USE OF MOLECULAR MARKERS AND ANTIOXIDANT PROPERTIES TO STUDY BIODIVERSITY AND TO DISCRIMINATE BLUEBERRIES

By

DHRUMIT SHAILESHBHAI BHATT

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biology

Memorial University of Newfoundland

September 2021

St. John's Newfoundland and Labrador Canada

This work is dedicated to my wife, brother, and parents

Princee Panchal, Dhruv Bhatt, Ila Bhatt, and Shailesh Bhatt

Abstract

Blueberries (*Vaccinium spp*.) have gained much attention worldwide with their potential health benefits and economic importance. Estimation of genetic diversity using molecular markers, antioxidant properties, and their association can reveal genotypes with important characteristics and help in berry improvement programs. Genetic diversity was estimated in blueberry hybrids, wild clones, and cultivars using expressed sequence tag- simple sequence repeat (EST-SSR), genomic (G)-SSR, and EST- polymerase chain reaction (PCR) markers. The antioxidant efficacy was estimated by total phenolic and flavonoid contents in blueberry hybrids, wild clones, and cultivars. Wide diversity existed among the genotypes, for antioxidant properties with the highest variation for 2, 2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging activity (20-fold) followed by the contents of total flavonoids (16-fold) and phenolics (3.8-fold). Although a group of 11 hybrids generated the maximum diversity for antioxidant activity (15-fold), wild clones collected from Quebec, Canada, had the maximum variation for total phenolic (2.8-fold) and flavonoid contents (6.9-fold). Extensive genetic diversity was evident from Shannon's index (0.34 for EST-SSRs, 0.29 for G-SSR, 0.26 for EST-PCR) and expected heterozygosity (0.23 for EST-SSR, 0.19 for G-SSR, 0.16 for EST-PCR). STRUCTURE analysis separated the genotypes into three groups, which were in agreement with principal coordinate and neighbor-joining analyses. Molecular variance suggested 19% variations among groups and 81% among genotypes within groups. Clustering based on biochemical data and molecular analysis did not coincide, indicating a random distribution of loci in the blueberry genome conferring antioxidant properties. However, the stepwise multiple regression analysis (SMRA) revealed that a total of 17 EST-SSR, G-SSR, and EST-PCR markers were associated with antioxidant properties. The outcome of this study will

provide valuable information for the selection of blueberry genotypes with desired traits, breeding, and germplasm conservation programs.

Acknowledgments

I owe my sincerest and deepest gratitude to my supervisor, Dr. Samir C. Debnath for providing me with opportunity, knowledge, supervision and financial support through my Master's program. Without his honest guidance and constant encouragement as a supervisor, I would not have been able to realize my ability to carry out research and author this manuscript. I also extend my gratitude towards my co-supervisor, Dr. Julissa Roncal, for her timely suggestions and encouragement. Also, I would like to thank the other member of my supervisory committee, Dr. Dawn Marshall, for her support and valuable comments.

I will eternally be grateful to my parents and my brother, who have been my support system from day one to this day. I express my earnest thanks to my wife Princee Panchal, who has been my constant source of encouragement and who has stood by me through thick and thin. I am very grateful to my parents-in-law for their inspiration and blessing to complete this program successfully. Support of my friends back home has been a source of my energy throughout my graduate study, for which I thank you Vishal, Krupa, Vidhi, Mihir, Hardik, Ushma, Harsh, Esha, and Eve. My friends in Canada, Amrita, Swapneel, Deep, Kunal, Divyang too have contributed towards keeping me aligned with my goals.

My sincere appreciation to the School of Graduate Studies, Department of Biology and St. John's Research and Development Centre (SJRDC) for a fellowship and resources throughout my program. I am indebted to the wonderful administrative and research team at SJRDC, who always made time to assist me with my research. I am grateful to Glen Chubbs and Darryl Martin, who helped me maintain plants in the greenhouse and perform laboratory analysis. Last but not least, my thanks go to my lab-mates, especially Amrita, Juran, and Sapan, for their valuable help and consultation.

Table of Contents

Abstract iii

Acknowledgments v

List of Tables x

List of Figures xii

List of Abbreviations xvi

CHAPTER 1 1

Introduction and Overview 1

1.1 Introduction 1

1.2 Overview 1

1.2.1 Biology and Types of Blueberry 1

1.2.2 Economic Importance of Blueberry 3

1.2.3 Medicinal Value of Blueberry 5

1.2.4 Phenolics in Plants 6

1.2.5 Antioxidant Activity of Blueberry Phenolics 7

1.3 Genetic Diversity 13

1.3.1 Crop Improvement, Biodiversity, and Wild Clones 14

1.3.2 Estimation of Genetic Diversity 15

1.3.2.1 Morphological Markers 16

1.3.2.2 Biochemical Markers 17

1.3.2.3 Molecular Markers 19

1.3.2.3.1 RFLP (Restriction Fragment Length Polymorphism) 21

1.3.2.3.2 RAPD (Random Amplified Polymorphic DNA) 21
1.3.2.3.3 AFLP (Amplified Fragment Length Polymorphism) 23
1.3.2.3.4 ISSR (Inter Simple Sequence Repeats) 24
1.3.2.3.5 Microsatellites or SSRs (Simple Sequence repeats) 25
1.3.2.3.5.1 EST-SSR (Expressed Sequence Tagged-Simple Sequence Repeats) 25
1.3.2.3.6 SNP (Single Nucleotide Polymorphism) 26
1.4 Research Rationale and Objectives 30
CHAPTER 2 31
Materials and Methods 31
2.1 Plant Materials 31
2.2 Molecular Analysis 34
2.2.1 DNA Extraction, PCR Amplification, and Electrophoresis 34
2.2.2 Data Collection and Statistical Analysis 40
2.2.3 Genetic Diversity Indices 42
2.2.3.1 Polymorphic Information Content (PIC) 42
2.2.3.2 Effective Multiplex Ratio (EMR) 43
2.2.3.3 Marker Index (MI) 43
2.2.3.4 Discrimination Power (D) 43
2.2.3.5 Resolving Power (R) 44
2.2.3.6 Percentage of Polymorphic Loci (PL) 44
2.2.3.7 Number of alleles (N_a) 45
2.2.3.8 Effective Number of Alleles (N_e) 45
2.2.3.9 Expected Heterozygosity/Nei's Gene Diversity Index (H_e) 45

2.2.3.10 Shannon's Information Index of Diversity (I) 46
2.2.4 Population Structure Analysis 46
2.2.4.1 STRUCTURE Analysis 46
2.2.4.2 Unweighted Neighbour-Joining Tree (NJ) 47
2.2.4.3 Principal Coordinate Analysis (PCoA) 47
2.2.3.4 Analysis of Molecular Variance (AMOVA) 47
2.3 Biochemical Analysis 48
2.3.1 Preparation of Leaf Extracts 48
2.3.2 Total Antioxidant Activity (TAA) 49
2.3.3 Total Phenolic Content (TPC) 49
2.3.4 Total Flavonoid Content (TFC) 50
2.3.5 Cluster Analysis of Antioxidant Properties 50
2.4 Correlation between Biochemical and Molecular Analysis 51
CHAPTER 3 53
Results 53
3.1 Molecular Analysis 53
3.1.1 Diversity Parameters 53
3.1.1.1 Analysis of Primer's Discriminatory Power 53
3.1.1.2 Analysis of Population Genetic Diversity 57
3.1.1.2.1 STRUCTURE Analysis 61
3.1.1.2.2 Unweighted Neighbour-Joining (NJ) Tree 65
3.1.1.2.3 Principal Coordinate Analysis (PCoA) 70
3.1.1.2.4 Analysis of Molecular Variance (AMOVA) 74

3.2 Biochemical Analysis 76
3.2.1 Total Antioxidant Activity (TAA) 76
3.2.3 Total Phenolic Content (TPC) 83
3.2.3 Total Flavonoid Content (TFC) 83
3.2.4 Relationship among Antioxidant Properties 84
3.2.5 Cluster Analysis of Antioxidant Properties 86
3.2.5.1 Agglomerative Hierarchical Clustering (AHC) 86
3.2.5.2 Principal Component Analysis (PCA) 87
3.3 Relationship between Biochemical and Molecular Analysis 88
3.3.1 Mantel Test 88
3.3.2 Stepwise Multiple Regression Analysis (SMRA) 90
CHAPTER 4 95
Discussion 95
CHAPTER 5 103
Conclusions 103
CHAPTER 6 105
References 105

List of Tables

Table 1.1 Anthocyanins commonly present in blueberries. 9

Table 1.2 Comparison among different types of markers 28

Table 2.1 Parentage of lowbush, half-high, and highbush blueberry cultivars and hybrids 32

Table 2.2 Wild blueberry clones designated as BC1 to BC36, collected from Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB).....33

Table 2.3 EST-SSR, genomic (G-) SSR and EST-PCR primer pairs: their sequences and annealing temperatures.....35

Table 3.1 Discriminatory power indices of EST-SSR, genomic (G-) SSR and EST-PCR markers for diversity analysis in blueberry hybrids, wild clones, and cultivars. 55

Table 3.2 Population genetic diversity parameters of blueberry hybrid groups (Cross 1: HB1-11; Cross 2: HB12-28), wild clones collected from Canadian provinces Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), and six cultivars (CV). 58

Table 3.3 Genetic differentiation among blueberry genotypes by AMOVA based on seven groups where four groups comprised of wild clones collected from four Canadian provinces; the fifth

group with all six cultivars; sixth group contained 11 hybrids from the first cross (HB1-11); and seventh group contained 17 hybrids from the second cross (HB12-28).....75

Table 3.4 Total antioxidant activity (TAA), phenolic (TPC), and flavonoid contents (TFC) of groups of wild clones collected from four Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), cultivars (CV): Fundy, Polaris, Patriot, Chippewa, St. Cloud, Northblue, hybrids: Cross 1 and Cross 2. GAE = gallic acid equivalents, CE = catechin equivalent, fl = fresh leaf. 77

Table 3.5 Total antioxidant activity (TAA), phenolic (TPC), and flavonoid contents (TFC) of individual wild clones collected from four Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), cultivars (CV): Fundy, Polaris, Patriot, Chippewa, St. Cloud, Northblue, hybrids: Cross 1 and Cross 2. GAE = gallic acid equivalents, CE = catechin equivalent, fl = fresh leaf. 78

Table 3.6 Pearson correlation coefficients of antioxidant properties. The respective p-values are shown in brackets. The significance level alpha is 0.05.....86

Table 3.7. Association of EST-SSR, G-SSR and EST-PCR markers with biochemical components in blueberry leaf extract as revealed by SMRA.....92

List of Figures

Figure 1.1 Percentage share of farm gate value of fresh and processed fruits of Canada in 2019. .4

Figure 1.2 Blueberry production share by region worldwide in 2019. 5

Figure 1.3 Classification and structure of phenolic compounds. 7

Figure 1.4 Reduction of molecular oxygen with electrons. 8

Figure 2.1 Various blueberry plants maintained in a greenhouse at St. John's Research and Development Centre (St. John's RDC), Agriculture Agri-Food Canada, St. John's, Newfoundland and Labrador, Canada. 31

Figure 2.2 An example of banding pattern of blueberry genotypes expressed by A(1) & A(2) EST-SSR (NA961); B(1) & B(2) G-SSR (VCC_J1); and C(1), C(2) EST-PCR primer (NA27). M1 and M2 are 100bp and 1K ladders, respectively. Lanes 1 to 70 represents genotypes in order, as shown in Table 2.2. 42

Figure 3.1 (a) ΔK is reaching its maximum value at K = 2 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 2) based on EST-SSR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label).....61

Figure 3.2 (a) ΔK is reaching its maximum value at K = 2 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 2) based on genomic

(G)-SSR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label). 62

Figure 3.3 (a) ΔK is reaching its maximum value at K = 3 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 3) based on EST-PCR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label). 63

Figure 3.4 (a) ΔK is reaching its maximum value at K = 3 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 3) based on the combination of EST-SSR, genomic (G) - SSR, and EST-PCR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label). 65

Figure 3.5 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by EST-SSR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label). 66

Figure 3.6 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by genomic (G) - SSR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label). 67

Figure 3.7 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by EST-PCR markers among blueberry genotypes. The

colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label). 69

Figure 3.8 Dendrogram produced the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by the combination of EST-SSR, genomic (G) - SSR, and EST-PCR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label). 70

Figure 3.9 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by EST-SSR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label). 71

Figure 3.10 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by genomic (G) - SSR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).....72

Figure 3.11 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by genomic EST-PCR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).....73

Figure 3.12 2D principle coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by the combination of EST-SSR, genomic (G) -SSR and EST-PCR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).....74

Figure 3.13 Linear regression between antioxidant properties in wild clones, cultivars, and hybrids. (a) Total antioxidant activity by DPPH (mg GAE/g fl) and total phenolic content (mg GAE/g

fl). (b) Total antioxidant activity by DPPH (mg GAE/g fl) and total flavonoid content (mg CE/g fl). (c) Total phenolic content (mg GAE/g fl) and total flavonoid content (mg CE/g fl). 85

Figure 3.14 AHC dendrogram by UPGMA method based on Euclidean dissimilarity distance matrix of antioxidant activity, phenolic, and flavonoid content data of 70 genotypes.....87

Figure 3.15 The PCA bi-plot based on biochemical characteristics of 70 genotypes and correlation among quantitative variables. 88

Figure 3.16 The scatter plot (a, b, c, d) and histogram (e, f, g, h) of Mantel test between biochemical and genetic matrices (EST-SSR, G-SSR, EST-PCR and combined primers, respectively)....90

List of Abbreviations

6-PGD	6- phosphogluconate dehydrogenase
AFLP	Amplified fragment length polymorphism
AHC	Agglomerative hierarchical clustering
AMOVA	Analysis of molecular variance
ANOVA	Analysis of variation
AP-PCR	Arbitrarily-primed PCR
BHA	Butylated hydroxyl anisole
BHT	Butylated hydroxytoluene
BP	Base Pair
CA	Cold-acclimated
CAPS	Cleaved amplified polymorphic sequence
cDNA	complimentary DNA
CE	Catechin equivalent
DAF	DNA amplification fingerprinting
DAMD	Directly amplified minisatelltie DNA
DArP	Diversity arrays technology

DNA	Deoxyribonucleic Acid
dNTP	deoxy nucleotide phosphate
DPPH	2, 2-diphenyl-1-picrylhyrazyl
EDTA	Ethylene diamine tetra acetic acid
EMR	Effective multiplex ratio
EST	Expressed sequence tags
FC	Folin-Ciocalteu
FL	Fresh leaves
GAE	Gallic acid equivalent
HRB	High resolution blend
IDH	Isocitrate dehydrogenase
ISSR	Inter-simple sequence repeat
Kb	Kilo base
MALDI-TOF	Matrix-assisted laser description ionization-time of flight
MCMC	Markov Chain Monte Carlo
MDH	Malate dehydrogenase
MI	Marker index

MP-PCR	Microsatellite-primed PCR
NA	Non-acclimated
NDGA	Nordihydroguaiaretic acid
NGS	Next-generation sequencing
NJ	Neighbour-joining
NM	nanometer
PAGE	Polyacrylamide gel electrophoresis
PAL	Phenylalanine lyase L.
PCA	Principal component analysis
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PG	Propyl gallate
PGI	Phosphoglucose isomerase
PIC	Polymorphic information content
PL	Percentage of polymorphic loci
PPF	Photosynthetic photon flux
RAMP	Randomly amplified microsatellite polymorphism

RAPD	Randomly amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RFLP	Restriction fragment length polymorphism
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
SAMPL	Selective amplified microsatellite polymorphic locus
SCAR	Sequence characterized amplified region
SCoT	Start codon targeted
SD	Standard deviation
SMRA	Stepwise Multiple Regression Analysis
SNP	Single nucleotide polymorphism
SPAR	Single primer amplification reaction
SRAP	Sequence-related amplified polymorphism
SSAP	Sequence-specific amplification polymorphism
SSCP	Single strand conformation polymorphism
SSR	Simple sequence repeats
STS	Sequence-tagged sites

TBHQ	Tertiary butyl hydroquinone
TRAP	Target region amplification polymorphism
UPGMA	Unweighted pair-group method with arithmetic mean
US	United States of America
UV	Ultraviolet
VNTR	Variable number tandem repeats

CHAPTER 1

Introduction and Overview

1.1 Introduction

Blueberry belongs to genus *Vaccinium* L. (family: Ericaceae; tribe: Vacciniae; subfamily: Vaccinoiodae), which includes diploids such as *V. corymbosum L.*, *V. tenellum* Ait., *V. myrtilloides Michx*, *V. darrowi* Camp, *V. pallidum* Ait., *V. elliottii* Chapm, *V. boreale*; tetraploids such as *V. corymbosum*, *V. angustifolium* Ait, *V. hirsutum* Buckley, *V. myrsinites* Lam., *V. simulatum* Small, and hexaploids such as *V. constablaei* Gray and *V. ashei* Reade. These blueberry are native to North America (Camp, 1945; Longley, 1927; Vander Kloet, 1988). The cultivars in section *Cyanococcus*, *V. corymbosum* L. (highbush blueberry), *V. ashei* Reade (rabbiteye blueberry), and *V. angustifolium* Ait. (lowbush blueberry) have the most commercial value (Song & Hancock, 2011).

1.2 Overview

1.2.1 Biology and Types of Blueberry

Blueberries can be divided into five major groups that are grown commercially:

(1) Lowbush blueberry (*V. angustifolium*; 2n = 4x = 48, *V. boreale* Hall; 2n = 2x = 24 and Aald., *V. myrtilloides* Michx.; 2n = 2x = 24)

They are also known as sweet blueberry (Song & Hancock, 2011). It is a dwarf (1- 2 feet), woody, usually deciduous shrub that grows in the wild. Leaves are pale to dark green, with an elliptical shape; 5–20 mm × 16–40 mm in size; margins are uniformly serrated; surfaces are glabrous/smooth predominantly, or hairy. Shoots of lowbush blueberry are erect, forming dense, extensive colonies;

twigs are green to glaucous, glabrous or hairy. Woody rhizomes are on average 4.5 mm in diameter and grow around 6 cm underground. Stems are smooth and vary in color from tan to red (Flinn & Pringle, 1983). Flowers are self-incompatible, bell-shaped, and usually white or pinkish-white, borne in short, few-flowered terminals or axillary racemes (Mohr & Kevan, 1987). Fruits are intermediate in size with orbicular, oval, and blue to dark blue with or without waxy coating (Camp, 1945). The pedicel scar is medium, and the calyx end is closed. They are the cold-hardy bushes and are intolerant to harsh summer heat with an average annual minimum temperate range (AAMTR) from -17.8 to -12.3 ℃ (Song & Hancock, 2011).

(2) Northern highbush blueberry (*V. corymbosum*; 2n = 4x = 48)

The crown-forming northern highbush plants usually are 6-8 ft tall. They are naturally found in areas between Nova Scotia to Wisconsin, Georgeo and Alabama (Song & Hancock, 2011). Stems are angular to terete; glabrous to densely pubescent. Leaves are oval to narrow elliptical; 20 - 30 mm wide and 40 - 80 mm long. Leaf-blades are pubescent or glabrous with entire or sharply serrate margins. The flowers are cylindrical; white, white-pink, or pink with green or glaucous calyx (Camp, 1945; Vander Kloet, 1980). Berries are blue, dull black, or even black.

(3) Southern highbush blueberry (2n = 4x = 48)

The southern highbush blueberries were developed from hybridization of *V. corymbosum* with *V. darrowi* Camp and *V. ashei*. They were explicitly hybridized for superior fruit, soil adaptability, heat tolerance, and low winter chilling (Ratnaparkhe, 2007). The flowers are white, bell-shaped. Berries' colours range from powder-blue to medium-blue, with an intense flavor. Some southern highbush cultivars are self-fertile, but the berries grow larger if two varieties are planted together.

(4) Half-high blueberry (2n = 4x = 48)

They are hybridization between highbush and lowbush blueberries (Galletta & Ballington, 1996b). Leaves are narrow to broadly elliptical with entire or serrate margins; 30 - 60 mm long. Pubescence

of leaves depends on the leaf characters of the parent type of *V. corymbosum* (Camp, 1945). Fruits are dark-blue, dull, or even black.

(5) Rabbiteye blueberry (*V. ashei*; 2n = 6x = 72)

Rabbiteyes are wild selections and hybrid cultivars of *V. ashei* (Ratnaparkhe, 2007). They are usually 6-13 ft tall and can withstand dry periods better than other blueberries (Song & Hancock, 2011). Leaves usually are deciduous or sometimes evergreen. Berries are black to dull in color with 8 - 18 mm in diameter.

1.2.2 Economic Importance of Blueberry

Native Americans burned sections of forest to promote blueberry cultivation (Russell, 1976). Highbush blueberries are grown in few states in the USA, Canada, New Zealand, Australia, South America, Europe, and Asia (Strik, 2005; Strik & Yarborough, 2005). Though highbush and lowbush blueberries are used for most commercial productions, rabbiteyes are also commercially cultivated, mainly for their late-ripening fruits, in southeast regions of North America, the Pacific Northwest, and Chile. The northern highbush varieties are cultivated in New York, New Jersey, Oregon, Michigan, North Carolina, Washington in the USA, and British Columbia in Canada. Whereas southern highbush varieties are grown in Florida, California, and Georgia. Lowbush blueberries are predominantly produced in Maine in the USA and Newfoundland, Nova Scotia, New Brunswick, and Quebec in Canada (Hancock et al., 2008; Strik, 2005). Fifty percent of highbush production is sold as fresh fruit, and the other fifty percent of it used by different processing industries to make pie, muffin, and pancake fillings, syrups, preservatives, jams, yogurts, and ice cream. Blueberry juice is not a very popular choice among consumers because it has a strong taste and deep colour (Song & Hancock, 2011). Due to its high medicinal and economic values, the breeders worldwide have been manipulating *Vaccinium spp*. for high productivity, the

better quality of fruits, and resistance to different abiotic and biotic stresses. Another aspect other than its medicinal value, which attracts crop breeders, *Vaccinium spp*. also possesses ornamental value because of its colourful flowers and leaves and its distinguished arrangement. The vast majority of Canadian berry revenue comes from blueberry production. In the year 2019, blueberry earned $270.7 million in revenue, which is 23% (Figure 1.1) of the total fruit farm gate value for Canada (Statistics Canada, 2020). With an increase in Canada's exports of low-bush blueberries (23.4%), high-bush blueberries (9.3%), and strawberries (12.6%), the country's total fresh fruit exports increased in value by 2.8%, reaching $843 million in 2019 (Statistics Canada, 2020). Blueberries continue to be Canada's top fruit export by volume (65.5% of export dollars) and value (53.4% of export tonnage), out of which low-bush blueberries were the main contributor, accounting for 53.3% of the total blueberry value and 60.6% of total blueberry tonnage (Statistics Canada, 2020). The United States remains Canada's No. 1 export terminus, accounting for 64.8% of all fruit exports by value and 65.7% by tonnage in 2019 (Statistics Canada, 2020). USA and Canada remains one the largest blueberry producer in the world (Figure 1.2).

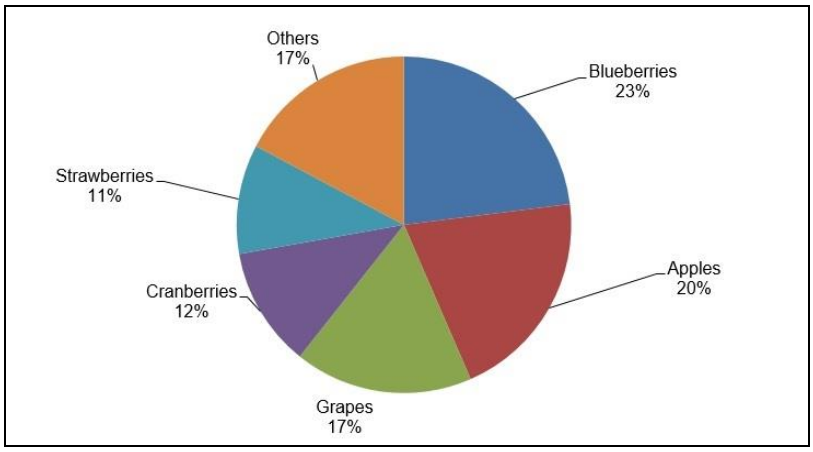


Figure 1.1 Percentage share of farm gate value of fresh and processed fruits of Canada in 2019.

Source: (Statistics Canada, 2020).

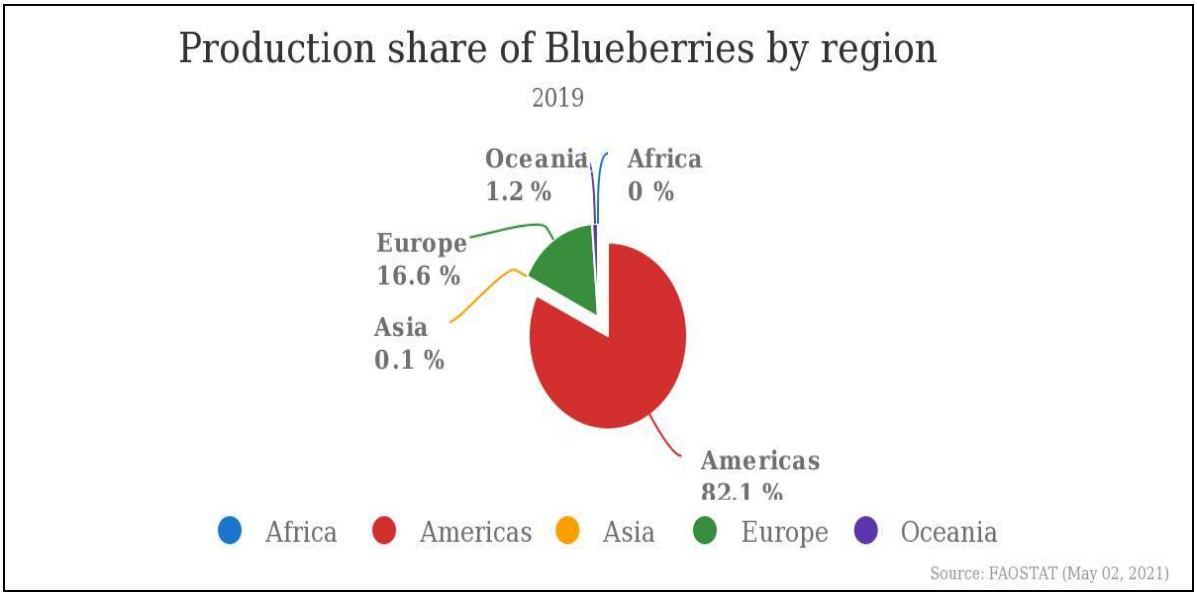


Figure 1.2 Blueberry production share by region worldwide in 2019.

Source: (FAOSTAT, 2019).

1.2.3 Medicinal Value of Blueberry

Fresh blueberries are a rich source of essential nutritional components and consist of 85% water, 15.3% carbohydrates, 0.7% protein, and 0.5% fat (Hancock et al., 2003). Apart from these primary components, blueberries (*Vaccinium spp.*) also contain high levels of anthocyanins, flavonols, catechins, or proanthocyanidins. Proanthocyanidins or condensed tannins are present in a very high concentration in blueberry, which shows bacterial anti-adhesion activity against uropathogenic *Escherichia coli* (Howell et al., 1998; Ofek et al., 1991). The addition of blueberries to the diet has significantly increased serum anti-oxidation activities, lowering blood pressure and blood cholesterol, which may have links to the prevention of cardiovascular diseases and atherosclerosis (Basu et al., 2010; Mazza et al., 2002; Prior et al., 2009; Stull et al., 2010; Wu et al., 2002; Wu et al., 2010). Blueberry have anti-cancerous properties, which have the capacity to influence a number of factors in the carcinogenesis process, including inhibition of oxidation, induction of apoptosis,

anti-proliferation, and reduction of invasion and metastasis (Howell, 2008). Similar observations made by various studies (Adams et al., 2010; Bomser et al., 1996; Chen et al., 2005; Liu et al., 2010; Matchett et al., 2005; Samad et al., 2014; Schantz et al., 2010; Schmidt et al., 2006; Srivastava et al., 2007; Yun et al., 2009). The contents of blueberries have shown to assist the brain in numerous ways, including protection from oxidative stresses and inflammatory reactions (Lau et al., 2007; Sweeney et al., 2002), directly altering cell signaling involved in neuronal communication (Joseph et al., 2007), calcium buffering ability (Shukitt-Hale et al., 2008), bone protection (Shen et al., 2012), and effects on memory (Joseph et al., 2003; Joseph et al., 1999). Blueberry extract reportedly reduced cataracts and significantly reduced lipid oxidation in blood in animal models (Kolosova et al., 2004). It is anti-diabetic and it scavenges free radicals to improve oxygen delivery to eyes in humans (Calò & Marabini, 2014).

1.2.4 Phenolics in Plants

Blueberries are most famous for their antioxidant phytochemicals, mainly phenolic metabolites, which provide significant health benefits other than the primary nutrients such as carbohydrates, proteins, fats, minerals, and vitamins. Phenolic compounds are the largest category of phytochemicals and the most widely distributed throughout the plant. The structure of a phenolic compound consists of aromatic rings with one or more hydroxyl groups attached (Figure 1.3). The plant phenolics are secondary metabolites that are produced naturally (Pridham, 1960) as well as in response to various stresses such as metal toxicity, drought, chilling, wounding, and nutrient deficiency (Winkel-Shirley, 2002) via the shikimate acid pathway from aromatic amino acids, l-phenylalanine and/or l-tyrosine (Cheynier et al., 2013; Heleno et al., 2015; Maeda & Dudareva, 2012; Rice-Evans et al., 1996). They perform a range of activities from acting as a signal in plant-microbe interactions, pigmentation, protection against UV light, pollination, and antioxidants to

protect against pathogenic attacks (Naczk & Shahidi, 2006; Olsen et al., 2010). They mainly include flavonoids, phenolic acids, tannins, and other phenolic compounds (Naczk & Shahidi, 2006). Most plant phytochemicals are classified into flavonoids and non-flavonoids (Działo et al., 2016).

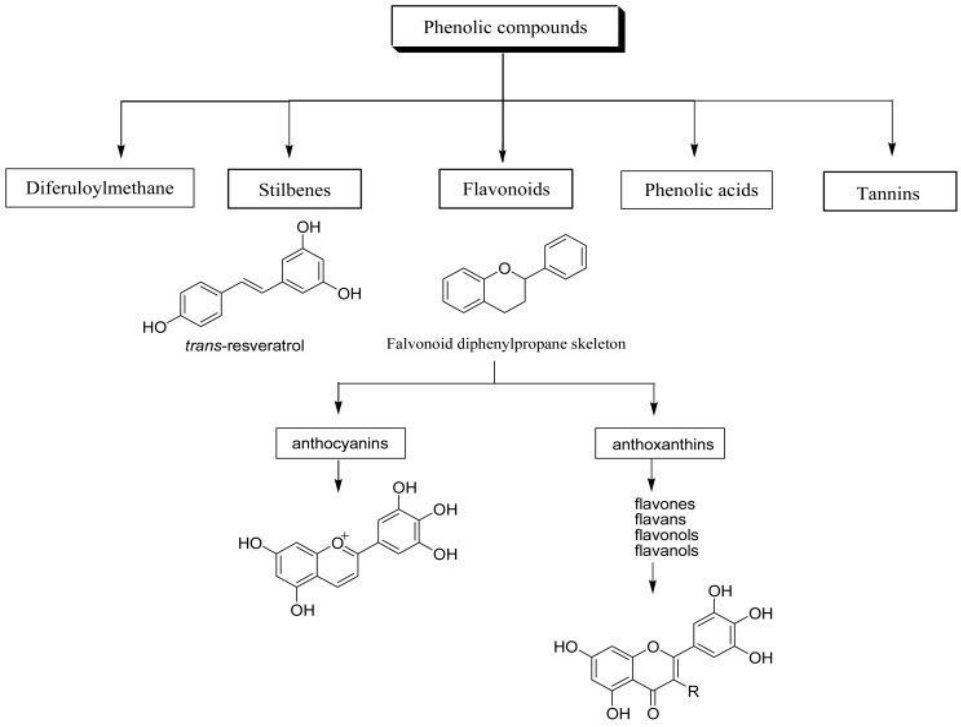


Figure 1.3 Classification and structure of phenolic compounds.

Source: (Han et al., 2007).

1.2.5 Antioxidant Activity of Blueberry Phenolics

Years of scientific research in the quest for "superfood" (Ferlemi & Lamari, 2016) suggests that it is now no longer a secret that fruits and vegetables contain metabolites that possess antioxidant activity, which is positively linked to its beneficial effects on health (Cao et al., 1996). Blueberry is well known for high levels of such metabolites and its ability to negate the harmful effects of free radicals and reactive oxygen species in the body. Blueberry possesses high levels of

hydroxycinnamic acids such as coumaric acid, caffeic acid derivatives, chlorogenic acid benzoic acids (Häkkinen et al., 1999; Määttä-Riihinen et al., 2004; Mattila et al., 2006; Taruscio et al., 2004) and flavonoids such as flavonols (quercetin derivatives), anthocyanidins, proanthocyanidins, catechins and their glycosides (Gavrilova et al., 2011; Giovanelli et al., 2013; Häkkinen et al., 1999). Anthocyanidins such as cyanidin, delphinidin, petunidin, peonidin, and malvidin derivatives are most commonly found in blueberries (Cardeñosa et al., 2016; Gao & Mazza, 1994; Giovanelli & Buratti, 2009). Anthocyanins are bioactive flavonoids responsible for the vibrant colours of fruits, leaves, and other parts of blueberry plants (Ferlemi & Lamari, 2016). The molecular structures of different kinds of anthocyanins present in blueberries are shown in Table 1.1, along with the colours they represent (Routray & Orsat, 2011). Proanthocyanidins are widely distributed in berries (Zifkin et al., 2012), they bind strongly with carbohydrates and proteins and act as potent free radical scavengers (Kähkönen et al., 2001).

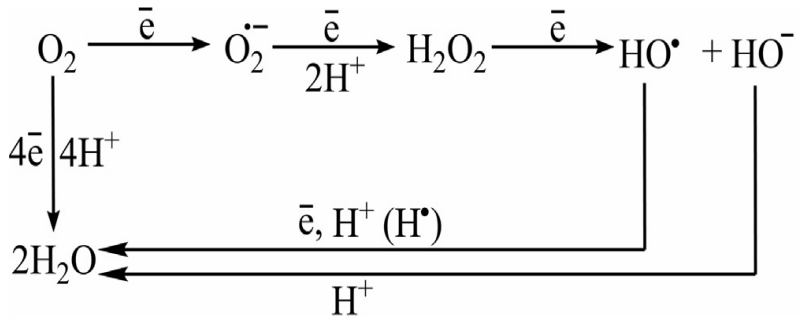
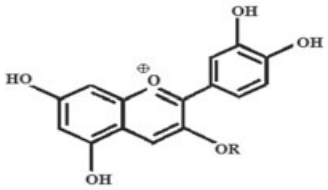
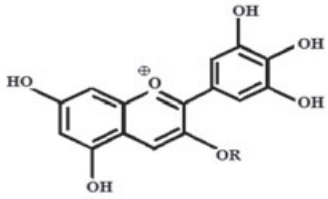
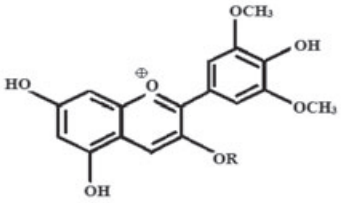
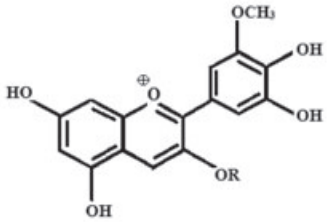
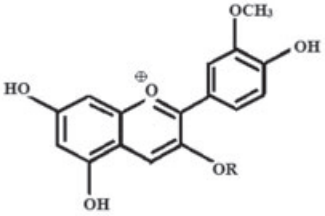


Figure 1.4 Reduction of molecular oxygen with electrons.

