and it is not cost effective (Semagn et al., 2006) unless you characterize the identified sequences.

Previous studies show that some molecular markers were used to fingerprint strawberry genotypes including AFLPs, RAPDs and ISSRs (Milella et al., 2006). According to Govan et al., (2008), all of those systems rely on arbitrary primer pair for the amplification of unknown DNA fragments. So, there have been difficulties with reproducibility and standardization of those markers compared to SSR markers. Thus, robust marker sets are lacking for strawberry. Govan et al., (2008) reported that SSR marker sets are standard for strawberry cultivar identification, genetic diversity analysis since they are derived from the known sequence.

Compared to other fruit crops, strawberry is becoming an attractive target for breeders and horticulturists as it has small fruit size, economically viable, and exhibits phenotypic variability. Approximately 100 breeding programs worldwide focus on strawberry varietal improvement, but breeders are struggling with difficulties to find out elite varieties as strawberry is propagated from stolon (Govan et al., 2008). Substantial losses can be observed due to wrong varietal selection. It requires a robust and reliable system of breeding to protect the genetic integrity and breeder's right.

Molecular markers especially SSR markers are becoming popular because of their known sequence and high reproducibility (Davis et al., 2006). Approximately, 200 SSR markers are developed for *Fragaria* species (Sargent et al., 2006). Govan et al., (2008) developed 10 SSR markers which are ideal for genetic fingerprinting in all stages of strawberry plant and are easily transferable to other *Fragaria* species.

Some markers can be associated with the trait of interest for genetic improvement. Several such associations include disease resistance, heat shock protein, auxin-binding protein, the gene for lipid transfer protein etc. (Lewers et al., 2005). Sometimes the use of molecular markers can be lacking due to the unavailability of genetic information about desired traits and associated markers. However, it is an important selection aid in a breeding system for disease resistance and fruit quality in both public and private sector (Whitaker, 2011)

In strawberry, the most significant improvement has been observed in the area of disease resistance. The first marker-trait association was found for the resistance to *Phytophthora fragariae, the* causal organism of red stele root rot in strawberry (Haymes, 1997). The second marker-trait association was found for the anthracnose fruit rot of strawberry against the causal agent *Colletotrichum acutatum* (Denoyes-Rothan, 2005).

Singh et al., (2015) demonstrated the effectiveness of SSR to resolve population structure from wild and cultivated accessions of pomegranate. SSR markers are becoming valuable genetic markers in plant studies including linkage mapping, marker-assisted selection, gene flow characterization, etc. (Lewers et al., 2005). Microsatellite analysis such as neighbor-joining (NJ) tree analysis, principal component analysis (PCA) and model-based population structure analysis

corroborated the genetic relationships among wild-type and cultivated type pomegranate (Singh et al., 2015).

According to Govan et al., (2008), individual markers can be multiplexed for SSR markers and can save both money and time. Many research groups are developing SSR markers from strawberry cultivars, and these markers can be easily transferable from one cultivar to another cultivar (Bassil et al., 2006) and it has been proven. The high cost of SSR can be eliminated by the high level of transferability among cross-species (Ashley et al., 2003).

Genotyping is the process of determining genotype variation among individuals by examining the DNA sequence of an individual. Genotyping by SSR markers can be an ideal, cost-effective, reliable, and convenient tool for cultivar fingerprinting and identification which will be greatly beneficial for growing strawberry industry.

Both genotyping and phenotyping are interrelated and are important to develop any improved cultivar. Phenotyping is the process of determining phenotypic differences between individuals. Phenotype refers to the observable traits such as morphological, biochemical or physiological. Plant phenotype is the product of the interaction between genotype and environment. Rapid developments are taking place in the field of image analysis based phenotyping. Phenotyping for horticultural and commercial traits is very important to translate genomic knowledge through marker-assisted breeding to enhance breeding efficiency (Mathey et al., 2013). A breeding program will be successful when it identifies any genotype with optimum phenotypic traits. Most breeding programs utilize the traditional breeding program to develop new phenotypic trait. Nowadays, marker-assisted breeding integrates traditional breeding in trait development due to decreased costs, increased efficiency, and availability of markers for that trait (Bliss, 2010).

According to Mathey et al., (2013), the lack of phenotypic data can hinder the use of statistical methods for identifying an association between phenotype and genotype (Bassil et al., 2010) hence phenotyping is a meaningful way to improve genetic resources.

Individual plant traits can influence the performance of plant such as productivity or ecology (McGill et al., 2006). Plant height, leaf economic spectrum like leaf venation, leaf structure, gas exchange rates are getting more attention in trait network (Sack et al., 2013). Leaf morphology has remarkable phenotypic variation throughout the plant kingdom. Leaf shape may influence photosynthesis, yield, biotic and abiotic stress resistance, etc. Andres et al., (2014) found that leaf shape is an important factor for cotton yield. Another leaf trait, leaf petiole also contributes to photosynthesis indirectly by supporting the leaf blade, and it delivers to its position which is more appropriate for photosynthesis (Kozuka et al., 2005). The hairy like structure on leaves, called leaf trichome also plays a significant role in crop improvement. It influences the leaf physiological response to abiotic stresses. Dalin et al., (2008), found that trichomes influence the energy, carbon and water balance of plants. The glandular and non-glandular trichomes can help against ozone stress (Li et al., 2018). Hence, observation of different leaf traits has a significant role to find out phenotypic diversity.

In our study, we observed four phenotypic traits in strawberry genotypes which includes total antioxidant content in leaves, leaf petiole size, leaf shape, and trichome density. These traits and microsatellite analysis were utilized to find out a better genotype in this study.

CHAPTER III: MATERIALS AND METHODS

3.1 Plant material:

Thirty-three strawberry genotypes were utilized as shown in Table1 for this research. Most of the plants were raised in the greenhouse or net house. Some plants were grown at Delaware State University Smyrna Outreach & Research Centre (SORC) farm. All the genotypes were collected from several sources representing extensive geographical regions of the world. Among these genotypes twenty six are octaploid (8n=56) and seven (bold in Table 1) were diploid (2n=14). *Fragaria vesca* is the ancestor of these seven genotypes while *Fragaria chiloensis* & *Fragaria virginiana* are the ancestors of other 26 cultivars which are also known as a modern cultivar.

Table 1. Strawberry genotypes utilized in the study

Genotype No.	**Genotype Name**	**Genotype No.**	**Genotype Name**
1	Clancy Strawberry Junebearer	18	Honeoye strawberry Junebearer
2	**Mignonette**	19	Albion
3	Annapolis	20	Cardinal
4	Wendy	21	Valley Sunset
5	Sparkle Strawberry Junebearer	22	Earliglow Strawberry Junebearer
6	Tribute Strawberry Everbearer	23	Seascape
7	Cavendish	24	Guardian
8	Mesabi	25	Sureccrop Strawberry Junebearer
9	**Red wonder**	26	Tennessee Beauty
10	Red Chief	27	Ozark beauty strawberry everbearer
11	Record	28	Jewel
12	**Attila**	29	Cabot
13	**Reine des Vallees**	**30**	**Fragola Quattro Stagioni**
14	Yambu	31	Sequoia
15	**Vanilla yellow**	**32**	**Alexandria**
16	Eversweet strawberry everbearer	33	Elan Hybrid
17	Wonderful Pineberry Strawberry		

3.2 SSR primers:

Initially, 72 SSR primers were collected from different published papers (Govan et al., 2008; Brunings et al., 2010; K. S. Lewers, Hokanson, & Bassil, 2005). Majority of these primers were designed from genome sequences but not been characterized for specific traits. Standard PCR was done with genomic DNA utilizing five strawberry genotypes in Molecular Plant Breeding lab at Delaware State University. Sixty primers were amplified which were sent to West Virginia State University for SSR analysis using AATI (Advanced Analytical Technology Inc.) fragment analyzer. After analyzing in capillary electrophoresis, 35 primers (Table 2) were selected which showed proper amplification with 100% polymorphism.

Table 2. List of SSR primers utilized in the genomic analysis

Sl. No.	Primer Number	Primer Name	Forward Sequence	Reverse Sequence
1	primer 1	ARSFL_11	GCG AAG CAT AAC TGG CAG TAT CTG	GCG GGC CTA GGT GAT CTT GGA
2	primer 2	ChFaM-023	AGG AGA AGA CCG GCT GTG TA	TGC CTA TAG CTG TGG CTG TG
3	primer 5	EMFn170	CAG TTT GCC CAA CAA CAA GG	TTG ATG GCA ACA AAT CAC G
4	primer 6	EMFn181	CCA AAT TCA AAT TCC TCT TTC C	GCC GAA AAA CTC AAA CTA CCC
5	primer 7	EMFn182	GCA ACA AAG GAG GTT AGA GTC G	TGG TGA GTG CTC ATT GTT CC
6	primer 8	EMFv104	TGG AAA CAT TCT TAC ATA GCC AAA	CAG ACG AGT CCT TCA TGT GC
7	primer 9	EMFvi136	GAG CCT GCT ACG CTT TTC TAT G	CCT CTG ATT CGA TGA TTT GCT
8	primer 10	EMFvi166	ACC GAC AGC TGA GTT AGA GGA G	AGT CAT AGG ACC CCA CTT CAA A
9	primer 11	ARSFL_9	GCG AGG CGA TCA TGG AGA GA	GCG TTT CCT ACG TCC CAA TAA ATC
10	primer 12	ARSFL_10	GCG TCA GCC GTA GTG ATG TAG CAG	GCG CCA GCC CCT CAA ATA TC
11	primer 13	ARSFL_4	GCG GTC GCA TTG AGT TGG AGG ATA	GCG TAG CCA AAC ACC GAT CTA CC

12	primer 14	ARSFL_12	GCG GAA CCA AGC CAA TAA GAT G	GCG ACC ACG ACA GTT TCT CAC TCT
13	primer 15	ARSFL_15	GCG GGC TGT CCA CAC TCC TTT CT	GCG ATG CGT AAG TCT CTT CAA ATA
14	primer 16	ARSFL_17	GCG CAT CAC AAT CGC CAT AGA AAC	GCG AAC ACG CCT TCA ACA ACC AC
15	primer 17	ARSFL_97	CAA GCA ATC CAA CAG CTC AA	ACG CCT CTA AGC ACT TCC TG
16	primer 19	ARSFL_35	TGG GAT CTG CTT AGG CTT TT	AAG CCA CTT TTT ACC CCT CAA
17	primer 23	ARSFL_33	TTC AAC AAT GGC TTG AGC TG	TGA ACC TTA TGC CTC CTG CT
18	primer 24	FAC-001	AAA TCC TGT TCC TGC CAG TG	TGG TGA CGT ATT GGG TGA TG
19	primer 25	ARSFL_31	CGA CCC AGC GAC TAC ATT G	ACT TTA ACC GCC ACC AAC TG
20	primer 26	FAC-007	GAC GGA CCG ACA CTA AAC TTT G	CTA GCT GAC CTC ATT GCT CTG T
21	primer 27	ARSFL_100	TGA TGT ATT GCA TTT CGT GCT	CTA TCT CCC GGT GCT TTG AC
22	primer 31	ARSFL_101	CAG CTA AAA CCC TGC TCT CG	GTG ACG ATA GGC CGT GAA AC
23	primer 32	ARSFL_30	TTC GAA GAT TGG AGA AGA AAG G	AAG CCA CTT TTT ACC CCT CAA
24	primer 33	FAC-002	TCA TCC TCT TTC ACC TCC ACT T	TCA AAA GAC TTG GAA ATG TTG C
25	primer 34	FAC-004d	GCC AAT GTT CGA TGT TTC ACT A	TCC TTG GGT CGA TCA CAT AAA T
26	primer 35	FAC-016	TTT CAA GAC AGG AGC AAG ATC A	AGT GGT GGT CCC ATC TTT CTT A
27	primer 36	ARSFL_133	AAA CTT GAT TGG CGG AGA GA	TTC TGT TTT GAG GCC CAG AC
28	primer 43	ARSFL_92	TCC GGT GAC GAA TCT AAA GG	GAA GAA CAA GCA CCA CCA CA
29	primer 44	ARSFL_96	AGT CTA GGC TGC TTG GGT TG	CCA AGG GAA GAA CAG ACA TGA
30	primer 45	ARSFL_1	GCG GAC CCA TAG CAC ACT GTT GAC	GCG CCT TCC CTT GAT ACA ACT TAC
31	primer 47	ARSFL_3	GCG GGT GCT TAG GTT TTC ACA ACT	GCG CAA GTG GTA TTT AAG GGT TAG
32	primer 48	ARSFL_7	GCG CGC ATA AGG CAA CAA AG	GCG AAT GGC AAT GAC ATC TTC TCT
33	primer 49	FG1a/b	TGGTTTGCCGGTAGCAAAT AGCAGCA	TGACACACACTCTCTCTGTC TGATCCCT
34	primer 50	FG2a/b	TGAACTGGTCCATCGGTGC TGAAA	TGATCACACAATACGCATTA CCAAGCCT
35	primer 51	FG7a/b	GCAGTGCTACATCGACTCA GGTCCAA	ACCAAGGAAGTGCCGAAGT GGGTTT

3.3 DNA isolation:

Leaf samples were collected from all genotypes and stored in -80 freezer. Initially DNA was isolated using CTAB method and later on moved to use of Kit. Leaf tissue was collected and promptly placed in liquid nitrogen. Approximately 1 gm of leaf tissue was taken in a chilled mortar with liquid nitrogen for each sample to make it into a fine powder. The ground tissue was placed in a 15mL centrifuge tube. 5mL of preheated (60°C) CTAB isolation buffer was added. The composition of CTAB is Hexadecyle Trimethyl-Ammonium Bromide, Sodium Chloride, 2M Tris-HCl (p^H 8.0), 0.5M EDTA (p^H 8.0), 2-mercaptoethanol, and distilled water. The samples were then incubated for 30 minutes at 60°C. Then 5mL of a mixture of Chloroform: Isoamyl alcohol (24:1) was added to the above mixture in each sample and shaken vigorously to release the pressure. The samples were centrifuged for 15 minutes at 1301x g Relative centrifuge force (RCF) and the only supernatant was transferred into a clean 15mL tube. 2.5 mL of isopropyl alcohol was used to precipitate the DNA. The samples were placed on ice for several hours to overnight. Then the samples were centrifuged for 15minutes at 956 x g RCF. The supernatant was discarded and the pellet was washed with 5mL of DNA wash solution. DNA wash solution is made up of 7.5M Ammonium Acetate (p^H 7.7), 95% ethanol.