Source: (Lushchak, 2014).

Table 1.1 Anthocyanins commonly present in blueberries.

Anthocyanin	Basic structure	Colour
Cyanidin		Orange-red
Delphinidin		Blue-red
Malvidin		Blue-red
Petunidin		Blue-red
Peonidin		Orange-red

Source: (Routray & Orsat, 2011).

Blueberries are very well known for their high antioxidant activities. These antioxidant activities of blueberries depend on their phytochemical components, structures, and redox potential (Prior et

al., 1998; Wang, 2007). Phenolic compounds donate an electron or a hydrogen atom to a free radical with one or more unpaired electrons (Halliwell, 1991), convert it into a neutralized, non-harmful molecule, and thus act as antioxidant molecules *in vitro* and *in vivo* (Skrovankova et al., 2015).

The proof that free radicals are harmful first appeared during their study in 1954, where damage to lung tissue was observed as a result of excessive oxygen (Gerschman et al., 2001). The free radicals are more reactive than paired ones, and their reaction with biomolecules can be detrimental to biological species. The reactive oxygen species are generated due to the chemical and physical activation of dioxygen (O_2) present in the body (Perl-Treves & Perl, 2002). Most of the oxygen taken up by the organism is subjected to a reduction in a mitochondrial respiratory chain reaction to H_2O. If the reduction is carried out by one electron, then it produces superoxide radical ($O_2^{\cdot}$-), and if a second electron is introduced to $O_2^{\cdot}$-, it produces hydrogen peroxide (H_2O_2), which is not a free radical. Furthermore, the reduction of H_2O_2 by one electron generates H_2O and hydroxyl radical ($^{\cdot}OH$), one of the most prominent oxidants in the biological system. $O_2^{\cdot}$-, H_2O_2 and $^{\cdot}OH$ are called reactive oxygen species (ROS) and are produced naturally in organisms as a part of the regulatory process. However, ROS imbalance leads to oxidative stress, causing structural and functional changes to biomolecules (Castro & Freeman, 2001). Free radicals have been the main reason leading to DNA (deoxyribonucleic acid), protein, and lipid damage (Wang et al., 1996) and numerous other diseases, including cancer, cardiovascular diseases, neurodegenerative, diabetes, arthritis, and several other diseases (Halliwell, 1994, 1992, 1991; Yu, 1994). There are a number of reasons leading to an imbalance in ROS generation involved directly or indirectly. These include abiotic factors such as chill (Einset et al., 2007) and high temperature (Bruskov et al., 2002); xenobiotic compounds such as pollutants, for example, chlorine (Winder, 2001), ozone (Bromberg, 2016), and phosgene (Li & Pauluhn, 2017); organic compounds (Kumagai et al., 2012; Lin et al.,

2017; Moghe et al., 2015); and certain heavy metals (Cuypers et al., 2010; Gupta & Sandalio, 2011; Mahboob et al., 2001; Vaziri, 2008; Wu et al., 2013).

It is now well recognized that there exists a strong and positive relationship between total phenolic and anthocyanin content and antioxidant activity (Giovanelli et al., 2013; Moyer et al., 2002; Sellappan et al., 2002). The overall antioxidant activity may be a function of different phytochemicals, working additively or synergistically. The phytochemical compounds and their antioxidant activities in berries depend on the species, cultivars and varieties, degree of maturity, plant tissues, growing seasons and locations, environmental conditions and time of harvest, and postharvest conditions.

It is well-known that the content of phenolics and their profiles varies with the species, variety, cultivar, and genotype of blueberries. Antioxidant phenolics varied greatly between commercial and non-commercial blueberry species (Cardeñosa et al., 2016; Moyer et al., 2002). Rabbiteye blueberries contained a higher average polyphenolic concentration than northern and southern highbush blueberries (Sellappan et al., 2002; Wang et al., 2015). Lowbush blueberries comprised consistently higher levels of total phenolic content than highbush blueberries and blackberries (Kalt et al., 2001; Koca & Karadeniz, 2009). Similar observation made by Giovanelli & Buratti (2009) and Prior et al. (1998) where wild blueberry clones have much higher concentrations of total phenolics ranging from 2.99-6.00 mg/g than cultivated highbush blueberries, which ranged from 1.81-3.90 mg/g. Total phenolic content varied 2.1x (Prior et al., 1998), 3.4x (Ehlenfeldt & Prior, 2001), 2.4x (Howard et al., 2003) among highbush blueberry cultivars, 2.2x among southern highbush blueberries cultivars, and 3.4x among rabbiteye (Sellappan et al., 2002). The levels of phenolic compounds in blueberries vary significantly with inter-species variability and intra-species variations.

Lowbush blueberries contain a higher level of anthocyanins than highbush blueberries (Kalt et al., 2001; Vendrame et al., 2016). Half-high blueberry cultivars had higher levels of anthocyanidins than highbush cultivars (Li et al., 2017). Sellappan et al. (2002) reported that the average anthocyanin content in rabbiteye blueberries (1.14 mg/g) was higher than southern highbush blueberries (0.84 mg/g). The anthocyanin content among highbush blueberry cultivars varied 5x (Howard et al., 2003), 3.7x (Sellappan et al., 2002), and 2.2x (Ehlenfeldt & Prior, 2001) and among rabbiteye blueberry cultivars varied around 9x (Sellappan et al., 2002). Total anthocyanin content among ten cultivars and hybrids of lowbush blueberries ranged from 1.1 to 2.6 mg/g (Gao & Mazza, 1994). Variation in the levels of phenolic and anthocyanin compounds within the same species is mainly due to differences in the berry genotypes, differences in the growth and maturity stages of fruits, or plants' growth conditions.

Many studies reported vast differences in antioxidant activity among various types, species, cultivars, and genotypes of blueberries (Cardeñosa et al., 2016; Castrejón et al., 2008; Kalt et al., 1999). Lowbush blueberries have significantly higher antioxidant activity than highbush, rabbiteye, and southern highbush blueberries (Ehala et al., 2005; Kalt et al., 2001; Prior et al., 1998; Sellappan et al., 2002). Moyer et al. (2002) and Wang et al. (2015) reported that rabbiteye blueberries possessed the highest antioxidant activity among the genotypes of nine *Vaccinium spp*. The high concentration of anthocyanin content in the thick skin of rabbiteye blueberry could be responsible for such high antioxidant activity (Skrovankova et al., 2015). Antioxidant activity ranged between 19.7-38.3 μmol Trolox equivalent (TE)/g for the Georgia-grown rabbiteye and 8.11-26.5 TE/g for the southern highbush blueberry cultivars (Sellappan et al., 2002). Prior et al. (1998) reported that the total antioxidant activity ranged from 13.9 to 45.9 μmol TE/g among northern and southern highbush, rabbiteye, and lowbush blueberry genotypes harvested single year, with significant overlap in antioxidant activity values among the genotypes of different species. In a study, Howard

et al. (2003) reported that antioxidant activity determined as oxygen radical scavenging capacity (ORAC) of blueberry genotypes ranged from a low of 20.5 to 60.3 nmol TE/g, reflecting a 2.9x difference. Similar high variations were also reported in other studies involving blueberry cultivars and wild clones, where antioxidant activity varied 1.8x (Kalt et al., 1999), 2.5x (Prior et al., 1998), 3.3x (Sellappan et al., 2002), 4.7x (Connor et al., 2002), 5.2x (Moyer et al., 2002), and 6.8x (Ehlenfeldt & Prior, 2001). The variances in total phenolic and anthocyanin content between cultivars and wild clones at different maturity stages could be responsible for the changes in the antioxidant activity of blueberry.

1.3 Genetic Diversity

Genetic diversity is defined as the genetic variation within and among populations of a plant species, which helps plant breeders in selection for plant improvement. Therefore, it is important to conserve diversity for a sustainable human future. Chances are, one may find most of the diversity of a species be found within individual populations, or in some cases, among several different populations. The knowledge of diversity can help us determine what to conserve and where to conserve (Rao & Hodgkin, 2002). Assessment of genetic diversity in any particulate group of individuals can help organize and classify accessions and identify any subgroups of the principal group with potential efficacy in breeding programs (Mohammadi & Prasanna, 2003). There are many applications of correct analysis of genetic diversity, including identification of genetic variations among cultivars (Cox et al., 1986), that can help in the detection of different paternal groupings to have offspring with supreme genetic variability for further selection (Barrett & Kidwell, 1998) and introduction of desired genes into germplasm (Thompson et al., 1998).

1.3.1 Crop Improvement, Biodiversity, and Wild Clones

The domestication and improvement of crops have been in practice for many years so that humans can reap their benefits. Before discovering any sophisticated techniques, earlier civilizations had to rely on their visible diversity (phenotypes) for selection and breeding. There have been instances where crop productions were improved through such practices, such as maize, soybeans, wheat, and sorghum, during the 19th century (Kronstad, 1986). Genetic diversity makes this possible because breeders can choose a particular trait or a group of traits for breeding and further improvement. Genetic improvement is an essential part of crop improvement that was first recognized by Vavilov (1926). The recent improvements in DNA-based marker technologies for berry crops, especially for characters that do not allow visible characterization, would help scientists and breeders run crop improvement exercises with high accuracy and in a time-dependent manner (Debnath, 2016).

With numerous crop improvement programs in many plant species, there has been concern over genetic diversity's exhaustion, in turn affecting biodiversity. It is of prime importance that genetic diversity is maintained for species to survive and adapt to the rapidly changing world environment. The improvement and propagation of cultivars with higher agronomic values have replaced wild clones with heterogeneous cultivars. This loss of genetic diversity is contributed to by selective propagation and breeding and agricultural practices, climate change, urbanization, natural disasters, and movement of people on a large scale due to war (Richards et al., 1997). Increased awareness of berry fruits' health benefits has also increased its demand, which led to commercial cultivation. However, only select species with high value were cultivated to meet the farmers' demand, thereby ignoring other species. It was observed that wild blueberry cultivars have higher genetic diversity than the improved cultivars (Boches et al., 2006). The heterogeneity is declining

as a result of selective breeding in wild clones of southern highbush blueberries as well as cultivated highbush blueberry (Brevis et al., 2008). With the help of genetic diversity and mapping technology, one can determine the level of diversity between cultivated and wild-type cultivars, and efforts can also be made to introduce novel genetic diversity. There have been significantly few efforts in this direction and conservation efforts that require more attention. In Canada, Agriculture and Agri-Food Canada is commitment to the Canadian Biodiversity Strategy in response to the Convention on Biological Diversity (Debnath, 2016; Pairment, 1995). The Canadian Clonal Genebank introduced in 1989 and located in Harrow, Ontario, is responsible for conservation, characterization, virus indexing, and distribution of trees and small fruit (Debnath, 2016). A berry improvement program was established to develop blueberry cultivars at the Atlantic Cool Climate Crop Research Centre of Agriculture and Agri-Food Canada in Newfoundland and Labrador in 1999 (Debnath, 2000). As a part of this program, blueberry wild clones collected from various parts of North America are maintained at St. John's Research and Development Center (SJRDC), St. John's, Newfoundland and Labrador, for further breeding programs to generate better quality blueberry plants.

1.3.2 Estimation of Genetic Diversity

Markers are essential for the accurate measurement of genetic diversity. The first use of genetic markers to assess genetic diversity was documented in the early 19th century (Sax, 1923; Wexelsen, 1933). Markers measure the relatedness and differences at different levels, depending on the type of markers employed, between the individuals of a population or, in some cases, among different populations. There are mainly three types of marker systems, (1) morphological, (2) biochemical, and (3) DNA based markers, the first two had been traditionally used and still being used for berry genetic diversity analysis in combination with the third class of marker systems. The choice of

marker depends on its specific strengths and limitations, e.g., the abundance of markers, the nature of inheritance, etc.

1.3.2.1 Morphological Markers

Before the 1980s, genetic variations were evaluated by anatomy, morphology, embryology, and physiology (Debnath, 2016). The chemical contents analysis is essentially the profiling of compounds (metabolites) present in plants in response to various growth conditions. These metabolites act as markers of quality traits that help during breeding experiments. The selection of secondary metabolites, which are unique and able to differentiate between varieties, largely depends on the plant's ability to produce such metabolites. Such markers should be independent of external factors, such as the environment (Debnath, 2016). However, morphological traits of leaf, flower, and fruit were the traditional choices for the variation analysis (Dale, 1996). These markers have a limitation in distinguishing cultivars closely related, where morphological indices alone are not helpful. In addition to that, morphological characters are often affected by environmental effects. Examples of such studies based on morphology were done in *Fragaria* (Harrison et al., 1997; Harrison et al., 2000), *Rubus* (Ryabova, 2007; Vander Kloet & Dickinson, 1999), and *Vaccinium* (Česonienė et al., 2013) species. The constraints of morphological markers were overcome and often complemented with biochemical markers. Phylogenetic relationship studies, genetic diversity, and taxonomy using isozymes have been in practice for many years (Bretting & Widrlechner, 1995; Staub et al., 1996). Enzymes having different forms with the same catalytic function, but that are different at molecular levels are known as isozymes (Kumar et al., 2009).

1.3.2.2 Biochemical Markers

The development of electrophoretic assays, especially for isozymes for the separation of protein on the gel, has enabled the understanding of hereditary variation within organismal genomes (Markert & Møller, 1959). Scientists tried to screen isozyme markers to resolve zymotypes in blueberry where they used four enzyme systems viz., malate dehydrogenase (MDH), phosphoglucose isomerase (PGI), α-esterase, and peroxidase (Hill & Vander Kloet, 1983). While the first three systems showed uniformity across *Vaccinium spp*. section *Cyanococcus* and other sections, peroxidase system could produce slight variation among zymograms when mature leaves were used. The first use of such markers in inheritance studies in blueberry was documented by Vorsa et al. (1988). In his study, the extract prepared from leaf tissue of di-, tetra-, and hexaploid species of blueberry, *Vaccinium* section *Cyanococcus*, were analyzed electrophoretically for isozyme polymorphism using 12 enzyme systems. Out of which aldolase, shikimate dehydrogenase, triosephosphate isomerase, glucose-6-phosphate dehydrogenase, phosphoglucomutase, aconitase, alcohol dehydrogenase, and aspartate aminotransferase showed no definitive banding pattern. Four enzymes out of twelve, which showed good banding were malate dehydrogenase (MDH), phosphoglucose isomerase (PGI), 6-phosphogluconate dehydrogenase (6-PGD), and isocitrate dehydrogenase (IDH). The ratio of allozyme separation in the progeny of controlled diploid crosses supported single-gene Mendelian inheritance. Banding patterns of all allozymes suggested that the structure of these four enzymes was dimeric. The preliminary data of this study suggest that allozyme analyses can be useful in taxonomic studies in blueberry.

Tetrasomic inheritance in highbush blueberry was also reported in the following year (Krebs & Hancock, 1989). Efforts were made to find population genetic structure, genetic relationships, and

allozymic diversity among blueberry populations (Bruederle & Vorsa, 1994; Bruederle et al., 1991; Hokanson & Hancock, 1998). In 1991, Bruederle and his co-workers tried to analyze the distribution of genetic variation within and among the diploids with the help of allozyme data (Bruederle et al., 1991). Population genetic analyses were conducted on three species: *V. elliottii*, *V. myrtilloides*, and *V. tenellum*. These species presented high levels of variation within populations, where 18 loci were polymorphic. The number and frequency of alleles varied among taxa. Average values of proportion of polymorphic loci, alleles per polymorphic locus, and expected heterozygosity were 48.2, 2.8, and 0.148, respectively. High intra-specific genetic identities (> 0.96) indicated that these taxa are homogeneous.

Bruederle & Vorsa (1994) employed allozyme markers to evaluate the genetic relationships of diploid blueberry populations. Samples gathered from 25 sites were subjected to starch gel electrophoresis, and data were recorded at 11 polymorphic loci that were previously known to have high taxonomic information content. Data were scrutinized for similarity and standard measures of population genetic structure. Based on high allelic frequencies five genetic lineages were discriminated. Four years later, Karen Hokanson and Jim Hancock examined native Michigan populations of diploid *V. myrtilloides* and the tetraploids, *V. angustifolium*, and *V. corymbosum* for allozymic diversity. The level of heterozygosity and the number of alleles per locus were higher in the tetraploids, *V. corymbosum* (75.6%; 3.6), and *V. angustifolium* (57.1%; 3.4) than diploids, *V. myrtilloides* (21.7%; 2.9).

The limitation of isozyme markers is that there are fewer markers available as compared to DNA-based methods. Staining protocols are available for the limited number of enzyme loci, and also, with toxic staining ingredients and the inability to distinguish bands of two subunits of an enzyme, it makes its application reasonably limited. (Meerow, 2003; Tanksley & Orton, 1983).

1.3.2.3 Molecular Markers

Molecular markers are defined as a stretch of nuclear, chloroplast, and mitochondrial DNA that are highly heritable and carry conserved information (Kumar et al., 2009). The limitation of biochemical marks led to the development of DNA-based markers (Kan & Dozy, 1978), which can detect polymorphism in a nucleotide sequence. Polymorphism is defined as the presence of a stretch of DNA, which often codes for a gene simultaneously in the same population of two genotypes (Kumar et al., 2009). The sequences flanking such stretch of DNA creates a unique pattern on agarose or acrylamide gels upon electrophoresis called fingerprints. The commercially available linear stretches of DNA with a known size that carry the marker allow DNA fingerprints to be compared (Westermeier, 2016). DNA-based markers can be dominant (which cannot differentiate between alleles of a gene) or co-dominant (ones that can differentiate between alleles of a gene). The comparison of genetic material independent of an environment is the crucial feature of DNA-based markers (Weising et al., 1995), and it has a far better resolving capacity of genetic variability as compared to isozymes (Altukhov & Salmenkova, 2002). The ability to make billions of copies of short segments of desired DNA in no time using polymerase chain reaction (Mullis et al., 1986) has enabled scientists to use many sophisticated techniques for population genetic studies and diversity analysis (Bell et al., 2009). One of the first use of molecular markers was reported by Botstein et al. (1980). DNA polymorphism was detected using restriction fragment length polymorphism (RFLP) by constructing genetic maps. These markers have their advantages and disadvantages. To achieve the best genetic diversity results within a population, the marker system and statistical tool selection are the key factors after the generation of the dataset (Debnath, 2016). Some of the marker systems that are being used extensively from past decades include Simple Sequence Repeats (SSR; Hamada & Kakunaga, 1982; Tautz & Renz, 1984), Variable Number

Tandem Repeats (VNTR) or Directly Amplified Minisatellite DNA (DAMD; Heath et al., 1993; Jeffreys et al., 1990; Jeffreys et al., 1985), Sequence-Tagged Sites (STS; Olson et al., 1989), RAPD (Randomly Amplified Polymorphic DNA; Williams et al., 1990), SSCP (Single Strand Conformation Polymorphism; Hayashi, 1992; Hayashi & Yandell, 1993), Cleaved Amplified Polymorphic Sequence (CAPS) or Polymerase Chain Reaction-RFLP (PCR-RFLP; Konieczny & Ausubel, 1993), Sequence Characterized Amplified Region (SCAR; Paran & Michelmore, 1993), Inter-Simple Sequence Repeats (ISSR; Gupta et al., 1994; Meyer et al., 1993; Zietkiewicz et al., 1994), Randomly Amplified Microsatellite Polymorphism (RAMP; Wu et al., 1994), Single Primer Amplificaiton Reactions (SPAR; Gupta et al., 1994), Amplified Fragment Length Polymorphism (AFLP; Vos et al., 1995), Microsatellite Primed-PCR (MP-PCR; Weising et al., 1995), Single Nucleotide Polymorphism (SNP; Brookes, 1999; Cho et al., 1999; Primmer et al., 2002), Diversity Arrarys Technology (DArT; Jaccoud et al., 2001), Sequence-Related Amplified Polymorphism (SRAP; Li & Quiros, 2001), Target Region Amplificaiton Polymorphism (TRAP; Hu & Vick, 2003) and Start Codon Targeted (SCoT; Collard & Mackill, 2009). The marker techniques can be categorized into three types: (1) techniques that use hybridization and not PCR (e.g., RFLP), (2) techniques that use PCR application (e.g., AFLP, RAPD, Arbitrarily Primed (AP)-PCR, ISSR, SSR, Expressed Sequence Tags (EST), CAPS, SCAR, STS, and Sequence-Specific Amplification Polymorphism (SSAP)) and (3) DNA micro-chip based techniques. There are two subcategories in PCR-based techniques: (1) non-specific sequence PCR-based techniques such as RAPD & AFLP and (2) sequence-specific PCR-based techniques such as SSRs & STSs (Agarwal et al., 2008; Debnath, 2016). The selection of marker is to be determined by the purpose. Therefore, it is of prime importance to classify criteria through which a marker system is shortlisted. Ideally, a molecular marker should be polymorphic and co-dominant; also, it should be present frequently along the genome, readily available, and reproducible (Kumar et al., 2009).

1.3.2.3.1 RFLP (Restriction Fragment Length Polymorphism)

RFLP technique, the first marker technology to be employed, is the stepping stone of marker history. It was first used to detect DNA polymorphism in adenoviruses (Sambrook et al., 1975), later, in human genome mapping (Botstein et al., 1980) and plants (Helentjaris et al., 1986). The technique uses bacterial restriction enzymes (e.g., EcoR1), which recognize and cleave DNA at a specific size, which can be 4-8 base pairs (bp) in size, thereby creating a unique set of DNA fragments. Individuals under study can produce a different set of fragments as one or more sites may be altered, replaced, or relocated because of mutation, evolutionary processes and unequal crossing over (Schlötterer & Tautz, 1992). These fragments from individuals can be subjected to electrophoresis are then denatured and transferred to a nitrocellulose membrane using a technique called 'southern blotting' (Southern, 1975). This DNA immobilized membrane is then allowed to hybridize with the radioactively labeled probe specific to the DNA of interest. The membrane can then be visualized under ultra-violet radiation. The recent modification replaces the tedious Southern blot with PCR if the flanking regions of target sites are known. RFLP was used to examine variations and organelle inheritance in northern highbush blueberry cultivars and *V. ashei* using chloroplast and mitochondrial DNA, where DNA was cleaved using 23 restriction enzymes (Haghighi & Hancock, 1992). In another study, RFLP was used to differentiate 15 woody plants including *Vaccinium spp*., where polymorphic sites were exploited using *DpnII* and *HhaI* restriction enzymes (Bobowski et al., 1999).

1.3.2.3.2 RAPD (Random Amplified Polymorphic DNA)

The foundation of the RAPD technique is an amplification of genomic DNA based on the difference in the sequence. The polymorphism in the sequence may have been caused by mutations

(Williams et al., 1990). The technique does not require any sequence information, and therefore any random set of primers can be used and applied to many species. Because of its relatively fast speed and higher efficiency than RFLP, genetic maps were developed in many plants such as alfalfa (Kiss et al., 1993), faba beans (Torres et al., 1993), and apple (Hemmat et al., 1994). Because it produces false-positive bands, and because of its non-reproducibility, and being dominant, RAPD is a less preferred tool for genome-wide analysis. There are other variants of RAPD, such as AP-PCR (Welsh & McClelland, 1990) and DAF (DNA Amplification Fingerprinting; Caetano-Anollés & Bassam, 1993) that were developed based on specific applications.

Many researchers have used RAPD to assess the genetic relationship among blueberry populations, wild accessions, cultivars, and clones (Albert et al., 2004; Albert et al., 2005; Aruna et al., 1995; Aruna et al., 1993; Burgher et al., 2002; Debnath, 2005; Levi & Rowland, 1997). Thirty RAPD primers were used to screen 26 wild lowbush clones, clustered according to their geographical location (Burgher et al., 1998; Burgher et al., 2002). In 2005, Debnath used RAPD markers for characterization and assessment of genetic relatedness of three different *Vaccinium spp*. that included 13 wild cranberries (*V. macrocarpon* Ait) clones collected from Newfoundland and Labrador, Canada; and lowbush blueberry (*V. angustifolium* Ait) and lingonberry (*V. vitis-idaea* L) (Debnath, 2005). It resulted in informative, reproducible, and polymorphic banding patterns in 13 cranberry clones and 114 polymorphic bands out of 134 (Debnath, 2005). In 2007, Debnath reported very high genetic diversity (90%) among five cranberry cultivars and 43 wild clones collected from four Canadian provinces (Debnath, 2007). RAPD was used to explore the extent of genetic variation or closeness among 15 improved cultivars and four wild selections (Aruna et al., 1993). The genetic distance between improved cultivars was lower compared to four wild accessions, which may be due to inbreeding. In another study, Aruna et al. (1995) were able to

distinguish the genetic identity of two wild selections, ‘Ethel’ and ‘Satilla.’ They were also able to identify unknown samples following fingerprinting. Furthermore, no genetic variation among morphological variants of the cultivar ‘Brightwell’ was identified. A high proportion of genetic variation is invaluable in breeding, which was also observed within *V. uliginosum* L. (Albert et al., 2005), *V. mytrillus* L. (Albert et al., 2004), and *V. vitis-idaea* L. (Garkava-Gustavsson et al., 2005; Persson & Gustavsson, 2001) populations. In a recent genetic diversity study among 45 blueberry cultivars, 210 polymorphic bands were observed using RAPD markers (Cho et al., 2017). The cluster analysis of which divided them into two main clusters with a similarity value of 0.65. Cluster I consisted of four rabbiteye cultivars (Alapaha, Pink Lemonade, Titan, and Vernon) and a northern highbush cultivar, Ashworth. Whereas Cluster II comprised 31 northern highbush cultivars, eight southern highbush blueberry cultivars, and Northland half-highbush blueberry cultivars.

1.3.2.3.3 AFLP (Amplified Fragment Length Polymorphism)

AFLP technique was developed (Vos et al., 1995) to overcome the drawback of reproducibility in the RAPD technique. It is a PCR-based technique that uses the application of both RFLP and RAPD. This technique uses adaptors of a known DNA sequence added to each end of the fragments of genomic DNA generated by digestion through restriction enzymes and can be used as PCR primer sites. The different lengths of fragments generated are run through either capillary electrophoresis or PAGE (polyacrylamide gel electrophoresis). SAMPL (Selective Amplified Microsatellite Polymorphic Locus) is a variant of AFLP which uses an AFLP primer and a primer complementary to microsatellite sequences (Witsenboer et al., 1997). Thirty-two clones and their spatial distribution were identified using AFLP among 112 *V. myrtillus* populations by Albert et

al. (2004). Genetic diversity was measured among animal dispersed population of *V. membranaceum* (Douglas ex Torr.) using AFLP (Yang et al., 2008).

1.3.2.3.4 ISSR (Inter Simple Sequence Repeats)

They are DNA fragments of roughly 100-3000bp in size and are flanked by inversely placed microsatellite regions. It was first used by Zietkiewicz et al. (1994), where primers were designed for microsatellites to amplify inter-simple sequence DNA sequences. Primers used can be anchored at 3' or 5' end with a couple of degenerate bases into flanking regions or unanchored based on the application (Meyer et al., 1993; Zietkiewicz et al., 1994). Primers with overlap on flanking regions avoid self-priming and formation of smear (Reddy et al., 2002). It can generate up to 60 fragments of DNA upon electrophoresis on agarose or polyacrylamide gel that ranges from 200-2000 bp in size. Other variations of this technique include SPAR (Gupta et al., 1994) and DAMD (Heath et al., 1993). Three ISSR markers were used to analyze genetic diversity which separated 15 highbush blueberry cultivars, two rabbiteye blueberry cultivars and one southern lowbush selection from the wild (Levi & Rowland, 1997). Similarly, high polymorphism (80% polymorphic loci) was reported in ten highbush blueberry and three rabbiteye blueberry cultivars using six ISSR markers (Garriga et al., 2013). Thirteen ISSR primers were used to estimate the genetic variations among 43 wild lowbush blueberry clones from four Canadian provinces and in the cultivar 'Fundy' (Debnath, 2009). Fifteen primers generated 356 polymorphic ISSR-PCR bands while assessing genetic variability among 43 wild lingonberries (*V. vitis-idaea* ssp. *minus* Lodd.) clones collected from four Canadian provinces (Debnath, 2007). A high degree of genetic diversity was observed among the wild collections, as indicated by cluster analysis. AMOVA (Analysis of Molecular Variance) suggested that 10% of the total variation was due to geographical distributions. In another study, 34 clones and eight cultivars of lingonberry were subjected to diversity analysis using four ISSR

primers. They generated 113 polymorphic bands and could sufficiently differentiate lingonberry (Debnath & Sion, 2009).

1.3.2.3.5 Microsatellites or SSRs (Simple Sequence repeats)

SSR or microsatellite markers are tandem repeats or simple sequence repeats of DNA that flank the DNA region (coding or non-coding) and are present throughout the whole genome. The repeats can be di-, tri-, tetra-, or penta-nucleotide and are created during DNA replication because of strand slippage (Schlötterer & Tautz, 1992). If the sequence of these repeats is known, then primers can be designed complementary to the sequence and can be amplified with the help of PCR. The SSRs can be screened through various online sequence databases, and primers can be designed (Varshney et al., 2005). Because of their high variability, SSRs are highly preferred in distinguishing closely related cultivars (Vosman et al., 1992) and population studies (Smith & Devey, 1994). SSR markers were used to differentiate 69 accessions of blueberry and assess genetic diversity in cultivated and wild highbush blueberries (Boches et al., 2006). Genetic relationships and the effects of hybridization on genetic diversity were studied in southern highbush blueberry cultivars using SSR markers (Brevis et al., 2008). Similar studies were conducted to assess genetic diversity and population structure (Bian et al., 2014). SSR markers were used to estimate genetic diversity in southeastern *Vaccinium spp*. (Bassil et al., 2018; Bidani et al., 2017).

1.3.2.3.5.1 EST-SSR (Expressed Sequence Tagged-Simple Sequence Repeats)

SSRs are highly polymorphic, but they are mostly species-specific, and their development is expensive and time-consuming. With the availability of many genomic sequences of various species, one can look for a more specific and transferable SSR. To be able to use the application of SSR and still be able to make interspecies assessments, a search for gene-based SSRs that are

transferable across the species is required rather than random SSRs (Gupta et al., 2003). EST databases were sought as an excellent alternative to search for SSRs for a number of species (Boches et al., 2005; Boches et al., 2004; Graham et al., 2004; Gupta et al., 2003; Kantety et al., 2002; Kota et al., 2001). Thirty EST-PCR and EST-PCR-derived CAPSs markers were used to separate and identify genetic relationships among 15 highbush, two rabbiteye cultivars, and one wild selection each from *V. darrowi* and *V. corymbosium* (Rowland et al., 2003c). EST-PCR primers that were designed from highbush blueberries were used in cranberries (Rowland et al., 2003a), lowbush blueberries (Bell et al., 2008), and rabbiteye blueberries (Rowland et al., 2010). A study involving 28 wild blueberry clones, six half-high cultivars, and two selections of highbush blueberry used ten EST-PCR and two EST-SSR primers to investigate the genetic structure and diversity (Debnath, 2014). A similar study was done using 40 EST-SSR primers to differentiate 30 cultivars of blueberry (Mei et al., 2016).

1.3.2.3.6 SNP (Single Nucleotide Polymorphism)

SNP refers to change in a single base in a DNA sequence. SNPs are the most widely distributed molecular markers through genomes (Brookes, 1999). Their frequency of occurrence in the genome varies among different species; for example, maize has one SNP at every 60-120 bp (Ching et al., 2002), whereas humans have one at every 1000 bp (The International SNP Map Working Group, 2001). SNP can be present in coding or non-coding regions, where it can lead to change in the DNA sequence, thereby changing amino-acid sequences (Sunyaev et al., 1999), or staying identical, does not lead to change in amino-acid sequence. The former can result in a change in phenotypic expression (Richard & Beckmann, 1995). Many approaches were employed towards the discovery of SNP, including hetero-duplex analysis (Sorrentino et al., 1992), SSCP (Hecker et al., 1999), and Next-generation sequencing (NGS; Davey et al., 2011). The availability of NGS

and the availability of EST libraries have made it possible to evaluate genetic diversity at the DNA sequence level (Buetow et al., 1999; Soleimani et al., 2003). SNP genotyping is still not favorable despite being highly abundant and co-dominant because it is expensive and requires specialization. Traditionally, SNP assays include denaturing gradient gel electrophoresis (Cariello et al., 1988), SSCP (Suzuki et al., 1990), ligation chain reaction (Kälin et al., 1992), single base sequencing (Cotton, 1993), allele-specific oligonucleotide hybridization (Malmgren et al., 1996), primer extension, oligonucleotide ligation and invasive cleavage (Sobrino et al., 2005). Other approaches which involve SNP detection with sophisticated instruments include Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF; Pusch et al., 2002; Ross et al., 1998; Storm & Darnhofer-Patel, 2003), pyrosequencing (Ahmadian et al., 2000; Alderborn et al., 2000; He et al., 2003), Taqman allelic discrimination (Li et al., 2004), real-time PCR (Nurmi et al., 2001) and micro-array chips (Hacia et al., 1999). A genetic association map of an interspecific diploid blueberry population was constructed with the help of SNP and other primer systems (Rowland et al., 2014).

To be able to make the right choice of the marker based on study material and application, one must consider many factors such as availability, pros-cons, cost, etc. Table 1.2 (Adhikari et al., 2017; Kumar et al., 2009) summarizes the comparison among different markers.

Table 1.2 Comparison among different types of markers

	Type of marker system					
	RFLP	**RAPD**	**AFLP**	**ISSR**	**SSRs**	**SNP**
Requirement of DNA (ng)	10K	20	500-1K	50	50	50
Quality of DNA	Pure	Pure	Medium	Medium	Medium	Pure
Part of DNA gauged	Low copy-coding region	Whole genome	Whole genome	Whole genome	Whole genome	Whole genome
PCR required	No	Yes	Yes	Yes	Yes	Yes
Abundance	High	Very high	Very high	Medium	Medium	High
Level of polymorphism	Medium	High	Very high	High	High	High
Marker index	Low	Medium	High	Medium	Medium	Medium
Inheritance	Co-dominant	Dominant	Dominant	Dominant	Co-dominant	Co-dominant
Ease of utility	Laborious	Easy	Difficult	Easy	Easy	Difficult
Automation	Low	Medium	Medium	Medium	High	High
Reproducibility	High	Intermediate	High	Moderate-High	High	High

Table 1.2 cont'd

Development Cost	Low	Low	Moderate	Moderate	High	High
Radioactive detection	Yes	No	Yes	No	No	No
Cost of Analysis	High	Low	Moderate	Low	Low	High
Prior sequence knowledge required	No	No	No	Yes	Yes	Yes

Source: (Adhikari et al., 2017; Kumar et al., 2009)

1.4 Research Rationale and Objectives

In summary, blueberry is an economically and medicinally important plant. Genetic diversity studies should provide information about genotypes and their hybrids. This can be helpful to plant breeders to design breeding experiments to develop better varieties. The farmers can then select the genotype, which yields them high market returns. Much of the blueberry's antioxidant health aspect has been focused on antioxidants derived from berry fruits, thereby leaving out the potential of leaf antioxidants in health-promoting abilities. Previous studies on genetic diversity in blueberry only focussed on either genetic or biochemical diversity as separate objectives. There are no studies available that investigated association of molecular markers and biochemical traits.

Therefore, the present study was designed to bridge this gap. The general idea was to determine the extent of genetic diversity present among various blueberry genotypes and their association with phytochemical traits. Thus, the objectives of this dissertation are:

(1) Study genetic diversity among wild, cultivars, and hybrid blueberries (*Vaccinium spp.*) using EST-SSR, G-SSR, and EST-PCR molecular markers.

(2) Estimate total antioxidant activity, total flavonoid and phenolic content in leaves of these blueberry genotypes.

(3) Compare and analyse association of molecular markers with biochemical traits.

CHAPTER 2

Materials and Methods

2.1 Plant Materials

A total of 28 blueberry hybrids (Table 2.1), 36 lowbush blueberry wild clones (Table 2.2), and one lowbush, four half-high, and one highbush blueberry cultivars (Table 2.1) were used in the study. While the hybrids designated as HB1 to HB28 were selected from crosses between lowbush wild clones and half-high/highbush blueberries; the wild clones, named as BC1 to BC36, were collected from four Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC) and New Brunswick (NB). Each clone implied one wild plant selected based on vigor, berry colour, berry size and yield per plant, and free from diseases and insects. The distance between any two plants in the same place was above 10 m. Each selected plant was phenotypically different from the other and was considered an individual clone (Debnath, 2014).

Figure 2.1 Various blueberry plants maintained in a greenhouse at St. John's Research and Development Centre (St. John's RDC), Agriculture Agri-Food Canada, St. John's, Newfoundland and Labrador, Canada.

Table 2.1 Parentage of lowbush, half-high, and highbush blueberry cultivars and hybrids

Genotypes	
Lowbush cultivar	
Fundy (FUN)	An open-pollinated seedling of cultivar Augusta (Hall et al., 1988)
Half-high cultivar	
Patriot (PAT)	US3 (Dixi x Michigan lowbush No. 1) × Earliblue (Finn et al., 1990)
Chippewa (CHIP)	B18A (G65 × Ashworth) x US3 (Dixi × Michigan lowbush No. 1) (Okie, 1997)
St. Cloud (STC)	B19 (G65 × Ashworth) × US3 (Dixi × Michigan lowbush No. 1) (Finn et al., 1990)
Northblue (NOB)	B10 (G65 × Ashworth) × US3 (Dixi × Michigan lowbush No. 1) (Luby et al., 1986)
Highbush cultivar	
Polaris (POL)	B15 (G65 × Ashworth) × Bluetta (Okie, 1997)
Hybrids	
Cross 1 (HB1 – 11)	Wild clone × Half-high/highbush cultivar (Debnath, personal communication)
Cross 2 (HB12 – 28)	Wild clone × Half-high/highbush cultivar (Debnath, personal communication)

Table 2.2 Wild blueberry clones designated as BC1 to BC36, collected from Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB)

Clone	No. of clones	Province	Community	Latitude (N)	Longitude (W)
BC1 - 5, 9	6	NL	North River	47°32'	53°18'
BC6 - 8, 10	4	NL	Sears town	47°35'	53°17'
BC11 - 20	10	PE	Blooming Point	46°23'	62°58'
BC21, 27, 28	3	QC	Longue-River	48° 33'	69° 14'
BC22, 23, 26	3	QC	Baie-Trinite	49° 25'	67° 18'
BC24	1	QC	Pointe-Lebel	49° 09'	68° 13'
BC25	1	QC	Baie-Comeau	49° 13'	68° 08'
BC29 - 36	8	NB	Little Shemogue	46°06'	64°01'
Total	36				

The lowbush blueberry cultivar Fundy (FUN) was selected from open-pollinated seedlings of cultivar Augusta and was developed as a cultivar at Kentville Research and Development Centre, AAFC, NS, Canada (Hall et al., 1988). The parentage of highbush blueberry cultivar Polaris (POL), four half-high blueberry cultivars, Patriot (PAT), Chippewa (CHIP), St. Cloud (STC), and Northblue (NOB), and 28 hybrids are presented in Table 2.1. The hybrids were developed from two cross combinations between selected wild clones and half-high and/or highbush blueberries.

Hybrids HB1 to HB11 were from Cross No. 1, and HB12 – 28 were from Cross No. 2 (Debnath, personal communication; parentage are not disclosed). The lowbush blueberries were less than 0.5 m in height, and highbush blueberries were 2 – 2.5 m tall. Half-high and hybrid blueberries used in the study were of intermediate height between lowbush and highbush blueberries. All genotypes were grown and maintained in a greenhouse in 6 L plastic pots containing 2 peat: 1 perlite mixture, under natural light conditions (maximum PPF 90 µmol m^{-2} s^{-1}) at 20 ± 2 °C, 85% relative humidity. Standard cultural practices were followed to maintain the plants (Debnath, 2014). In a replication of three from each plant, fresh young leaves were collected, shock-frozen with liquid nitrogen, and stored at – 80 °C until DNA and phenolic extraction for molecular and biochemical analyses.

2.2 Molecular Analysis

2.2.1 DNA Extraction, PCR Amplification, and Electrophoresis

Two hundred mg of stored young leaf sample was used to isolate DNA using DNeasy Plant Mini Kit (Qiagen GmbH, Hilden, Germany), with some modifications (Debnath, 2014). Frozen leaves were homogenised with 550 µL buffer AP1 with the help of FastPrep - 24 Tissue and Cell Homogenizer (MP Biomedicals, Santa Ana, CA, USA). Four µL of RNase A was added to the homogenized mixture and incubated for 60 min at 65 ℃ and contents were mixed several times at regular intervals. Further steps were followed as per manufacturer instructions. The purity and concentration of DNA were measured spectrophotometrically (Ultrospec 2000, Pharmacia Biotech, Cambridge, UK). DNA with an A260/A280 absorbance ratio of 1.7–2.1 was diluted (concentration: 10 ng mL^{-1}) and used as template DNA for PCR reactions (Debnath, 2014). A total of ten EST-SSR (CA23, CA112, CA169, CA236, CA421, CA483, CA787, NA800, NA961 and NA1040), eight genomic SSR (G-SSR; VCC_B3, VCC_I2, VCC_I8, VCC_J1, VCC_J3, VCC_J9, VCC_K4,

and VCC_S10) and eight EST-PCR primer pairs (CA21, CA54, CA227, CA287, CA791, CA1029, CA1423, and NA27; Table 2.3) that were found effective for blueberries (Bell et al., 2008; Boches et al., 2005) were obtained from Integrated DNA Technologies, Inc. (IDT), Coralville, Iowa, USA. The EST-PCR primer pairs were developed from floral buds of cold-acclimated (CA) and non-acclimated (NA) highbush blueberries (Rowland et al., 2003b). SSR markers were the derivatives of SSR-enriched genomic libraries and EST libraries constructed from highbush blueberries (Boches et al., 2005). The annealing temperature of all primers was standardized using temperature gradient PCR that ranged from 49 ℃ for the EST-SSR primer pair CA112 to 62 ℃ for EST-SSR primer pair NA800 and G-SSR primer pairs VCC_B3, VCC_B12, VCC_J9 and VCC_K4 (Table 2.3).

Table 2.3 EST-SSR, genomic (G-) SSR and EST-PCR primer pairs: their sequences and annealing temperatures

Primer pair[1]	Forward and Reverse primer sequence (5'-3')	Annealing temperature (℃)
EST-SSR		
CA23	F: GAGAGGGTTTCGAGGAGGAG R: GTTTAGAAACGGGACTGTGAGACG	60
CA112	F: TCCACCCACTTCACAGTTCA R: GTTTATTGGGAGGGAATTGGAAAC	49

Table 2.3 cont'd

CA169	F: TAGTGGAGGGTTTTGCTTGG R: GTTTATCGAAGCGAAGGTCAAAGA	52
CA236	F: GTTAAGCTTTTAGATGAGTTGATGG R: GTTTAACCAGTCCCAGACCCAAAT	54
CA421	F: TCAAATTCAAAGCTCAAAATCAA R: GTTTAAGGATGATCCCGAAGCTCT	58
CA483	F: GTCTTCCTCAGGTTCGGTTG R: GAACGGCTCCGAAGACAG	58
CA787	F: TCCTCGTTCTCTCCCTCTCA R: GTTTCGCTGAAGTTGGAGTCCTT	60
NA800	F: CAATCCATTCCAAGCATGTG R: GTTTCCCTAGACCAGTGCCACTTA	62

Table 2.3 cont'd

NA961	F: TCAGACATGATTGGGGAGGT R: GTTTGGAATAATAGAGGCGGTGGA	56
NA1040	F: GCAACTCCCAGACTTTCTCC R: GTTTAGTCAGCAGGGTGCACAA	58
G-SSR		
VCC_B3	F: CCTTCGATCTTGTTCCTTGC R: GTTTGATGCAATTGAGGTGGAGA	62
VCC_K4	F: CCTCCACCCCACTTTCATTA R: GCACACAGGTCCAGTTTTTG	62
VCC_ S10	F: ATTTGGTGTGAAACCCCTGA R: GTTTGCGGCTATATCCGTGTTTGT	60

Table 2.3 cont'd

EST-PCR		
CA21	F: TCCGATAACCGTTACCAAGC R: TATACAGCGACACGCCAAAA	52
CA54	F: CCGGTGAACTTCCACTTGTT R: AGATACTACTGGGGGTGGGG	52
CA227	F: TGGAGACTGGAGTGATGCAA R: TTTGCAAGAACCATGCTGAG	56
CA287	F: AGGGCTTTCCCTCAATCACT R: CCTTGTTGTTCCTTCCTTCG	58
CA791	F: AGAGCCAAAAGAAGGGGAAG R: TCAAAATTTTCCGGACCAG	56
CA1029	F: GAAGTTTTCCGTTCTCTGCAA R: CTGCAGCTAGGACCGAAGAG	52

Table 2.3 cont'd

CA1423	F: TCATAGCCAATACACTCGAACC R: GCCCCACCTTTAGCAAACTC	60
NA27	F: CGCTCGCTCCATTGTTTC R: TATGCATGAAGCTTGCCGTA	60

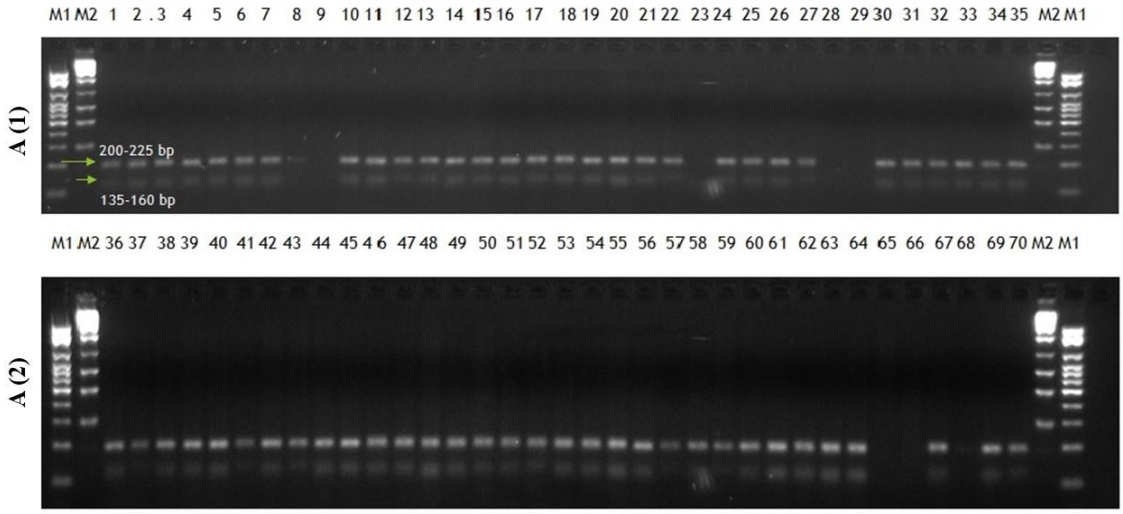
[1]Above markers were developed from EST libraries generated from floral buds of cold-acclimated (designated as prefix CA) and non-acclimated (designated as prefix NA) highbush blueberry plants (Rowland et al., 2003b).

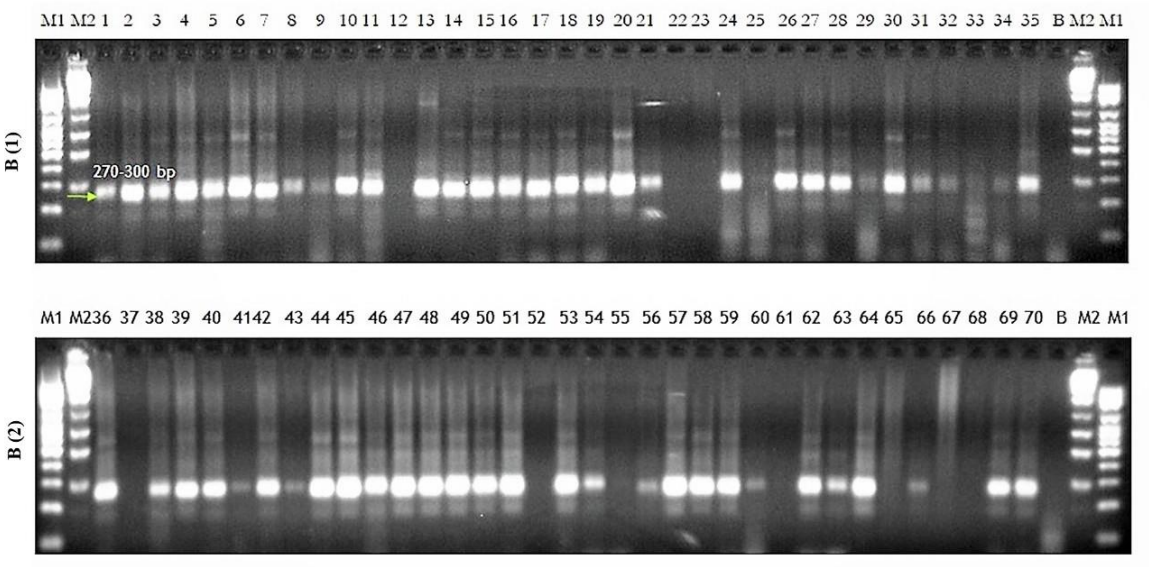
The PCR was carried out in an optimized amplification reaction mixture (25 µL) containing 20 ng of template DNA, 1× PCR buffer (1.5 mM MgCl2, pH 8.7; Qiagen), 200 µM of each deoxynucleotide triphosphate (dNTP), 0.2 µM of each of the 20 forward and reverse primers and 0.63 unit of Taq DNA polymerase (Qiagen). Mastercycler EP Gradient S (Eppendorf AG, 22331 Hamburg, Germany) was used to amplify DNA, which was programmed for a 10-min initial "hot start" denaturation step at 94 ℃, and then 40 cycles of 40 s of denaturation step at 92 ℃, 70 s annealing step at appropriate annealing temperature and 2 min extension step at 72 ℃. The final extension step was at 72 ℃ for 10 min, and then the sample was held at 4 ℃. The amplified DNA products were separated by electrophoresis using 2% agarose 3:1 high-resolution blend (HRB) (Ameresco, Solon, OH) gel pre-casted with 2× Tris-borate EDTA buffer and 1× GelRed nucleic acid stain (Biotium Inc., Hayward, CA, USA) solution along with a low range 100 bp DNA ladder and a midrange 1 kb DNA ladder (Norgen Bioteck Corp., Thorold, ON, Canada). UV light enabled visualization, scoring, and recording banding patterns in a trans-illuminating gel documentation

system (InGenius 3, Syngene, Beacon House, Cambridge, UK). The length of the DNA fragment was calculated by Gene Tools software (Syngene) by comparison with standard size marker mobility (Tailor et al., 2017).

2.2.2 Data Collection and Statistical Analysis

EST-SSR, G-SSR, and EST-PCR markers are co-dominant in nature that can discriminate between homozygous and heterozygous individuals. In polyploid plants like blueberry, it is very difficult to perform the co-dominant scoring of alleles in heterozygote samples. As it is extremely difficult to calculate the number of alleles present at a particular locus from band intensities, the only suggested way of scoring is to record the absence or presence of an allele denoted as 0 or 1 respectively in a matrix and all bands from one primer were treated as alleles at one locus (Esselink et al., 2003; Horvath et al., 2011).


M1 M2 1 2 .3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 M2 M1
A (1)
200-225 bp
135-160 bp
M1 M2 36 37 38 39 40 41 42 43 44 45 4 6 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 M2 M1
A (2)


M1 M2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 B M2 M1
B (1)
270-300 bp
M1 M236 37 38 39 40 4142 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 B M2 M1
B (2)

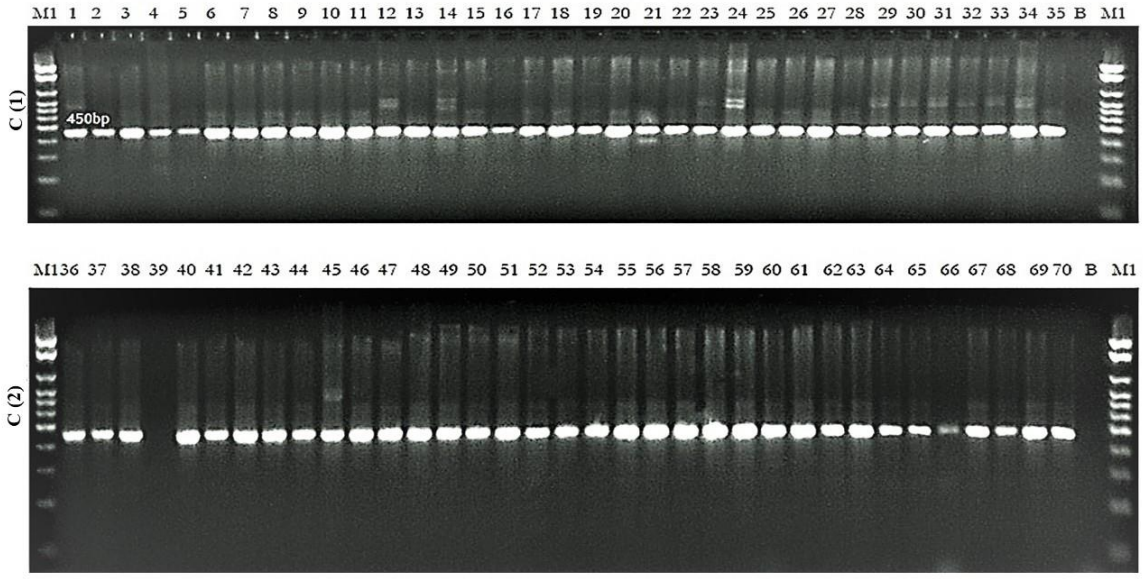


Figure 2.2 An example of banding pattern of blueberry genotypes expressed by A(1) & A(2) EST-SSR (NA961); B(1) & B(2) G-SSR (VCC_J1); and C(1), C(2) EST-PCR primer (NA27). M1 and M2 are 100bp and 1K ladders, respectively. Lanes 1 to 70 represents genotypes in order, as shown in Table 2.2.

Indices including the polymorphic information content (PIC) (Botstein et al., 1980), effective multiplex ratio (EMR; Powell et al., 1996), discrimination power (D) (Tessier et al., 1999), and resolving power (R) (Prevost & Wilkinson, 1999) were calculated using a program, Online Marker Efficiency Calculator (iMEC) (Amiryousefi et al., 2018). These indices give an idea about the primer's ability to discriminate among genotypes and the primer system's overall utility.

2.2.3 Genetic Diversity Indices

2.2.3.1 Polymorphic Information Content (PIC)

The PIC is a value that reflects the marker's ability to detect polymorphism within the population. It gives an approximation of a primer's discrimination capacity based on the number of alleles that are expressed and respective allelic frequencies. This was used to evaluate the level of

informativeness of each primer (high, PIC > 0.5; moderate, 0.5 > PIC > 0.25; low, PIC < 0.25) (Botstein et al., 1980). The PIC value of an l-allele locus is calculated as:

$$PIC=1-\sum_{i=1}^{n} Pi^2-\sum_{i=1}^{n-1}.\sum_{j=i+1}^{n} 2Pi^2Pj^2$$

where Pi and Pj are the population frequency of the i^{th} and j^{th} allele, $i \neq j$ and n = number of alleles.

2.2.3.2 Effective Multiplex Ratio (EMR)

The associated value EMR was calculated as the product of a total number of polymorphic loci and the portion of polymorphic loci (Powell et al., 1996) as:

$$EMR= n \cdot (n_p/n_p+n_{np}),$$

where n is total number of loci, n_p is the number of polymorphic loci, and n_{np} is the number of non-polymorphic loci. Therefore, the higher the value of EMR, the more competent the primer is.

2.2.3.3 Marker Index (MI)

MI is another associated statistical parameter used to evaluate a marker system (Powell et al., 1996). It was calculated as:

$$MI = PIC \cdot EMR$$

2.2.3.4 Discrimination Power (D)

Primer discrimination power, D, is defined as the probability that two randomly chosen individuals have different banding patterns and are, therefore, differentiable (Tessier et al., 1999). This was calculated as:

$$D = 1 - C,$$

where C is the confusion probability. For the i^{th} pattern of the given j^{th} primer, present at frequency pi in a set of varieties, $C = \Sigma ci = \Sigma pi\ Npi\text{-}1/N-1$ where for N individuals, C is equal to the sum of all ci for all of the patterns generated by the primer.

2.2.3.5 Resolving Power (R)

Resolving power, R, was calculated as:

$$R = \sum Ib,$$

where Ib or band informativeness is denoted on a scale of 0 or 1 and is described as $Ib = 1 - (2 \times |0.5 - p|)$; p is defined as the samples' portion of the observed band. The resolving power or the primer's capability to differentiate between genotypes could be represented by the sum of these adjusted values for all generated bands (Prevost & Wilkinson, 1999).

To compare diversity among blueberry genotypes, indices such as percentage of polymorphic loci (PL), observed (N_a) and the effective number of alleles (N_e) (Hartl & Clark, 1997), expected heterozygosity/Nei gene diversity index (H_e) (Nei, 1973b) and Shannon's information index of diversity (I) (Shannon & Weaver, 1949) were calculated using GenAlEx version 6.5 (Peakall & Smouse, 2012).

2.2.3.6 Percentage of Polymorphic Loci (PL)

It is a portion of loci that are polymorphic regardless of their frequency multiplied by 100.

$$PL = 100.\frac{n_p}{n}$$

Where n_p is the number of polymorphic loci, and n is the total number of loci.

2.2.3.7 Number of alleles (N_a)

It is determined by directly counting the average number of alleles per loci.

2.2.3.8 Effective Number of Alleles (N_e)

It is an estimate of the number of alleles with equal frequencies in a population. It allows the comparisons of allelic diversity among loci with a wide range of allele frequency distributions (Brown & Weir, 1983). The effective number of alleles (N_e) is calculated as:

$$N_e = 1/ (1-H_e),$$

where H_e is expected heterozygosity.

2.2.3.9 Expected Heterozygosity/Nei's Gene Diversity Index (H_e)

He is defined as the probability of an individual is heterozygous at a given locus in the population (Liu, 1998) and is calculated using following formula:

$$He=1- \sum_{i=1}^{n} Pi^2$$

where, Pi is the frequency for the i^{th} allele among a total of n alleles. It is also known as Nei's gene diversity index (Nei, 1973a).

2.2.3.10 Shannon's Information Index of Diversity (I)

Shannon-Weaver index (I; Shannon & Weaver, 1949) is within the population index (Hennink & Zeven, 1990). It is described as the index for information:

$$I = \sum_{i=1}^{z} X_i \,.^2\log X_i$$

where X_i is an allele frequency at the i^{th} allele in z number of groups.

2.2.4 Population Structure Analysis

2.2.4.1 STRUCTURE Analysis

A Bayesian clustering approach for population structure analysis was used for 70 genotypes using STRUCTURE program ver. 2.3.4 (https://web.stanford.edu/group/pritchardlab/structure_software/release_versions/v2.3.4/html/structure.html). The software uses Markov Chain Monte Carlo (MCMC) estimation to determine the number of sub-populations (ΔK) (Pritchard et al., 2000). Using this model, some populations (K) were presumed to be present, and each of them was characterized by a set of allele frequencies at every locus. Genotypes in the sample were assigned to clusters (populations) or jointly to more populations if their genotypes indicated that they were admixed. Every locus was thought to be independent, and each K population was presumed to follow Hardy-Weinberg equilibrium. The ΔK method (Evanno et al., 2005) was used to determine the most likely number of K. The number of genetically different clusters (K) was kept to range 1 to 10, with five independent runs followed by a burn-in length of 100,000 and 100,000 iterations. A web application (http://royfrancis.github.io/pophelper/) of software POPHELPER, an R package, was used to estimate the number of population clusters and their visualization (Francis, 2017).

2.2.4.2 Unweighted Neighbour-Joining Tree (NJ)

The software DARwin 6.0.9 (Perrier & Jacquemoud-Collet, 2006) was used to depict phylogenetic trees with EST-SSR, G-SSR, and EST-PCR markers using the unweighted neighbour-joining method (NJ). Jaccard's coefficient (Sneath & Sokal, 1973) was used to calculate the dissimilarity matrix with 30,000 bootstraps.

$$d_{ij} = \frac{b+c}{a+(b+c)}$$

where *dij*: dissimilarity between units *i* and *j*,

a: number of variables where *i* = presence and *j* = presence

b: number of variables where *i* = presence and *j* = absence

c: number of variables where *i* = absence and *j* = presence

2.2.4.3 Principal Coordinate Analysis (PCoA)

Principal Coordinate Analysis (PCoA) is a multivariate technique, that was performed using GenAlex 6.503 (Peakall & Smouse, 2012) to validate clusters of STRUCTURE and NJ. This analysis can reveal and plot key patterns within a multivariate data set. For multidimensional data sets, each consecutive axis explains proportionately less of the total variation, such that when the groups are distinct. The first two or three axes will typically reveal most of the separation among them. The procedure in GenAlEx is based on an algorithm published by Orlóci (1978).

2.2.3.4 Analysis of Molecular Variance (AMOVA)

A hierarchical analysis of molecular variance (AMOVA; Excoffier et al., 1992) was performed using GenAlex 6.503 (Peakall & Smouse, 2006). The blueberry genotypes were divided into seven

groups, out of which four groups comprised of wild clones collected from four Canadian provinces (NL, PE, QC, and NB). Group 5 consisted of six blueberry cultivars, group 6, 11 hybrids from the first cross (HB1-11), and group 7, 17 hybrids from the second cross (HB12-28). This analysis allowed structuring variations by EST-SSR, G-SSR and EST-PCR markers into among- and within-the group variation components and provides a measure of inter-community genetic distances as the proportion of total EST and genomic variations residing among blueberry genotypes of any two communities (called Phi statistics; Excoffier et al., 1992).

2.3 Biochemical Analysis

2.3.1 Preparation of Leaf Extracts

Five hundred mg of shock-frozen leaves of each genotype were homogenized in a FastPrep-24 Tissue and Cell Homogenizer (MP Biomedicals, Irvine, CA, USA) containing a solution of 80% aqueous acetone and 0.2% formic acid (1:4 g/mL) (Ghosh et al., 2018; Goyali et al., 2015). The homogenate was retained at 4 °C with slow agitation for 30 min, followed by centrifugation at 13,000 rpm at 4 °C for 15 min using a Allegra 64R (Beckman Coulter Inc., Palo Alto, CA, US), and the supernatant was collected. Extraction was performed twice more with the pallets, and the supernatant was mixed with the original crude extract. The extracts were saved in an ultralow freezer (Thermo Scientific, Burlington, ON, Canada) for further determination of antioxidant capacity and the total phenolic and flavonoid contents. All chemical analyses were conducted thrice with each sample, and mean values were used for analysis.

2.3.2 Total Antioxidant Activity (TAA)

The free radical scavenging activity was estimated as percentage inhibition of DPPH (2, 2-diphenyl-1-picrylhydrazyl) radicals (Ghosh et al., 2018). An aliquot of diluted extract or gallic acid standard solution (5 mg/mL; ≥98% purity) was added to 1.7 mL DPPH methanolic solution (0.06 mM), mixed thoroughly, and kept in the dark for 20 min at room temperature. The mixture's absorbance was examined at 517 nm using Ultrospec 4300 pro, UV/Visible Spectrophotometer (Amersham Biosciences Corp. San Francisco, CA, USA). Blank was prepared using aqueous acetone (80%) mixed with the DPPH solution. The gallic acid at (≥98% purity) standard curve was prepared, and the linearity of the gallic acid standard curve ($r^2 = 0.98$) was obtained in the range of 20–80 µg/mL. The results were expressed as milligrams gallic acid equivalents (GAE) per gram of fresh leaf (mg GAE/g fl). Following formula was used to calculate the percentage inhibition (Mishra et al., 2012):

$$\% \text{ Radical scavenging activity} = [(\text{Absorbance}_{(\text{Blank})} - \text{Absorbance}_{(\text{Extract})})/\text{Absorbance}_{(\text{Blank})}] \times 100$$

2.3.3 Total Phenolic Content (TPC)

An optimized Folin-Ciocalteu (FC) method (Goyali et al., 2015; Singleton & Rossi, 1965) was used to estimate total phenolic content. One hundred microliters of FC reagent was added to the diluted leaf extract (100 µL; 200 µg/mL), and 200 µL of 20% saturated (w/v) sodium carbonate was added to it after 5 min followed by 1.5 mL distilled water. The mixture was incubated in the dark for 35 min at room temperature and centrifuged at 4000 × g for 10 min in Allegra 64R centrifuge. The absorbance of the gallic acid standard solution (5 mg/mL) and test samples were measured with Ultrospec 4300 pro spectrophotometer at 725 nm wavelength after 3 min. The absorbance values were recorded linearly of the standard calibration curve ($r^2 = 0.98$) for gallic

acid standard solution (5 mg/mL; ≥98% purity) taken in a range of 2.5–10 μg/mL, and outcomes were presented as mg GAE/g fl.

2.3.4 Total Flavonoid Content (TFC)

Total flavonoid content was estimated using a colorimetric assay (Zhishen et al., 1999) following Goyali et al. (2015). 500 μL of sample extract was added to 2mL distilled water followed by 150 mL of 5% (w/v) sodium nitrate to which 150 mL of 10% (w/v) aluminum chloride was added after 5 min. 1 M sodium hydroxide solution (1 mL) was added to the mixture after 6 min of incubation at room temperature. The mixture was diluted with 1.2 mL distilled water, and the absorbance was measured at 510 nm using Ultrospec 4300 pro. Catechin solution (1 mg/mL) was used in a range of 20–200 μg/mL for standard curve calibration (r^2 = 0.99), and the total flavonoid content was calculated as milligrams of catechin equivalents per gram of fresh leaves (mg CE/g fl).

2.3.5 Cluster Analysis of Antioxidant Properties

Data for antioxidant activity, phenolic, and flavonoid contents are presented as mean value ± standard deviation (SD) of three replications. The TAA result among groups was statistically evaluated by variance analysis (ANOVA), and Tukey's test was employed for comparing treatment means at a critical difference (P) of ≤ 0.05. Because the residuals of TPC and TFC among groups followed non-normal distribution, a violation of one of the preconditions of ANOVA, the observations were statistically evaluated by a non-parametric Kruskal-Wallis test, with a significance value fixed at ≤ 0.05 (McKight & Najab, 2010). Similarly, the residuals of TAA, TPC, and TFC among individuals followed non-normal distribution; the observations were statistically evaluated by a non-parametric Kruskal-Wallis test, with a significance value fixed at ≤ 0.05. The correlation coefficient (r), coefficient of determination (r^2), and linear regression between TPC and

TFC, TPC and TAA, and TFC and TAA were analyzed at a confidence interval of 95%. To eliminate the effects of difference scales of measurement, the biochemical data were standardized by subtracting mean values from the original values followed by division with SD (An et al., 2015). Euclidean dissimilarity distance matrix was generated using these standardized values. An Agglomerative Hierarchical Clustering (AHC) method, an algorithm that works from the dissimilarities between various individuals and group them, was used to generate an Unweighted Pair-Group Method with Arithmetic mean (UPGMA) dendrogram based on Euclidean dissimilarity distance matrix of antioxidant activity and phenolic and flavonoid contents data. Principal Component Analysis (PCA), a multivariate technique that analyzes data matrices of several correlated quantitative dependent variables, was performed for individuals based on biochemical data as well as for biochemical components (Abdi & Williams, 2010). The above analysis was performed using XLSTAT 2020 Version: 22.3.21.0 (©Addinsoft, New York, NY, USA).

2.4 Correlation between Biochemical and Molecular Analysis

A mantel test (Louati et al., 2019; Mantel, 1967) was performed to check the correlation between Euclidean dissimilarity distance matrix generated using standardized values of biochemical data in XLSTAT 2020 Version: 22.3.21.0 (©Addinsoft, New York, NY, USA) and genetic distance matrices of EST-SSR, G-SSR, EST-PCR, and all primers combined generated using GenAlEx version 6.5 (Peakall & Smouse, 2012).

Association between EST-SSR, G-SSR, and EST-PCR markers and biochemical attributes of blueberry leaf extracts (TAA, TPC, and TFC) was estimated by stepwise multiple regression analysis (SMRA; Efroymson, 1960) using SPSS version 27 (IBM Corp., Armonk, NY, USA). The biochemical components were treated as the dependent variable, and molecular markers were

treated as the independent variable. The F value criteria was set between 0.045-0.099 for inclusion or removal of independent variables for regression (Afifi et al., 2003).

CHAPTER 3

Results

3.1 Molecular Analysis

3.1.1 Diversity Parameters

3.1.1.1 Analysis of Primer's Discriminatory Power

For ten EST-SSR primers, the polymorphic information content (PIC) values varied between 0.03 for the marker CA23 to 0.96 for CA112 with an average of 0.35 (Table 3.1), which suggested that CA112 is the most informative and CA23 the poorest among EST-SSR primer pairs. All other EST-SSR primers except CA787 and NA961 fell into a moderate category with PIC values between 0.5 and 0.25 (Botstein et al., 1980). While markers CA236 and NA961 had the highest effective multiplex ratio (EMR) value (1.80), CA112 was the poorest with an EMR value of 0.08. Marker index (MI) ranged from 0.03 to 0.65 for the markers CA23 and CA421, respectively. The discrimination power (D) was the highest for CA236 (0.91) and the lowest for CA23 (0.03). The resolving power (R) ranged from 0.03 (CA23) to 2.91 (CA421). EST-SSR primer CA23 is the poorest primer pair among all primers in the category with the lowest MI, D, and R values. The highest values for MI and R were observed in CA421, followed by CA236 (Table 2.4). However, the latter was the best for D values (0.91), and it was followed by CA483 (0.90).

All eight G-SSR primer pairs fell into the moderate informative category with PIC values ranging from 0.26 to 0.37 except VCC_12 and VCC_J1, which had a PIC value of 0.21. For G-SSR, the EMR values ranged from 0.63 (VCC_B3) to 2.44 (VCC_J9), and the MI ranged from 0.18 (VCC_I2 and VCC_J1) to 0.63 (VCC_J9 and VCC_K4). A similar trend was also observed for D

and R's values, where VCC_I2 and VCC_J1 possessed the lowest values (0.27 and 0.29, respectively), and VCC_K4 acquired the highest values (0.94 and 3.11, respectively). Therefore, VCC_K4 becomes the most powerful primer pair with the highest MI, D, and R values, while VCC_I2 and VCC_J1, the poorest with the lowest values of MI, D, and R among G-SSR primer pairs (Table 3.1).