3.4 DNA clean up with RNase:

Ribonuclease is a type of nuclease which helps to degrade RNA into small pieces and make DNA pure during DNA isolation. In a 1.5 µL centrifuge tube, 85µL DNA was taken with 10 µL 10x RNase buffer. 5µL RNase (cat#EN0531, Thermo Fisher) enzyme was added to make it 100µL in total. It was incubated at 37°C for 30 minutes. The incubated solution was run on 1% agarose gel to check the integrity of the DNA.

3.5 Protocol for DNA kit (OMEGA, cat. #D5511-02):

Fresh and frozen tissue samples were ground in liquid nitrogen. 50mg ground tissue was transferred into a nuclease-free 1.5 mL microcentrifuge tube. 400µL SP1 buffer and 5µL RNase A were added and vortexed to mix thoroughly at maximum speed. Samples were incubated at 65 °C for 10 minutes. Samples were mixed twice during incubation by inverting the tube. 140 µL SP2 buffer was added and vortexed to mix thoroughly. Samples were kept on ice for 5minutes. Then samples were centrifuged at maximum speed for 10 minutes. A homogenizer minicolumn was inserted into a 2mL collection tube. The supernatant was transferred carefully to the homogenizer minicolumn. The tube was immediately centrifuged for 2 minutes at maximum speed. Cleared lysate was transferred to a new 1.5mL microcentrifuge tube. Care was taken to keep the insoluble pellet undisturbed. 1.5 mL volumes of SP3 buffer was added and the precipitate may form at this point. A HiBind DNA Mini Colum was inserted into a 2mL collection tube. 650 µL sample was transferred to the HiBind DNA Mini Column. The tube was centrifuged for 1min at maximum speed. The filtrate was discarded and the collection tube was reused. HiBind DNA Mini Column was transferred to a new 2mL collection tube. 650 µL SPW Wash Buffer, diluted with 100% ethanol was added and centrifuged for 1minute at maximum speed. The filtrate was discarded and reused the collection buffer. This step was repeated for another time. The empty HiBind DNA Mini Column was centrifuged for 2 minutes at maximum speed to make the column dry. The HiBind DNA Mini Column was transferred into a clean 1.5 mL microcentrifuge tube. 50-100 µL of preheated elution buffer (65°C for 10 min) was added to the minicolumn. The DNA containing tube was incubated at room temperature for 3-5 minutes and then was centrifuged at maximum speed for 1 minute. To elute the remaining DNA this step was repeated. The eluted DNA was stored at -20°C.

3.6 Quantification of DNA:

The nanodrop Spectrophotometer 2000 (Thermo Fisher Scientific Inc.) was used for quantification. The nanodrop machine was blanked with elution buffer since the DNA samples were resuspended in the elution buffer. 1 µL DNA was used for quantification. Nanodrop calculates the amount of DNA in nanogram (ng) per microliter (µL). The concentrated DNA samples were used as the stock solution. The concentrated solution was diluted for PCR amplification and it was calculated using the equation $C_1V_1=C_2V_2$. 25ng/ µL concentrated solution was made from the concentrated DNA solution and it was again quantified from nanodrop.

3.7 Standard PCR amplification:

To test the use of the selected SSR primers, Polymerase Chain Reaction (PCR) was done for five genotypes using 72 SSR primers and analyzed in 2% Agarose gel. Consistent amplification was obtained from 60 of the 72 primer sets. PCR was done in Plant Molecular Breeding (PMB) lab at Delaware State University by using BIO-RAD T100^{TM} Thermal Cycler. Each 25 µL PCR reaction contained 16dH_2O, 5 µL of 5X Green Gotaq reaction buffer (cat#M3001) containing 7.5Mm concentration of $MgCl_2$, 0.5 µL of forward and reverse primers each, 1 µL dNTPs, 0.25 µL of Taq polymerase and 2 µL of DNA which is taken as a template. The PCR protocol was set for 35 cycles for all primers.

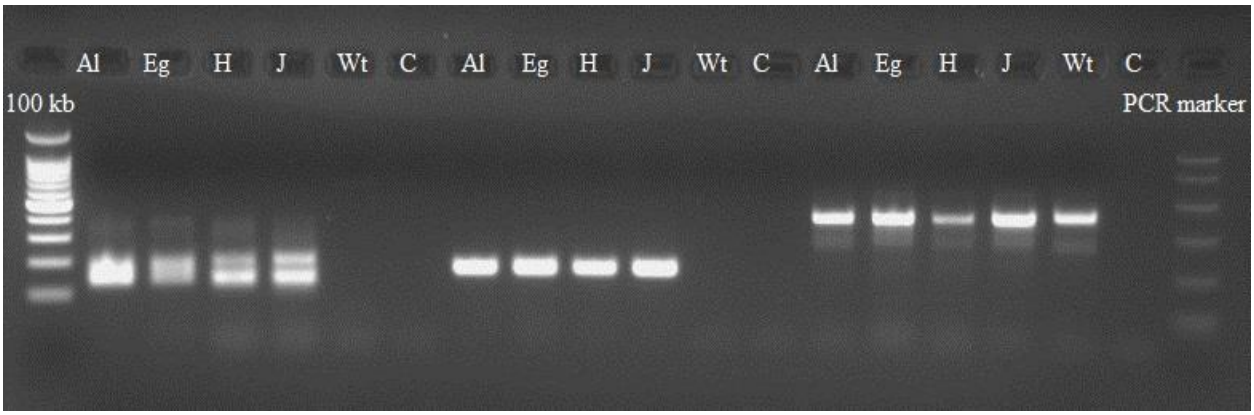

Figure 1. Representative gel picture from regular PCR; Gel picture of SSR markers on 2% agarose gel with standard PCR amplification. The marker shown here are ChFAM 4, FAC 012, & FAC 016 among five selected strawberry genotypes (Al, Eg, H, J and Wt) and negative control (C).

For the PCR program denaturation and extension temperature was 94°C and 72°C, respectively but annealing temperature varied according to primers melting point temperature. PCR products were electrophoresed (Fig 1) for size fractionation. Based on this PCR analysis only 60 SSR primer sets were advanced for AATI analysis.

3.8 Further analysis of SSR markers using fragment analyzer:

The above 60 alleles were further analyzed on 33 genotypes using AATI fragment analyzer. PCR reaction consisted of 50 ng of DNA, 0.20 µM of mixed primers (forward and reverse), 1X buffer consists of 10mM Tris-HCl where pH was 8.2, 50mM KCl, Triton 0.1%, BSA 1mg/ml, 0.2mM dNTPs, 1.5 mM $MgCl_2$, 1U Taq polymerase in 10µL reaction volume was used for PCR reaction. GeneAmp PCR 9700 System thermal cycler (Applied Biosystems Inc.) was used for amplification.

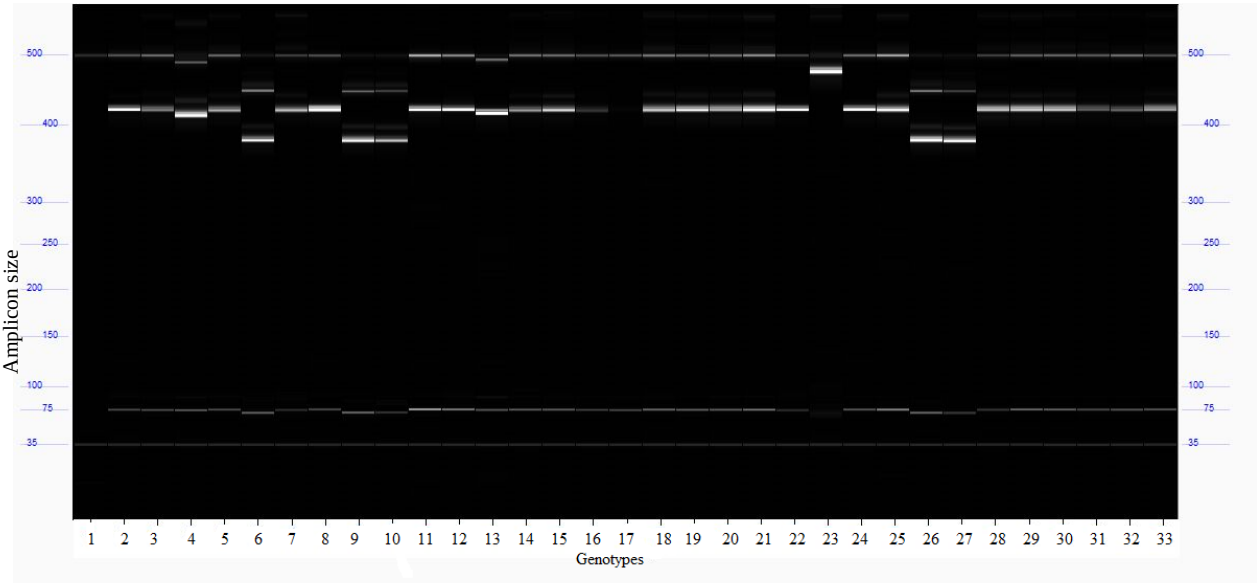

Figure 2. Representative AATI gel picture; amplification pattern of the allele FAC-016 with 33 strawberry genotypes

This machine was programmed to 94 °C for 2minutes, 35 cycles at 94 °C for the 30s, 50-65 °C for 30s, 72 °C for 1 min, and an extension step at 72 °C for 10 minutes. A high throughput DNA fragment analyzer (AdvanCE^{TM} FS) of Advanced Analytical Technologies, Inc. (AATI), Ames, IA 50010 was used to separate amplified products. Amplified PCR products were diluted to 1:11. This dilution depends upon the concentrations of products. The dilution and injection voltage was adjusted to prevent overloading of the PCR product on the fragment analyzer. 22µL of 1X TE dilution buffer was loaded into each well of the sample plate including 2µL of PCR product. A drop of mineral oil was overlaid on each sample to prevent evaporation. 96 capillary automated systems with capillaries of 80cm length were used to separate samples based on the fragment size. Polymer and other reagents are taken from dsDNA kit DNF-900 of Advanced Analytical Technologies (AATI). DNF-900 dsDNA reagent kit was used to separate the amplicon efficiently which ranges from 35bp and 500bp and this kit can resolve 1bp differences

between various alleles. After capillary electrophoresis, the data was processed by PRO Size™ 2.0, the software of AATI. The lower marker was 35bp, the upper marker was 500bp and the data was normalized. Then it was calibrated to the 75-400bp range. Fig 2 shows the amplification of 33 genotypes for primer FAC-016 by using AATI analyzer. The amplification and the polymorphism are more apparent in figure 2 compared to regular PCR amplification shown in figure 1 because of higher resolution.

3.9 Genetic diversity analysis:

Population structure and identification of admixed individuals into different clusters (K) was performed by using the model-based Structure version 2.2 (Prichard et al., 2000). The software uses the posterior probability of the data for a given K, Pr (X|K). The number of populations (K) was determined with a K of 2-6 following the admixture model with correlated alleles. To estimate each value of K five independent runs of 100,000 Markov Chain Monte Carlo (MCMC) generations and 100,000 generations of burn-in were used. Adhoc statistic ΔK (Evanno et al., 2005) was used to determine optimal K value. ΔK values were estimated from the software Structure Harvester which was used to evaluate the number of Ks. Genetic distances were calculated according to Crossa et al., (2004). A dendrogram was build based on the genetic distances using the software MEGA (Tamura et al., 2011). Neighbor-joining algorithm was used to build this dendrogram. Molecular diversity and population structure were analyzed using the analysis of variance (AMOVA) in the program Arlequin 2.0 (Schneider et al., 2000). The genetic variance partitioned between and among the strawberry groups was identified by AMOVA. Genetic diversity and heterozygosity (h) were used to estimate molecular genetic diversity. PopGene version 1.31 (Yeh et al., 1999) was used to estimate F_{IS} (the inbreeding coefficient) and F_{ST} (the proportion of genetic variance in a subpopulation relative to the total population) based

on Wright's F-statistics (Weir et al., 1984). To estimate kinship (K) matrix, marker set was used. TASSEL 3.0 uses a proportion of alleles shared between each pair of accessions in the study which was used to estimate kinship as well. Q matrix was adapted from K-5 which was derived from the analysis obtained from Structure version 2.2.