When eight EST-PCR primers were used, PIC was the highest for CA227 and CA1423 (0.37) and the lowest for NA27. The EMR values ranged from 0.86 (CA287) to 3.57 (CA54). The MI was the highest for CA227 (1.16) and the lowest for NA27 (0.03). The trends for D and R values are almost similar. The highest value (0.96) for D was observed in CA21, CA791, and CA54, while NA27 attained the lowest value (0.03). The highest value (4.31) of R was observed in CA791, followed by CA54 (4.11), while the lowest being 0.03 for NA27 (Table 3.1).

Among three groups of primers, the average values of EMR, MI, D, and R suggest that the EST-PCR primer system is the most effective primer system compared to the G-SSR and EST-SSR primer systems (Table 3.1).

Table 3.1 Discriminatory power indices of EST-SSR, genomic (G-) SSR and EST-PCR markers for diversity analysis in blueberry hybrids, wild clones, and cultivars.

Primers	PIC	EMR	MI	D	R
EST-SSR					
CA23	0.03	0.99	0.03	0.03	0.03
CA112	0.96	0.08	0.08	0.08	0.09
CA169	0.37	1.11	0.41	0.69	0.97
CA236	0.33	1.80	0.60	0.91	2.17
CA421	0.37	1.74	0.65	0.81	2.91
CA483	0.34	0.96	0.33	0.90	1.40
CA787	0.20	0.87	0.17	0.24	0.26
NA800	0.37	1.10	0.41	0.70	1.57
NA961	0.16	1.80	0.29	0.19	0.40
NA1040	0.37	0.86	0.32	0.82	0.40
Mean	0.35	1.13	0.33	0.54	1.02

Table 3.1 cont'd

G-SSR					
VCC_I2	0.21	0.86	0.18	0.27	0.29
VCC_B3	0.34	0.63	0.21	0.90	0.86
VCC_I8	0.37	1.01	0.38	0.74	0.31
VCC_J1	0.21	0.86	0.18	0.27	0.29
VCC_J3	0.36	0.79	0.29	0.85	0.49
VCC_J9	0.26	2.44	0.63	0.34	1.11
VCC_K4	0.30	2.13	0.63	0.94	3.11
VCC_S10	0.33	1.53	0.51	0.91	2.20
Mean	0.30	1.28	0.38	0.65	1.08
EST-PCR					
CA21	0.28	1.47	0.41	0.96	2.94
CA54	0.28	3.57	0.99	0.96	4.11
CA227	0.37	3.11	1.16	0.80	3.03

Table 3.1 cont'd

CA287	0.32	0.86	0.28	0.92	0.74
CA791	0.27	2.16	0.57	0.96	4.31
CA1029	0.29	1.56	0.45	0.95	1.80
CA1423	0.37	2.69	1.00	0.71	2.00
NA27	0.03	0.99	0.03	0.03	0.03
Mean	0.28	2.05	0.61	0.79	2.37

PIC, polymorphic information content; EMR, effective multiplex ratio; MI, marker index; D, discrimination power; R, resolving power.

3.1.1.2 Analysis of Population Genetic Diversity

For EST-SSR primers, the percentage of polymorphic loci (PL) among seven populations ranged from 29% for cultivars to 79% for PE clones and hybrid group Cross 2 with a mean of 62% (Table 3.2). The observed (N_a) and effective number of alleles (N_e) were also highest in hybrid group Cross 2 (N_a = 10.08; N_e = 1.48) and lowest in cultivars (N_a = 3.75; N_e = 1.30). Nei's gene diversity or expected heterozygosity (H_e) and Shannon's information index (I) were also highest in the Cross 2 group (H_e = 0.28, I = 0.42), while H_e was the lowest in cultivars and hybrid group Cross 1 (H_e = 0.17). Hybrid group Cross 1 has the lowest value for I (0.25), followed by cultivars (0.26). The average values for H_e and I were 0.23 and 0.34, respectively (Table 3.2).

Table 3.2 Population genetic diversity parameters of blueberry hybrid groups (Cross 1: HB1-11; Cross 2: HB12-28), wild clones collected from Canadian provinces Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), and six cultivars (CV).

	Sample size	PL (%)	N_a	N_e	H_e	I
EST-SSR						
Cross 1	11	46	7.56	1.30	0.17	0.25
Cross 2	17	79	10.08	1.48	0.28	0.42
NL	10	71	6.44	1.48	0.27	0.40
PE	10	79	6.42	1.44	0.26	0.39
QC	8	71	4.99	1.42	0.24	0.35
NB	8	58	4.89	1.42	0.23	0.34
CV	6	29	3.75	1.30	0.17	0.26
Mean		62	6.30	1.41	0.23	0.34
G-SSR						
Cross 1	11	44	6.25	1.30	0.17	0.24
Cross 2	17	64	8.38	1.37	0.21	0.32

Table 3.2 cont'd

NL	10	60	5.64	1.29	0.18	0.28
PE	10	48	5.75	1.28	0.17	0.25
QC	8	60	4.02	1.44	0.24	0.35
NB	8	48	4.50	1.38	0.20	0.29
CV	4	32	2.97	1.31	0.18	0.27
Mean		51	5.36	1.34	0.19	0.29
EST-PCR						
Cross 1	11	50	5.06	1.28	0.16	0.24
Cross 2	17	76	7.92	1.33	0.20	0.31
NL	10	53	4.20	1.22	0.14	0.22
PE	10	66	4.22	1.27	0.17	0.27
QC	8	55	3.53	1.26	0.16	0.25
NB	8	47	3.15	1.23	0.14	0.21
CV	4	47	2.75	1.28	0.18	0.28
Mean		56	4.40	1.27	0.16	0.26

PL, percentage of polymorphic loci; N_a, number of alleles; N_e, effective number of alleles; H_e, Nei's genetic diversity or expected heterozygosity; I, Shannon's information index.

For G-SSR primer pairs, the PL, like for EST-SSRs, was highest in hybrid group Cross 2 (64%) and lowest in cultivars (32%) with an average of 51%. The number of observed alleles was the highest in hybrid group Cross 2 (8.38), followed by Cross 1 (6.25). The QC clones had the lowest observed alleles (4.02) and had the highest effective alleles (N_e = 1.44). The average number of alleles was found the lowest in cultivars (N_a = 2.97) and PE (N_e = 1.28) populations. The QC clones had the highest H_e (0.24) and I (0.35) values, and it was followed by hybrid group Cross 2 (H_e = 0.21; I = 0.32). The H_e value was lowest in the Cross 1 group and PE wild clones (H_e = 0.17). Hybrid group Cross 1 also showed the lowest I value (0.24), followed by PE clones (0.25) (Table 3.2).

For the EST-PCR primer system, the values of PL (75.86%), Na (7.92), N_e (1.33), H_e (0.20), and I (0.31) were the highest for the HS2 population. The lowest level of polymorphism (46.55%) was observed in the NB population and cultivars. The number of alleles (N_a = 2.75) was also lowest for cultivars. The effective number of alleles (N_e = 1.22) was minimum in the NL population. NB population also had the lowest values of 0.14 and 0.21 for Nei's gene diversity or expected heterozygosity (H_e) and Shannon's information index (I), respectively (Table 3.2).

3.1.1.2.1 STRUCTURE Analysis

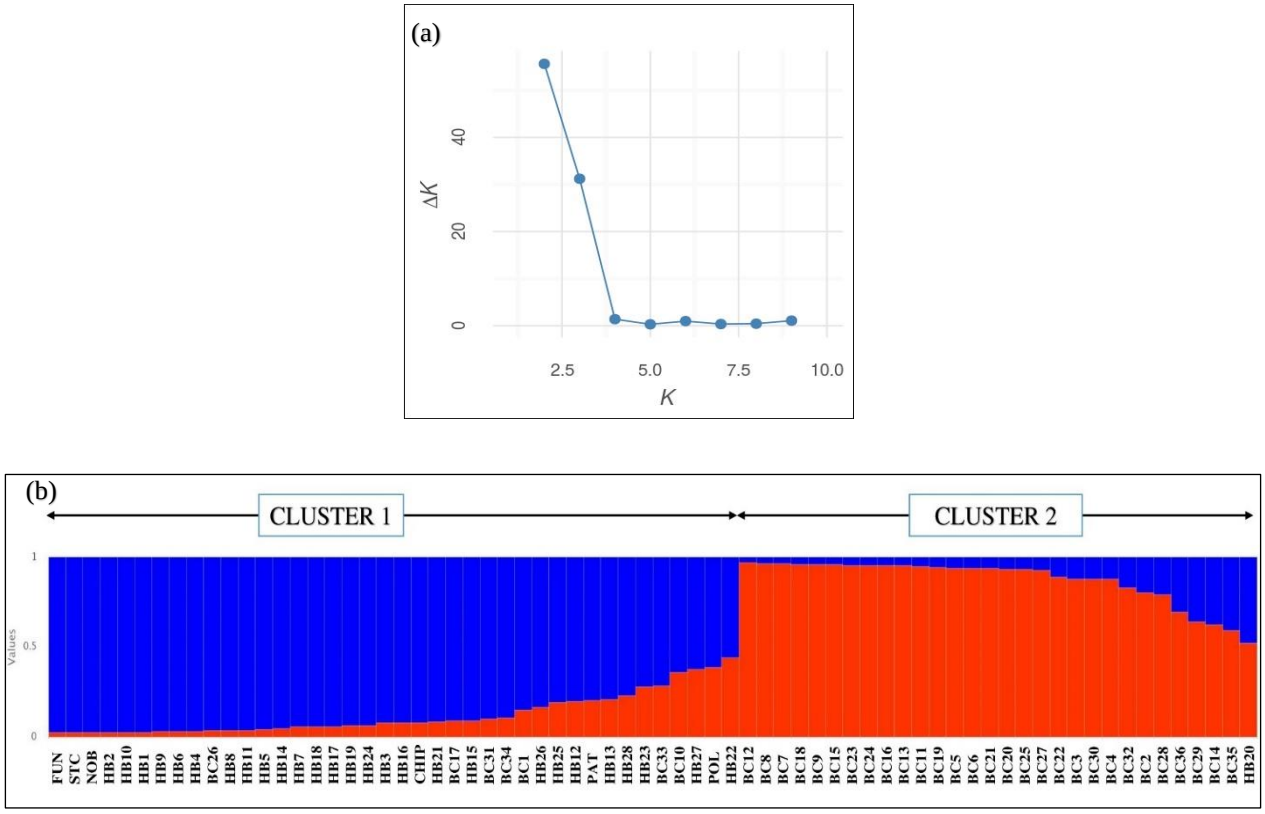


Figure 3.1 (a) ΔK is reaching its maximum value at K = 2 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 2) based on EST-SSR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label).

A Bayesian clustering approach was implemented to define population structures while ignoring any prior geographical information related to sampled individuals. The STRUCTURE analysis using EST-SSR (Figure 3.1a) and G-SSR (Figure 3.3a) markers suggested the number of clusters K = 2, as indicated by the modal value of ΔK predominant peak at K = 2. For EST-SSR, while Cluster 1 comprised of 40 genotypes including all six cultivars (Fundy, FUN; Patriot, PAT; Chippewa, CHIP; St. Cloud, STC; Northblue, NOB; Polaris, POL) and all hybrids except HB20, Cluster 2 contained 30 genotypes comprising of hybrid HB20 and all wild clones except BC1, 10, 17, 26, 31, 33 and 34. HB20 had, however, around 48% admixture from Cluster 1. Cultivar POL

clones BC10 and 33, and hybrids HB22, 23, and 27, although grouped in Cluster 1, had 28 – 44% admixture from Cluster 2 (Figure 3.1b).

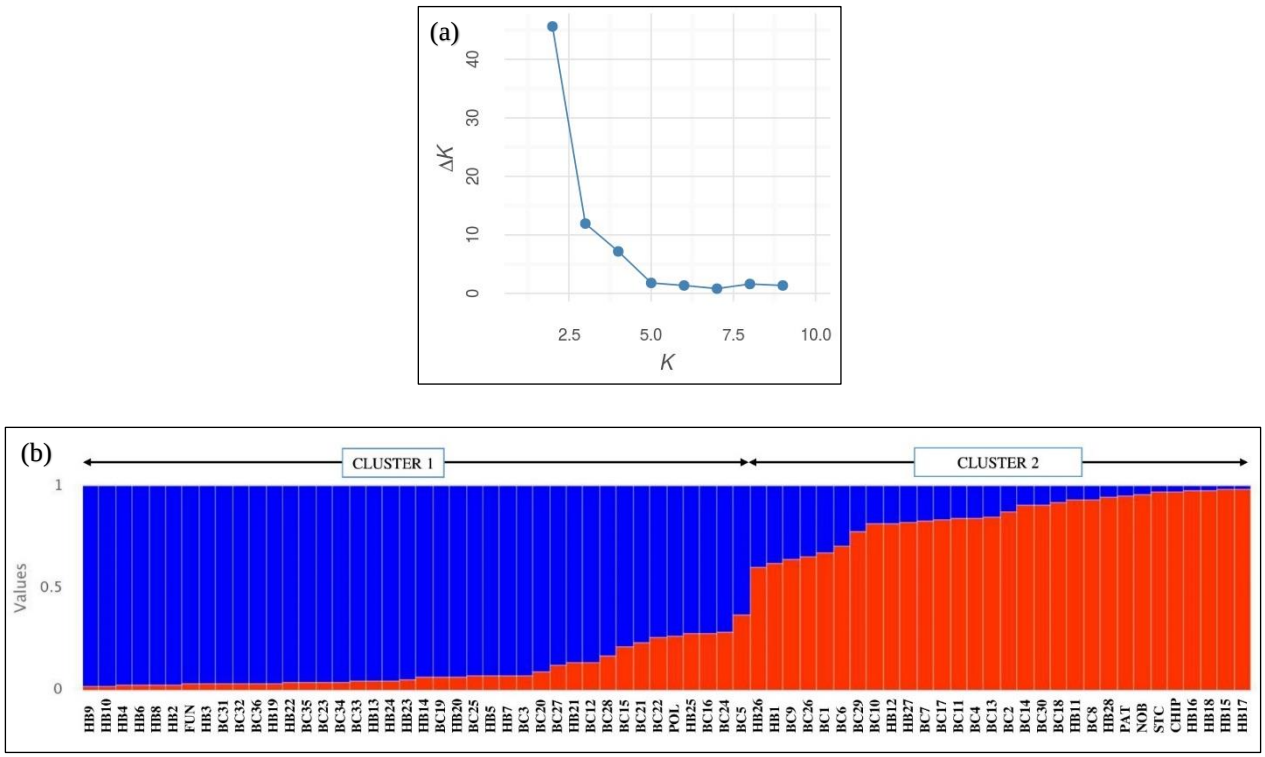
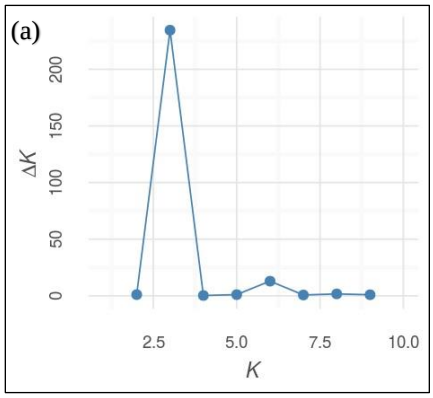


Figure 3.2 (a) ΔK is reaching its maximum value at K = 2 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 2) based on genomic (G)-SSR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label).

With G-SSR primers, the first cluster contained two NL (BC3 and 5), five PE (BC 12, 15, 16, 19, 20), seven QC (BC 21 – 25, 27, 28), and six NB clones (BC 31– 36); two cultivars Fundy and Polaris, and 18 hybrids (HB 2 – 10, 13, 14, 19 – 25). The rest of the genotypes were structured into cluster 2. While as many as eight genotypes of Cluster 1 had around 21 – 38% admixture from Cluster 2 (BC5, 15, 16, 21, 22, 24; POL, HB 25); in Cluster 2, hybrids HB1 and 26, and clones BC 1, 6, 9, 26 and 29 were with around 30 – 38% admixture from Cluster 1 (Figure 3.2b).



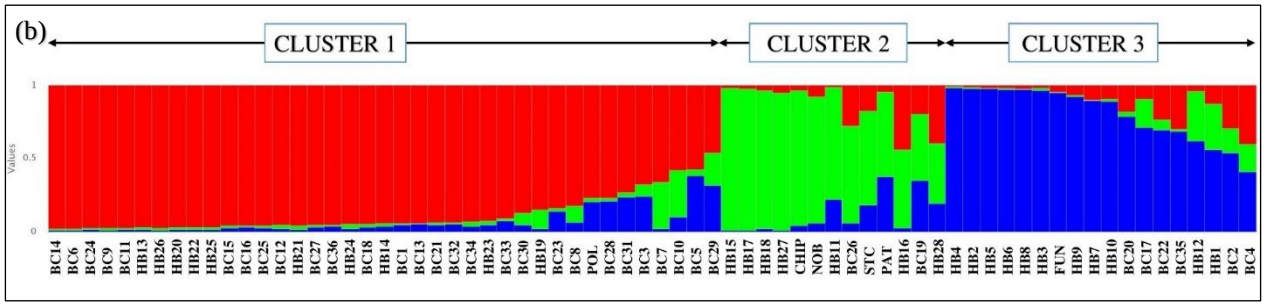


Figure 3.3 (a) ΔK is reaching its maximum value at K = 3 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 3) based on EST-PCR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label).

STRUCTURE analysis for EST-PCR primers suggested three clusters (K = 3), as indicated by the modal value of ΔK (Figure 3.3a). The 70 blueberry genotypes are loosely separated according to their phenotype into three clusters (Figure 3.3b). Cluster 1 consists eight NL (BC1, 3, 5 – 10), seven PE (BC11 – 16, 18), six QC (BC21, 23 – 25, 27, 28) and seven NB wild clone (BC29 – 34, 36); the highbush cultivar Polaris and ten hybrids (HB13, 14, 19 - 26). Polaris showed an admixture of around 4% and 21% from Cluster 2 and Cluster 3, respectively. Other genotypes that showed admixture from Cluster 2 (around 15 – 35%) include BC7, 10 and 29, and hybrids HB19, and from cluster 3 (around 13 – 38%) were BC3, 5, 23, 28, 29, and 31. Cluster 2 consisted of 13 genotypes, including four half-high cultivars (PAT, CHIP, STC, NOB), clones BC19 and 26, and hybrids HB11, 15 – 18, 27, and 28. In Cluster 2, the half-high cultivars, St. Cloud and Patriot, and hybrid HB28 showed an admixture from Cluster 1 (around 19%, 5%, and 39%, respectively) and Cluster

3 (around 19%, 38%, and 19%, respectively). In this cluster, hybrid HB16 showed an admixture of around 43% from Cluster 1. In the same cluster, the lowbush wild clone BC19 showed an admixture of around 20% from Cluster 1 and 32% from Cluster 3. Cluster 3 consisted of two NL (BC2, 4), two PE (BC17, 20), one QC (BC22), and one NB wild clones (BC35); the lowbush cultivar Fundy and 11 hybrids (HB1 – 10, 12). Clones BC2 and BC4 had admixtures from Cluster 1 (about 29% and 41%, respectively) and Cluster 2 (about 17% and 23%, respectively). The hybrids HB1 and 12 had around 14% and 3.6%, respectively; cluster 2, around 30%, and 32%.

The combination of all three types of markers suggested three clusters (K = 3), as indicated by the modal value of ΔK in the Bayesian clustering approach (Figure 3.4a). The classification was consistent with phenotypic groupings as all wild clones except BC17 and 26 were in Cluster 1. Cluster 1 also included the highbush cultivar Polaris and ten hybrids (HB13, 14, 19 - 26). In Cluster 1, clones; BC22, BC30 and BC31 showed admixtures from cluster 2 (around 3%, 31% and 3%, respectively) and cluster 3 (around 30%, <1% and 31%, respectively). Cluster 2 comprised of two wild clones (BC17, 26), all four half-high cultivars (PAT, CHIP, STC, NOB), and nine hybrids (HB1, 11, 12, 15 – 18, 27, 28). In this cluster, while clones BC17 and 26 had admixture from Cluster 1, clones; BC22, BC30 and BC31 showed admixtures from cluster 2 (around 3%, 31% and 3%, respectively) and cluster 3 (around 30%, <1% and 31%, respectively). Cluster 2 comprised of two wild clones (BC17, 26), all four half-high cultivars (PAT, CHIP, STC, NOB), and nine hybrids (HB1, 11, 12, 15 – 18, 27, 28). In this cluster, while clones BC17 and 26 had admixtures from Cluster 1 (around 22% and 6%, respectively) and cluster 3 (around 16% and 23%, respectively), hybrid HB1 showed about 3% admixture from Cluster 1 and about 38% from Cluster 3. The lowbush cultivar Fundy was grouped with nine other hybrids (HB2 – 10) in Cluster 3, with no significant admixture from other groups (Figure 3.4b).

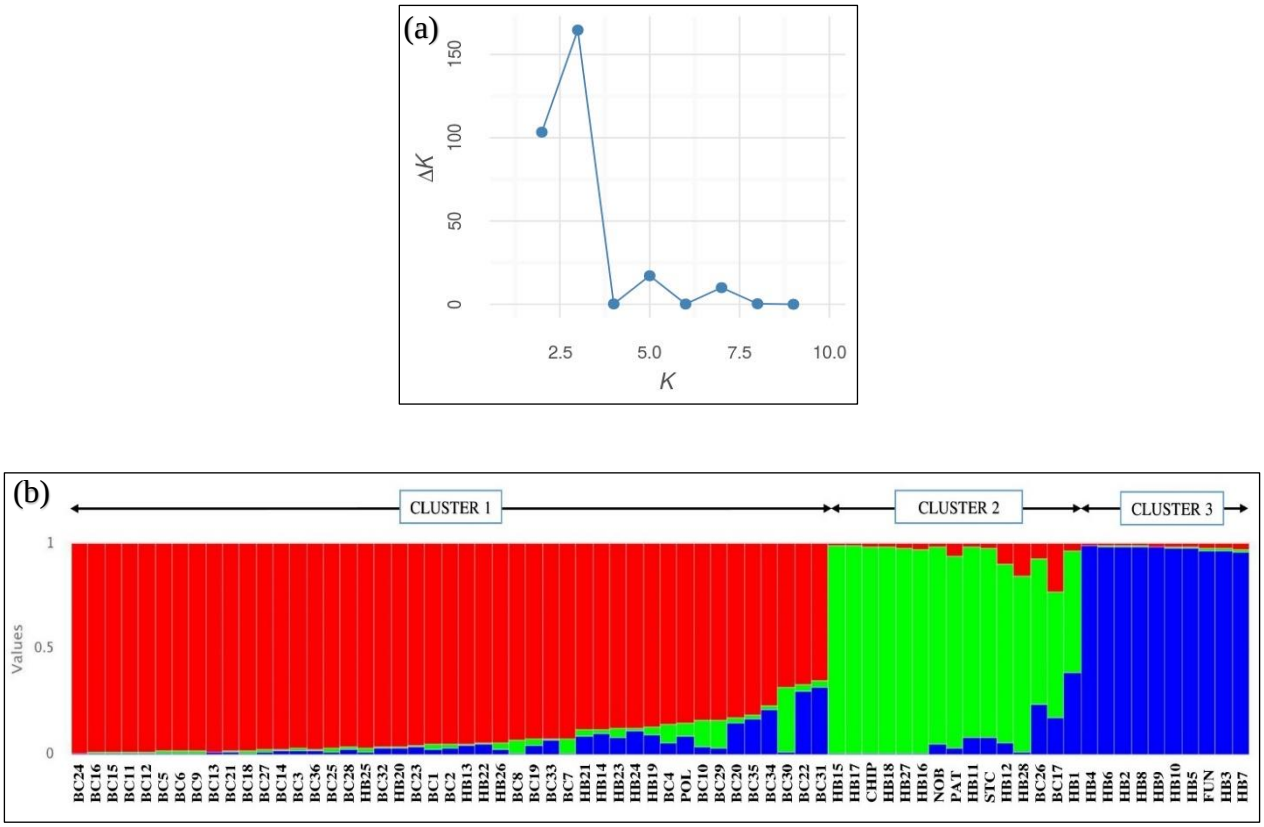


Figure 3.4 (a) ΔK is reaching its maximum value at K = 3 following ad-hoc method; (b) Distribution of blueberry genotypes as per STRUCTURE analysis (K = 3) based on the combination of EST-SSR, genomic (G) - SSR, and EST-PCR primer pairs. The genotypes are represented as vertical bars, and the colour represents different clusters (see Tables 2.1 and 2.2 for genotype label).

3.1.1.2.2 Unweighted Neighbour-Joining (NJ) Tree

The NJ analysis displayed inter-individual distances graphically. Blueberry genotypes were resolved with statistical confidence based on Jaccard's dissimilarity coefficients (Sneath & Sokal, 1973).

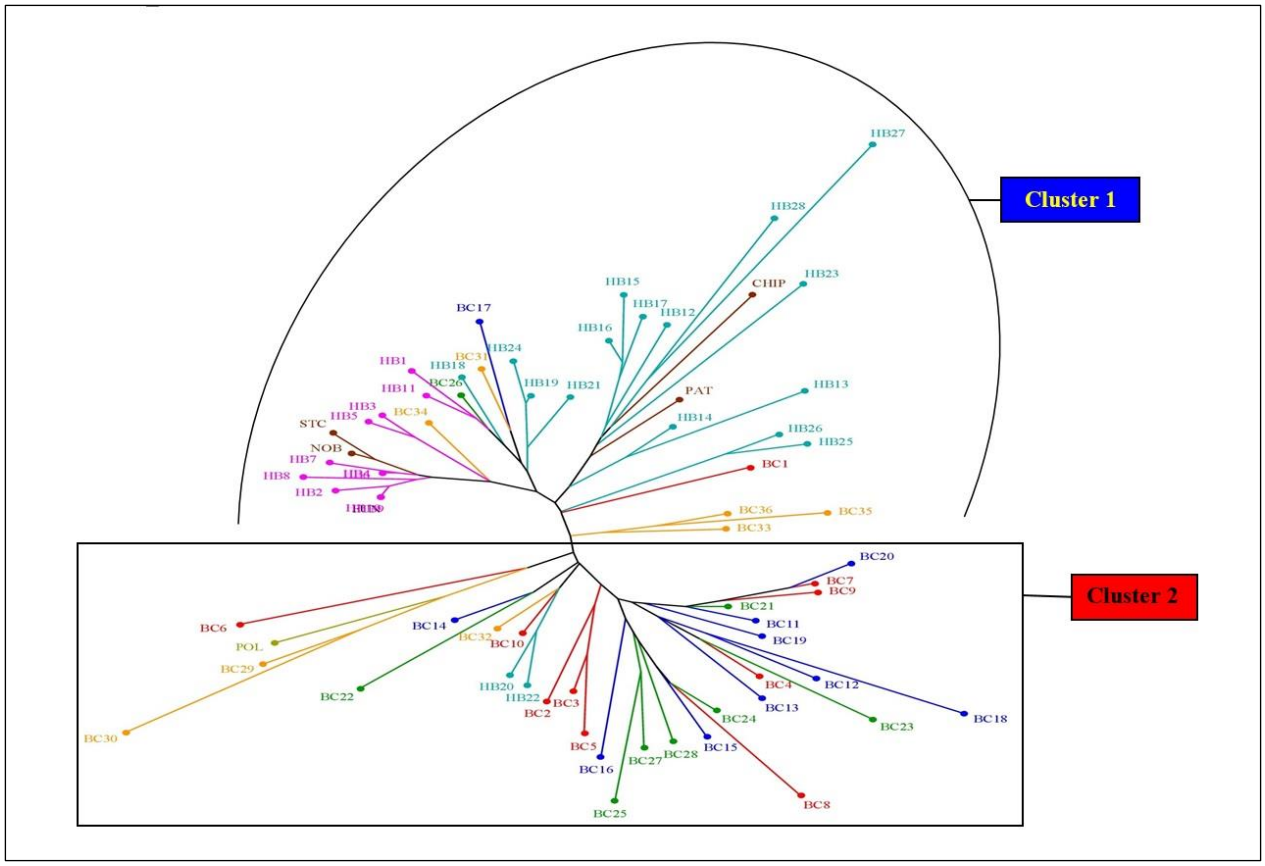


Figure 3.5 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by EST-SSR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label).

As was in STRUCTURE analysis (Figure 3.1b), two main clusters were observed for EST-SSR primer pairs in NJ analysis (Figure 3.5). Cluster 1 comprised of 39 genotypes, including one NL (BC1), one PE (BC17), one QC (BC26), and five NB wild clones (BC31, 33 - 36); all lowbush and half-high cultivars (FUN, PAT, CHIP, STC, NOB) and all hybrids except HB20 and 22. Cluster 2 had 31 genotypes: nine NL (BC2 -10), nine PE (BC11 - 16, 18 - 20), seven QC (BC21 - 25, 27, 28) and three NB wild clones (BC29, 30, 32); highbush cultivar Polaris and two hybrids (HB20, 22). The NJ tree resembled clusters of STRUCTURE analysis (Figure 3.1b), with few exceptions. The lowbush wild clones BC35 and 36, which are part of cluster 1 of the NJ tree (Figure 3.5), are grouped into cluster 2 of STRUCTURE analysis (Figure 3.1b). The highbush cultivar Polaris and

hybrid HB22 in Cluster 2 of the NJ tree (Figure 3.5) are part of Cluster 1 of STRUCTURE analysis (Figure 3.1b).

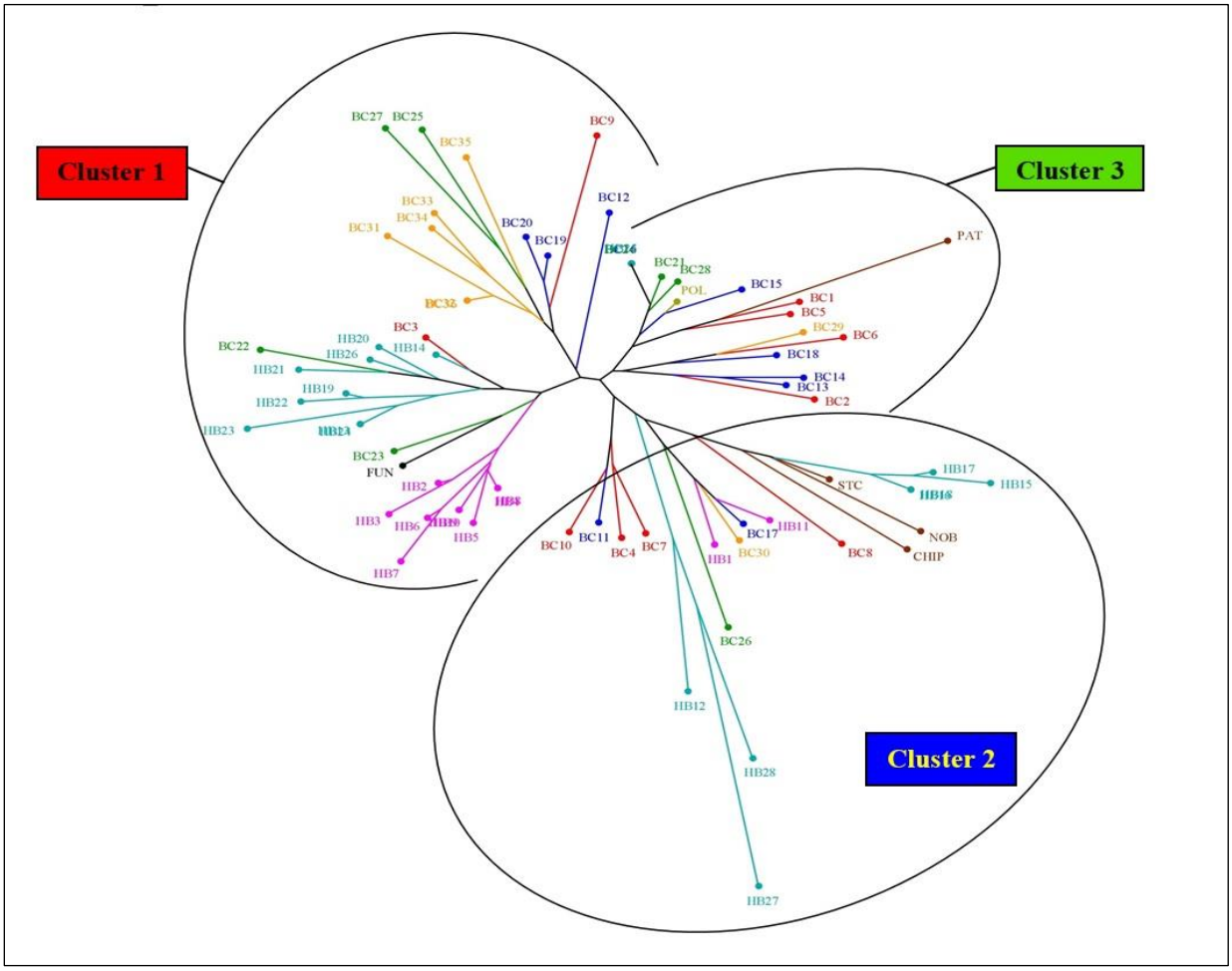


Figure 3.6 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by genomic (G) - SSR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label).

Unlike STRUCTURE analysis for G-SSR (Figure 3.2b), the genotypes were grouped into three clusters in NJ analysis (Figure 3.6). Cluster 1 of NJ tree consisted of two NL (BC3, 9), three PE (BC12, 19, 20), four QC (BC22, 23, 25, 27) and six NB clones (BC31 - 36); lowbush cultivar Fundy and 18 hybrids (HB2-10, 13, 14, 19 - 24, 26). Cluster 2 possessed four NL (BC4, 7, 8,10), two PE (BC11, 17), one QC (BC26) and one NB wild clones (BC30); three half-high cultivars

(CHIP, STC, NOB) and nine hybrids (HB1, 11, 12, 15 - 18, 27, 28). Cluster 3 consisted of four NL (BC1, 2, 5, 6), five PE (BC13 - 16, 18), three QC (BC21, 24, 28), and one NB clones (BC29); highbush and half-high cultivars Polaris and Patriot, and the hybrid HB25. The eight genotypes from cluster 1 of STRUCTURE (Figure 3.2b) analysis (BC5, 15, 16, 21, 24, 28, POL, HB25) are part of cluster 3 of the NJ tree (Figure 3.6). Similarly, genotypes BC9 and HB26 that are part of Cluster 1 of NJ analysis (Figure 3.6) are part of Cluster 2 of STRUCTURE analysis (Figure 3.2b). The genotypes from Cluster 2 of STRUCTURE analysis (Figure 3.2b; BC1, 2, 6, 13, 14, 18, 29, PAT) are grouped into Cluster 3 of the NJ tree (Figure 3.6).

The NJ tree of EST-PCR divided 70 genotypes into three clusters (Figure 3.7). Cluster 1 consisted of three NL (BC1, 6, 9), eight PE (BC11 - 14, 15, 16, 18, 20), five QC (BC21, 23 - 25, 27) and six NB clones (BC30 - 34, 36); highbush cultivar Polaris; and ten hybrids (HB13, 14, 19 - 26). As many as four sub-clusters were identified in this cluster; nine hybrids from the second cross formed a sub-cluster with one NL (BC1) and two PE clones (BC11, 16) cultivar Polaris. Cluster 2 contained seven NL (BC2 – 5, 7, 8, 10), two PE (BC17, 19) and one QC clones (BC26); three half-high cultivars (CHIP, STC, NOB) and seven hybrids (HB1, 11, 15 - 18, 27). Cluster 2 was divided into three sub-clusters, where Northblue grouped with five hybrids in a sub-sub-cluster (Figure 3.7). Cluster 3 consisted of two QC (BC22, 28) and two NB clones (BC29, 35), lowbush cultivar Fundy, half-high cultivar Patriot and 11 hybrids (HB2 – 10, 12, 28). Cluster 3 can be resolved into two main sub-clusters, where Fundy formed a group with nine hybrids (HB2 - 10) and BC35 in a sub-cluster. The other two hybrids (HB12, 28) of this cluster were grouped with cultivar Patriot and clones BC22, 28, and 29 in the second sub-cluster (Figure 3.7).