3.10 Antioxidant assay:

Living organisms have a large number of antioxidants, including macro and micromolecules, and enzymes, which represent the total antioxidant activity of the system, and play a central role in preventing oxidative stress. Strawberries are claimed, to be having plentiful amounts of antioxidants and one of the goals of this project is to measure total antioxidants in the genotypes under study. According to the Sigma-Aldrich's (catalog # CS0790), the principle of the antioxidant assay is formation of a ferryl myoglobin radical from metmyoglobin and hydrogen peroxide, which oxidizes the ABTS (2,2¢-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) to produce a radical cation, ABTS·+ , a soluble chromogen that is green in color and can be determined spectrophotometrically at 405 nm.

HX-FeIII + H2O2 → *X-[FeIV=O] + H2O

ABTS + *X-[FeIV=O] → ABTS**+ + HX-FeIII

In this equation, HX-FeIII is metmyoglobin and *X-[FeIV=O] is ferryl myoglobin. Antioxidants suppress the production of the radical cation in a concentration-dependent manner, and the color intensity decreases proportionally. Trolox, a water-soluble vitamin E analog, serves as a standard or control antioxidant (https://www.sigmaaldrich.com).

Antioxidant assay kit (Sigma-Aldrich, cat. # CS0790) was used for the antioxidant assay. Fresh and frozen leaf tissue was homogenized using mortar and pestle. Then homogenized tissue was placed in ice-cold 1x assay buffer. Around 0.5 mL of buffer was used for ~100gm of tissue. Then the lysate tissue was centrifuged for 15 minutes at 12,000 ×g for 15 minutes at 4°C. The supernatant was removed and kept it on the ice and was stored in -80 for long-term use. Trolox was used as a standard for total antioxidant assay. The first step to assay total antioxidant was to prepare a Trolox standard for a standard curve followed by preparation of ABTS (2,2′-azino-bis 3-ethylbenzthiazoline-6-sulfonic acid) substrate working solution. In 96 well plates, after loading Trolox standard solution (10 µL of a Trolox standard and 20 µL of Myoglobin working solution), test samples (10 µL of the test sample and 20 µL of Myoglobin working solution) were loaded. ABTS substrate working solution (150 µL) was added to each well. Then all samples were incubated for 5minutes at room temperature. Stop solution (100 µL) was added to each well to stop the reaction. Then endpoint absorbance was read at 405 nm using a colorimeter, a plate reader (BioTek Synergy HTX, Software: Gen5) from Dr. Milleti's lab, Biology department at Delaware State University.

3.11 Statistical analysis of antioxidant levels:

A standard curve was prepared by plotting the average absorbance of each Trolox Standard as a function of the final Trolox concentration (mM). The antioxidant concentration of the test sample was calculated using the following equation obtained from the linear regression of the standard curve.

X (mM) = (Y (A405) – Intercept)/slope

X (mM) = Antioxidant concentration [(mM) relative to the concentration of the Trolox standard].

Y (A405) = the average absorbance of the Test Sample at 405 nm

Intercept = intercept of the Y axis by the standard curve

Slope = Slope of the standard curve

R software was used for total antioxidant analysis. Summary statistics and histogram are calculated from R software. Ascending order of total antioxidant was arranged in excel.

3.12 Leaf morphological analysis by Image capturing:

Leaf morphological traits are essential to know about the economic traits and how the traits impact on photosynthesis as well as productivity. In this study, leaves were categorized based on three different traits. Imaging facilities of OSCAR at Delaware State University was utilized for image capturing to see the leaf shape, trichome density, and petiole size. Fresh leaves were collected in 50 mL tubes with some water for each genotype. Then the leaf image was taken by using Nikon SMZ18 Stereo Zoom Microscope. Images from the dorsal and ventral surface were taken to see the petiole size and trichrome density. Full leaf shape image was taken by using DSLR, Canon EOS Rebel T2i. For each trait, 0 .75X zoom pattern was applied for all the genotypes.

3.13 Analysis of leaf image:

Leaf samples were arranged according to their petiole sizes (small, medium, large), leaf shape, and trichome density. All genotypes were grouped together according to their total antioxidant content, leaf shape, petiole size and trichome density to see, whether there is an association among those traits or not.

CHAPTER IV: RESULTS

4.1 Microsatellite alleles:

There are different ploidy levels in strawberry. Allele numbers are varied due to polyploidy in strawberry. The varied allele numbers for the 35 SSR markers studied presented in Table 3.

Table 3. Details of total alleles and Shanon Index; amplified alleles (na), number of effective alleles (ne), Shannon index (I) and number of SSR markers used in the current study.

Locus	na*	ne*	I*
ARSFL_11	4	2.9811	1.1706
ChFaM-023	2	1.9139	0.6705
EMFn170	2	1.3846	0.4506
EMFn181	3	1.8318	0.6981
EMFn182	5	2.3873	1.1201
EMFv104	3	1.8392	0.7739
EMFvi136	4	2.6732	1.1137
EMFvi166	2	1.198	0.3046
ARSFL_9	2	1.0689	0.1461
ARSFL_10	4	2.5903	1.06
ARSFL_4	2	1.9978	0.6926
ARSFL_12	4	2.7638	1.15
ARSFL_15	2	1.9321	0.6755
ARSFL_17	2	1.4322	0.4792
ARSFL_97	2	1.7076	0.6049
ARSFL_35	3	1.1834	0.3285
ARSFL_33	4	1.4376	0.6416
FAC-001	6	3.3081	1.4067
ARSFL_31	3	2.9569	1.0914
FAC-007	3	2.9468	1.0896
ARSFL_100	5	2.2124	1.0555
ARSFL_101	4	1.5669	0.7298
ARSFL_30	4	1.4681	0.6179
FAC-002	3	1.1353	0.2771
FAC-004d	5	2.9801	1.1944
FAC-016	4	1.4851	0.6754
ARSFL_133	3	1.5531	0.6397
ARSFL_92	4	2.0516	0.8656

ARSFL_96	2	1.8221	0.6435
ARSFL_1	2	1.4506	0.4896
ARSFL_1	3	1.218	0.3795
ARSFL_3	3	2.2456	0.8815
ARSFL_7	2	1.9122	0.67
FG1a/b	6	4.7868	1.6635
FG2a/b	4	2.9196	1.1327
FG7a/b	4	2.5496	1.1369

In total, 120 alleles were amplified from 33 genotypes. On an average 3.43 alleles from each marker was found. Shannon indices (I*) are the measurement of polymorphism among all those alleles (Singh et al., 2015.). I* ranges from 0.1461 for the marker ARSFL_9 to 1.6635 for FG1a/b (Table 3). Marker FG1 a/b is more polymorphic compared to other markers. It is also visible from the observed and effective allele numbers from the table. The second highest polymorphic marker is FAC001.

4.2 Genetic diversity and population structure

NJ tree was derived using MEGA 5.0 software. NJ tree represents phylogenetic distances. Based on the NJ tree we can observe that the cultivars under study were classified into two distinct groups, for convenience we call them red and blue groups (Fig.3). Mignonette, Red wonder, Attila, Reine des Vallees, Vanilla yellow, Fragola Quattro Stagioni, & Alexandria were found in the red cluster. They all are alpine type strawberries. *Fragaria vesca* is the ancestor of these genotypes. These genotypes were found in Europe as cultivated for hundred years. They are the diploid type, and their fruit size is small but sweet. Other twenty-six genotypes are found in the blue cluster. All these genotypes are octoploid and they were derived from a cross between *Fragaria chiloensis* and *Fragaria virginiana.*

A number of population (K) is characterized by a set of allele frequency at each locus and an individual having more than 70% of its genome is assigned as a group or cluster (Yoon et al., 2012.) In our experiment, we have got three groups (Fig 4) where group one (blue in color) comprised of seven genotypes which are same as red cluster I found from NJ tree and their ancestry is from *Fragaria vesca.* Fig 4 also represents that eight genotypes fall under group II where all genotypes are octoploid in nature and their ancestry is from *F. chiloensis* or *F. virginiana*. Group III (18 genotypes) represents the admixed that represents their ancestors are similar to either group I or Group II.

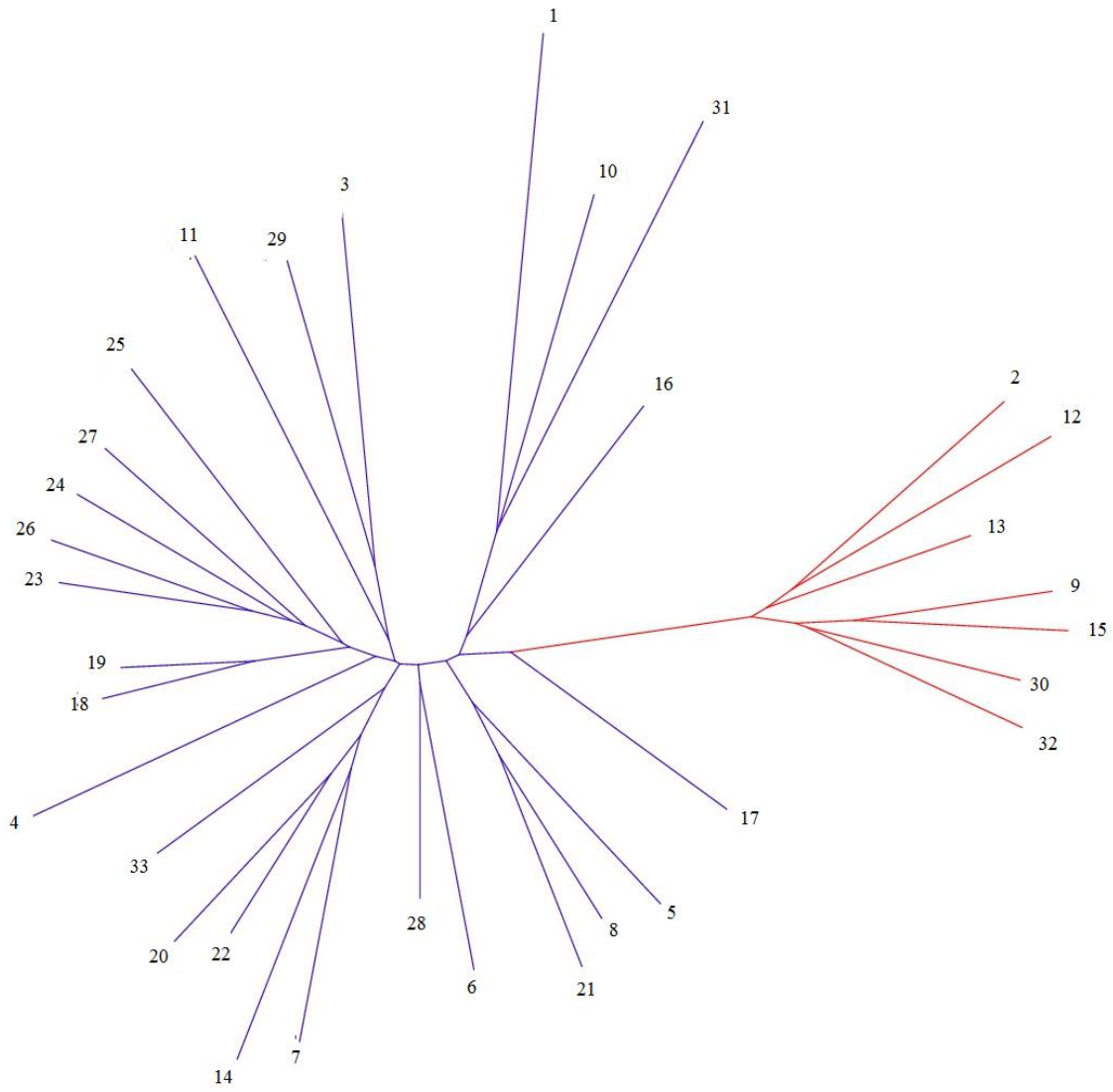

Figure 3. Neighbor-joining (NJ) tree; depicting estimated relationship among 33 strawberry genotypes. Two different colors for two different groups. Red color represents the diploid genotypes, and blue color represents the cultivated strawberry genotypes. The Number of tree branches refers to the accession numbers.

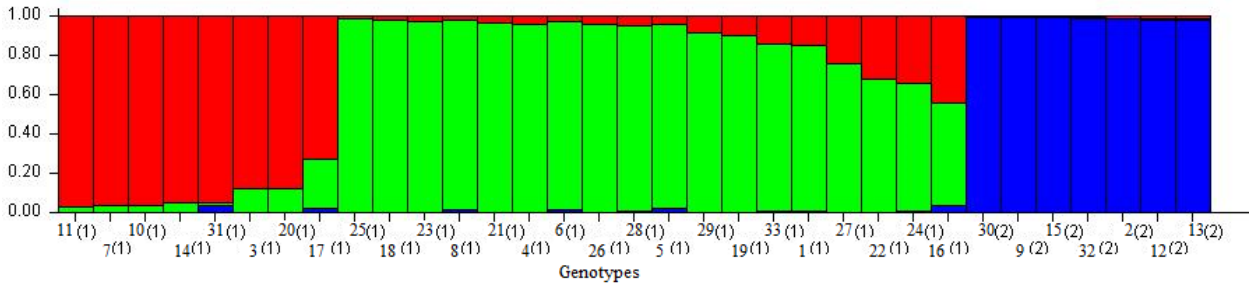

Figure 4. Shared ancestry among 33 genotypes by population structure analysis. The number in parenthesis refers to cluster and number outside parenthesis refers to the accession number.

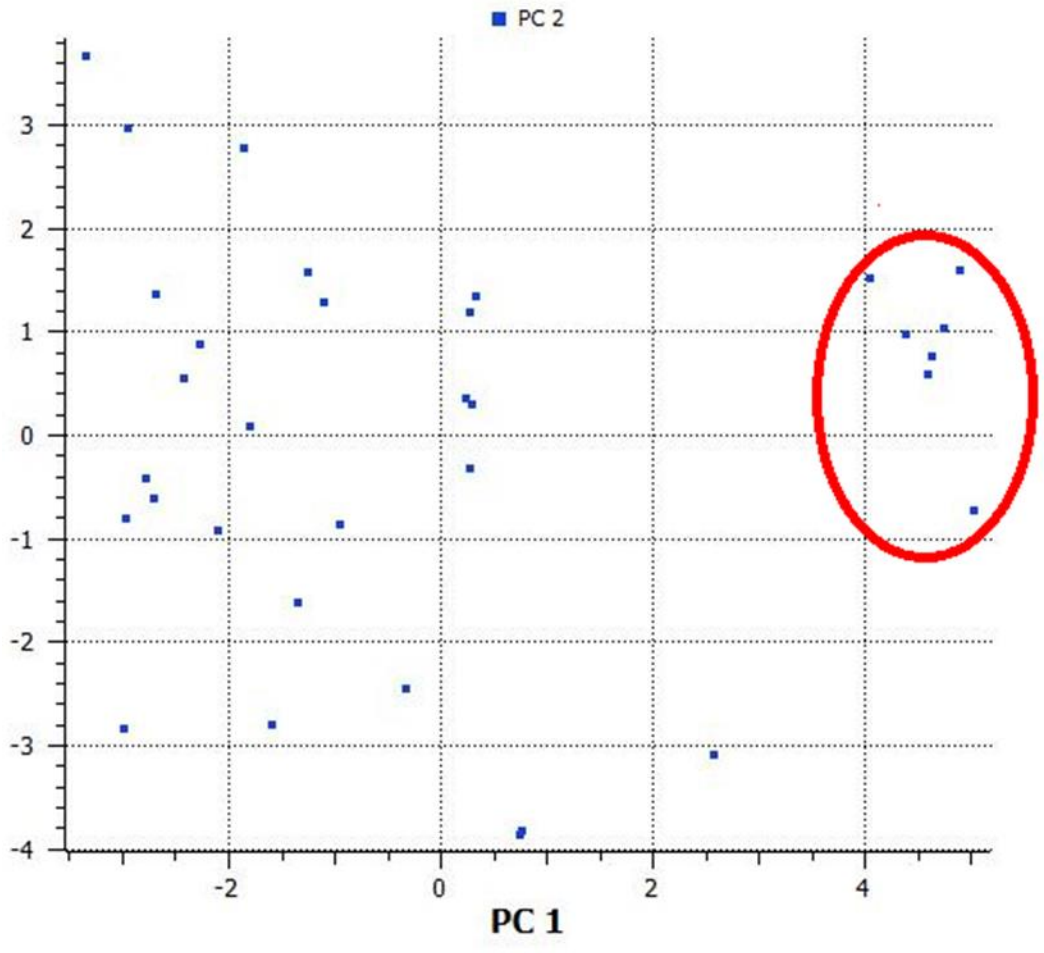

Figure 5. Principal component analysis (PCA) of wild-type and cultivated type of strawberry

PCA revealed closeness of cluster I and cluster II.

NJ tree and PCA are very important to track the unknown accessions (Singh et al., 2015). PCA can help visualizing the maximum variability among populations. PCA revealed closeness of cluster I and cluster II.

The results obtained from NJ tree and population structure completely agreed with PCA analysis. Population structure analysis based on K. Population structure analysis is performed using the model-based assumptions and was used to estimate clusters K-2 to K-6. The results were analyzed for mean ±SD LnP (K) (Table 4). The ΔK value was calculated using Structure Harvester. The highest ΔK value for the particular cluster among all clusters represents the most genetic diversity. So, cluster K-3 was the most appropriate cluster size where ΔK (19.82) value is the highest (Table 4 &Fig.6).

Table 4. The Evanno table output

#K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
2	3	-1619.83	1.4572	NA	NA	NA
3	3	-1600.03	2.3459	19.8	46.5	**19.82166**
4	3	-1626.73	30.287	-26.7	7.5	0.247631
5	3	-1660.93	27.508	-34.2	15.66667	0.56953
6	3	-1710.8	22.558	-49.8667	NA	NA

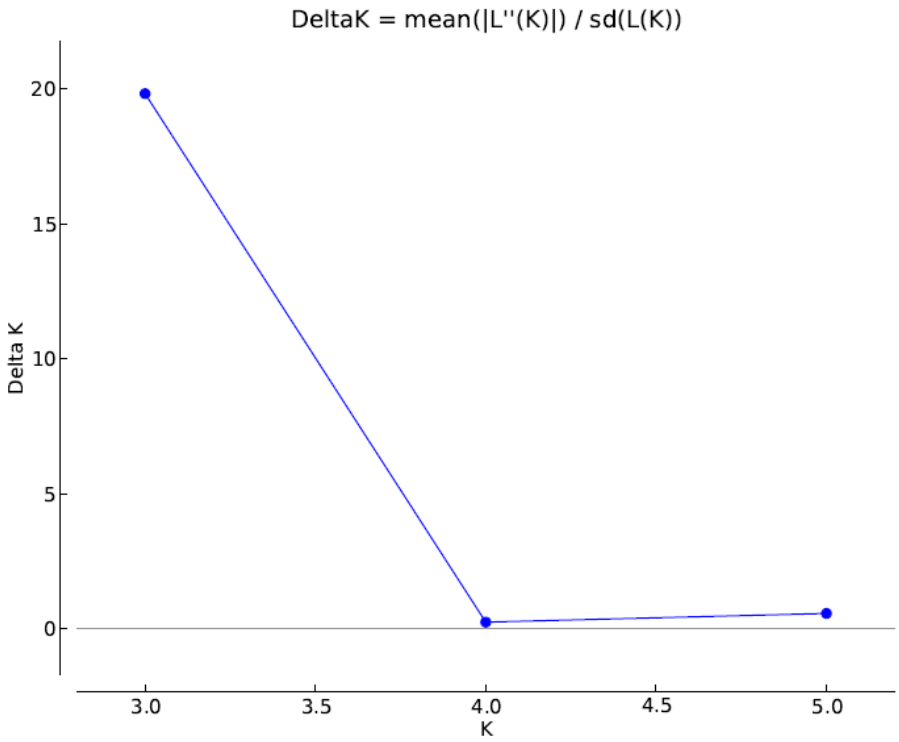

Figure 6. Relationship between K & ΔK

To analyze variance among and within cluster we used AMOVA (Table 5). Variation among population means variation within main two clusters while variation within population means variation within all genotypes. The percentage of variation among the population and within the population was 19.98 and 80.02. The variance was lower among population (sub-populations) compared to within population.

Table 5. Analysis of molecular variance (AMOVA)

Source of variation	df	Sum of squares	Variance components	Percentage of variation
Among Population	1	20.935	0.80318 Va	19.98
Within Population	64	205.838	3.21622 Vb	80.02
Total	65	226.773	4.0194	

4.3 Gene flow and population differentiation

Gene flow is the movement of genes or alleles from one population to another population. Gene flow can be a parameter to estimate genetic diversity. Molecular genetic diversity was estimated from heterozygosity (h) (Table 6). The more heterozygosity represents, the more genetic diversity.

Table 6. Level of expected heterozygosity

Locus	Obs_Hom	Obs_Het	Exp_Hom*	Exp_Het*	Nei**	Ave_Het
ARSFL_11	0.1875	0.8125	0.3249	0.6751	0.6646	0.5058
ChFaM-023	0.2121	0.7879	0.5152	0.4848	0.4775	0.25
EMFn170	0.8	0.2	0.7175	0.2825	0.2778	0.3316
EMFn181	0.6061	0.3939	0.5389	0.4611	0.4541	0.4275
EMFn182	0.5	0.5	0.409	0.591	0.5811	0.2359
EMFv104	0.3871	0.6129	0.5362	0.4638	0.4563	0.3971
EMFvi136	0.6129	0.3871	0.3638	0.6362	0.6259	0.4467
EMFvi166	1	0	0.8322	0.1678	0.1653	0.1021
ARSFL_9	1	0	0.9345	0.0655	0.0644	0.0399
ARSFL_10	0.2903	0.7097	0.376	0.624	0.6139	0.5066
ARSFL_4	0.1667	0.8333	0.4921	0.5079	0.4994	0.4596
ARSFL_12	0.875	0.125	0.3517	0.6483	0.6382	0.4516
ARSFL_15	0.1875	0.8125	0.5099	0.4901	0.4824	0.4168
ARSFL_17	0.6296	0.3704	0.6925	0.3075	0.3018	0.1814
ARSFL_97	0.4138	0.5862	0.5783	0.4217	0.4144	0.2371
ARSFL_35	0.8333	0.1667	0.8424	0.1576	0.155	0.0988
ARSFL_33	0.8621	0.1379	0.6903	0.3097	0.3044	0.3485
FAC-001	0.2903	0.7097	0.2909	0.7091	0.6977	0.4481
ARSFL_31	0.4194	0.5806	0.3273	0.6727	0.6618	0.2976
FAC-007	0.625	0.375	0.3289	0.6711	0.6606	0.5269
ARSFL_100	0.7037	0.2963	0.4416	0.5584	0.548	0.5872
ARSFL_101	0.6364	0.3636	0.6326	0.3674	0.3618	0.2189
ARSFL_30	0.6875	0.3125	0.6761	0.3239	0.3188	0.1882
FAC-002	1	0	0.879	0.121	0.1191	0.0752
FAC-004d	0.3333	0.6667	0.3243	0.6757	0.6644	0.294
FAC-016	0.8438	0.1562	0.6682	0.3318	0.3267	0.2004
ARSFL_133	0.7	0.3	0.6379	0.3621	0.3561	0.3885
ARSFL_92	0.3077	0.6923	0.4774	0.5226	0.5126	0.2661
ARSFL_96	0.3125	0.6875	0.5417	0.4583	0.4512	0.463
ARSFL_1	1	0	0.6833	0.3167	0.3107	0.1814
ARSFL_3	0.0625	0.9375	0.4365	0.5635	0.5547	0.557

ARSFL_7	0.2143	0.7857	0.5143	0.4857	0.477	0.2495
FG1a/b	0.0303	0.9697	0.1967	0.8033	0.7911	0.6154
FG2a/b	0.303	0.697	0.3324	0.6676	0.6575	0.5125
FG7a/b	0.6429	0.3571	0.3812	0.6188	0.6078	0.3553
Mean	0.5466	0.4534	0.5359	0.4641	0.4565	0.3351
St. Dev	0.2982	0.2982	0.1887	0.1887	0.1857	0.1546

* Expected homozygosity and heterozygosity were computed using Levene (1949)

** Nei's (1973) expected heterozygosity

The number of polymorphic loci is: 35

The percentage of polymorphic loci is: 100.00 %

The heterozygosity ranges from 0.0655 for the allele ARSFL_9 to 0.8033 for FG1a/b allele (table 6). All the observed alleles obey the Hardy Weinberg equilibrium. The combination of $F_{ST,}$ expected heterozygosity, and gene flow estimates the comprehensive overview of gene flow. The highest combination for F_{ST} and Nm was found for ChFaM-023, FAC-001, ARSFL-31, FAC-004d, ARSFL_7, FG7a/b (Table 7). These SSR markers indicate their essential role in population differentiation and gene flow among all genotypes.

Table 7. Summary of F-Statistics and Gene Flow for All Loci

Locus	**Fis**	**Fit**	**Fst**	**Nm***
ARSFL_11	0.1154	0.1462	0.2346	0.8159
ChFaM-023	-1	-0.3333	0.3333	0.5
EMFn170	0.623	0.6462	0.0614	3.82
EMFn181	0.0489	0.2324	0.1929	1.046
EMFn182	0.2226	0.5258	0.6122	0.1584
EMFv104	-0.379	-0.2606	0.0859	2.6606
EMFvi136	0.4403	0.5411	0.1801	1.1384
EMFvi166	1	1	0.0612	3.8333
ARSFL_9	1	1	0.0213	11.5
ARSFL_10	0.0064	0.1271	0.1215	1.8082
ARSFL_4	-0.262	-0.1988	0.05	4.7479

ARSFL_12	0.8229	0.8557	0.1853	1.0989
ARSFL_15	0.6179	-0.5086	0.0675	3.4515
ARSFL_17	0.3125	-0.1351	0.1351	1.6
ARSFL_97	0.6296	-0.2394	0.2394	0.7941
ARSFL_35	0.1005	-0.0478	0.0478	4.9762
ARSFL_33	0.4596	0.4759	0.0302	8.0389
FAC-001	0.0181	0.3536	0.3417	0.4816
ARSFL_31	0.2097	0.3886	0.4946	0.2555
FAC-007	0.5445	0.6159	0.1568	1.3443
ARSFL_100	0.5588	0.5669	0.0183	13.4319
ARSFL_101	0.0541	0.0545	0.103	2.1765
ARSFL_30	0.0216	0.0821	0.1015	2.213
FAC-002	1	1	0.0309	7.8333
FAC-004d	0.4791	0.2663	0.504	0.246
FAC-016	0.501	0.5459	0.0899	2.5303
ARSFL_133	0.0346	0.1268	0.0955	2.3677
ARSFL_92	0.4095	-0.0985	0.2206	0.8833
ARSFL_96	0.7444	-0.6774	0.0384	6.26
ARSFL_1	1	1	0.1351	1.6
ARSFL_3	1	1	0.019	12.9247
ARSFL_7	0.9167	-0.3143	0.3143	0.5455
FG1a/b	0.5938	-0.2593	0.2099	0.9412
FG2a/b	0.137	0.3084	0.1986	1.009
FG7a/b	0.3604	0.6579	0.4652	0.2874
Mean	0.0292	0.1839	0.207	0.9577

The Fst is the proportion of genetic variation of subpopulation to the total population. The overall F_{ST} was 0.19983 among the collections studied.