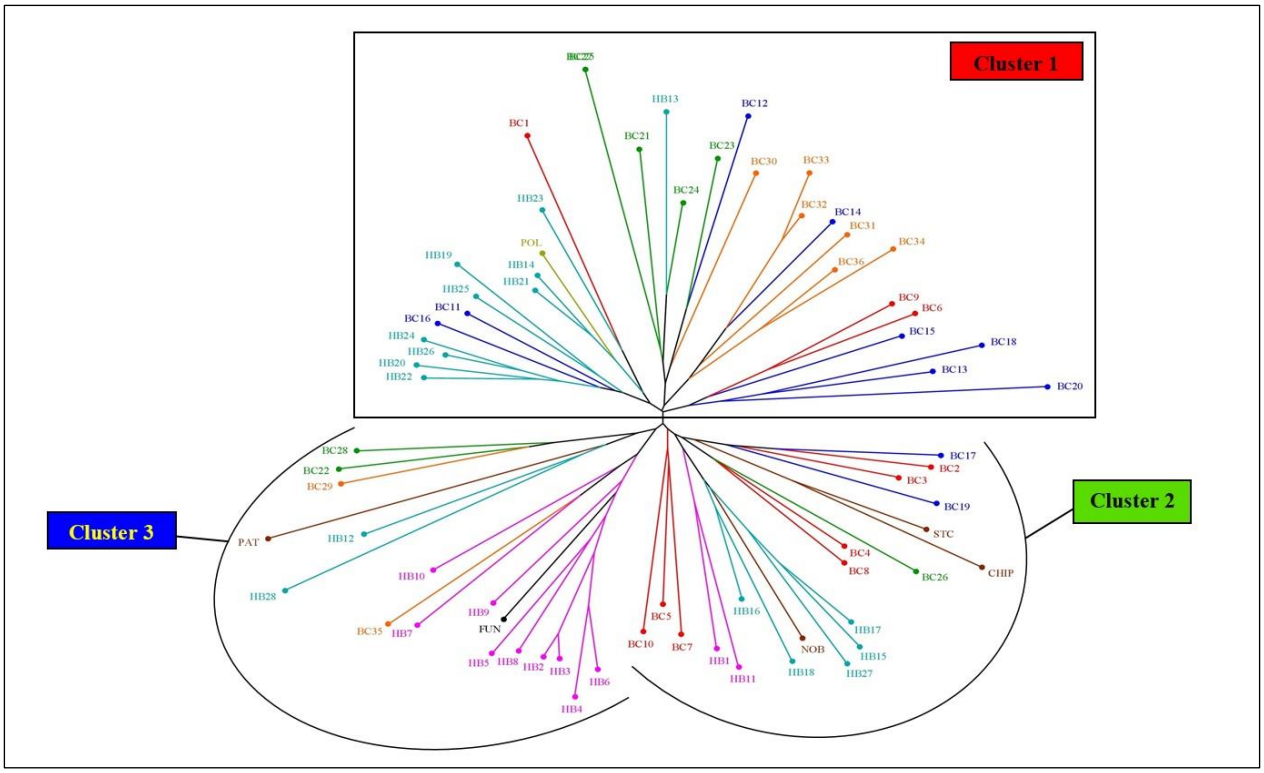


Figure 3.7 Dendrogram produced by using the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by EST-PCR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label).

The NJ tree, generated by using all three types of markers, divided 70 blueberry genotypes into three major clusters (Figure 3.8). Like Cluster 1 of STRUCTURE analysis (Figure 3.4), Cluster 1 of NJ tree also consisted of all ten NL (BC1 - 10), nine PE (BC11 - 16, 18 - 20), seven QC (BC21 – 25, 27, 28) and two NB clones (BC29, 30). Apart from these wild clones, Cluster 1 also contained seven hybrids (HB19 - 22, 24 - 26) but in an isolated sub-cluster (Figure 3.8). Cluster 2 contained PE clone BC17 and QC clone BC26, all four half-high cultivars (PAT, CHIP, STC, NOB) and nine hybrids (HB1, 11, 12, 15 - 18, 27, 28) corroborating the same clustering in STRUCTURE analysis where all these genotypes were grouped in Cluster 2 (Figure 3.4). Cluster 2 can be resolved into three sub-clusters where Northblue, St. Cloud, and Chippewa formed a sub-cluster with four hybrids (HB15-18), and Patriot formed another sub-cluster with three hybrids (HB12, 27, 28). The

third sub-cluster of this cluster consisted of hybrids HB1 and 11 and clones BC17 and 26 (Figure 3.8). Cluster 3 comprised of six NB wild clones (BC31 - 36), cultivars Fundy and Polaris, and 12 hybrids (HB2 - 10, 13, 14, 23). In this cluster, three distinct sub-clusters were noticed: (i) six NB wild clones (BC31 - 36), (ii) Polaris with three hybrids from Cross 2 (HB13, 14, 23), and (iii) Fundy with nine hybrids from the first cross (HB2 – 10; Figure 3.8).

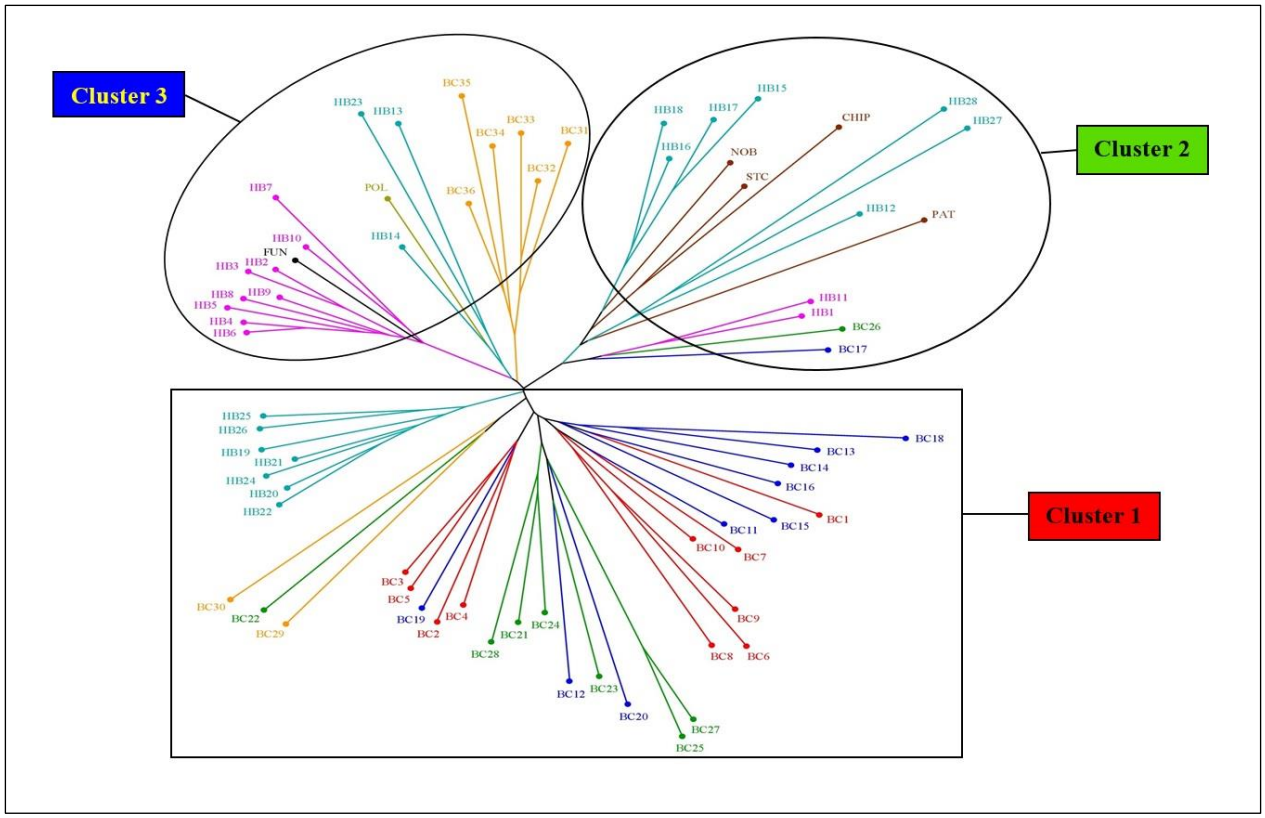


Figure 3.8 Dendrogram produced the unweighted neighbour-joining (NJ) method based on genetic dissimilarity produced by the combination of EST-SSR, genomic (G) - SSR, and EST-PCR markers among blueberry genotypes. The colour of the branches indicates different groups (see Tables 2.1 and 2.2 for genotype label).

3.1.1.2.3 Principal Coordinate Analysis (PCoA)

The PCoA revealed the genetic relationship of 70 genotypes supporting Bayesian inferences from the STRUCTURE analysis and the unweighted neighbour-joining analyses for most of the genotypes. The PCoA for EST-SSR (Figure 3.9) confirmed the STRUCTURE (Figure 3.1b) and

NJ groupings (Figure 3.5) as most of the lowbush wild clones are on the left side of the axis (Cluster 2) except for BC1, BC10, BC17, BC26, BC32, and BC34. All six blueberry cultivars and all hybrids except HB20 and 22 were also placed on the right side of the axis (Cluster 1, Figure 3.9).

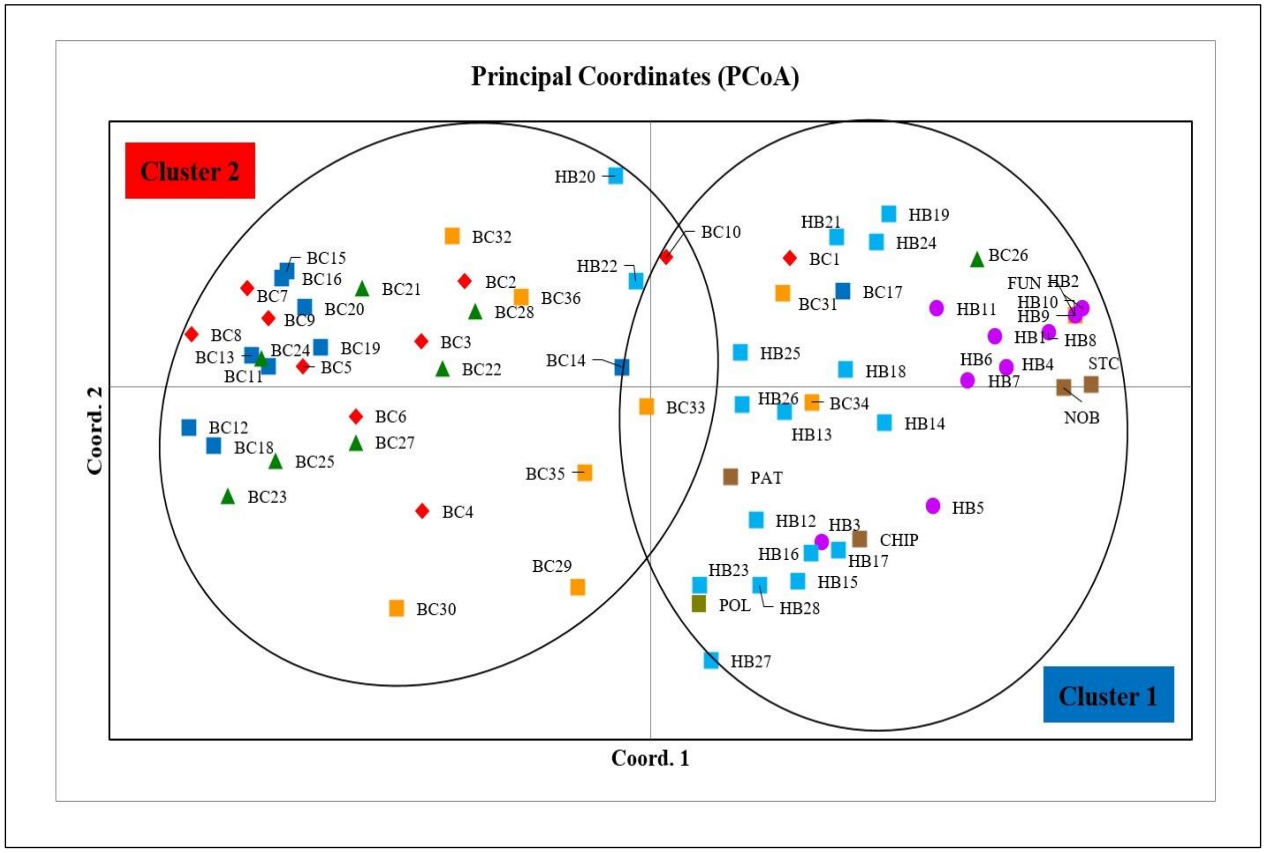


Figure 3.9 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by EST-SSR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).

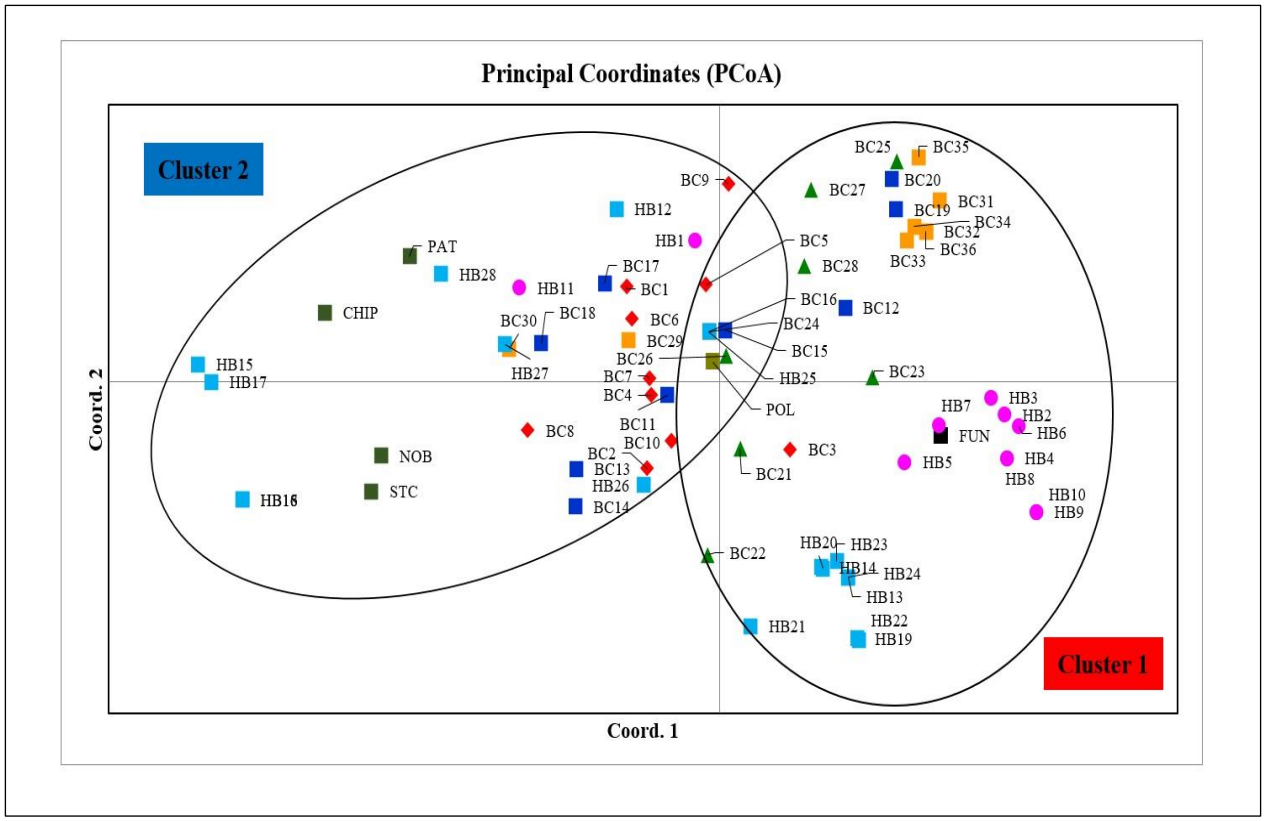


Figure 3.10 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by genomic (G) - SSR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).

The separation of the majority of genotypes in PCoA for G-SSR primers (Figure 3.10) is also aligned with the STRUCTURE (Figure 3.2b) and the NJ groupings (Figure 3.6). Most of these genotypes that are on the right side of the axis of the PCoA graph (Cluster 1, Figure 3.10) are also represented in Cluster 1 of STRUCTURE analysis (Figure 3.2b). The majority of these genotypes are also present in Cluster 1 of the NJ tree (Figure 3.6). Similarly, genotypes present on the left side of the PCoA graph are also assembled in Cluster 2 of STRUCTURE and NJ groupings. The genotypes from Cluster 3 of the NJ tree (Figure 3.6) are seen in the PCoA graph's center close to the central axis (Figure 3.10).

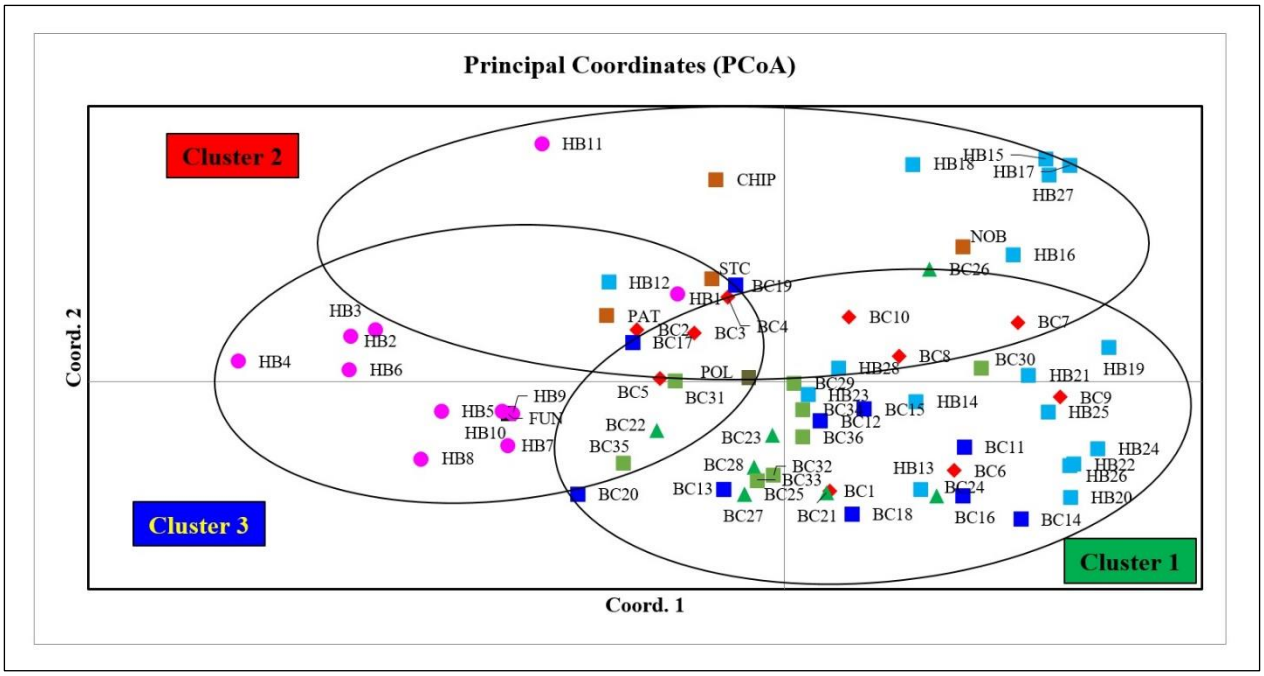


Figure 3.11 2D principal coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by genomic EST-PCR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).

PCoA of EST-PCR primers (Figure 3.11) resembled the results of STRUCTURE (Figure 3.3b) and NJ analyses (Figure 3.7) for the majority of the genotypes. The bottom right quadrant of the PCoA graph contains all genotypes from Cluster 1 of STRUCTURE and NJ analyses. The genotypes of Cluster 2 of STRUCTURE and NJ analyses are found in the top two quadrants of the PCoA graph. Similarly, genotypes of cluster 3 of STRUCTURE and NJ analyses are seen grouped together at the far-left end of the PCoA graph (Figure 3.1).

PCoA graph for combined analysis of all markers (Figure 3.12) also confirmed clustering patterns of STRUCTURE (Figure 3.4b) and NJ analyses (Figure 3.8) for most of the blueberry genotypes where cultivar Fundy and nine hybrids (HB2 – 10) from cluster 3 of STRUCTURE and NJ tree are found in the lower-left quadrant of PCoA graph (Cluster 3; Figure 3.12). The individuals from cluster 2 of STRUCTURE and NJ tree are found in the upper-left quadrant of the PCoA graph.

Similarly, individuals from cluster 1 of STRUCTURE and NJ tree are found on the right side of the central axis of the PCoA graph (Figure 3.12).

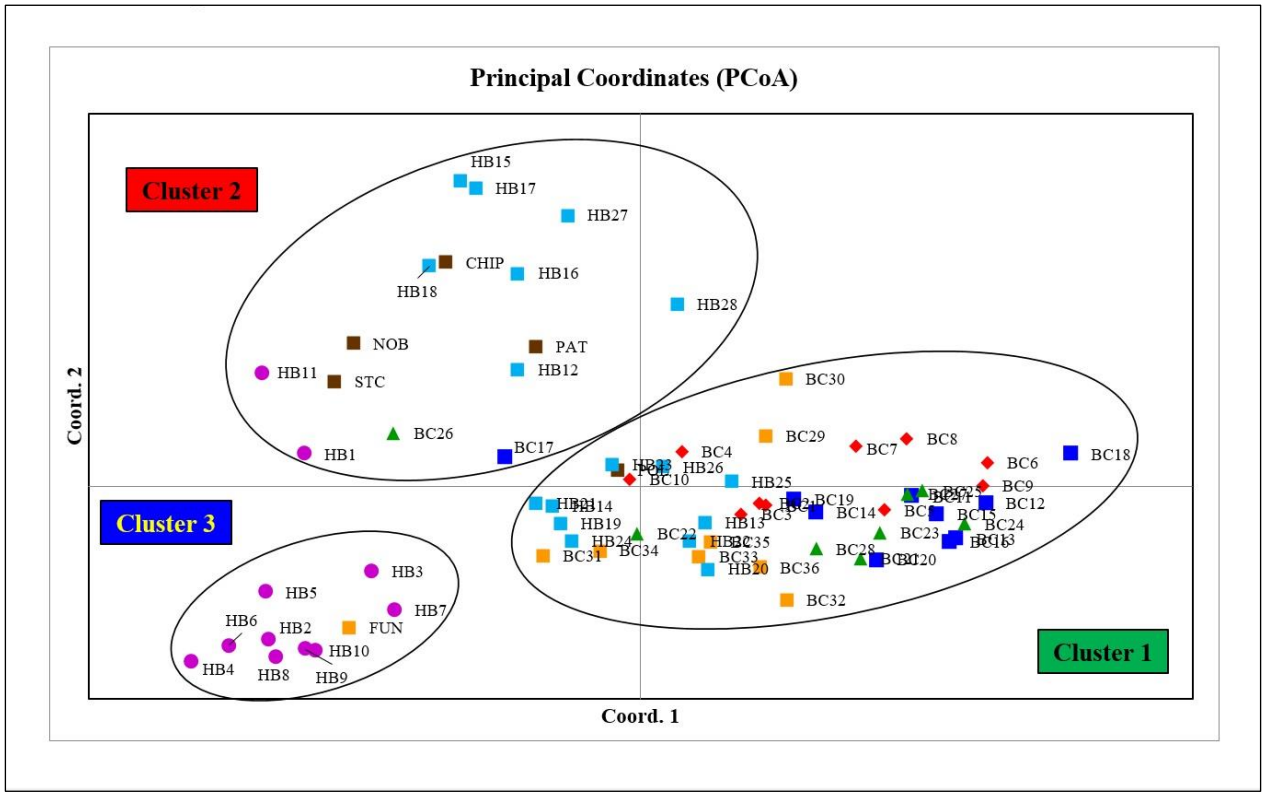


Figure 3.12 2D principle coordinate analysis (PCoA) plot of blueberry genotypes using genetic distance matrix produced by the combination of EST-SSR, genomic (G) -SSR and EST-PCR markers. The colour and shape of the points indicate different groups (see Tables 2.1 and 2.2 for genotype label).

3.1.1.2.4 Analysis of Molecular Variance (AMOVA)

There were significant differences among the seven blueberry groups ($p \leq 0.0001$), demonstrating a high genetic diversity level. This was confirmed with relatively high values for total differentiation (PhiPT: 0.228, 0.182, 0.167, and 0.186 for EST-SSR, G-SSR, EST-PCR, and combined primer pairs, respectively) for all groups showing little similarity among them. In the present study, AMOVA analysis of EST-SSR, G-SSR, EST-PCR, and all primers combined

showed a variance of 23%, 18%, 17%, and 19%, respectively, among groups. These values for variation among genotypes within these groups were 77%, 82%, 83%, and 81%, respectively (Table 3.3).

Table 3.3 Genetic differentiation among blueberry genotypes by AMOVA based on seven groups where four groups comprised of wild clones collected from four Canadian provinces; the fifth group with all six cultivars; sixth group contained 11 hybrids from the first cross (HB1-11); and seventh group contained 17 hybrids from the second cross (HB12-28).

Source of variation	Marker type	Degrees of freedom	Sum of squares	Mean square	Variance components	Percentage of variation
Among groups	EST-SSR	6	67.133	11.189	0.847	23%
	G-SSR	6	50.291	8.382	0.586	18%
	EST-PCR	6	111.626	18.604	1.256	17%
	Combined	6	229.051	38.175	2.689	19%
Within the groups	EST-SSR	63	180.681	2.868	2.868	77%
	G-SSR	63	165.437	2.626	2.626	82%
	EST-PCR	63	394.688	6.265	6.265	83%
	Combined	63	740.806	11.759	11.759	81%
Total	EST-SSR	69	247.814		3.715	100%
	G-SSR	69	215.729		3.212	100%
	EST-PCR	69	503.314		7.521	100%
	Combined	69	969.857		14.448	100%

Stat (PhiPT)	Value	P
EST-SSR	0.228	0.0001
G-SSR	0.182	
EST-PCR	0.167	
Combined	0.186	

3.2 Biochemical Analysis

3.2.1 Total Antioxidant Activity (TAA)

The antioxidant activity was highly diverse ($p < 0.015$) in the present material (Table 3.4 and 3.5). The TAA data were expressed as mg GAE/g fl. The values varied 19.93 fold between genotypes BC22 (5.82 ± 0.03) and HB6 (0.29 ± 0.10). The variation was the highest in Cross 1 (15.34 times) and the lowest in Cross 2 (2.04 times), followed by cultivars (2.09 times) among all groups. The NB wild clones (4.29 ± 1.06) had the highest average TAA value, followed by CV (4.23 ± 1.10) and Cross 2 (4.08 ± 0.89) among all the groups. Cross 1 had the lowest average TAA value (2.55 ± 1.28) among all the groups. The highest TAA value was observed in NL clones in genotype BC6 (5.82 ± 0.03) and the lowest in BC9 (2.45 ± 0.07). The TAA values for PE clones ranged from 1.14 ± 0.08 for BC12 to 5.45 ± 0.09 for BC13. In wild QC clones, the TAA was the highest in BC22 (5.82 ± 0.03) and the lowest for BC28 (1.14 ± 0.05). The value of TAA among NB wild clones ranged from 1.84 ± 0.05 for BC29 to 5.13 ± 0.07 for BC34. The value of TAA among NB wild clones ranged from 1.84 ± 0.05 for BC29 to 5.13 ± 0.07 for BC34. The average TAA value for eight NB wild clones was the highest (4.29 ± 1.06) among all groups (Table 3.4). For cultivars, the highest TAA value was observed in the highbush cultivar, POL (5.19 ± 0.03), and the lowest in half-high cultivar, STC (2.48 ± 0.08). The value for TAA in Cross 1 ranged from 0.29 ± 0.10 for HB6 to 4.48 ± 0.04 for HB11 and in Cross 2 from 2.52 ± 0.07 for HB16 to 5.13 ± 0.05 for HB21. Two NL (BC6: 5.82 ± 0.03 mg GAE/g fl, BC13: 5.45 ± 0.09 mg GAE/g fl) and one QC (BC22: 5.82 ± 0.03 mg GAE/g fl) clones possessed more TAA than all cultivars (Table 3.5).

Table 3.4 Total antioxidant activity (TAA), phenolic (TPC), and flavonoid contents (TFC) of groups of wild clones collected from four Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), cultivars (CV): Fundy, Polaris, Patriot, Chippewa, St. Cloud, Northblue, hybrids: Cross 1 and Cross 2. GAE = gallic acid equivalents, CE = catechin equivalent, fl = fresh leaf.

Clones/Cultivars/Hybrids (No.)		TAA (mg GAE/g fl)	TPC (mg GAE /g of fl)	TFC (mg CE/g fl)
NL (10)	Mean ± SD	3.94 ± 1.02 ab	0.16 ± 0.05 b	5.26 ± 2.49 b
	Min – Max (Fold)	2.45-5.82 (2.37)	0.09-0.24 (2.60)	2.00-9.99 (4.99)
PE (10)	Mean ± SD	3.34 ± 1.36 ab	0.12 ± 0.02 ab	2.14 ± 0.79 a
	Min – Max (Fold)	1.14-5.45 (4.77)	0.09-0.15 (1.58)	1.22-3.77 (3.09)
QC (8)	Mean ± SD	3.35 ± 1.64 ab	0.09 ± 0.04 a	1.88 ± 1.59 a
	Min – Max (Fold)	1.14-5.82 (5.09)	0.06-0.18 (2.79)	0.81-5.56 (6.85)
NB (8)	Mean ± SD	4.29 ± 1.06 a	0.11 ± 0.02 ab	1.74 ± 0.46 a
	Min – Max (Fold)	1.84-5.13 (2.79)	0.08-0.15 (1.95)	1.19-2.70 (2.26)
CV (6)	Mean ± SD	4.23 ± 1.10 ab	0.13 ± 0.03 ab	1.84 ± 0.84 a
	Min – Max (Fold)	2.48-5.19 (2.09)	0.09-0.16 (1.78)	0.64-2.83 (4.43)
Cross 1 (11)	Mean ± SD	2.55 ± 1.28 b	0.11 ± 0.03 ab	1.39 ± 0.18 a
	Min – Max (Fold)	0.29-4.48 (15.34)	0.07-0.16 (2.17)	1.04-1.71 (1.64)
Cross 2 (17)	Mean ± SD	4.08 ± 0.89 ab	0.10 ± 0.01 a	1.41 ± 0.29 a
	Min – Max (Fold)	2.52-5.13 (2.04)	0.08-0.12 (1.51)	0.87-1.95 (2.24)
All (70)	Mean ± SD	3.67± 0.25	0.11 ± 0.03	2.19 ± 1.69
	Min – Max (Fold)	0.29-5.82 (19.93)	0.06-0.24 (3.83)	0.64-9.99 (15.63)

Values are means ± SD values of at least three replicates.

Table 3.5 Total antioxidant activity (TAA), phenolic (TPC), and flavonoid contents (TFC) of individual wild clones collected from four Canadian provinces: Newfoundland and Labrador (NL), Prince Edward Island (PE), Quebec (QC), and New Brunswick (NB), cultivars (CV): Fundy, Polaris, Patriot, Chippewa, St. Cloud, Northblue, hybrids: Cross 1 and Cross 2. GAE = gallic acid equivalents, CE = catechin equivalent, fl = fresh leaf.

Clones/Cultivars/Hybrids	TAA (mg GAE/g fl)	TPC (mg GAE /g of fl)	TFC (mg CE/g fl)
BC1	4.45 ± 0.07	0.18 ± 0.01	6.55 ± 0.39
BC2	4.58 ± 0.07	0.24 ± 0.01	9.99 ± 0.11
BC3	4.03 ± 0.09	0.20 ± 0.02	6.99 ± 0.12
BC4	2.75 ± 0.08	0.14 ± 0.01	3.93 ± 0.13
BC5	4.57 ± 0.05	0.20 ± 0.02	6.03 ± 0.02
BC6	5.82 ± 0.03	0.20 ± 0.01	7.07 ± 0.05
BC7	4.20 ± 0.07	0.11 ± 0.00	3.49 ± 0.03
BC8	3.53 ± 0.12	0.09 ± 0.01	2.00 ± 0.01
BC9	2.45 ± 0.07	0.11 ± 0.01	2.41 ± 0.07
BC10	3.04 ± 0.08	0.13 ± 0.00	4.18 ± 0.01
BC11	4.88 ± 0.05	0.15 ± 0.01	2.92 ± 0.13

Table 3.5 cont'd

BC12	1.14 ± 0.08	0.11 ± 0.01	1.56 ± 0.02
BC13	5.45 ± 0.09	0.14 ± 0.00	2.65 ± 0.12
BC14	2.77 ± 0.05	0.10 ± 0.01	1.37 ± 0.04
BC15	1.79 ± 0.10	0.10 ± 0.01	1.22 ± 0.02
BC16	2.85 ± 0.07	0.11 ± 0.00	1.66 ± 0.09
BC17	4.14 ± 0.08	0.10 ± 0.01	1.89 ± 0.12
BC18	4.40 ± 0.05	0.15 ± 0.00	3.77 ± 0.09
BC19	2.86 ± 0.03	0.10 ± 0.00	2.35 ± 0.17
BC20	3.14 ± 0.08	0.09 ± 0.00	1.96 ± 0.11
BC21	2.56 ± 0.07	0.06 ± 0.01	0.81 ± 0.02
BC22	5.82 ± 0.03	0.18 ± 0.01	5.56 ± 0.04
BC23	3.25 ± 0.08	0.07 ± 0.01	1.35 ± 0.02
BC24	4.71 ± 0.07	0.10 ± 0.01	2.15 ± 0.08
BC25	4.62 ± 0.03	0.09 ± 0.01	2.26 ± 0.02
BC26	3.26 ± 0.07	0.09 ± 0.00	1.11 ± 0.01

Table 3.5 cont'd

BC27	1.41 ± 0.05	0.08 ± 0.00	0.89 ± 0.07
BC28	1.14 ± 0.05	0.08 ± 0.00	0.93 ± 0.03
BC29	1.84 ± 0.05	0.08 ± 0.00	1.19 ± 0.02
BC30	3.85 ± 0.05	0.09 ± 0.00	1.43 ± 0.04
BC31	4.82 ± 0.07	0.11 ± 0.01	1.82 ± 0.04
BC32	4.85 ± 0.05	0.11 ± 0.01	1.92 ± 0.04
BC33	4.71 ± 0.07	0.11 ± 0.00	1.59 ± 0.01
BC34	5.13 ± 0.07	0.10 ± 0.00	2.70 ± 0.05
BC35	4.60 ± 0.05	0.10 ± 0.01	1.85 ± 0.03
BC36	4.49 ± 0.03	0.15 ± 0.01	1.45 ± 0.02
FUN	5.12 ± 0.07	0.15 ± 0.01	2.83 ± 0.02
PAT	4.55 ± 0.05	0.13 ± 0.01	1.76 ± 0.04
POL	5.19 ± 0.03	0.16 ± 0.01	1.92 ± 0.04
CHIP	4.73 ± 0.07	0.12 ± 0.02	2.69 ± 0.02
STC	2.48 ± 0.08	0.09 ± 0.00	1.20 ± 0.02

Table 3.5 cont'd

NOB	3.28 ± 0.05	0.10 ± 0.00	0.64 ± 0.02
HB1	1.95 ± 0.09	0.16 ± 0.01	1.04 ± 0.03
HB2	2.75 ± 0.08	0.09 ± 0.00	1.57 ± 0.02
HB3	4.07 ± 0.10	0.07 ± 0.01	1.71 ± 0.02
HB4	3.68 ± 0.05	0.11 ± 0.00	1.40 ± 0.02
HB5	2.94 ± 0.17	0.10 ± 0.01	1.42 ± 0.02
HB6	0.29 ± 0.10	0.13 ± 0.01	1.22 ± 0.02
HB7	2.41 ± 0.10	0.09 ± 0.00	1.23 ± 0.02
HB8	2.00 ± 0.33	0.14 ± 0.01	1.48 ± 0.06
HB9	2.74 ± 0.08	0.10 ± 0.00	1.42 ± 0.02
HB10	0.79 ± 0.10	0.10 ± 0.00	1.32 ± 0.02
HB11	4.48 ± 0.04	0.16 ± 0.01	1.49 ± 0.02
HB12	4.82 ± 0.09	0.09 ± 0.01	1.95 ± 0.03
HB13	3.13 ± 0.10	0.10 ± 0.02	1.27 ± 0.01
HB14	5.11 ± 0.05	0.12 ± 0.00	1.67 ± 0.02

Table 3.5 cont'd

HB15	4.47 ± 0.07	0.09 ± 0.00	1.40 ± 0.02
HB16	2.52 ± 0.07	0.09 ± 0.00	0.91 ± 0.03
HB17	4.79 ± 0.07	0.10 ± 0.00	1.30 ± 0.04
HB18	4.09 ± 0.03	0.10 ± 0.00	1.45 ± 0.02
HB19	2.74 ± 0.05	0.08 ± 0.00	0.87 ± 0.04
HB20	4.69 ± 0.09	0.10 ± 0.00	1.37 ± 0.02
HB21	5.13 ± 0.05	0.10 ± 0.00	1.40 ± 0.02
HB22	4.62 ± 0.03	0.10 ± 0.00	1.27 ± 0.01
HB23	4.01 ± 0.04	0.10 ± 0.00	1.82 ± 0.02
HB24	3.25 ± 0.08	0.09 ± 0.00	1.19 ± 0.01
HB25	2.76 ± 0.10	0.09 ± 0.00	1.22 ± 0.01
HB26	3.66 ± 0.07	0.10 ± 0.00	1.52 ± 0.03
HB27	4.79 ± 0.09	0.09 ± 0.00	1.56 ± 0.02
HB28	4.72 ± 0.06	0.10 ± 0.00	1.74 ± 0.04

Values are means ± SD values of at least three replicates. There was no significant difference among individuals in TAA, TFC and TPC according to the Kruskal-Wallis test.