F_{IS} is the inbreeding coefficient which is based on Wright's F-statistics (Weir & Cockerham, 1984) and Nm is the migratory allele presented in table 7.

4.4 Phenotypical traits:

We started phenotyping with 42 genotypes; however analysis was performed for only 33 genotypes in concurrence with genotyping. There were four morphological traits which were observed in this research; total antioxidant, leaf shape, petiole size, and trichome density.

4.5 Total antioxidant content:

Leaf samples were used to assay total antioxidant from all genotypes. Five genotypes were found to have relatively higher antioxidant content. The high antioxidant containing genotypes are Early Glow, Wendy, Elan Hybrid, Clancy, & Record (Table 8).

Table 8. Total antioxidant content (arranged in ascending order) in mM Trolox equivalents for 33 strawberry genotypes.

No.	Genotype name	TA (mM)
25	Sureccrop	0.44724636
23	Seascape	0.518998126
20	Cardinal	0.52198794
26	Tennessee Beauty	0.531589661
3	Annapolis	0.544455967
27	Ozark	0.552905481
15	Vanilla yellow	0.555017859
10	Red Chef	0.556938203
13	Reine des Vallees	0.558282444
28	Jewel	0.560582832
30	Fragola Quattro Stagioni	0.562162776
17	Wonderful Pineberry	0.570678932
29	Cabot	0.571925364
7	Cavendish	0.57276462
6	Tribute	0.583177145
19	Albion	0.583392539
9	Red wonder	0.586383976
24	Guardian	0.589873986
31	Sequoia	0.590749891
12	Attila	0.592616136
14	Yambu	0.59548293
2	Migno nette	0.598225081

21	Valley Sunset	0.602338307
16	Eversweet	0.602961523
18	Honeoye	0.603834025
8	Mesabi	0.606326889
5	Sparkle	0.606950105
32	Alexandria	0.841665344
22	**Earliglow**	**1.081045337**
4	**Wendy**	**1.092948209**
33	**Elan Hybrid**	**1.092948209**
1	**Clancy**	**1.108289689**
11	**Record**	**1.131037401**

The highest antioxidant containing genotype is Record. It contains 1.13 mM Trolox equivalents where the lowest antioxidant containing genotype is Surecrop which contains only 0.447246 mM Trolox equivalents.

Table 9. Summary of statistics

Mean	SD	IQR	0%	25%	50%	75%	100%
0.658054	0.1986696	0.04555158	0.4472464	0.5582824	0.586384	0.603834	1.131037

N = 33

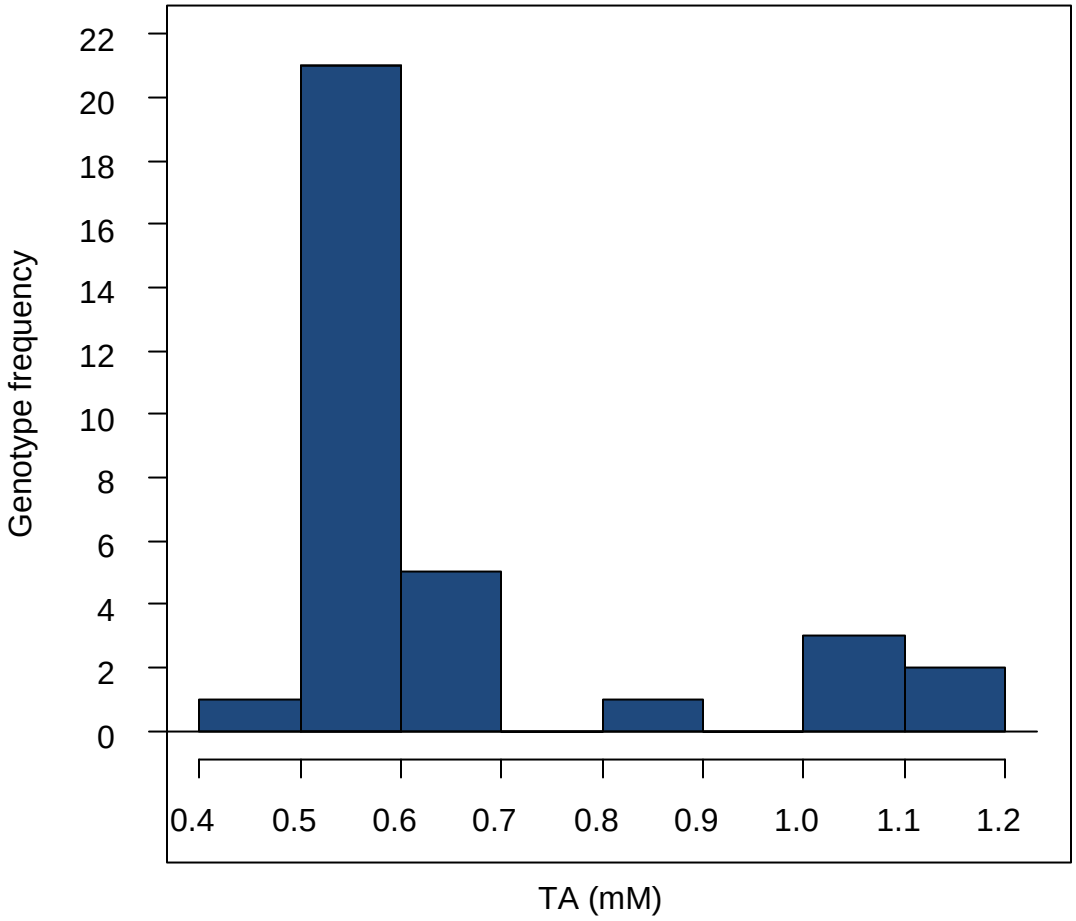

Figure 7. The relationship between total antioxidant and genotype frequency.

The majority of the genotypes (21) studied have the total antioxidant ranges from 0.5 to 0.6 mM trolox equivalents. The second major group (5 genotypes) contains 0.6 to 0.7 mM trolox equivalents. Another five genotypes contain total antioxidant content ranges from 1.0 to 1.2 mM. One genotype contains 0.4 to 0.5 mM and another contains 0.8 to 0.9 mM trolox equivalents.

4.6 Genotype classification based on the leaf shape:

In general, genotypes were classified according to their similarity in leaf shape. The strawberry genotypes in this study were classified into five groups according to their leaf shape (Table 10 & Fig 8). Strawberry leaves are lobed in their structure. The leaves of the group I look like oblique

in shape where one side has asymmetrical leaf base and another side is lower than the other side (Fig 8a).

Table 10. Genotype classification based on the leaf shape

Group I (Oblique)	**Group II (Orbicular)**	**Group III (Ovate or egg-shaped)**	**Group IV (Obovate)**	**Group V (Oblanceolate)**
Record	Ever sweet	Attila	yambu	Migno nette
Honeoye	Mesabi	Albion	Valley sunset	Reine des valles
Seascape	Sure crop	Fragola Quattro stagioni	Clancy	Cardinal
Red wonder	Tennessee	Sequioa		
Ozark	Guardian	Vanilla yellow		
Cabot	Annapolis			
Cavendish	Early glow			
Jewel	Sparkle			
Elan hybrid	Tribute			
Alexandria	Red chef			
Wonderful pineberry				
Wendy				

Fig 8a) Group I (Oblique)

Figure 8. Genotype classification based on the leaf shape; 8a) Group I (Oblique), 8b) Group II (Orbicular), 8c) Group III (Ovate or egg-shaped), 8d) Group IV (Obovate or ovate), 8d) Group IV (Obovate or ovate)

In group two leaf shape looks like orbicular which means leaves are circular in shape (Fig 8b).

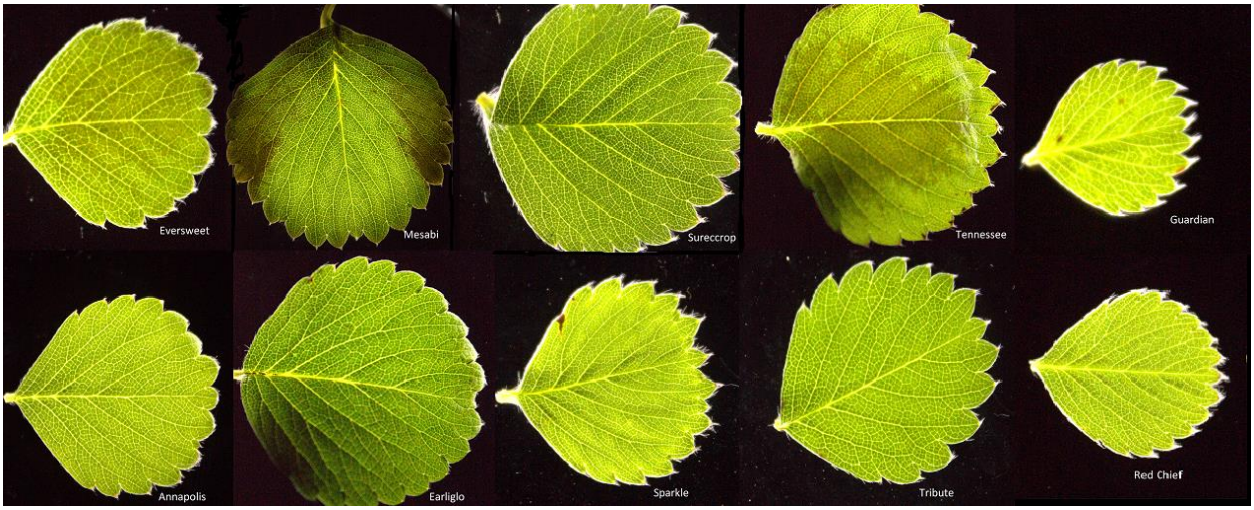

Fig 8b) Group II (Orbicular)

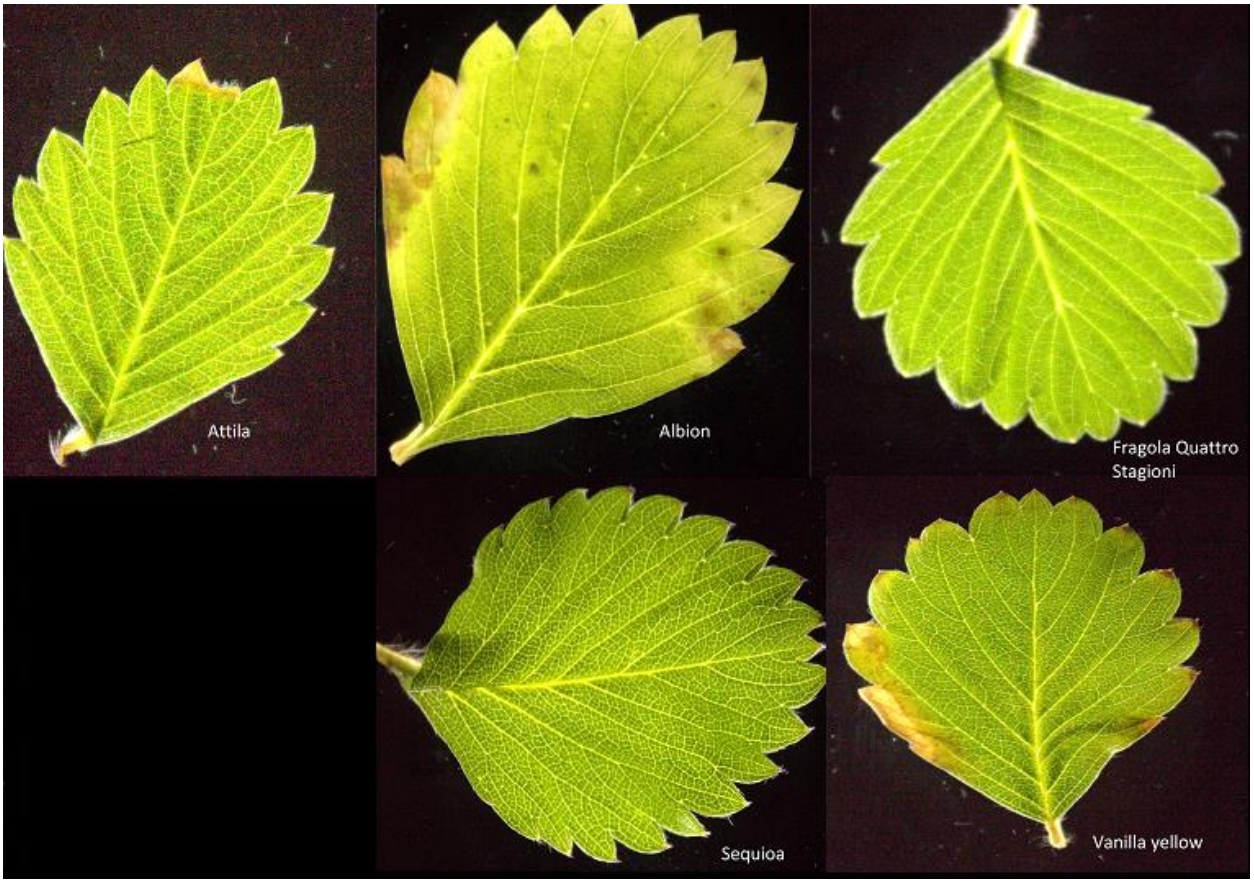

Fig 8c) Group III (Ovate or egg-shaped)

In group three leaves are oval or egg-shaped, with a tapering point which is called ovate in shape (Fig 8c).

Group four fall in between obovate or ovate in shape (Fig 8d).

Fig 8d) Group IV (Obovate or ovate)

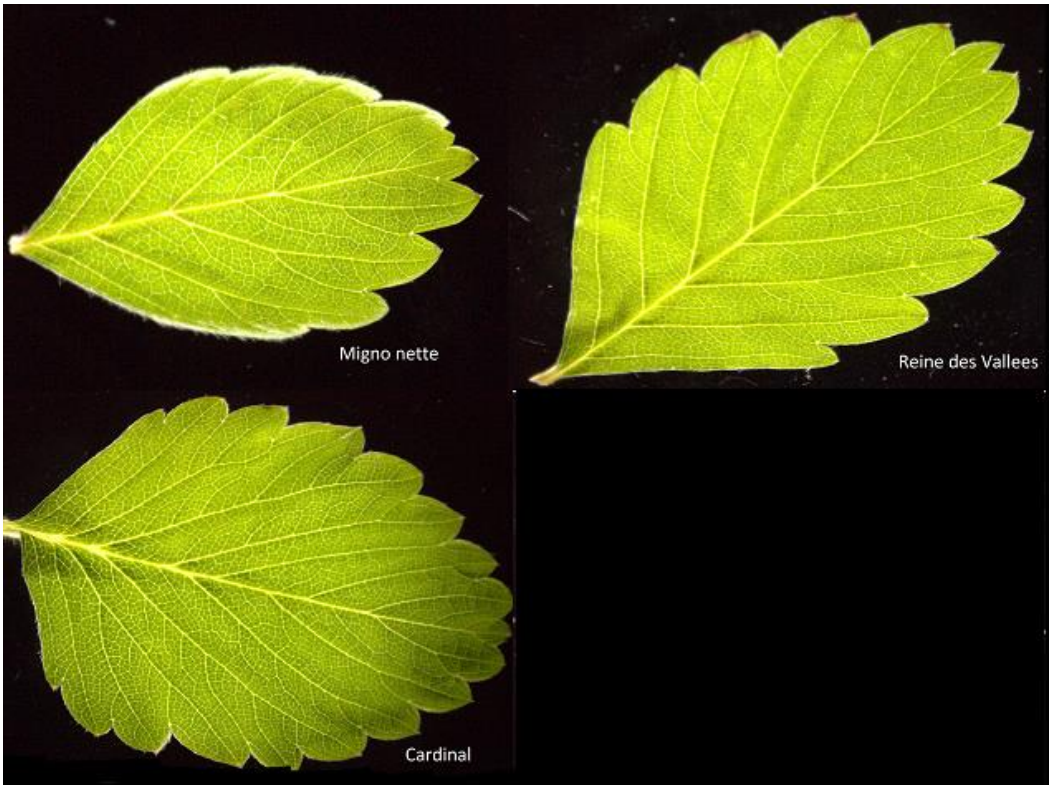

Fig 8e) Group V (Oblanceolate)

Similarly, group five falls in between oblanceolate or ovate in shape.

4.7 Genotype classification based on the leaf petiole size:

The structure that attaches leaf base to the stem is petiole. Petiole serves to transport photosynthates synthesized in plants to the rest of the plant and absorbed water and nutrients by the roots to the leaf blades. In some plants, it also contributes to photosynthesis and leaf senescence and fall in deciduous plants. Care was taken while harvesting petiole to detach from the base of the petiole that attaches to the stem to avoid any errors.

Based on the petiole size the genotypes are categorized into three; large, medium, & small. Seven genotypes fall into the large petiole size group, nine genotypes fall into the medium-size category, and seventeen genotypes fall in small petiole size group (Fig 9a, 9b, 9c & Table 11).

Table 11. Genotype classification based on the petiole size

Group I (large)	**Group II (medium)**	**Group III (small)**
Alexandria	Cavendish	Albion
Cardinal	Clancy	Attila
Earliglow	Fragola Quattro Stagioni	Cabot
Jewel	Guardian	Elan hybrid
Mesabi	Honeoye	Eversweet
Tennessee beauty	Ozark	Mignonette
Wonderful pineberry	Sequioa	Record
	Tribute	Red chef
	Vanilla yellow	Red wonder
		Reine des valles
		Seascape
		Sparkle
		Surecrop
		Valley sunset
		Wendy
		Yambu

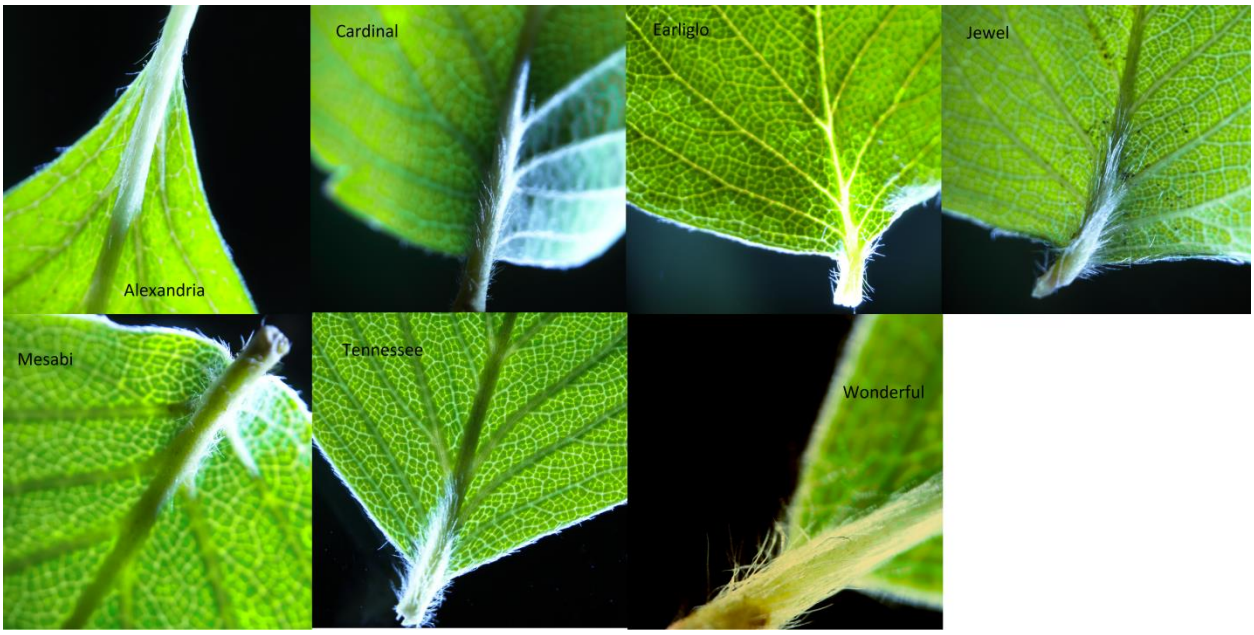

Fig 9a) Group I (large petiole size)

Figure 9. Genotype classification based on the leaf petiole size; 9a) Group I (large petiole size), 9b) Group II (medium petiole size), 9c) Group III (small petiole size)

Fig 9b) Group II (medium petiole size)

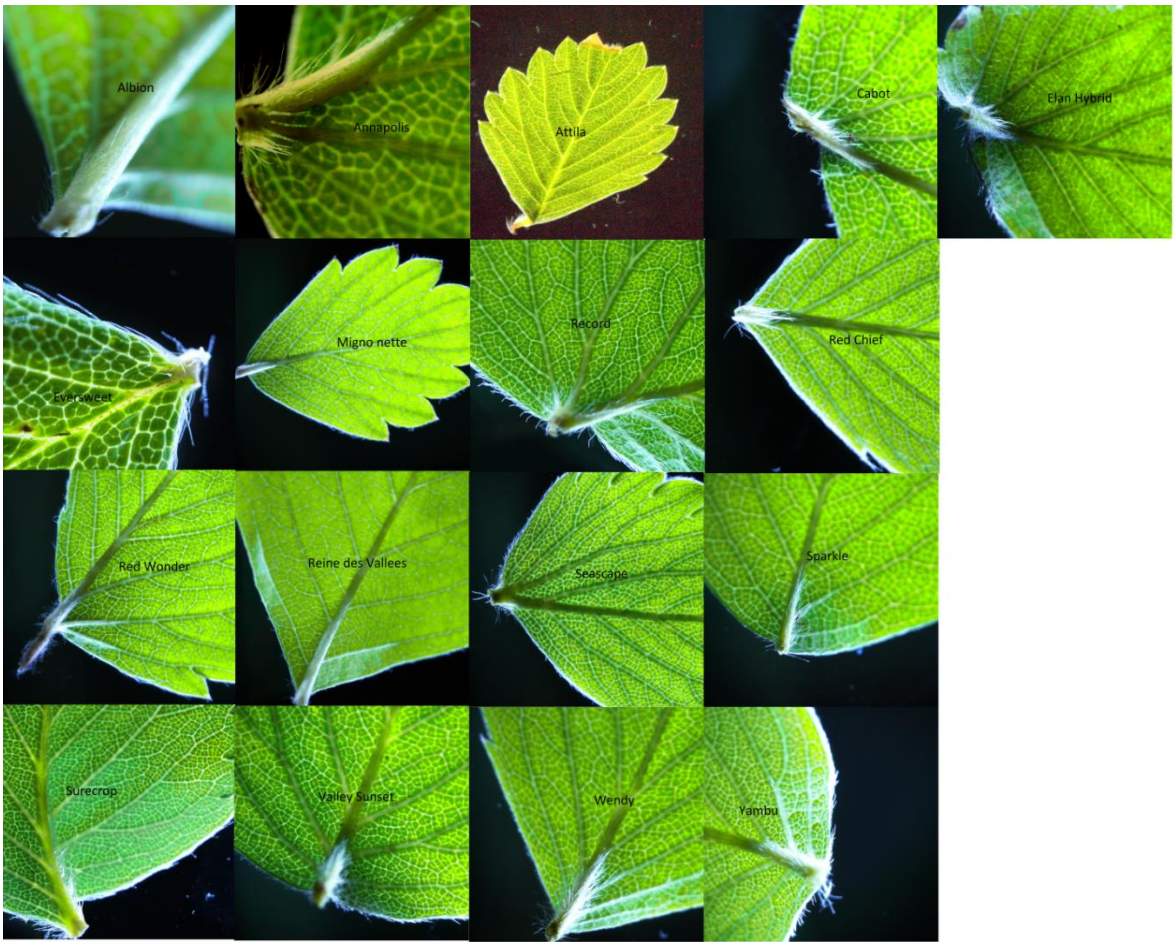

Fig 9c) Group III (small petiole size)

4.8 Genotype classification based on the leaf trichome density:

Small hair like structure from the epidermis of a plant is known as trichome. Based on the trichome level, all genotypes under this study were classified into three groups. These are high density, medium density, and low-density trichome level genotype group. Again all leaves were observed under similar zoom level. Classifications were done based on eye perception. Twelve genotypes fall into high-density group, six genotypes are under medium density group, and thirteen genotypes are under low-density group (Table 12).

Table 12. Genotype classification based on the leaf trichome density

Group I (high)	Group II (medium)	Group III (low)
Annapolis	Fragola quattro stagioni	Albion
Cavendish	Guardian	Alexandria
Clancy	Mignonettee	Cabot
Valley sunset	Tennessee beauty	Cardinal
Elan hybrid	Wonderful pineberry	Mesabi
Eversweet	Yambu	Ozark
Honeoye		Record
Jewel		Reine des valles
Red chef		Seascape
Red wonder		Sparkle
Surecrop		Wendy
Tribute		Earliglow
		Vanilla yellow

One representative leaf is shown in enlarged picture for each category. Sure crop (Fig 10a) is the representative of high-density trichome level, Yambu (Fig 10b) is the representative of medium density trichome, and Wendy (Fig 10c) is the representative of low density trichome.