3.2.3 Total Phenolic Content (TPC)

The TPC estimation using FC reagent is a well-accepted method. The relevant results for TPC are presented in Table 3.4 and 3.5, which were highly diverse (p <0.0003). The variation for TPC values ranged between 3.83 times between genotypes BC2 (0.24 ± 0.01) and BC21 (0.06 ± 0.01) among all genotypes. The ten clones collected from NL had the highest average TPC value (0.16 ± 0.05) followed by CV (0.13 ± 0.03) and PE wild clones (0.12 ± 0.02) among all groups. In NL clones, the TPC values ranged from 0.09 ± 0.01 for BC8 to 0.24 ± 0.01 for BC2. Among PE clones, the lowest TPC was observed in BC20 (0.09 ± 0.01) and the highest in BC11, 18 (0.15 ± 0.01 and 0.15 ± 0.00, respectively). The TPC for eight QC wild clones varied from 0.06 ± 0.01 (BC21) to 0.18 ± 0.01 (BC22). In NB wild clones, the highest TPC was observed for BC36 (0.15 ± 0.01) and lowest for BC29 (0.08 ± 0.00). The value for TPC among cultivars was the highest for highbush cultivar, POL (0.15 ± 0.01) and the lowest for half-high cultivar, STC (0.09 ± 0.00). The TPC value for Cross 1 and Cross 2 ranged from 0.07 ± 0.01 (HB3) to 0.16 ± 0.01 (HB1, 11) and from 0.08 ± 0.00 (HB19) to 0.12 ± 0.00 (HB14), respectively. Among all genotypes, four NL clones (BC2, 0.24 ± 0.01; BC3, 0.20 ± 0.02, BC5, 0.20 ± 0.02; BC6, 0.20 ± 0.01) and one QC clone (BC22, 0.18 ± 0.01) possessed higher TPC than the cultivars.

3.2.3 Total Flavonoid Content (TFC)

The results for TFC are displayed as mg CE/g fl (Table 3.4 and 3.5). The wild clones, cultivars, and hybrids showed wide variation (15.63 times) among themselves for TFC (p <0.0003) from 0.64 ± 0.02 to 9.99 ± 0.11 for the cultivar NOB and NL wild clone BC2, respectively. The wild clones from NL (5.26 ± 2.49), PE (2.14 ± 0.79), and QC (1.88 ± 1.59) possessed the highest average TFC values among all groups. The average values of TFC were the lowest in Cross 2 (1.41 ± 0.29)

and Cross 1 (1.39 ± 0.18). For NL clones, the TFC value was the highest in BC2 (9.99 ± 0.11) and the lowest in BC8 (2.00 ± 0.01). The values of TFC for PE clones ranged from 3.77 ± 0.09 for BC18 to 1.22 ± 0.02 for BC15. The highest and the lowest values of TFC for QC clones were found in BC22 (5.56 ± 0.04) and BC21 (0.81 ± 0.02). The TFC for eight NB wild clones ranged from 1.19 ± 0.02 (BC29) to 2.70 ± 0.05 (BC34). The value for TFC among cultivars was the highest for lowbush cultivar, FUN (2.83 ± 0.02), and the lowest for half-high cultivar, NOB (0.64 ± 0.02). The TFC value for Cross 1 and Cross 2 ranged from 1.04 ± 0.03 (HB1) to 1.71 ± 0.02 (HB3) and from 0.87 ± 0.04 (HB19) to 1.95 ± 0.03 (HB12), respectively. While eight NL clones (BC1 – 7, 10) were found superior with higher TPC ranging from 3.49 ± 0.03 mg CE/g fl to 9.99 ± 0.11 mg CE/g fl than those of the cultivars, there was one PE clone (BC18, 3.77 ± 0.09 mg CE/g fl) and one QC clone (BC22, 5.56 ± 0.04) that had higher TFC than the cultivars (Table 3.5).

3.2.4 Relationship among Antioxidant Properties

To evaluate relationship among TAA, TPC and TFC, linear regression was performed. There was a significant relationship observed between TAA and TPC (r^2 = 0.124, Figure 3.13a); TAA and TFC (r^2 = 0.149, Figure 3.13b); and TFC and TPC (r^2 = 0.682, Figure 3.13c). The Pearson correlation coefficient between TFC and TPC (r = 0.826) was significantly higher followed by TAA and TFC (r = 0.387); and TAA and TPC (r = 0.352) (Table 3.6).

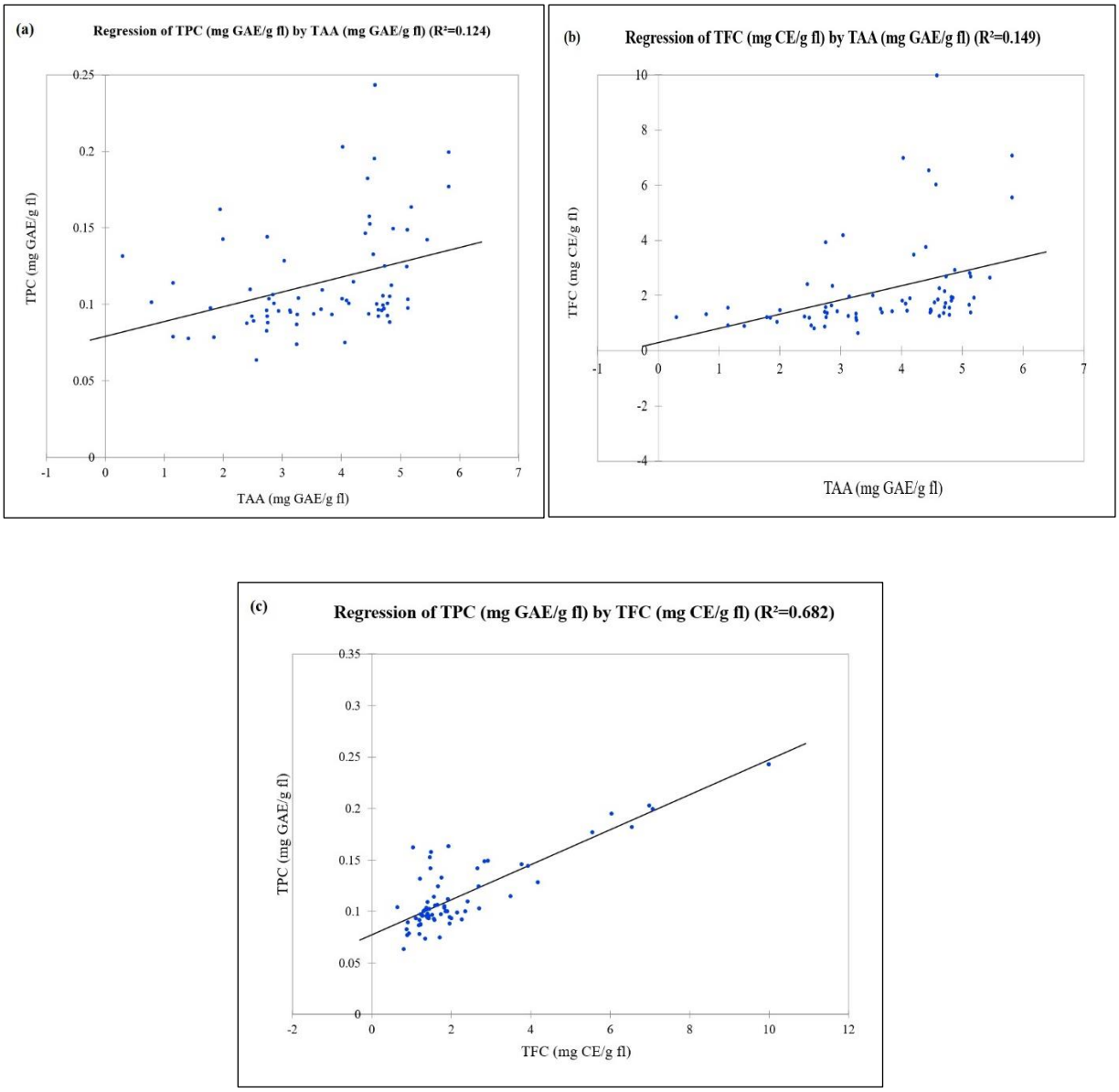


Figure 3.13 Linear regression between antioxidant properties in wild clones, cultivars, and hybrids. (a) Total antioxidant activity by DPPH (mg GAE/g fl) and total phenolic content (mg GAE/g fl). (b) Total antioxidant activity by DPPH (mg GAE/g fl) and total flavonoid content (mg CE/g fl). (c) Total phenolic content (mg GAE/g fl) and total flavonoid content (mg CE/g fl).

Table 3.6 Pearson correlation coefficients of antioxidant properties. The respective p-values are shown in brackets. The significance level alpha is 0.05.

Variables	TAA (mg GAE/g fl)	TFC (mg CE/g fl)	TPC (mg GAE/g fl)
TAA (mg GAE/g fl)	1 (0)	0.387 (0.001)	0.352 (0.003)
TFC (mg CE/g fl)	0.387 (0.001)	1 (0)	0.826 (<0.0001)
TPC (mg GAE/g fl)	0.352 (0.003)	0.826 (<0.0001)	1 (0)

3.2.5 Cluster Analysis of Antioxidant Properties

3.2.5.1 Agglomerative Hierarchical Clustering (AHC)

A UPGMA dendrogram was generated based on the Euclidean dissimilarity distance matrix of antioxidant activity, phenolic, and flavonoid content data. The dendrogram (Figure 3.14) shows the presence of two clades (I & II) sub-divided into seven major groups (Clade I: G1, 2; Clade II: G3-7), which are neither grouped according to genotype nor geographic distribution. Group 1 consists of four NL (BC1, 3, 5, 6) and one QC (BC22) wild clones. Group 2 only has one NL wild clone (BC2). NL wild clones BC4 and 10 make into group 3. Group 4 comprised of 19 genotypes including one NL (BC7), two QC (BC24, 25), five NB (BC31-35) wild clones; two half-high cultivars (PAT, CHIP); and nine individuals from Cross 2 (HB12, 14, 15, 17, 20-22, 27, 28). The largest group 5 consisted two NL (BC8, 9), six PE (BC14-17, 19, 20), five QC (BC21, 23, 26-28), two NB (BC 29, 30) wild clones; two half-high cultivars (STC, NOB); 14 hybrids with six belonging to cross1 (HB2-5, 7, 9) and eight to cross2 (HB13, 16, 18, 19, 23, 24-26). Group 6 comprised of three PE (BC11, 13, 18) and one each from NB (BC36), lowbush cultivar (FUN), and

cross1 (HB11) individuals. Group 7 predominately consists of individuals from cross1 (HB1, 6, 8, 10) and one PE wild clone (BC12).

3.2.5.2 Principal Component Analysis (PCA)

The PCA distribution could differentiate genotypes based on biochemical characteristics, but no clear cluster was detected among 70 genotypes (Figure 3.15). The biplot axes explain 94.23% of the information, out of which y-axes (F1) and x-axes account for 69.30% and 24.93%, respectively.

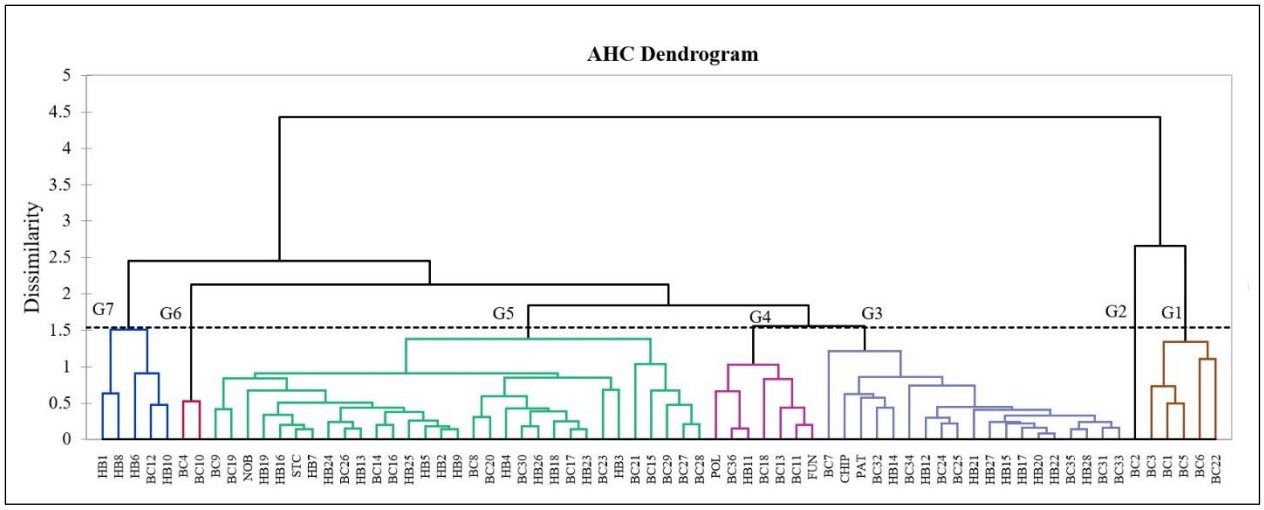


Figure 3.14 AHC dendrogram by UPGMA method based on Euclidean dissimilarity distance matrix of antioxidant activity, phenolic, and flavonoid content data of 70 genotypes.

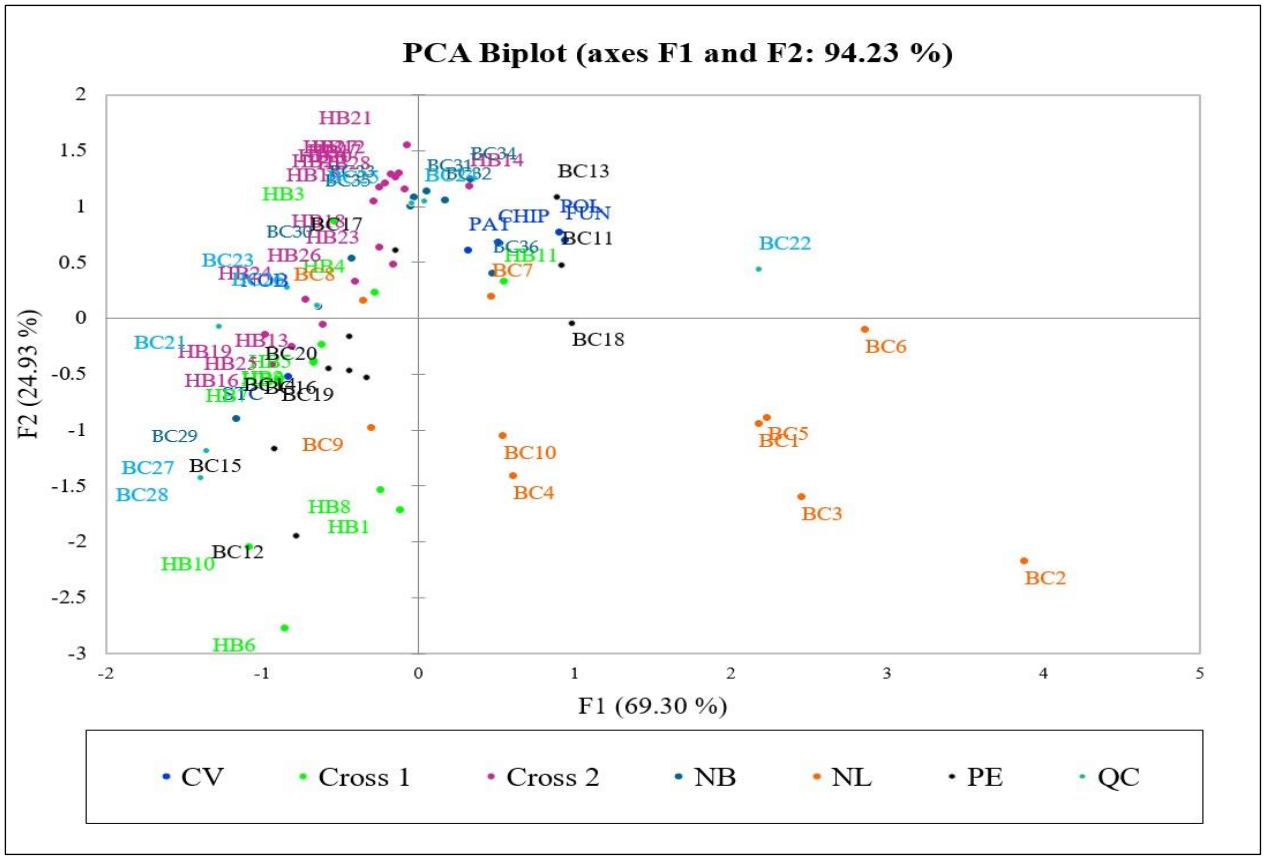
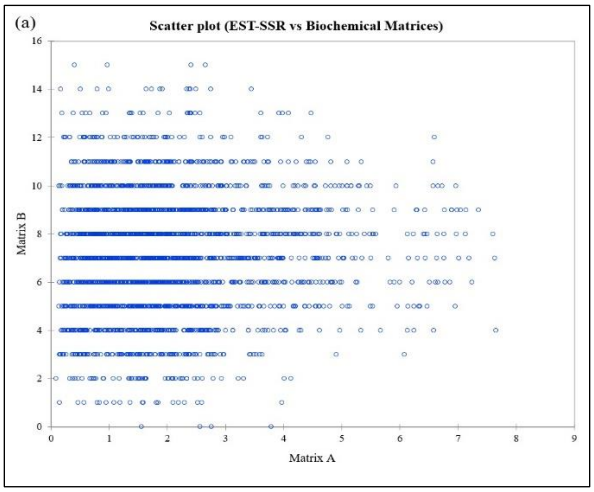


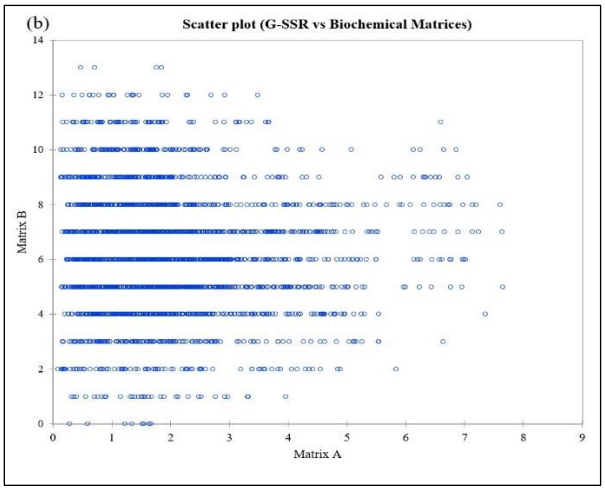
Figure 3.15 The PCA bi-plot based on biochemical characteristics of 70 genotypes and correlation among quantitative variables.

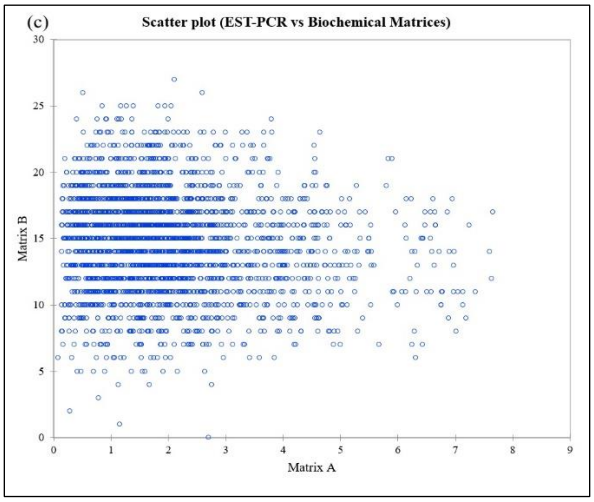
3.3 Relationship between Biochemical and Molecular Analysis

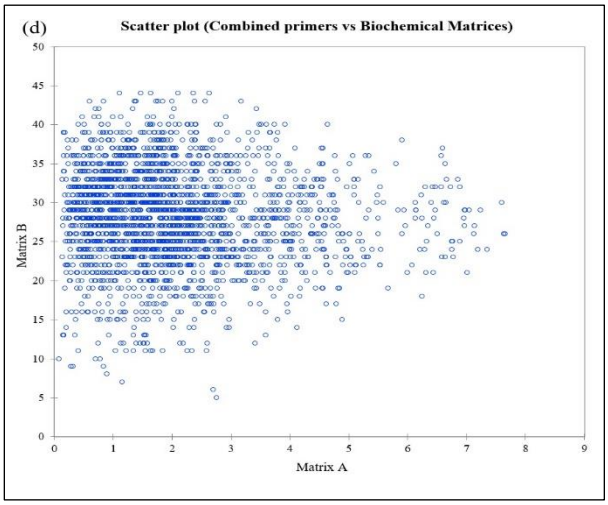
3.3.1 Mantel Test

The Mantel test was used to check the correlation between biochemical and genetic distances of EST-SSR, G-SSR, EST-PCR, and all primers combined. Figure 3.16 shows there was no significant correlation between biochemical and genetic distances as indicated by scatter plots (a, b, c, d) and poor correlation coefficient values (e, f, g, h), r(AB), of 0.046, -0.042, -0.018, and -0.064 for matrices of biochemical and EST-SSR, G-SSR, EST-PCR, and combined primers, respectively.


(a)
Scatter plot (EST-SSR vs Biochemical Matrices)
Matrix B
Matrix A


(b)
Scatter plot (G-SSR vs Biochemical Matrices)
Matrix B
Matrix A


(c)
Scatter plot (EST-PCR vs Biochemical Matrices)
Matrix B
Matrix A


(d)
Scatter plot (Combined primers vs Biochemical Matrices)
Matrix B
Matrix A

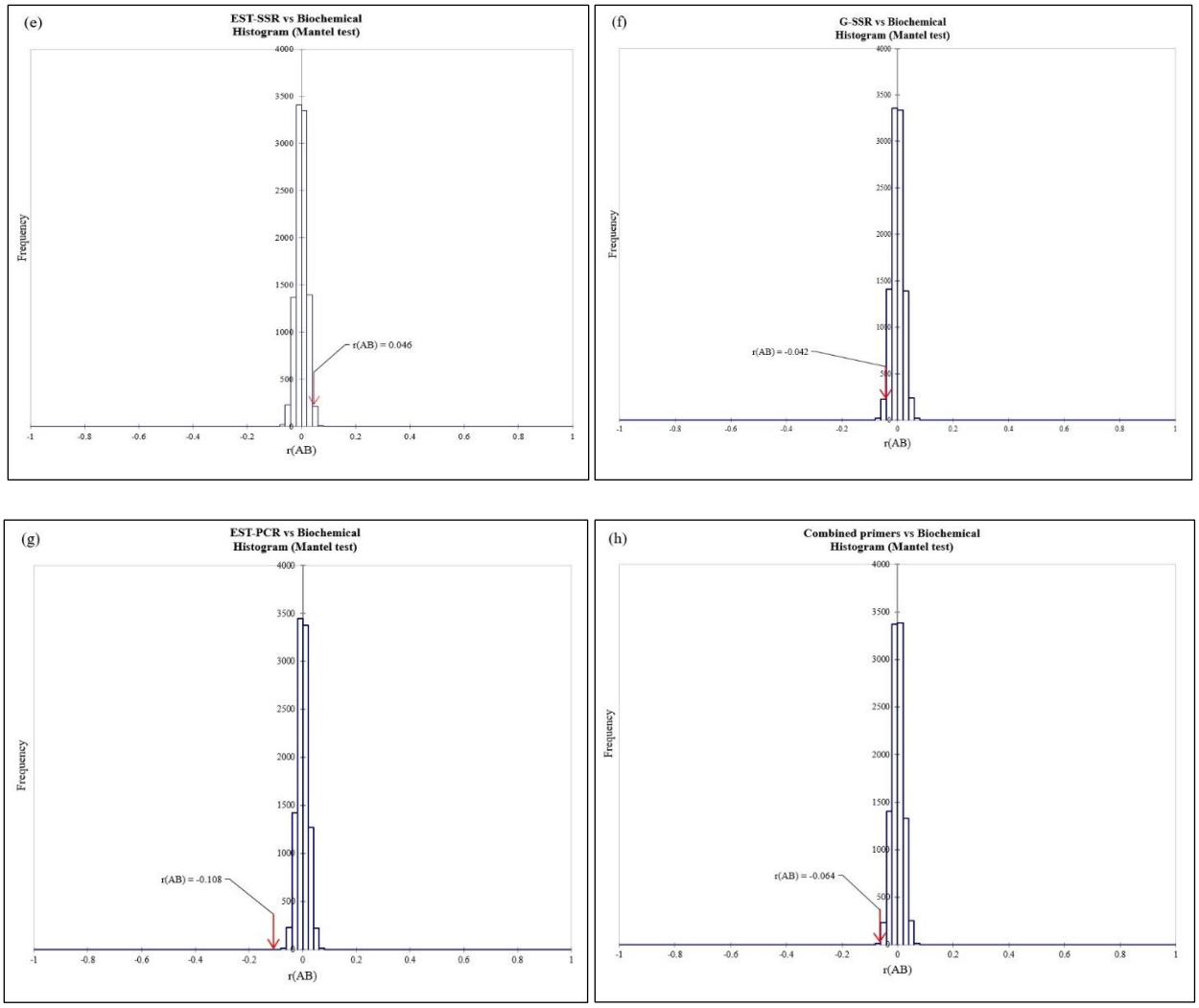


Figure 3.16 The scatter plot (a, b, c, d) and histogram (e, f, g, h) of Mantel test between biochemical and genetic matrices (EST-SSR, G-SSR, EST-PCR and combined primers, respectively).

3.3.2 Stepwise Multiple Regression Analysis (SMRA)

Results of SMRA between polymorphic EST-SSR, G-SSR, and EST-PCR markers, with 24, 25, and 58 alleles, respectively, and the biochemical traits in 70 genotypes are represented in Table 7. Alleles showing significant association based on multiple correlation co-efficient (R^2) were considered. SMRA identified 17 alleles associated with various biochemical components. According to SMRA, calculations of the fraction of variation for each primer pair (R^2; Table 3.7)

revealed that a combination of four alleles accounted for 33% of TAA among 70 blueberry genotypes. Out of these four, only VCC_I2_1 showed a positive, statistically significant ($t = 2.343$, $p < 0.022$) correlation with TAA as explained by a beta coefficient value of 0.242. The other three alleles, VCC_S10_1, NA800_1, and CA791_4, had a statistically significant but negative correlation to TAA (Table 3.7). TPC variation (69%) was explained by 11 alleles, out of which seven alleles, CA1423_6, VCC_K4_6, CA54_7, CA23_1, VCC_K4_1, CA1029_1, and VCC_K4_9, were positively correlated, while other four alleles, CA54_3, VCC_I8_1, CA791_7, and CA21_7 were negatively correlated with TPC (Table 3.7). VCC_K4_6 was the highest contributor with beta coefficient value of 0.585 and positively significant ($t = 5.554$, $p < 0.000$). Six alleles, VCC_K4_6, CA54_7, CA23_1, CA1029_1, VCC_K4_1, and CA1423_1, out of seven alleles were positively correlated with TFC. Allele CA791_1 was negatively correlated with TFC. VCC_K4_6 ($t = 8.369$, $p < 0.000$) was the key contributor with a proportion of variation of 21% with a beta coefficient value of 0.818 (Table 3.7). Five alleles, VCC_K4_6, CA54_7, CA23_1, VCC_K4_1, and CA1029_1, were also associated and positively correlated with TPC as well as TFC (Table 3.7).

Table 3.7. Association of EST-SSR, G-SSR and EST-PCR markers with biochemical components in blueberry leaf extract as revealed by SMRA.

Traits	Alleles	Type of Marker	R	R^2	R^2 change	F change	Standardized beta coefficients	t value	*p* value
TAA	VCC_S10_1	G-SSR	0.378	0.143	0.143	11.363	-0.493	-4.659	0.000
	+ NA800_1	EST-SSR	0.455	0.207	0.064	5.428	-0.288	-2.757	0.008
	+ CA791_4	EST-PCR	0.525	0.276	0.068	6.225	-0.283	-2.774	0.007
	+ VCC_I2_1	G-SSR	0.576	0.332	0.056	5.490	0.242	2.343	0.022
TPC	CA1423_6	EST-PCR	0.335	0.112	0.112	8.585	0.281	3.563	0.001
	+ VCC_K4_6	G-SSR	0.472	0.223	0.111	9.576	0.585	5.554	0.000
	+ CA54_3	EST-PCR	0.549	0.302	0.078	7.405	-0.189	-2.372	0.021
	+ CA54_7	EST-PCR	0.603	0.363	0.062	6.309	0.310	3.878	0.000
	+ CA23_1	EST-SSR	0.652	0.425	0.061	6.826	0.368	3.519	0.001

Table 3.7 cont'd

	+ VCC_K4_1	G-SSR	0.697	0.485	0.061	7.444	0.253	3.370	0.001
	+ VCC_I8_1	G-SSR	0.727	0.529	0.043	5.670	-0.279	-3.535	0.001
	+ CA791_7	EST-PCR	0.757	0.573	0.045	6.380	-0.244	-3.253	0.002
	+ CA21_7	EST-PCR	0.785	0.617	0.043	6.805	-0.225	-2.829	0.006
	+ CA1029_1	EST-PCR	0.816	0.665	0.048	8.519	0.251	3.266	0.002
	+ VCC_K4_9	G-SSR	0.830	0.689	0.024	4.414	0.172	2.101	0.040
TFC	VCC_K4_6	G-SSR	0.461	0.212	0.212	18.340	0.818	8.369	0.000
	+ CA54_7	EST-PCR	0.583	0.340	0.127	12.893	0.378	5.434	0.000
	+ CA23_1	EST-SSR	0.667	0.445	0.105	12.476	0.455	4.662	0.000
	+ CA1029_1	EST-PCR	0.738	0.544	0.100	14.211	0.453	5.831	0.000

Table 3.7 cont'd

+ CA1029_1	EST-PCR	0.738	0.544	0.100	14.211	0.453	5.831	0.000
+ VCC_K4_1	G-SSR	0.797	0.635	0.091	15.917	0.305	4.385	0.000
+ CA791_1	EST-PCR	0.822	0.676	0.041	7.917	-0.269	-3.415	0.001
+ CA1423_1	EST-PCR	0.837	0.700	0.025	5.116	0.165	2.262	0.027

+ Denotes the inclusion of alleles (s) in the preceding step (s) in the SMRA.

CHAPTER 4

Discussion

The study presented here provides insight into genetic diversity with respect to genetic relationship and structure and biochemical properties in two groups of selected hybrids between lowbush and half-high blueberries, wild blueberry clones, half-high, highbush, and lowbush blueberry cultivars. The antioxidant properties of blueberries are well known for their medicinal value in negating the harmful effects of free radicals (Osawa, 1994). The leaves of blueberry wild clones and cultivars can have higher antioxidant activity (Ehlenfeldt & Prior, 2001; Feng et al., 2017), polyphenolics, and proanthocyanidins than fruits (Percival & MacKenzie, 2007; Riihinen et al., 2008).

The antioxidant activity depends on the synergistic and antagonistic interaction of various compounds and environmental factors (Hassimotto et al., 2005). There is no standard agreed method for estimating antioxidant activity because of its complexity (Frankel & Meyer, 2000). In the present study, we used the DPPH radical scavenging method, as it is sensitive and cheaper than other known procedures (Giovanelli & Buratti, 2009). Out of all the groups, the NB wild clones had the highest TAA, followed by CV, Cross 2, and NL wild clones. The TPC and TFC were the highest in NL wild clones, followed by CV. The wild clones from NL and NB can prove to be an important resource for improving antioxidant properties in the blueberry breeding program. The phenolics are an abundantly available secondary metabolite derived from phenylalanine via a secondary metabolic pathway catalyzed by phenylalanine lyase L (PAL). Various biotic and abiotic factors can cause stress in source plants and trigger higher activity of PAL (Taiz & Zeiger, 2006). Low levels of lights in NL province could have contributed to higher levels of TAA, TPC, and TFC. The leaf maturity can have a significant impact on the phytochemical composition of

blueberry. In their study, Riihinen et al. (2008) reported that the red leaves of *V. corymbosum* possessed higher levels of quercetin and kaempferol, p-coumaric, and caffeic acids than the green leaves. This could be the case because solar radiation increases these compounds as a part of the photo-protective mechanism (Riihinen et al., 2008). On top of that, the red leaves contain a small amount of anthocyanin, while green leaves do not have any anthocyanin content (Ferlemi & Lamari, 2016). Therefore, TPC and TFC may not sufficiently explain total antioxidant activity as they are the cocktail of various compounds and their activities. The DPPH value is calculated by the addition of various antioxidant compounds, which depend on the chemical used during the extraction of leaves (Lin et al., 1996). However, a contradicting observation that the young leaves from different varieties of blackberries, raspberries and strawberries possessed higher TPC and TAA than older leaves was reported (Wang & Lin, 2000). There was a positive correlation between TAA with TPC and TFC which was also reported in previous studies involving blueberries (Ghosh et al., 2018; Goyali et al., 2013, 2015).