Fig 10a) High-density trichrome group

Figure 10. Genotype classification based on the leaf trichome density; 10a) Group I, 10b) Group II, 10c) Group III

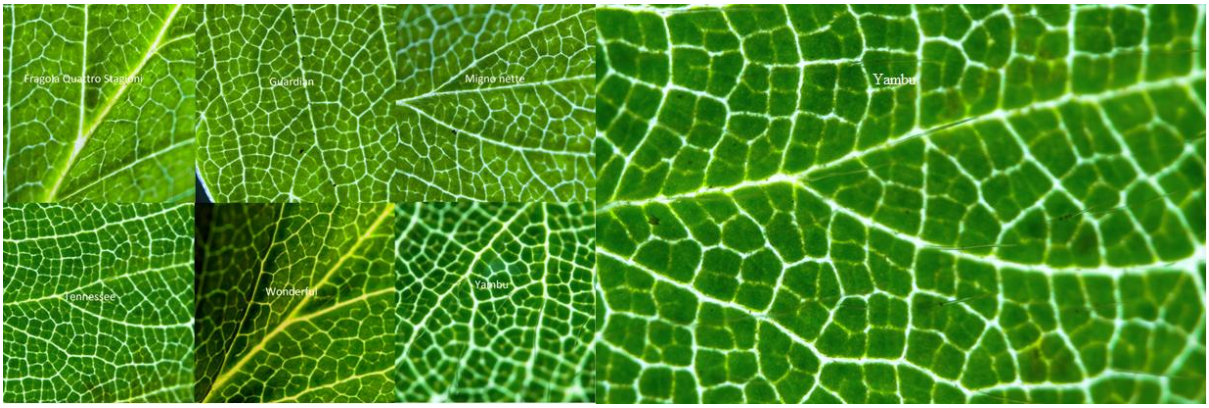

Fig 10b) Medium density trichrome group

Fig 10c) Low-density trichrome group

In our study, we associated all genotypes according to total antioxidant content, leaf shape, trichome density & petiole size (Table 13). We found only few genotypes are good with high trichome density and high total antioxidant content.

Table 13. The relationship among all genotypes according to total antioxidant content, Leaf shape, trichome density and petiole size.

Genotype No.	Genotype name	TA (mM)	Leaf shape	Trichrome density	Petiole size
25	Sureccrop	0.44724636	orbicular	high	small
23	Seascape	0.518998126	oblique	low	small
20	Cardinal	0.52198794	oblanceolate	low	large
26	Tennessee Beauty	0.531589661	orbicular	Medium	large
3	Annapolis	0.544455967	orbicular	high	small
27	Ozark	0.552905481	oblique	low	medium
15	Vanilla yellow	0.555017859	Egg-shaped	low	medium
10	Red Chef	0.556938203	orbicular	high	small
13	Reine des Vallees	0.558282444	oblanceolate	low	small
28	Jewel	0.560582832	oblique	high	large
30	Fragola Quattro Stagioni	0.562162776	ovate	medium	medium
17	Wonderful Pineberry	0.570678932	oblique	medium	large
29	Cabot	0.571925364	oblique	low	small
7	Cavendish	0.57276462	oblique	high	medium
6	Tribute	0.583177145	orbicular	high	medium
19	Albion	0.583392539	ovate	low	small
9	Red wonder	0.586383976	oblique	high	small
24	Guardian	0.589873986	orbicular	medium	medium
31	Sequoia	0.590749891	ovate	NA	medium
12	Attila	0.592616136	ovate	NA	small
14	Yambu	0.59548293	Obovate	medium	small
2	Migno nette	0.598225081	oblanceolate	medium	small
21	Valley Sunset	0.602338307	Obovate	NA	small
16	Eversweet	0.602961523	orbicular	high	small
18	Honeoye	0.603834025	oblique	high	medium
8	Mesabi	0.606326889	orbicular	low	large
5	Sparkle	0.606950105	orbicular	low	small
32	Alexandria	0.841665344	oblique	low	large
22	**Earliglow**	**1.081045337**	orbicular	**low**	**large**
4	**Wendy**	**1.092948209**	oblique	**low**	**small**
33	**Elan Hybrid**	**1.092948209**	oblique	**high**	**small**
1	**Clancy**	**1.108289689**	Obovate	**high**	**medium**
11	**Record**	**1.131037401**	oblique	**low**	**small**

CHAPTER V: DISCUSSION

Our study focused on genotyping of strawberry germplasm by SSR markers, genetic diversity analysis and phenotyping of these accessions by leaf traits. We have utilized thirty-three strawberry genotypes in this study. We provided a detailed analysis of population genetic structure for cultivated and wild-type strawberries. With the advantage of next-generation sequencing technologies and with the availability of latest bioinformatics tool, a significant amount of genomic information has been generated and made it available to the public. As a result, several uncharacterized SSR markers are available for strawberry similar to many crop species. Additionally, not all the identified markers will yield polymorphism among genotypes. Hence, there is a need for characterizing these markers and make them available for genetic analysis and for trait identification. Majority of our SSR markers had not been characterized previously. In this study, we started with 72 markers (have varying repeat elements and many of them are uncharacterized), after the initial screen only 60 were advanced, and with AATI analysis only 35 markers were found to be good.

In our research, we found 120 alleles from 35 SSR primers. Results of NJ tree and structure analysis showed clear divergence among 33 genotypes and these accessions were grouped into 5 clusters. Introgressive hybridization is of a great interest to plant breeders because it may produce new genotypes (Yoon et al., 2012). Increased genetic diversity leads to new adaptation and new ecotypes. A new combination of genes resulting from hybridization and introgression between wild-type and cultivated type is very important in a breeding program to domesticate crop species (Jarvis et al.,1999).

Strawberry is vegetatively propagated plant and it can be easily misidentified based on phenotype (Sargent et al., 2006). Less genetic divergence is found in strawberry due to its

vegetative propagation. Because of its complicated ploidy level, most of the alleles were shared through cultivated species (Yoon et al., 2012). According to Yoon et al., (2012), genetic characterization is difficult because of its limited information about genome structure.

The use of SSR is beneficial to solve the population structure in many other crops and it is proven. Novel SSR markers were developed by Sargent et al., (2006) to assess the genetic diversity and population structure. Based on the previous studies, the microsatellite markers are more powerful to resolve population structure than SNPs (Emanuelli et al., 2013; Ohashi & Tokunaga, 2003), they are also easy to use and develop usable PCR markers than SNPs.

SSRs has proven to be very locus-specific, highly reproducible, and very highly polymorphic (Powell et al., 1996) markers. To study germplasm characterization and conservation, assessment of genetic diversity is very important. SSR primers can help to differentiate between wild and cultivated species (Singh et al., 2015). From our study, we found the same result. Before starting the experiment, we didn't know that there are seven diploid cultivated type strawberries among 33 genotypes. We identified that from NJ tree analysis. These results can be used for the improved strawberry breeding program.

Another objective of our research project was phenotyping. Phenotyping is important for the improved breeding program along with marker-assisted breeding. Genotype with the a desirable trait is important for releasing a cultivar (Mathey et al., 2013). Phenotypic traits of our research were leaf total antioxidant content, leaf shape, petiole size, and trichome density.

Strawberry fruits are highly nutritious and are a very good source of antioxidant, but little is known about the linkage between antioxidant and molecular markers (Debnath et al., 2012). According to them, multiple genetic and environmental factors affect the production and

accumulation of bioactive compounds. From our research, we found that five genotypes were containing relatively higher levels of total antioxidant content. We used leaf samples for our research as we didn't get enough fruit from all genotypes. The antioxidant content may be higher in fruit compared to leaf samples. However, we assume the trend of antioxidant levels will be the same.

Leaf traits have some economic importance such as photosynthesis, herbivore resistance, stress tolerance etc. Leaf is the primary source of photosynthesis and its morphology influences the photosynthesis and yield. We observed leaf shape in our study among the genotypes and were clustered them into five different groups. Leaf shape affects flowering rates, yield, disease resistance, the efficacy of foliar chemical spray etc. (Andres et al., 2014). Leaf shape is also used by herbivores to detect the sources of palatable foliage for food (Yager et al., 2016). In rice, appropriate curliness in leaf can be useful for maintaining the leaf in upright positon and to support to receive more sunlight (Min et al., 2015). So, leaf shape study can be helpful in the improvement of strawberry.

Leaf petiole has a stem-like role because it supports the leaf blade and has an axial structure. We studied different petiole sizes. It might be a useful trait for strawberry breeding to consider especially if the breeding program aims to modify the plant architecture for various cultivation purposes. However, we didn't see any specific trend in relation to total antioxidant content.

Trichome contributes to herbivory resistance to plants (Dalin et al., 2008). According to Dalin et al., (2008), trichome density affects the interaction between insects, mites, aphids and other plant pests. It also affects the availability and effectiveness of feeding behavior of predators and parasitoids of herbivore. Leaf trichome also affects the leaf physiological responses and helps to absorb radiation. Higher trichome density contributes to better abiotic stress tolerance, for

example drought stress resistance in plants by reducing plant water loss. We have observed trichome density (high, medium, and low) among all strawberry genotypes. We found that eleven genotypes have relatively higher trichome density. Among those, two genotypes also contain high total antioxidant levels as well. Polyphenol is an antioxidant which has defensive mechanism against herbivore (Daayf et al., 2012). We may assume that Clancy and Elan hybrid may contain a high amount of polyphenol as it contains a high amount of total antioxidant.

We also investigated to observe presence of any direct correlation between leaf total antioxidant levels and leaf traits observed under this study. From this analysis, we have not found any significant correlation between these traits and antioxidant levels yet.

This research helped to address multiple scientific questions;

1) The strawberry accessions were newly acquired as part of germplasm collections at Delaware State University. Genetic diversity analysis of these accessions helps in designing new strawberry breeding program.
2) Characterization of recently developed SSR markers for the use of strawberry and other fruit researchers, and
3) Finally, trying to establish an association between genomic, phenomic markers and antioxidant levels.

REFERENCES

Acheson, R., & Williams, D. R. R. (1983). Does consumption of fruit and vegetables protect against stroke? *The Lancet*, *321*(8335), 1191–1193. https://doi.org/10.1016/S0140-6736(83)92467-4

Adams, R. P., & Demeke, T. (1993). Systematic Relationships in Juniperus Based on Random Amplified Polymorphic DNAs (RAPDs). *Taxon*, *42*(3), 553. https://doi.org/10.2307/1222534

Armour JAL, Alegre SA, Miles S, Williams LJ, Badge RM (1999). Minisatellites and mutation processes in tandemly repetitive DNA. In: Goldstein DB, Schlotterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp. 24–33.

Armstrong, B. K., Mann, J. I., Adelstein, A. M., & Eskin, F. (1975). Commodity consumption and ischemic heart disease mortality, with special reference to dietary practices. *Journal of Chronic Diseases*, *28*(9), 455–469. https://doi.org/10.1016/0021-9681(75)90056-9

Ashley, M. V., Wilk, J. A., Styan, S. M. N., Craft, K. J., Jones, K. L., Feldheim, K. A., … Ashman, T. L. (2003). High variability and disomic segregation of microsatellites in the octoploid Fragaria virginiana Mill. (Rosaceae). *Theoretical and Applied Genetics*, *107*(7), 1201–1207. https://doi.org/10.1007/s00122-003-1370-5

Andres, R. J., Bowman, D. T., Kaur, B., & Kuraparthy, V. (2014). Mapping and genomic targeting of the major leaf shape gene (L) in Upland cotton (Gossypium hirsutum L.). *Theoretical and applied genetics*, *127*(1), 167-177.

Barrett, B. A., & Kidwell, K. K. (1998). AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop science*, *38*(5), 1261-1271.

Bassil, N. V., Gunn, M., Folta, K., & Lewers, K. (2006). Microsatellite markers for Fragaria from "Strawberry Festival" expressed sequence tags. *Molecular Ecology Notes*, *6*(2), 473–476. https://doi.org/10.1111/j.1471-8286.2006.01278.x

Bliss, F. A. (2010). Marker-assisted breeding in horticultural crops. *Acta Horticulturae*, (859), 339–350. https://doi.org/10.17660/ActaHortic.2010.859.40

Bors, W., & Saran, M. (1987). Radical scavenging by flavonoid antioxidants. *Free Radical Research Communications*, *2*(4–6), 289–294. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3504810

Brunings, A., Moyer, C., Peres, N., & Folta, K. (2010). Implementation of simple sequence

repeat markers to genotype Florida strawberry varieties. *Euphytica*. Retrieved from http://link.springer.com/article/10.1007/s10681-009-0112-4

Capocasa, F., Diamanti, J., Tulipani, S., Battino, M., & Mezzetti, B. (2008). Breeding strawberry (Fragaria X ananassa Duch) to increase fruit nutritional quality. In *BioFactors* (Vol. 34, pp. 67–72). https://doi.org/10.1002/biof.5520340107

Cook, N. C., & Samman, S. (1996). Flavonoids-Chemistry, metabolism, cardioprotective effects, and dietary sources. *The Journal of Nutritional Biochemistry*, *7*, 66–76. https://doi.org/10.1016/S0955-2863(95)00168-9

Cox, T. S., Murphy, J. P., & Rodgers, D. M. (1986). Changes in genetic diversity in the red winter wheat regions of the United States. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(15), 5583–5586. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16593738

Crossa J, Franco J. 2004. Statistical methods for. Euphytica. 137 classifying genotypes:19–37.

Daayf, F., El Hadrami, A., El-Bebany, A. F., Henriquez, M. A., Yao, Z., Derksen, H., … Adam, L. R. (2012). Phenolic Compounds in Plant Defense and Pathogen Counter-Defense Mechanisms. In *Recent Advances in Polyphenol Research* (pp. 191–208). Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118299753.ch8

Dalin, P., Ågren, J., Björkman, C., Huttunen, P., & Kärkkäinen, K. (2008). Leaf trichome formation and plant resistance to herbivory. In *Induced plant resistance to herbivory* (pp. 89-105). Springer, Dordrecht.