The biochemical analysis in the present study provided important information about the diversity of antioxidant properties. However, biochemical characteristics by themselves are not enough for the presence of genetic diversity. The DNA marker system provides a precise and reliable method for further analysis of variability. The extent of genetic diversity between and within populations is often the outcome of a combination of factors such as gene flow, genetic drift, inbreeding, mutation, and the selection effect (Hartl & Clark, 1997). It is very expensive, time-consuming, and laborious to develop species-specific molecular markers. Because of these constraints, we used EST-SSR, G-SSR, and EST-PCR markers developed for highbush blueberries (Boches et al., 2005; Rowland et al., 2003b). Our report is apparently the first to use these three types of markers in assessing genetic diversity in a group of hybrids obtained by crossing lowbush with half-high

blueberries. Microsatellite markers have also been used for hybrid identification in closely related wild *Petunia* species (Turchetto et al., 2015). Although G-SSR markers are highly abundant in the plant genome and are attractive due to their reproducibility and polymorphic nature, most of them lack close linkage to transcribed regions and do not have a specific genic function. On the other hand, SSR markers derived from EST sequences are associated with the genome's transcribed or expressed regions (Varshney et al., 2005). The single-pass sequence of cDNA (complimentary DNA) clones that are picked randomly is the source of EST-SSR and EST-PCR markers (Adams et al., 1991). All primer pairs used in this study showed an elevated polymorphism that confirmed the high degree of genetic diversity in the blueberry genome of the current material.

In the present study, the discriminatory power of EST-SSR, G-SSR, and EST-PCR primer pairs were compared by PIC, EMR, MI, D, and R. These values help in determining the effectiveness of a specific primer pair in the analysis of genetic diversity. Although PIC values for EST-SSR (average 0.35) and G-SSR primer pairs (average 0.30) were higher than EST-PCRs (average 0.28), EMR, MI, D, and R values for EST-PCR markers were the highest (average 2.05, 0.61, 0.79 and 2.37, respectively) followed by G-SSR (average 1.28, 0.38, 0.65 and 1.08, respectively) and EST-SSR (average 1.13, 0.33, 0.54 and 1.02, respectively). The highest PIC value for EST-SSR primer pair CA112 (0.96) combined with very low values for EMR, MI, D, and R (0.08 for EMR, MI, D and 0.09 for R; Table 2.4) proved that this primer pair is not worthy for analyzing present blueberry hybrids, wild clones, and cultivars. On the other hand, the EST-PCR primer pair CA227 with its highest MI value among all primers (1.16) was the best for overall utility, and it was followed by CA1423 (MI = 1.00) and CA54 (MI = 0.99) to study the present material. These three EST-PCR primer pairs may be very valuable for analyzing blueberry hybrids. However, the moderate to high

values for most of the primer pairs could be attributed to their effectiveness in studying the genetic diversity of the present material.

In the present study, the mean allele number for EST-SSR, G-SSR, and EST-PCR were 6.30, 5.36, and 4.40, respectively, which is comparable or lower than the previous studies involving SSR and/or EST-PCR primer pairs in blueberries ((22.4; Boches et al., 2006); (4.8; Bell et al., 2008); (18.5; Hinrichsen et al., 2008); (8.33; Garriga et al., 2013); (17; Debnath, 2014); (14.24; Bian et al., 2014); (10; Liu et al., 2014); (20.5; Tailor et al., 2017); (14.33; Bassil et al., 2018)). The average Shannon's index (I) of 0.34 for EST-SSR, 0.29 for G-SSR, and 0.26 for EST-PCR are lower than those recorded for *V. vitis-idaea* (0.57; Persson & Gustavsson, 2001) and *V. myrtillus* (0.55; Albert et al., 2003), *V. uliginosum* (0.65; Albert et al., 2005), *V. corymbosum* (0.62; Bian et al., 2014) and *Vaccinium spp*. (1.93; Debnath, 2014); (2.56; Tailor et al., 2017). Average He values for EST-SSR (0.23), G-SSR (0.19) and EST-PCR (0.16) were also less than those reported in previous studies with blueberries (0.88; Hinrichsen et al., 2008); (0.81; Liu et al., 2014); (0.87; Bian et al., 2014) ; (0.86; Debnath, 2014); (0.80; Tailor et al., 2017); (0.56; Bassil et al., 2018). The lower values of the diversity parameter could indicate genetic erosion resulting from selective farming and deforestation (Debnath, 2016).

We used three complementary methods: STRUCTURE, NJ tree, and PCoA to study population structure and genotype relationships in the wild, cultivated, and hybrid blueberries using 26 PCR-based marker pairs. Genotype identification using DNA markers are favored due to their consistency and reliability, and as they are unaffected by the environment (Debnath et al., 2012). The STRUCTURE analysis of EST-SSR and G-SSR divided genotypes into two major groups. Whereas, EST-PCR and the combined STRUCTURE analysis divided the genotypes into three major groups, with some admixtures. EST-PCR have shown higher values for discrimination

capacity as well as various diversity parameters, which is an indication of its ability to effectively resolve various genotypes in to groups and sub-groups and higher genomic coverage than EST-SSR and G-SSR. This was also confirmed by PCoA and NJ analyses for most genotypes. Admixtures in the wild blueberry clones that were observed in the present material with STRUCTURE analysis might be due to the consequence of a glacial bottleneck and quick colonization of these blueberries along with increased regional gene flow due to migration of human beings and trade in agriculture (Aldrich & Doebley, 1992). Although the hybrids were distributed in all three clusters, most of them formed distinct sub-groups either alone or with lowbush or half-high cultivars. This might be as they had been developed through crossing between lowbush and half-high blueberries and share the genes from both parents. However, most of the wild lowbush blueberries except the NB clones and the half-high cultivars were grouped based on their phenotypes. While lowbush blueberries are less than 0.5 m tall, the suckering to crown forming half-high blueberry plants are 0.5 to 1.0 m tall. Highbush blueberry plants are crown forming and 2.0 m or higher in height (Galletta & Ballington, 1996a). In the present study, most of the wild clones, although collected from four different provinces, did not group based on their collection place. Although there is a wide genetic variation among the wild clones, there is no pattern of differentiation based on their collection places. This was also observed in AMOVA analysis, where in the combined analysis, the variation among groups was 19%, and most of the variations (81%) was among the genotypes within provinces, cultivars, or hybrid groups. Similar observations were also reported by Debnath (2014) and Tailor et al. (2017), who worked with different sets of wild lowbush blueberry clones and observed that wild clones were grouped into different clusters. The present study showed no clear difference between the wild, cultivated, and hybrid blueberries, indicating that diversity-wise, the present genotypes are all heterogeneous in nature. This might be due to a smaller variation among different groups than the variation among

the genotypes within a group. The population size could also be reason as small populations are more likely to lose genetic diversity quickly than large populations due to stochastic sampling error (i.e., genetic drift). STRUCTURE, NJ, and PCoA analyses along with AMOVA analysis were complementary to each other and, thus, instead of using one method, a number of procedures are more informative for drawing valid conclusions. Similarly, using more than one type of molecular markers are always better than using a single type of molecular marker (Debnath, 2014; Tailor et al., 2017). In our study, STRUCTURE, NJ, PCoA and AMOVA analyses with EST-SSR, G-SSR and EST-PCR markers have well discriminated the blueberry wild clones, cultivars and hybrids between wild and cultivated blueberries that are part of our current germplasm repository for the cool climates of Canada.

There are no reports available on the relationship between molecular markers and biochemical properties in blueberry. Our study found no parallels between genetic and biochemical data, as observed by phylogenetic trees, PCA-PCoA graphs, and Mantel test of correlation. The poor correlation between genetic clustering from biochemical data indicates varying genomic coverage in blueberries. Molecular markers span across the genome and most of which are not expressed at the phenotypic level. The noncoding regions of the genome that are not accessible to phenotypic expression might be the reason for the dissimilarity between the molecular and chemical diversity (Debnath et al., 2012). There are only three reports available on the comparative analysis of molecular markers with biochemical properties. In their study, Debnath & Sion (2009) reported no correlation between genetic diversity based on ISSR markers and chemical diversity based on antioxidant activity and anthocyanin content in lingonberry. Similar observations were also reported in strawberry (Debnath & Ricard, 2009) and cranberry using ISSR, EST-SSR, and EST-PCR markers (Debnath & An, 2019). We also studied the association of EST-SSR, G-SSR, and

EST-PCR markers with 70 genotypes of blueberry and found that only one marker is associated with TPC as revealed by DPPH assay and five markers associated with both TPC and TFC. This can be explained by the polyploid nature of blueberry and the distribution of associated alleles across the whole genome. However, our study is the first one to use SMRA to identify markers associated with antioxidant properties. This method can provide easy and reliable identification of favorable genotypes or populations in a breeding program at an early stage and has been used to associate molecular markers with traits in numerous species such as mulberry (Kar et al., 2008), buckthorn (Ruan et al., 2009), and Tunisian olive (Omri et al., 2020). This approach is a convenient and quick tool for marker-trait association without the need of mapping populations. The multi-genic control of TPC, TFC, and antioxidant traits can have practical use in future blueberry breeding programs.

Blueberries are of significant importance for their antioxidant phytochemicals, especially phenolic metabolites, that play a significant role in human health benefits and plant defense mechanism (Debnath & Goyali, 2020). Most plant phenolics are flavonoids and non-flavonoids. Flavonoids are of two types: anthocyanins and anthoxanthins. While anthocyanins are pigment molecules (red, blue, and purple), anthoxanthins are white to yellow or colourless molecules and include flavanols, flavonols, flavones, and isoflavones. Non-flavonoids comprise phenolic acids, lignans, and stilbenes. Tannin and lignin are other non-flavonoid subclasses (Debnath & Goyali, 2020). The flavonol quercetin is an important nutritional bioactive compound with its high bioaccessibility (~80%) (Gapski et al., 2019). Quercetin helps in protecting against osteoporosis, cancer, pulmonary and cardiovascular diseases, and aging (Boots et al., 2008). In blueberry, anthocyanins were found to possess the highest inhibition effects on in vitro colon cancer cell proliferation, followed by flavonol and tannins (Massarotto et al., 2016). Biomarker-based human clinical studies showed that regular and moderate consumption of blueberries and/or anthocyanins are associated with a

reduced risk of death, cardiovascular disease, and type 2 diabetes (Kalt et al., 2020). In another study with in vitro cell bioassays for anti-inflammatory and antioxidant activities, Grace et al. (2019) reported that the anthocyanin group of phenolics was mainly responsible for the bioactivity and the blueberry extract suppressed pro-inflammatory markers (interleukin-1β, cyclooxygenase-2, inducible nitric oxide synthase; interleukin-6; (Esposito et al., 2014)). Polyphenol-, anthocyanin- and proanthocyanidin-rich crude wild blueberry extract components were found to suppress mRNA biomarkers of acute inflammation, and malvidin-3-glucoside suppressed the effects of pro-inflammatory genes that are responsible for transcriptional regulation and cytokine-mediated inflammation (Esposito et al., 2014). It has been observed that the in vitro antioxidant assay with blueberries resulted in a strong correlation with those of total phenolics and total anthocyanin contents (Grace et al., 2019). In the current study, we measured total phenolics, flavonoids and antioxidant contents to study genetic and biochemical diversity in a blueberry germplasm and to identify blueberry genotypes with high bioactive components and wide diversity for using in an on-going breeding program. Identifying phenolic-rich cultivars for breeding species with high bioactive composition is an important approach to improve the nutritional quality in blueberries. Crossing between selected genotypes is expected to develop new cultivars combining superior health-promoting bioactive components with diverse adaptability under changed environment.

CHAPTER 5

Conclusions

Our study is the first of its kind to investigate the antioxidant activity, phenolic, and flavonoid contents along with genetic diversity analysis using three types of marker systems: EST-SSR, G-SSR, and EST-PCR among blueberry. The study identified two NL (BC2, 6) for TAA and one QC wild clone (BC22) for TPC and TFC, superior to cultivars and hybrids. These wild blueberry clones hold a key for designing future breeding exercises to generate cultivars with valuable antioxidant traits. The present study indicates that ten EST-SSR, eight G-SSR, and eight EST-PCR primer pairs could distinguish and report genetic variations at the molecular level among wild and cultivated lowbush, half-high and highbush blueberries and among hybrids between lowbush and half-high blueberries. The EST-PCR primer pair CA227 was the best to discriminate blueberry hybrids, clones, and cultivars, followed by EST-PCR primer pairs CA1423 and CA54. The alleles of CA1423 and CA54 also showed a strong positive correlation and association to TPC and TFC in SMRA. The utility of these primers across different blueberry species can help identify and characterize inter-species blueberry hybrids and select useful genotypes as a parent in a breeding program. The DNA fingerprinting with more than one type of molecular marker will allow better management of blueberry germplasm and conservation efforts. Clustering based on EST-SSR, G-SSR, EST-PCR, and combined primer data was different from antioxidant properties. These markers are spread across the genome, many of which are located in non-coding regions, explaining the poor correlation between genetic and biochemical data. However, these markers' potential utility is immense, as shown by our association study in the blueberry marker-biochemical relationship using SMRA, and can prove to be a valuable tool. Though three marker systems were

used in the present study, more markers and more sophisticated marker systems, such as SNPs, can be used for genetic assessment with higher number of genotypes in each groups especially cultivars and hybrids. A combinatorial approach of gene discovery through genomics, proteomics, and other supplementary branches of biotechnology will speed up the berry improvement programs. Biochemical analysis using more advanced techniques such as HPLC can be employed. Such techniques can be expanded to various maturity stages of leaves as well as fruits to broaden our understanding of antioxidants.

CHAPTER 6

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews. Computational statistics, 2*(4), 433-459.

Adams, L. S., Phung, S., Yee, N., Seeram, N. P., Li, L., & Chen, S. (2010). Blueberry phytochemicals inhibit growth and metastatic potential of MDA-MB-231 breast cancer cells through modulation of the phosphatidylinositol 3-kinase pathway. *Cancer Research, 70*(9), 3594-3605.

Adams, M., Kelley, J., Gocayne, J., Dubnick, M., Polymeropoulos, M., Xiao, H., . . . et, a. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science, 252*(5013), 1651-1656. doi:10.1126/science.2047873

Adhikari, S., Saha, S., Biswas, A., Rana, T., Bandyopadhyay, T. K., & Ghosh, P. (2017). Application of molecular markers in plant genome analysis: a review. *The Nucleus, 60*(3), 283-297.

Afifi, A., May, S., & Clark, V. A. (2003). *Computer-aided multivariate analysis*: CRC Press.

Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports, 27*(4), 617-631.

Ahmadian, A., Gharizadeh, B., Gustafsson, A. C., Sterky, F., Nyrén, P., Uhlén, M., & Lundeberg, J. (2000). Single-nucleotide polymorphism analysis by pyrosequencing. *Analytical Biochemistry, 280*(1), 103-110.

Albert, T., Raspé, O., & Jacquemart, A.-L. (2004). Clonal diversity and genetic structure in Vaccinium myrtillus populations from different habitats. *Belgian Journal of Botany*, 155-162.

Albert, T., Raspé, O., & Jacquemart, A.-L. (2003). Clonal structure in Vaccinium myrtillus L. revealed by RAPD and AFLP markers. *International Journal of Plant Sciences, 164*(4), 649-655.

Albert, T., Raspé, O., & Jacquemart, A.-L. (2005). Diversity and spatial structure of clones in Vaccinium uliginosum populations. *Canadian Journal of Botany, 83*(2), 211-218.

Alderborn, A., Kristofferson, A., & Hammerling, U. (2000). Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. *Genome Research, 10*(8), 1249-1258.

Aldrich, P., & Doebley, J. (1992). Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild Sorghum bicolor. *Theoretical and Applied Genetics, 85*(2-3), 293-302.

Altukhov, Y. P., & Salmenkova, E. (2002). DNA polymorphism in population genetics. *Russian Journal of Genetics, 38*(9), 989-1008.

Amiryousefi, A., Hyvönen, J., & Poczai, P. (2018). iMEC: Online Marker Efficiency Calculator. *Applications in Plant Sciences, 6*(6), e01159.

An, D., Bykova, N. V., & Debnath, S. C. (2015). EST-PCR, EST-SSR and ISSR markers to identify a set of wild cranberries and evaluating their relationships. *Canadian Journal of Plant Science*. doi:10.4141/cjps-2015-158

Aruna, M., Austin, M. E., & Ozias-Akins, P. (1995). Randomly amplified polymorphic DNA fingerprinting for identifying rabbiteye blueberry (Vaccinium ashei Reade) cultivars. *Journal of the American Society for Horticultural Science, 120*(5), 710-713.

Aruna, M., Ozias-Akins, P., Austin, M. E., & Kochert, G. (1993). Genetic relatedness among rabbiteye blueberry (Vaccinium ashei) cultivars determined by DNA amplification using single primers of arbitrary sequence. *Genome, 36*(5), 971-977.

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

Bassil, N., Bidani, A., Hummer, K., Rowland, L. J., Olmstead, J., Lyrene, P., & Richards, C. (2018). Assessing genetic diversity of wild southeastern North American Vaccinium species using microsatellite markers. *Genetic Resources and Crop Evolution, 65*(3), 939-950.

Basu, A., Du, M., Leyva, M. J., Sanchez, K., Betts, N. M., Wu, M., . . . Lyons, T. J. (2010). Blueberries Decrease Cardiovascular Risk Factors in Obese Men and Women with Metabolic Syndrome–3. *The Journal of nutrition, 140*(9), 1582-1587.

Bell, D. J., Rowland, L. J., Polashock, J. J., & Drummond, F. A. (2008). Suitability of EST-PCR markers developed in highbush blueberry for genetic fingerprinting and relationship studies in lowbush blueberry and related species. *Journal of the American Society for Horticultural Science, 133*(5), 701-707.

Bell, D. J., Rowland, L. J., Smagula, J., & Drummond, F. (2009). Recent advances in the biology and genetics of lowbush blueberry. *Maine Agricultural and Forest Experiment Station Technical Bulletin, 203*.

Bian, Y., Ballington, J., Raja, A., Brouwer, C., Reid, R., Burke, M., . . . Brown, A. (2014). Patterns of simple sequence repeats in cultivated blueberries (Vaccinium section Cyanococcus spp.) and their use in revealing genetic diversity and population structure. *Molecular Breeding, 34*(2), 675-689.

Bidani, A., Hummer, K. E., Olmstead, J. W., Lyrene, P. M., Rowland, L. J., & Bassil, N. V. (2017). Microsatellite markers for genetic analyses in southeastern Vaccinium species. *Acta Horticulturae*(1180), 373-378. doi:10.17660/ActaHortic.2017.1180.51

Bobowski, B., Hole, D., Wolf, P. G., & Bryant, L. (1999). Identification of roots of woody species using polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) analysis. *Molecular Ecology, 8*(3), 485-491.

Boches, P., Bassil, N., & Rowland, L. (2005). Microsatellite markers for Vaccinium from EST and genomic libraries. *Molecular Ecology Resources, 5*(3), 657-660.

Boches, P., Bassil, N. V., & Rowland, L. (2006). Genetic diversity in the highbush blueberry evaluated with microsatellite markers. *Journal of the American Society for Horticultural Science, 131*(5), 674-686.

Boches, P., Hummer, K., Bassil, N., & Rowland, J. (2004). *Cross-species amplification of SSRs in the genus Vaccinium.* Paper presented at the VIII International Symposium on Vaccinium Culture 715.

Bomser, J., Madhavi, D., Singletary, K., & Smith, M. (1996). In vitro anticancer activity of fruit extracts from Vaccinium species. *Planta Medica, 62*(03), 212-216.

Boots, A. W., Haenen, G. R., & Bast, A. (2008). Health effects of quercetin: from antioxidant to nutraceutical. *European Journal of Pharmacology, 585*(2-3), 325-337.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics, 32*(3), 314.

Bretting, P., & Widrlechner, M. P. (1995). Genetic markers and horticultural germplasm management. *HortScience, 30*(7), 1349-1356.

Brevis, P. A., Bassil, N. V., Ballington, J. R., & Hancock, J. F. (2008). Impact of wide hybridization on highbush blueberry breeding. *Journal of the American Society for Horticultural Science, 133*(3), 427-437.

Bromberg, P. A. (2016). Mechanisms of the acute effects of inhaled ozone in humans. *Biochimica et Biophysica Acta (BBA)-General Subjects, 1860*(12), 2771-2781.

Brookes, A. J. (1999). The essence of SNPs. *Gene, 234*(2), 177-186.

Brown, A. H., & Weir, B. S. (1983). Measuring genetic variability in plant populations. *Isozymes in plant genetics and breeding, part A*, 219-239.

Bruederle, L. P., & Vorsa, N. (1994). Genetic differentiation of diploid blueberry, Vaccinium sect. Cyanococcus (Ericaceae). *Systematic Botany*, 337-349.

Bruederle, L. P., Vorsa, N., & Ballington, J. R. (1991). Population genetic structure in diploid blueberry Vaccinium section Cyanococcus (Ericaceae). *American Journal of Botany*, 230-237.

Bruskov, V. I., Malakhova, L. V., Masalimov, Z. K., & Chernikov, A. V. (2002). Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. *Nucleic Acids Research, 30*(6), 1354-1363.

Buetow, K. H., Edmonson, M. N., & Cassidy, A. B. (1999). Reliable identification of large numbers of candidate SNPs from public EST data. *Nature Genetics, 21*(3), 323.

Burgher, K., Jamieson, A., Lu, X., & McRae, K. (1998). Genetic relationship of 26 wild blueberry clones (Vaccinium angustifolium), as determined by random amplified Polymorphic DNA analysis. *Proc Atlantic Plant Tissue Culture Assoc Annu Mtg, Fredericton, New Brunswick, Canada, 7*.

Burgher, K. L., Jamieson, A. R., & Lu, X. (2002). Genetic relationships among lowbush blueberry genotypes as determined by randomly amplified polymorphic DNA analysis. *Journal of the American Society for Horticultural Science, 127*(1), 98-103.

Caetano-Anollés, G., & Bassam, B. J. (1993). DNA amplification fingerprinting using arbitrary oligonucleotide primers. *Applied Biochemistry and Biotechnology, 42*(2-3), 189-200.

Calò, R., & Marabini, L. (2014). Protective effect of Vaccinium myrtillus extract against UVA- and UVB-induced damage in a human keratinocyte cell line (HaCaT cells). *Journal of Photochemistry and Photobiology B: Biology, 132*, 27-35.

Camp, W. (1945). The North American blueberries with notes on other groups of Vacciniaceae. *Brittonia, 5*(3), 203-275.

Cao, G., Sofic, E., & Prior, R. L. (1996). Antioxidant capacity of tea and common vegetables. *Journal of Agricultural and Food Chemistry, 44*(11), 3426-3431.

Cardeñosa, V., Girones-Vilaplana, A., Muriel, J. L., Moreno, D. A., & Moreno-Rojas, J. M. (2016). Influence of genotype, cultivation system and irrigation regime on antioxidant capacity and selected phenolics of blueberries (Vaccinium corymbosum L.). *Food Chemistry, 202*, 276-283.

Cariello, N. F., Scott, J. K., Kat, A. G., Thilly, W. G., & Keohavong, P. (1988). Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT Munich. *American Journal of Human Genetics, 42*(5), 726.

Castrejón, A. D. R., Eichholz, I., Rohn, S., Kroh, L. W., & Huyskens-Keil, S. (2008). Phenolic profile and antioxidant activity of highbush blueberry (Vaccinium corymbosum L.) during fruit maturation and ripening. *Food Chemistry, 109*(3), 564-572.

Castro, L., & Freeman, B. A. (2001). Reactive oxygen species in human health and disease. *Nutrition, 17*(2), 161-165.

Česonienė, L., Daubaras, R., Paulauskas, A., Žukauskienė, J., & Zych, M. (2013). Morphological and genetic diversity of European cranberry (Vaccinium oxycoccos L., Ericaceae) clones in Lithuanian reserves. *Acta Societatis Botanicorum Poloniae, 82*(3), 211–217-211–217.

Chen, P.-N., Chu, S.-C., Chiou, H.-L., Chiang, C.-L., Yang, S.-F., & Hsieh, Y.-S. (2005). Cyanidin 3-glucoside and peonidin 3-glucoside inhibit tumor cell growth and induce apoptosis in vitro and suppress tumor growth in vivo. *Nutrition and Cancer, 53*(2), 232-243.

Cheynier, V., Comte, G., Davies, K. M., Lattanzio, V., & Martens, S. (2013). Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiology and Biochemistry, 72*, 1-20.

Ching, A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O., Tingey, S., . . . Rafalski, A. J. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics, 3*(1), 19.

Cho, K. H., Park, S. J., Kim, S. J., Kim, S. H., Lee, H. C., Kim, M. Y., & Chun, J. A. (2017). Blueberry Cultivar Identification Using Random Amplified Polymorphic DNA and Sequence-characterized Amplified Region Markers. *HortScience, 52*(11), 1483-1489.

Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E., . . . Federspiel, N. (1999). Genome-wide mapping with biallelic markers in Arabidopsis thaliana. *Nature Genetics, 23*(2), 203.

Collard, B. C., & Mackill, D. J. (2009). Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reporter, 27*(1), 86.

Connor, A. M., Luby, J. J., Hancock, J. F., Berkheimer, S., & Hanson, E. J. (2002). Changes in fruit antioxidant activity among blueberry cultivars during cold-temperature storage. *Journal of Agricultural and Food Chemistry, 50*(4), 893-898.

Cotton, R. (1993). Current methods of mutation detection. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 285*(1), 125-144.

Cox, T., Murphy, J., & Rodgers, D. (1986). Changes in genetic diversity in the red winter wheat regions of the United States. *Proceedings of the National Academy of Sciences, 83*(15), 5583-5586.

Cuypers, A., Plusquin, M., Remans, T., Jozefczak, M., Keunen, E., Gielen, H., . . . Artois, T. J. (2010). Cadmium stress: an oxidative challenge. *BioMetals, 23*(5), 927-940.

Dale, A. (1996). Key and vegetative descriptions of thirty-two common strawberry varieties grown in North America. *Advances in strawberry research*.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics, 12*(7), 499.

Debnath. (2007). Inter simple sequence repeat (ISSR) to assess genetic diversity within a collection of wild lingonberry (Vaccinium vitis-idaea L.) clones. *Canadian journal of plant science, 87*(2), 337-344.

Debnath, S. (2000). Combined application of classical and biotechnological techniques in the development of small fruits important to Newfoundland and Labrador. *Can. J. Plant Sci, 80*, 233.

Debnath, S. (2005). Differentiation of Vaccinium cultivars and wild clones using RAPD markers. *Journal of Plant Biochemistry and Biotechnology, 14*(2), 173-177.

Debnath, S., Siow, Y., Petkau, J., An, D., & Bykova, N. (2012). Molecular markers and antioxidant activity in berry crops: Genetic diversity analysis. *Canadian Journal of Plant Science, 92*(6), 1121-1133.

Debnath, S. C. (2009). Development of ISSR markers for genetic diversity studies inVaccinium angustifolium. *Nordic Journal of Botany, 27*(2), 141-148. doi:10.1111/j.1756-1051.2009.00402.x

Debnath, S. C. (2016). Genetic Diversity and Erosion in Berries. In *Genetic Diversity and Erosion in Plants* (pp. 75-129): Springer.

Debnath, S. C. (2014). Structured diversity using EST-PCR and EST-SSR markers in a set of wild blueberry clones and cultivars. *Biochemical Systematics and Ecology, 54*, 337-347.

Debnath, S. C., & An, D. (2019). Antioxidant properties and structured biodiversity in a diverse set of wild cranberry clones. *Heliyon, 5*(4), e01493.

Debnath, S. C., & Goyali, J. C. (2020). In Vitro Propagation and Variation of Antioxidant Properties in Micropropagated Vaccinium Berry Plants—A Review. *Molecules, 25*(4), 788.

Debnath, S. C., & Ricard, E. (2009). ISSR, anthocyanin content and antioxidant activity analyses to characterize strawberry genotypes. *Journal of Applied Horticulture, 11*(2), 83-89.

Debnath, S. C., & Sion, M. (2009). Genetic diversity, antioxidant activities, and anthocyanin contents in lingonberry. *International Journal of Fruit Science, 9*(2), 185-199.

Efroymson, M. (1960). *Multiple regression analysis*. New York: John Wiley & Sons.

Ehala, S., Vaher, M., & Kaljurand, M. (2005). Characterization of phenolic profiles of Northern European berries by capillary electrophoresis and determination of their antioxidant activity. *Journal of Agricultural and Food Chemistry, 53*(16), 6484-6490.

Ehlenfeldt, M. K., & Prior, R. L. (2001). Oxygen radical absorbance capacity (ORAC) and phenolic and anthocyanin concentrations in fruit and leaf tissues of highbush blueberry. *Journal of Agricultural and Food Chemistry, 49*(5), 2222-2227.

Einset, J., Winge, P., & Bones, A. (2007). ROS signaling pathways in chilling stress. *Plant signaling & behavior, 2*(5), 365-367.

Esposito, D., Chen, A., Grace, M. H., Komarnytsky, S., & Lila, M. A. (2014). Inhibitory effects of wild blueberry anthocyanins and other flavonoids on biomarkers of acute and chronic inflammation in vitro. *Journal of Agricultural and Food Chemistry, 62*(29), 7022-7028.

Esselink, G., Smulders, M., & Vosman, B. (2003). Identification of cut rose (Rosa hybrida) and rootstock varieties using robust sequence tagged microsatellite site markers. *Theoretical and Applied Genetics, 106*(2), 277-286.

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.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics, 131*(2), 479-491.

FAOSTAT. (2019). Food and Agriculture Organization of United Nations. http://www.fao.org/faostat/en/#data/QC

Feng, C.-Y., Wang, W.-W., Ye, J.-F., Li, S.-S., Wu, Q., Yin, D.-D., . . . Wang, L.-S. (2017). Polyphenol profile and antioxidant activity of the fruit and leaf of Vaccinium glaucoalbum from the Tibetan Himalayas. *Food Chemistry, 219*, 490-495.

Ferlemi, A.-V., & Lamari, F. N. (2016). Berry leaves: an alternative source of bioactive natural products of nutritional and medicinal value. *Antioxidants, 5*(2), 17.

Finn, C., Luby, J., & Wildung, D. (1990). Half-high blueberry cultivars. *Fruit Varieties Journal, 44*(2), 63-68.

Flinn, M. A., & Pringle, J. K. (1983). Heat tolerance of rhizomes of several understory species. *Canadian Journal of Botany, 61*(2), 452-457.

Francis, R. M. (2017). POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources, 17*(1), 27-32.

Frankel, E. N., & Meyer, A. S. (2000). The problems of using one-dimensional methods to evaluate multifunctional food and biological antioxidants. *Journal of the Science of Food and Agriculture, 80*(13), 1925-1941.

Galletta, G., & Ballington, J. (1996a). Blueberries, cranberries and lingonberries. *Fruit breeding, 2*(1), 107.

Galletta, G. J., & Ballington, J. R. (1996b). Blueberries, cranberries and lingonberries. In J. Janick & J. N. Moore (Eds.), *Fruit breeding* (Vol. 2, pp. 1-107).

Gao, L., & Mazza, G. (1994). Quantitation and distribution of simple and acylated anthocyanins and other phenolics in blueberries. *Journal of Food Science, 59*(5), 1057-1059.

Gapski, A., Gomes, T. M., Bredun, M. A., Ferreira-Lima, N. E., Ludka, F. K., Bordignon-Luiz, M. T., & Burin, V. M. (2019). Digestion behavior and antidepressant-like effect promoted by acute administration of blueberry extract on mice. *Food Research International, 125*, 108618.

Garkava-Gustavsson, L., Persson, H., Nybom, H., Rumpunen, K., Gustavsson, B., & Bartish, I. (2005). RAPD-based analysis of genetic diversity and selection of lingonberry (Vaccinium vitis-idaea L.) material for ex situ conservation. *Genetic Resources and Crop Evolution, 52*(6), 723-735.

Garriga, M., Parra, P. A., Caligari, P. D., Retamales, J. B., Carrasco, B. A., Lobos, G. A., & García-Gonzáles, R. (2013). Application of inter-simple sequence repeats relative to simple sequence repeats as a molecular marker system for indexing blueberry cultivars. *Canadian Journal of Plant Science*, *93*(5), 913-921.

Gavrilova, V., Kajdzanoska, M., Gjamovski, V., & Stefova, M. (2011). Separation, characterization and quantification of phenolic compounds in blueberries and red and black currants by HPLC− DAD− ESI-MS n. *Journal of Agricultural and Food Chemistry*, *59*(8), 4009-4018.

Gerschman, R., Gilbert, D., Nye, S., Dwyer, P., & Fenn, W. (2001). Oxygen poisoning and X-irradiation: a mechanism in common. 1954. *Nutrition (Burbank, Los Angeles County, Calif.)*, *17*(2), 162.

Ghosh, A., Igamberdiev, A. U., & Debnath, S. C. (2018). Thidiazuron-induced somatic embryogenesis and changes of antioxidant properties in tissue cultures of half-high blueberry plants. *Scientific Reports*, *8*(1), 16978.

Giovanelli, G., Brambilla, A., & Sinelli, N. (2013). Effects of osmo-air dehydration treatments on chemical, antioxidant and morphological characteristics of blueberries. *LWT-Food Science and Technology*, *54*(2), 577-584.

Giovanelli, G., & Buratti, S. (2009). Comparison of polyphenolic composition and antioxidant activity of wild Italian blueberries and some cultivated varieties. *Food Chemistry*, *112*(4), 903-908.

Goyali, J. C., Igamberdiev, A. U., & Debnath, S. C. (2013). Morphology, phenolic content and antioxidant capacity of lowbush blueberry (Vaccinium angustifolium Ait.) plants as affected by in vitro and ex vitro propagation methods. *Canadian Journal of Plant Science*, *93*(6), 1001-1008.

Goyali, J. C., Igamberdiev, A. U., & Debnath, S. C. (2015). Propagation methods affect fruit morphology and antioxidant properties but maintain clonal fidelity in lowbush blueberry. *HortScience, 50*(6), 888-896.

Grace, M. H., Xiong, J., Esposito, D., Ehlenfeldt, M., & Lila, M. A. (2019). Simultaneous LC-MS quantification of anthocyanins and non-anthocyanin phenolics from blueberries with widely divergent profiles and biological activities. *Food Chemistry, 277*, 336-346.

Graham, J., Smith, K., MacKenzie, K., Jorgenson, L., Hackett, C., & Powell, W. (2004). The construction of a genetic linkage map of red raspberry (Rubus idaeus subsp. idaeus) based on AFLPs, genomic-SSR and EST-SSR markers. *Theoretical and Applied Genetics, 109*(4), 740-749.

Gupta, D. K., & Sandalio, L. M. (2011). *Metal toxicity in plants: perception, signaling and remediation*: Springer Science & Business Media.

Gupta, M., Chyi, Y.-S., Romero-Severson, J., & Owen, J. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *TAG Theoretical and Applied Genetics, 89*(7), 998-1006.

Gupta, P., Rustgi, S., Sharma, S., Singh, R., Kumar, N., & Balyan, H. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genetics and Genomics, 270*(4), 315-323.