Darrow, G. M. (1966). The strawberry. History, breeding and physiology. *The strawberry. History, breeding and physiology.*

Davis TM, DiMeglio LM, Yang R, Styan SMN, Lewers KS (2006) Assessment of SSR transfer from the cultivated strawberry to diploid strawberry species: functionality, linkage group assignment, and use n diversity analysis. J Am Soc Hortic Sci 131:506–512

Debnath, S. C., Siow, Y. L., Petkau, J., An, D., & Bykova, N. V. (2012). Molecular markers and antioxidant activity in berry crops: Genetic diversity analysis. *Canadian Journal of Plant Science*, *92*(6), 1121–1133. https://doi.org/10.4141/cjps2011-240

Denoyes-Rothan, B., Guérin, G., Lerceteau-Köhler, E., & Risser, G. (2005). Inheritance of Resistance to *Colletotrichum acutatum* in *Fragaria* × *ananassa*. *Phytopathology*, *95*(4), 405–412. https://doi.org/10.1094/PHYTO-95-0405

DiMeglio, L., Staudt, G., Yu, H., & Davis, T. (2014). A phylogenetic analysis of the genus

Fragaria (strawberry) using intron-containing sequence from the ADH-1 gene. *PloS One*. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0102237

Emanuelli, F., Lorenzi, S., Grzeskowiak, L., Catalano, V., Stefanini, M., Troggio, M., … Grando, M. S. (2013). Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biology*, *13*(1), 39. https://doi.org/10.1186/1471-2229-13-39

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, *14*(8), 2611-2620.

Felgines, C., Talavéra, S., Gonthier, M.-P., Texier, O., Scalbert, A., Lamaison, J.-L., & Rémésy, C. (2003). Strawberry anthocyanins are recovered in urine as glucuro- and sulfoconjugates in humans. *The Journal of Nutrition*, *133*(5), 1296–1301. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12730413

Folta, K. M., & Davis, T. M. (2006). Strawberry genes and genomics. *Critical Reviews in Plant Sciences*, *25*(5), 399–415. https://doi.org/10.1080/07352680600824831

Folta, K. M., & Kole, C. (2011). *Genetics, genomics and breeding of berries*. Science Publishers.

Geisseler, D., & Horwath, W. R. (2014). Strawberry production in California. *Fertilizer Research and Education Program (FREP)*, 4.

Goldstein, D. B., & Schlötterer, C. (1999). Microsatellitesevolution and applications (No. 574.87328 M5).

Goldstein, D. B., & Pollock, D. D. (1997). Launching Microsatellites: A Review of Mutation Processes and Methods of Phylogenetic Inference. *Journal of Heredity*, *88*, 335–342. Retrieved from http://www.oxfordjournals.org/our_journals/jhered/freepdf/88-335.pdf

Govan, C. L., Simpson, A. D. W., Johnson, A. A. W., Tobutt, A. K. R., & Sargent, A. D. J. (2008). A reliable multiplexed microsatellite set for genotyping Fragaria and its use in a survey of 60 F. 3 ananassa cultivars. https://doi.org/10.1007/s11032-008-9206-2

Gupta, P. K., Varshney, R. K., Sharma, P. C., & Ramesh, B. (1999). Molecular markers and their applications in wheat breeding. *Plant Breeding*. https://doi.org/10.1046/j.1439-0523.1999.00401.x

Hancock, J. F., Callow, P. A., & Shaw, D. V. (1994). Randomly amplified polymorphic DNAs in the cultivated strawberry, Fragaria x ananassa. *Journal of the American Society for*

Horticultural Science, *119*(4), 862–864.

Hancock, J. F., & Luby, J. J. (1993). Genetic Resources at Our Doorstep: The Wild Strawberries. *BioScience*, *43*(3), 141–147. https://doi.org/10.2307/1312017

Haymes, K., Henken, B., & Davis, T. (1997). Identification of RAPD markers linked to a Phytophthora fragariae resistance gene (Rpf1) in the cultivated strawberry. *Theoretical and Applied*. Retrieved from http://link.springer.com/article/10.1007/s001220050521

Hokanson, S., & Maas, J. (2001). Strawberry biotechnology. *Plant Breeding Reviews*. Retrieved from https://books.google.com/books?hl=en&lr=&id=ra1p_V9YTr0C&oi=fnd&pg=PA139&dq=11.%09S.C.+Hokanson+and+J.L.+Maas,+Strawberry+biotechnology,+Plant+Breeding+Reviews+21+(2001),+139–180&ots=95sIGhqfs-&sig=7HGA5knN03GallbZ_eAYRGUJsO4

Heinonen, I. M., Meyer, A. S., & Frankel, E. N. (1998). Antioxidant activity of berry phenolics on human low-density lipoprotein and liposome oxidation. *Journal of Agricultural and Food Chemistry*, *46*(10), 4107-4112.

Jarvis, D. I., & Hodgkin, T. (1999). Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular ecology*, *8*(s1).

Joshipura, K., Hu, F., & Manson, J. (2001). The effect of fruit and vegetable intake on risk for coronary heart disease. *Annals of Internal*. Retrieved from http://annals.org/aim/article/714567/effect-fruit-vegetable-intake-risk-coronary-heart-disease

Kozuka, T., Horiguchi, G., Kim, G. T., Ohgishi, M., Sakai, T., & Tsukaya, H. (2005). The different growth responses of the Arabidopsis thaliana leaf blade and the petiole during shade avoidance are regulated by photoreceptors and sugar. *Plant and Cell Physiology*, *46*(1), 213-223.

Levene, H. (1949). On a matching problem arising in genetics. *The annals of mathematical statistics*, *20*(1), 91-94.

Lewers, K. S., Hokanson, S. C., & Bassil, N. V. (2005). Strawberry GenBank-derived and Genomic Simple Sequence Repeat (SSR) Markers and Their Utility with Strawberry , Blackberry , and Red and Black. *Journal of the American Society of Horticulture Science*, *130*(1), 102–115.

Li, S., Tosens, T., Harley, P. C., Jiang, Y., Kanagendran, A., Grosberg, M., ... & Niinemets, Ü.

(2018). Glandular trichomes as a barrier against atmospheric oxidative stress: relationships with ozone uptake, leaf damage and emission of LOX products across a diverse set of species. *Plant, cell & environment*.

Luby, J., & Shaw, D. (2001). Does marker-assisted selection make dollars and sense in a fruit breeding program? *HortScience*. Retrieved from http://hortsci.ashspublications.org/content/36/5/872.short

Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. *Sinauer Associates; 1 edition. 980p*.

Mathey, M. M., Mookerjee, S., Gündüz, K., Hancock, J. F., Iezzoni, A. F., Mahoney, L. L., … Finn, C. E. (2013). Large-Scale Standardized Phenotyping of Strawberry in RosBREED. *Journal of the American Pomological Society*, *67*(4), 205–216. Retrieved from https://mypages.unh.edu/sites/default/files/tmdavis/files/mathey_et_al_2013.pdf

McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, *21*(4), 178–185. https://doi.org/10.1016/J.TREE.2006.02.002

Meyers, K., Watkins, C., & Pritts, M. (2003). Antioxidant and antiproliferative activities of strawberries. *Journal of Agricultural*. Retrieved from http://pubs.acs.org/doi/abs/10.1021/jf034506n

Milella, L., Saluzzi, D., Lapelosa, M., Bertino, G., Spada, P., Greco, I., & Martelli, G. (2006). Relationships between an Italian Strawberry Ecotype and its Ancestor using RAPD Markers. *Genetic Resources and Crop Evolution*, *53*(8), 1715–1720. https://doi.org/10.1007/s10722-005-1405-7

Min, G. U. O., Jian, Y. A. O., Juan, Z. H. U., Xiang-yun, F. A. N., Wei, W. A. N. G., Shu-zhu, T. A. N. G., ... & Chang-jie, Y. A. N. (2015). RL3 (t), responsible for leaf shape formation, delimited to a 46-kb DNA fragment in rice. *Rice Science*, *22*(1), 44-48.

Mueller, U. G., Lipari, S. E., & Milgroom, M. G. (1996). Amplified fragment length polymorphism (AFLP) fingerprinting of symbiotic fungi cultured by the fungus-growing ant Cyphomyrmex minutus. *Molecular Ecology*, *5*(1), 119–122. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9147688

Nei M, (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences, USA 70:3321-3323.

Ohashi, J., & Tokunaga, K. (2003). Power of genome-wide linkage disequilibrium testing by

using microsatellite markers. *Journal of Human Genetics*, *48*(9), 487–491. https://doi.org/10.1007/s10038-003-0058-7

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2000). Association mapping in structured populations. *The American Journal of Human Genetics*, *67*(1), 170-181.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., & Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*, *2*(3), 225–238. https://doi.org/10.1007/BF00564200

Record, I. R., Dreosti, I. E., & Mcinerney, J. K. (2017). Changes in plasma antioxidant status following consumption of diets high or low in fruit and vegetables or following dietary supplementation with an antioxidant mixture. https://doi.org/10.1079/BJN2000292

Roy, A., Frascaria, N., MacKay, J., & Bousquet, J. (1992). Segregating random amplified polymorphic DNAs (RAPDs) in Betula alleghaniensis. *Theoretical and Applied Genetics*, *85*(2–3), 173–180. https://doi.org/10.1007/BF00222856

Sack, L., Scoffoni, C., John, G. P., Poorter, H., Mason, C. M., Mendez-Alonzo, R., & Donovan, L. A. (2013). How do leaf veins influence the worldwide leaf economic spectrum? Review and synthesis. *Journal of Experimental Botany*, *64*(13), 4053-4080.

Sacks, F. M., & Kass, E. H. (1988). Low blood pressure in vegetarians: effects of specific foods and nutrients. *The American Journal of Clinical Nutrition*, *48*(3 Suppl), 795–800. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3414588

Sambrook, J., Williams, J., Sharp, P. A., & Grodzicker, T. (1975). Physical mapping of temperature-sensitive mutations of adenoviruses. *Journal of Molecular Biology*, *97*(3), 369–390. https://doi.org/10.1016/S0022-2836(75)80046-5

Sargent, D. J., Clarke, J., Simpson, D. W., Tobutt, K. R., Arús, P., Monfort, A., … Battey, N. H. (2006). An enhanced microsatellite map of diploid Fragaria. *Theoretical and Applied Genetics*, *112*(7), 1349–1359. https://doi.org/10.1007/s00122-006-0237-y

Schneider, S., Roessli, D., & Excoffier, L. (2000). Arlequin: a software for population genetics data analysis. *User manual ver*, *2*, 2496-2497.

Semagn, K., Bjornstad, A., & Ndjiondjop, M. N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, *5*(25), 2540–2568. https://doi.org/10.5897/AJB2006.000-5110

Singh, N. V., Abburi, V. L., Ramajayam, D., Kumar, R., Chandra, R., Sharma, K. K., ... &

Saminathan, T. (2015). Genetic diversity and association mapping of bacterial blight and other horticulturally important traits with microsatellite markers in pomegranate from India. *Molecular genetics and genomics*, *290*(4), 1393-1402.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, *28*(10), 2731-2739

Thompson, J. A., Nelson, R. L., & Vodkin, L. O. (1998). Identification of diverse soybean germplasm using RAPD markers. *Crop Science*, *38*(5), 1348-1355.

Ulrich, D., Komes, D., Olbricht, K., & Hoberg, E. (2007). Diversity of aroma patterns in wild and cultivated Fragaria accessions. *Genetic Resources and Crop*. Retrieved from http://link.springer.com/article/10.1007/s10722-006-9009-4

Wang, H., Cao, G., & Prior, R. L. (1996). Total Antioxidant Capacity of Fruits. Retrieved from https://pubs-acs-org.desu.idm.oclc.org/doi/pdf/10.1021/jf950579y

Wang, S. Y., & Jiao, H. (2000). Scavenging capacity of berry crops on superoxide radicals, hydrogen peroxide, hydroxyl radicals, and singlet oxygen. *Journal of Agricultural and Food Chemistry*, *48*(11), 5677-5684.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358. https://doi.org/10.2307/2408641

Whitaker, V. (2011). Applications of molecular markers in strawberry. *Journal of Berry Research*. Retrieved from http://content.iospress.com/articles/journal-of-berry-research/jbr013

Wolfe, A. D., & Liston, A. (1998). Contributions of PCR-Based Methods to Plant Systematics and Evolutionary Biology. In *Molecular Systematics of Plants II* (pp. 43–86). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-5419-6_2

Wu, F., Guan, Z., & Whidden, A. (2012). Strawberry industry overview and outlook. *University of Florida*.

Yager, K. G., Schaefer, H. M., & Gould, K. S. (2016). The significance of shared leaf shape in Alseuosmia pusilla and Pseudowintera colorata. *Botany*, *94*(7), 555-564.

Yeh YR, Boyle T. 1999. POPGENE Version 1.31. Microsoft Window-based Freeware for Population Genetic Analysis Quick User Guide.

Yoon, M. Y., Moe, K. T., Kim, D. Y., Rho, I. R., Kim, S., Kim, K. T., ... & Park, Y. J. (2012). Genetic diversity and population structure analysis of strawberry (Fragaria x ananassa Duch.) using SSR markers. *Electronic Journal of Biotechnology*, *15*(2), 6-6.

(https://www.statista.com/statistics/193288/us-total-strawberry-production-since-2000/