Hacia, J. G., Fan, J.-B., Ryder, O., Jin, L., Edgemon, K., Ghandour, G., . . . Robbins, C. M. (1999). Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nature Genetics, 22*(2), 164.

Haghighi, K., & Hancock, J. (1992). DNA restriction fragment length variability in the genomes of highbush blueberry. *HortScience, 27*(1), 44-47.

Häkkinen, S. H., Kärenlampi, S. O., Heinonen, I. M., Mykkänen, H. M., & Törrönen, A. R. (1999). Content of the flavonols quercetin, myricetin, and kaempferol in 25 edible berries. *Journal of Agricultural and Food Chemistry, 47*(6), 2274-2279.

Hall, I., Jamieson, A., & Brydon, A. (1988). Cumberland and Fundy lowbush blueberries. *Canadian Journal of Plant Science, 68*(2), 553-555.

Halliwell, B. (1994). Free radicals, antioxidants, and human disease: curiosity, cause, or consequence? *The lancet, 344*(8924), 721-724.

Halliwell, B. (1992). Reactive oxygen species and the central nervous system. *Journal of Neurochemistry, 59*(5), 1609-1623.

Halliwell, B. (1991). Reactive oxygen species in living systems: source, biochemistry, and role in human disease. *The American journal of medicine, 91*(3), S14-S22.

Hamada, H., & Kakunaga, T. (1982). Potential Z-DNA forming sequences are highly dispersed in the human genome. *Nature, 298*(5872), 396-398.

Han, X., Shen, T., & Lou, H. (2007). Dietary polyphenols and their biological significance. *International Journal of Molecular Sciences, 8*(9), 950-988.

Hancock, J. F., Luby, J. J., & Beaudry, R. (2003). Fruits of the Ericaceae. In L. Trugo, P. M. Finglas, & B. Caballero (Eds.), *Encyclopedia of Food Science, Food Technology and Nutrition* (pp. 2762-2768). London, UK: Academic Press.

Hancock, J. F., Lyrene, P., Finn, C. E., Vorsa, N., & Lobos, G. A. (2008). Blueberries and Cranberries. In J. F. Hancock (Ed.), *Temperate Fruit Crop Breeding: Germplasm to Genomics* (pp. 115-150). Dordrecht: Springer Netherlands.

Harrison, R., Luby, J., Furnier, G., & Hancock, J. (1997). Morphological and molecular variation among populations of octoploid Fragaria virginiana and F. chiloensis (Rosaceae) from North America. *American Journal of Botany, 84*(5), 612-612.

Harrison, R. E., Luby, J. J., Furnier, G. R., & Hancock, J. F. (2000). Differences in the apportionment of molecular and morphological variation in North American strawberry and the consequences for genetic resource management. *Genetic Resources and Crop Evolution, 47*(6), 647-657.

Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics* (Vol. 116). Sunderland, USA: Sinauer associates, Inc.

Hassimotto, N. M. A., Genovese, M. I., & Lajolo, F. M. (2005). Antioxidant activity of dietary fruits, vegetables, and commercial frozen fruit pulps. *Journal of Agricultural and Food Chemistry, 53*(8), 2928-2935.

Hayashi, K. (1992). PCR-SSCP: a method for detection of mutations. *Genetic Analysis: Biomolecular Engineering, 9*(3), 73-79.

Hayashi, K., & Yandell, D. W. (1993). How sensitive is PCR-SSCP? *Human Mutation, 2*(5), 338-346.

He, C., Chen, L., Simmons, M., Li, P., Kim, S., & Liu, Z. (2003). Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. *Animal Genetics, 34*(6), 445-448.

Heath, D. D., Lwama, G. K., & Devlin, R. H. (1993). PCR primed with VNTR core sequences yields species specific patterns and hypervariable probes. *Nucleic Acids Research, 21*(24), 5782-5785.

Hecker, K. H., Taylor, P. D., & Gjerde, D. T. (1999). Mutation detection by denaturing DNA chromatography using fluorescently labeled polymerase chain reaction products. *Analytical Biochemistry, 272*(2), 156-164.

Heleno, S. A., Martins, A., Queiroz, M. J. R., & Ferreira, I. C. (2015). Bioactivity of phenolic acids: Metabolites versus parent compounds: A review. *Food Chemistry, 173*, 501-513.

Helentjaris, T., Slocum, M., Wright, S., Schaefer, A., & Nienhuis, J. (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical and Applied Genetics, 72*(6), 761-769.

Hemmat, M., Weedon, N., Manganaris, A., & Lawson, D. (1994). Molecular marker linkage map for apple. *Journal of Heredity, 85*(1), 4-11.

Hennink, S., & Zeven, A. (1990). The interpretation of Nei and Shannon-Weaver within population variation indices. *Euphytica, 51*(3), 235-240.

Hill, N., & Vander Kloet, S. (1983). Zymotypes in Vaccinium section Cyanococcus and related groups. *Proceedings of the Nova Scotian Institute of Science, 33*.

Hinrichsen, P., Herminia Castro, M., Ravest, G., Rojas, G., Méndez, M., Bassil, N., & Muñoz, C. (2008). *Minimal microsatellite marker panel for fingerprinting blueberry cultivars.* Paper presented at the IX International Vaccinium Symposium 810.

Hokanson, K., & Hancock, J. (1998). Levels of allozymic diversity in diploid and tetraploid Vaccinium sect. Cyanococcus (blueberries). *Canadian journal of plant science, 78*(2), 327-332.

Horvath, A., Sánchez-Sevilla, J., Punelli, F., Richard, L., Sesmero-Carrasco, R., Leone, A., . . . Barreneche, T. (2011). Structured diversity in octoploid strawberry cultivars: importance of the old European germplasm. *Annals of Applied Biology, 159*(3), 358-371.

Howard, L. R., Clark, J. R., & Brownmiller, C. (2003). Antioxidant capacity and phenolic content in blueberries as affected by genotype and growing season. *Journal of the Science of Food and Agriculture, 83*(12), 1238-1247.

Howell, A. (2008). *Update on health benefits of cranberry and blueberry*. Paper presented at the IX International Vaccinium Symposium 810.

Howell, A. B., Vorsa, N., Marderosian, A. D., & Foo, L. Y. (1998). Inhibition of the adherence of P-fimbriated Escherichia coli to uroepithelial-cell surfaces by proanthocyanidin extracts from cranberries. *New England Journal of Medicine, 339*(15), 1085-1086.

Hu, J., & Vick, B. A. (2003). Target region amplification polymorphism: a novel marker technique for plant genotyping. *Plant Molecular Biology Reporter, 21*(3), 289-294.

Jaccoud, D., Peng, K., Feinstein, D., & Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research, 29*(4), e25-e25.

Jeffreys, A. J., Neumann, R., & Wilson, V. (1990). Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell, 60*(3), 473-485.

Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Hypervariable 'minisatellite'regions in human DNA. *Nature, 314*(6006), 67.

Joseph, J. A., Arendash, G., Gordon, M., Diamond, D., Shukitt-Hale, B., Morgan, D., & Denisova, N. (2003). Blueberry supplementation enhances signaling and prevents behavioral deficits in an Alzheimer disease model. *Nutritional Neuroscience, 6*(3), 153-162.

Joseph, J. A., Shukitt-Hale, B., Denisova, N. A., Bielinski, D., Martin, A., McEwen, J. J., & Bickford, P. C. (1999). Reversals of age-related declines in neuronal signal transduction, cognitive, and motor behavioral deficits with blueberry, spinach, or strawberry dietary supplementation. *Journal of Neuroscience, 19*(18), 8114-8121.

Joseph, J. A., SHUKITT-HALE, B., & Lau, F. C. (2007). Fruit polyphenols and their effects on neuronal signaling and behavior in senescence. *Annals of the New York Academy of Sciences, 1100*(1), 470-485.

Kähkönen, M. P., Hopia, A. I., & Heinonen, M. (2001). Berry phenolics and their antioxidant activity. *Journal of Agricultural and Food Chemistry, 49*(8), 4076-4082.

Kälin, I., Shephard, S., & Candrian, U. (1992). Evaluation of the ligase chain reaction (LCR) for the detection of points mutation. *Mutation Research Letters, 283*(2), 119-123.

Kalt, W., Cassidy, A., Howard, L. R., Krikorian, R., Stull, A. J., Tremblay, F., & Zamora-Ros, R. (2020). Recent research on the health benefits of blueberries and their anthocyanins. *Advances in Nutrition, 11*(2), 224-236.

Kalt, W., McDonald, J., Ricker, R., & Lu, X. (1999). Anthocyanin content and profile within and among blueberry species. *Canadian Journal of Plant Science, 79*(4), 617-623.

Kalt, W., Ryan, D. A., Duy, J. C., Prior, R. L., Ehlenfeldt, M. K., & Vander Kloet, S. (2001). Interspecific variation in anthocyanins, phenolics, and antioxidant capacity among genotypes of highbush and lowbush blueberries (Vaccinium section cyanococcus spp.). *Journal of Agricultural and Food Chemistry, 49*(10), 4761-4767.

Kan, Y. W., & Dozy, A. M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proceedings of the National Academy of Sciences, 75*(11), 5631-5635.

Kantety, R. V., La Rota, M., Matthews, D. E., & Sorrells, M. E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology, 48*(5-6), 501-510.

Kar, P. K., Srivastava, P. P., Awasthi, A. K., & Urs, S. R. (2008). Genetic variability and association of ISSR markers with some biochemical traits in mulberry (Morus spp.) genetic resources available in India. *Tree Genetics & Genomes, 4*(1), 75-83. doi:10.1007/s11295-007-0089-x

Kiss, G. B., Csanadi, G., Kalman, K., Kaló, P., & Ökrész, L. (1993). Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Molecular and General Genetics MGG, 238*(1-2), 129-137.

Koca, I., & Karadeniz, B. (2009). Antioxidant properties of blackberry and blueberry fruits grown in the Black Sea Region of Turkey. *Scientia Horticulturae, 121*(4), 447-450.

Kolosova, N., Lebedev, P., & Dikalova, A. (2004). Comparison of antioxidants in the ability to prevent cataract in prematurely aging OXYS rats. *Bulletin of Experimental Biology and Medicine, 137*(3), 249-251.

Konieczny, A., & Ausubel, F. M. (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal, 4*(2), 403-410.

Kota, R., Varshney, R. K., Thiel, T., Dehmer, K. J., & Graner, A. (2001). Generation and Comparison of EST-Derived SSRs and SNPs in Barley (Hordeum Vulgare L.). *Hereditas, 135*(2-3), 145-151.

Krebs, S., & Hancock, J. (1989). Tetrasomic inheritance of isoenzyme markers in the highbush blueberry, Vaccinium corymbosum. *Heredity, 63*(1), 11-18.

Kronstad, W. E. (1986). *Genetic diversity and plant improvement.* Paper presented at the Plant Breeding Symposium DSIR.

Kumagai, Y., Shinkai, Y., Miura, T., & Cho, A. K. (2012). The chemical biology of naphthoquinones and its environmental implications. *Annual Review of Pharmacology and Toxicology, 52*.

Kumar, P., Gupta, V., Misra, A., Modi, D., & Pandey, B. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics, 2*(4), 141.

Lau, F. C., Bielinski, D. F., & Joseph, J. A. (2007). Inhibitory effects of blueberry extract on the production of inflammatory mediators in lipopolysaccharide-activated BV2 microglia. *Journal of Neuroscience Research, 85*(5), 1010-1017.

Levi, A., & Rowland, L. (1997). Identifying blueberry cultivars and evaluating their genetic relationships using randomly amplified polymorphic DNA (RAPD) and simple sequence

repeat-(SSR-) anchored primers. *Journal of the American Society for Horticultural Science, 122*(1), 74-78.

Li, B., Kadura, I., Fu, D.-J., & Watson, D. E. (2004). Genotyping with TaqMAMA. *Genomics, 83*(2), 311-320.

Li, D., Li, B., Ma, Y., Sun, X., Lin, Y., & Meng, X. (2017). Polyphenols, anthocyanins, and flavonoids contents and the antioxidant capacity of various cultivars of highbush and half-high blueberries. *Journal of Food Composition and Analysis, 62*, 84-93.

Li, G., & Quiros, C. F. (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. *Theoretical and Applied Genetics, 103*(2-3), 455-461.

Li, W., & Pauluhn, J. (2017). Phosgene-induced acute lung injury (ALI): differences from chlorine-induced ALI and attempts to translate toxicology to clinical medicine. *Clinical and translational medicine, 6*(1), 19.

Lin, Y.-H., Arashiro, M., Clapp, P. W., Cui, T., Sexton, K. G., Vizuete, W., . . . Surratt, J. D. (2017). Gene Expression Profiling in Human Lung Cells Exposed to Isoprene-Derived Secondary Organic Aerosol. *Environmental Science & Technology, 51*(14), 8166-8175.

Lin, Y.-L., Juan, I.-M., Chen, Y.-L., Liang, Y.-C., & Lin, J.-K. (1996). Composition of polyphenols in fresh tea leaves and associations of their oxygen-radical-absorbing capacity with antiproliferative actions in fibroblast cells. *Journal of Agricultural and Food Chemistry, 44*(6), 1387-1394.

Liu, B.-H. (1998). *Statistical genomics: linkage, mapping, and QTL analysis* (0849331668). Retrieved from

Liu, J., Zhang, W., Jing, H., & Popovich, D. G. (2010). Bog bilberry (Vaccinium uliginosum L.) extract reduces cultured Hep-G2, Caco-2, and 3T3-L1 cell viability, affects cell cycle

progression, and has variable effects on membrane permeability. *Journal of Food Science, 75*(3), H103-H107.

Liu, Y.-c., Liu, S., Liu, D.-c., Wei, Y.-x., Liu, C., Yang, Y.-m., . . . Liu, W.-s. (2014). Exploiting EST databases for the development and characterization of EST-SSR markers in blueberry (Vaccinium) and their cross-species transferability in Vaccinium spp. *Scientia Horticulturae, 176*, 319-329.

Longley, A. (1927). Chromosomes in Vaccinium. *Science, 66*(1719), 566-568.

Louati, M., Ucarli, C., Arikan, B., Ghada, B., Salhi Hannachi, A., & Turgut-Kara, N. (2019). Genetic, Morphological, and Biochemical Diversity of Argan Tree (Argania spinosa L.)(Sapotaceae) in Tunisia. *Plants, 8*(9), 319.

Luby, J., Wildung, D., Stushnoff, C., Munson, S., Read, P., & Hoover, E. (1986). 'Northblue','Northsky', and'Northcountry'blueberries. *HortScience (USA)*.

Lushchak, V. I. (2014). Free radicals, reactive oxygen species, oxidative stress and its classification. *Chemico-Biological Interactions, 224*, 164-175.

Määttä-Riihinen, K. R., Kamal-Eldin, A., Mattila, P. H., González-Paramás, A. M., & Törrönen, A. R. (2004). Distribution and contents of phenolic compounds in eighteen Scandinavian berry species. *Journal of Agricultural and Food Chemistry, 52*(14), 4477-4486.

Maeda, H., & Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annual Review of Plant Biology, 63*, 73-105.

Mahboob, M., Shireen, K. F., Atkinson, A., & Khan, A. T. (2001). Lipid peroxidation and antioxidant enzyme activity in different organs of mice exposed to low level of mercury. *Journal of Environmental Science and Health, Part B, 36*(5), 687-697.

Malmgren, H., Gustavsson, J., Tuvemo, T., & Dahl, N. (1996). Rapid detection of a mutation hot-spot in the human androgen receptor. *Clinical Genetics, 50*(4), 202-205.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research, 27*(2 Part 1), 209-220.

Markert, C. L., & Møller, F. (1959). Multiple forms of enzymes: tissue, ontogenetic, and species specific patterns. *Proceedings of the National Academy of Sciences, 45*(5), 753-763.

Massarotto, G., Barcellos, T., Garcia, C. S. C., Brandalize, A. P. C., Moura, S., Schwambach, J., . . . Roesch-Ely, M. (2016). Chemical characterization and cytotoxic activity of blueberry extracts (cv. Misty) cultivated in Brazil. *Journal of Food Science, 81*(8), H2076-H2084.

Matchett, M. D., MacKinnon, S. L., Sweeney, M. I., Gottschall-Pass, K. T., & Hurta, R. A. (2005). Blueberry flavonoids inhibit matrix metalloproteinase activity in DU145 human prostate cancer cells. *Biochemistry and Cell Biology, 83*(5), 637-643.

Mattila, P., Hellström, J., & Törrönen, R. (2006). Phenolic acids in berries, fruits, and beverages. *Journal of Agricultural and Food Chemistry, 54*(19), 7193-7199.

Mazza, G., Kay, C. D., Cottrell, T., & Holub, B. J. (2002). Absorption of anthocyanins from blueberries and serum antioxidant status in human subjects. *Journal of Agricultural and Food Chemistry, 50*(26), 7731-7737.

McKight, P. E., & Najab, J. (2010). Kruskal-wallis test. *The corsini encyclopedia of psychology*, 1-1.

Meerow, A. W. (2003). *Molecular genetic characterization of new floricultural germplasm.* Paper presented at the V International Symposium on New Floricultural Crops 683.

Mei, G., Hui, X., Xin, W., & Long, H. (2016). Fingerprinting construction and genetic relationship analysis of blueberry based on EST-SSR markers. *Guizhou Agricultural Sciences, 44*(4), 110-115.

Meyer, W., Mitchell, T. G., Freedman, E., & Vilgalys, R. (1993). Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction

to distinguish strains of Cryptococcus neoformans. *Journal of Clinical Microbiology, 31*(9), 2274-2280.

Mishra, K., Ojha, H., & Chaudhury, N. K. (2012). Estimation of antiradical properties of antioxidants using DPPH assay: A critical review and results. *Food Chemistry, 130*(4), 1036-1043.

Moghe, A., Ghare, S., Lamoreau, B., Mohammad, M., Barve, S., McClain, C., & Joshi-Barve, S. (2015). Molecular mechanisms of acrolein toxicity: relevance to human disease. *Toxicological Sciences, 143*(2), 242-255.

Mohammadi, S., & Prasanna, B. (2003). Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Science, 43*(4), 1235-1248.

Mohr, N., & Kevan, P. (1987). *Pollinators and pollination requirements of lowbush blueberry (Vaccinium angustifolium Ait. and V. myrtilloides Michx.) and cranberry (V. macrocarpon Ait.) in Ontario with notes on highbush blueberry (V. corymbosum L.) and lignonberry (V. vitis-idaea L.).* Paper presented at the Proceedings of the Entomological Society of Ontario.

Moyer, R. A., Hummer, K. E., Finn, C. E., Frei, B., & Wrolstad, R. E. (2002). Anthocyanins, phenolics, and antioxidant capacity in diverse small fruits: Vaccinium, Rubus, and Ribes. *Journal of Agricultural and Food Chemistry, 50*(3), 519-525.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986). *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Paper presented at the Cold Spring Harbor Symposia on Quantitative Biology.

Naczk, M., & Shahidi, F. (2006). Phenolics in cereals, fruits and vegetables: Occurrence, extraction and analysis. *Journal of Pharmaceutical and Biomedical Analysis, 41*(5), 1523-1542.

Nei, M. (1973a). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA, 70*(12), 3321-3323.

Nei, M. (1973b). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, 70*(12), 3321-3323.

Nurmi, J., Kiviniemi, M., Kujanpää, M., Sjöroos, M., Ilonen, J., & Lövgren, T. (2001). High-throughput genetic analysis using time-resolved fluorometry and closed-tube detection. *Analytical Biochemistry, 299*(2), 211-217.

Ofek, I., Goldhar, J., Zafriri, D., Lis, H., Adar, R., & Sharon, N. (1991). Anti-Escherichia coli adhesin activity of cranberry and blueberry juices. *The New England journal of medicine, 324*(22), 1599-1599.

Okie, W. (1997). Register of New Fruit and Nut Varieties Brooks and Olmo List 38. *HortScience, 32*(5), 785-805.

Olsen, K. M., Hehn, A., Jugdé, H., Slimestad, R., Larbat, R., Bourgaud, F., & Lillo, C. (2010). Identification and characterisation of CYP75A31, a new flavonoid 3'5'-hydroxylase, isolated from Solanum lycopersicum. *BMC Plant Biology, 10*(1), 21.

Olson, M., Hood, L., Cantor, C., & Botstein, D. (1989). A common language for physical mapping of the human genome. *Science, 245*(4925), 1434-1435.

Omri, A., Abdelhamid, S., Benincasa, C., Araouki, A., Ayadi, M., Gharsallaoui, M., & Gouiaa, M. (2020). Genetic diversity and association of molecular markers with biochemical traits in Tunisian olive cultivars. *Genetic Resources and Crop Evolution*, 1-17.

Orlóci, L. (1978). *Multivariate analysis in vegetation research*. The Hauge: Dr W. Junk BV.

Osawa, T. (1994). Novel natural antioxidants for utilization in food and biological systems. In *Postharvest biochemistry of plant food-materials in the tropics* (Vol. 1, pp. 241-251). Tokyo, Japan: Japan Scientific Press.

Pairment, K. (1995). *Canadian biodiversity strategy: Canada's response to the Convention on Biological Diversity*. Ottawa/Hull (NCR): Environment Bureau, Agriculture and Agri-Food Canada.

Paran, I., & Michelmore, R. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Tag Theoretical and Applied Genetics, 85*(8), 985-993.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics, 28*(19), 2537.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Resources, 6*(1), 288-295.

Percival, D., & MacKenzie, J. L. (2007). Use of plant growth regulators to increase polyphenolic compounds in the wild blueberry. *Canadian journal of plant science, 87*(2), 333-336.

Perl-Treves, R., & Perl, A. (2002). Oxidative stress: an introduction. *Oxidative stress in plants*, 1-32.

Perrier, X., & Jacquemoud-Collet, J. (2006). DARwin software. France: CIRAD. Retrieved from http://darwin.cirad.fr/

Persson, H., & Gustavsson, B. (2001). The extent of clonality and genetic diversity in lingonberry (Vaccinium vitis-idaea L.) revealed by RAPDs and leaf-shape analysis. *Molecular Ecology, 10*(6), 1385-1397.

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.

Prevost, A., & Wilkinson, M. (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics, 98*(1), 107-112.

Pridham, J. B. (1960). Phenolics in plants in health and disease. *Phenolics in plants in health and disease.*

Primmer, C., Borge, T., Lindell, J., & Sætre, G. P. (2002). Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Molecular Ecology, 11*(3), 603-612.

Prior, R. L., Cao, G., Martin, A., Sofic, E., McEwen, J., O'Brien, C., . . . Krewer, G. (1998). Antioxidant capacity as influenced by total phenolic and anthocyanin content, maturity, and variety of Vaccinium species. *Journal of Agricultural and Food Chemistry, 46*(7), 2686-2693.

Prior, R. L., Wu, X., Gu, L., Hager, T., Hager, A., Wilkes, S., & Howard, L. (2009). Purified berry anthocyanins but not whole berries normalize lipid parameters in mice fed an obesogenic high fat diet. *Molecular Nutrition & Food Research, 53*(11), 1406-1418.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945-959.

Pusch, W., Wurmbach, J.-H., Thiele, H., & Kostrzewa, M. (2002). MALDI-TOF mass spectrometry-based SNP genotyping. *Pharmacogenomics, 3*(4), 537-548.

Rao, V. R., & Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture, 68*(1), 1-19.

Ratnaparkhe, M. B. (2007). Blueberry. In *Fruits and Nuts* (pp. 217-227).

Reddy, M. P., Sarla, N., & Siddiq, E. (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica, 128*(1), 9-17.

Rice-Evans, C. A., Miller, N. J., & Paganga, G. (1996). Structure-antioxidant activity relationships of flavonoids and phenolic acids. *Free Radical Biology and Medicine, 20*(7), 933-956.

Richard, I., & Beckmann, J. S. (1995). How neutral are synonymous codon mutations? *Nature Genetics, 10*(3), 259.

Richards, P., Ruivenkamp, G., & van der Drift, R. (1997). *Seeds and survival: Crop genetic resources in war and reconstruction in Africa*: Bioversity International.

Riihinen, K., Jaakola, L., Kärenlampi, S., & Hohtola, A. (2008). Organ-specific distribution of phenolic compounds in bilberry (Vaccinium myrtillus) and 'northblue'blueberry (Vaccinium corymbosum x V. angustifolium). *Food Chemistry, 110*(1), 156-160.

Ross, P., Hall, L., Smirnov, I., & Haff, L. (1998). High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nature Biotechnology, 16*(13), 1347.

Routray, W., & Orsat, V. (2011). Blueberries and their anthocyanins: factors affecting biosynthesis and properties. *Comprehensive Reviews in Food Science and Food Safety, 10*(6), 303-320.

Rowland, Dhanaraj, A. L., Polashock, J. J., & Arora, R. (2003a). Utility of blueberry-derived EST-PCR primers in related Ericaceae species. *HortScience, 38*(7), 1428-1432.

Rowland, L., Mehra, S., Dhanraj, A. L., Ogden, E. L., Sloven, J. P., & Ehlenfeldt, M. K. (2003b). Development of EST-PCR markers for DNA fingerprinting and genetic relationship studies in blueberry (Vaccinium, section Cyanococcus). *Journal of the American Society for Horticultural Science, 128*(5), 682-690.

Rowland, L. J., Mehra, S., Dhanaraj, A. L., Ogden, E. L., Slovin, J. P., & Ehlenfeldt, M. K. (2003c). Development of EST-PCR markers for DNA fingerprinting and genetic relationship studies in blueberry (Vaccinium, section Cyanococcus). *Journal of the American Society for Horticultural Science, 128*(5), 682-690.

Rowland, L. J., Ogden, E. L., Bassil, N., Buck, E. J., McCallum, S., Graham, J., . . . Haynes, K. G. (2014). Construction of a genetic linkage map of an interspecific diploid blueberry

population and identification of QTL for chilling requirement and cold hardiness. *Molecular Breeding, 34*(4), 2033-2048.

Rowland, L. J., Ogden, E. L., & Ehlenfeldt, M. K. (2010). EST-PCR markers developed for highbush blueberry are also useful for genetic fingerprinting and relationship studies in rabbiteye blueberry. *Scientia Horticulturae, 125*(4), 779-784.

Ruan, C.-J., Li, H., & Mopper, S. (2009). Characterization and identification of ISSR markers associated with resistance to dried-shrink disease in sea buckthorn. *Molecular Breeding, 24*(3), 255-268.

Russell, H. S. (1976). *long, deep furrow*: University Press of New England.

Ryabova, D. (2007). Population evaluation in crop wild relatives for in situ conservation: a case study for raspberry Rubus idaeus L. in the Leningrad region, Russia. *Genetic Resources and Crop Evolution, 54*(5), 973-980.

Samad, N. B., Debnath, T., Ye, M., Hasnat, M. A., & Lim, B. O. (2014). In vitro antioxidant and anti–inflammatory activities of Korean blueberry (Vaccinium corymbosum L.) extracts. *Asian Pacific Journal of Tropical Biomedicine, 4*(10), 807-815.

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.

Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. *Genetics, 8*(6), 552.

Schantz, M., Mohn, C., Baum, M., & Richling, E. (2010). Antioxidative efficiency of an anthocyanin rich bilberry extract in the human colon tumor cell lines Caco-2 and HT-29. *Journal of berry research, 1*(1), 25-33.

Schlötterer, C., & Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research, 20*(2), 211-215.

Schmidt, B. M., Erdman, J. W., & Lila, M. A. (2006). Differential effects of blueberry proanthocyanidins on androgen sensitive and insensitive human prostate cancer cell lines. *Cancer Letters, 231*(2), 240-246.

Sellappan, S., Akoh, C. C., & Krewer, G. (2002). Phenolic compounds and antioxidant capacity of Georgia-grown blueberries and blackberries. *Journal of Agricultural and Food Chemistry, 50*(8), 2432-2438.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, Illinois: University of Illinois Press.

Shen, C.-L., von Bergen, V., Chyu, M.-C., Jenkins, M. R., Mo, H., Chen, C.-H., & Kwun, I.-S. (2012). Fruits and dietary phytochemicals in bone protection. *Nutrition Research, 32*(12), 897-910.

Shukitt-Hale, B., Lau, F. C., & Joseph, J. A. (2008). Berry fruit supplementation and the aging brain. *Journal of Agricultural and Food Chemistry, 56*(3), 636-641.

Singleton, V. L., & Rossi, J. A. (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *American Journal of Enology and Viticulture, 16*(3), 144-158.

Skrovankova, S., Sumczynski, D., Mlcek, J., Jurikova, T., & Sochor, J. (2015). Bioactive compounds and antioxidant activity in different types of berries. *International Journal of Molecular Sciences, 16*(10), 24673-24706.

Smith, D. N., & Devey, M. E. (1994). Occurrence and inheritance of microsatellites in Pinus radiata. *Genome, 37*(6), 977-983.

Sneath, P. H., & Sokal, R. R. (1973). *Numerical taxonomy. The principles and practice of numerical classification*.

Sobrino, B., Bríon, M., & Carracedo, A. (2005). SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Science International, 154*(2-3), 181-194.

Soleimani, V., Baum, B., & Johnson, D. (2003). Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley. *Plant Molecular Biology Reporter, 21*(3), 281-288.

Song, G.-Q., & Hancock, J. F. (2011). Vaccinium. In *Wild crop relatives: Genomic and breeding resources* (pp. 197-221): Springer.

Sorrentino, R., Potolicchio, I., Ferrara, G. B., & Tosi, R. (1992). A new approach to HLA-DPB1 typing combining DNA heteroduplex analysis with allele-specific amplification and enzyme restriction. *Immunogenetics, 36*(4), 248-254.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology, 98*(3), 503-517.

Srivastava, A., Akoh, C. C., Fischer, J., & Krewer, G. (2007). Effect of anthocyanin fractions from selected cultivars of Georgia-grown blueberries on apoptosis and phase II enzymes. *Journal of Agricultural and Food Chemistry, 55*(8), 3180-3185.

Statistics Canada. (2020). *Statistical Overview of the Canadian Fruit Industry 2019*. Retrieved from Canada:

Staub, J. E., Serquen, F. C., & Gupta, M. (1996). Genetic markers, map construction, and their application in plant breeding. *HortScience, 31*(5), 729-741.

Storm, N., & Darnhofer-Patel, B. (2003). MALDI-TOF mass spectrometry-based SNP genotyping. In *Single Nucleotide Polymorphisms* (pp. 241-262): Springer.

Strik, B. (2005). Blueberry: an expanding world berry crop. *Chronica Horticulturae, 45*(1), 7-12.

Strik, B. C., & Yarborough, D. (2005). Blueberry production trends in North America, 1992 to 2003, and predictions for growth. *HortTechnology, 15*(2), 391-398.

Stull, A. J., Cash, K. C., Johnson, W. D., Champagne, C. M., & Cefalu, W. T. (2010). Bioactives in Blueberries Improve Insulin Sensitivity in Obese, Insulin-Resistant Men and Women–4. *The Journal of nutrition, 140*(10), 1764-1768.

Sunyaev, S., Hanke, J., Aydin, A., Wirkner, U., Zastrow, I., Reich, J., & Bork, P. (1999). Prediction of nonsynonymous single nucleotide polymorphisms in human disease-associated genes. *Journal of Molecular Medicine, 77*(11), 754-760.

Suzuki, Y., Orita, M., Shiraishi, M., Hayashi, K., & Sekiya, T. (1990). Detection of ras gene mutations in human lung cancers by single-strand conformation polymorphism analysis of polymerase chain reaction products. *Oncogene, 5*(7), 1037-1043.

Sweeney, M., Kalt, W., MacKinnon, S., Ashby, J., & Gottschall-Pass, K. T. (2002). Feeding rats diets enriched in lowbush blueberries for six weeks decreases ischemia-induced brain damage. *Nutritional Neuroscience, 5*(6), 427-431.

Tailor, S., Bykova, N. V., Igamberdiev, A. U., & Debnath, S. C. (2017). Structural pattern and genetic diversity in blueberry (Vaccinium) clones and cultivars using EST-PCR and microsatellite markers. *Genetic Resources and Crop Evolution, 64*(8), 2071-2082.

Taiz, L., & Zeiger, E. (2006). *Plant Physiology* (4th ed.): Sinauer Associates Inc., Sunderland.

Tanksley, S., & Orton, T. (1983). ISOZYMIC VARIATION AND PLANT BREEDERS'RIGHTS. *Isozymes in plant genetics and breeding, 1*, 425.

Taruscio, T. G., Barney, D. L., & Exon, J. (2004). Content and profile of flavanoid and phenolic acid compounds in conjunction with the antioxidant capacity for a variety of northwest Vaccinium berries. *Journal of Agricultural and Food Chemistry, 52*(10), 3169-3176.

Tautz, D., & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research, 12*(10), 4127-4138.

Tessier, C., David, J., This, P., Boursiquot, J.-M., & Charrier, A. (1999). Optimization of the choice of molecular markers for varietal identification in Vitis vinifera L. *Theoretical and Applied Genetics, 98*(1), 171-177.

The International SNP Map Working Group. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature, 409*(6822), 928.

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

Torres, A., Weeden, N., & Martin, A. (1993). Linkage among isozyme, RFLP and RAPD markers in Vicia faba. *Theoretical and Applied Genetics, 85*(8), 937-945.

Turchetto, C., Segatto, A. L. A., Beduschi, J., Bonatto, S. L., & Freitas, L. B. (2015). Genetic differentiation and hybrid identification using microsatellite markers in closely related wild species. *AoB PLANTS, 7*. doi:10.1093/aobpla/plv084

Vander Kloet, S., & Dickinson, T. A. (1999). The taxonomy of Vaccinium section Myrtillus (Ericaceae). *Brittonia, 51*(2), 231-254.

Vander Kloet, S. P. (1988). *The genus Vaccinium in North America*: Agriculture Canada.

Varshney, R. K., Graner, A., & Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology, 23*(1), 48-55. doi:10.1016/j.tibtech.2004.11.005

Vavilov, N. I. (1926). Studies on the origin of cultivated plants.

Vaziri, N. D. (2008). Mechanisms of lead-induced hypertension and cardiovascular disease. *American Journal of Physiology-Heart and Circulatory Physiology, 295*(2), H454-H465.

Vendrame, S., Del Bo, C., Ciappellano, S., Riso, P., & Klimis-Zacas, D. (2016). Berry fruit consumption and metabolic syndrome. *Antioxidants, 5*(4), 34.

Vorsa, N., Manos, P. S., & Heemstra, M. I. v. (1988). Isozyme variation and inheritance in blueberry. *Genome, 30*(5), 776-781.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. v. d., Hornes, M., . . . Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research, 23*(21), 4407-4414.

Vosman, B., Arens, P., Rus-Kortekaas, W., & Smulders, M. (1992). Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics, 85*(2-3), 239-244.

Wang, H., Cao, G., & Prior, R. L. (1996). Total antioxidant capacity of fruits. *Journal of Agricultural and Food Chemistry, 44*(3), 701-705.

Wang, L.-J., Wu, J., Wang, H.-X., Li, S.-S., Zheng, X.-C., Du, H., . . . Wang, L.-S. (2015). Composition of phenolic compounds and antioxidant activity in the leaves of blueberry cultivars. *Journal of Functional Foods, 16*, 295-304.

Wang, S. (2007). Antioxidant capacity and phenolic content of berry fruits as affected by genotype, preharvest conditions, maturity, and postharvest handling. *FOOD SCIENCE AND TECHNOLOGY-NEW YORK-MARCEL DEKKER-, 168*, 147.

Wang, S. Y., & Lin, H.-S. (2000). Antioxidant activity in fruits and leaves of blackberry, raspberry, and strawberry varies with cultivar and developmental stage. *Journal of Agricultural and Food Chemistry, 48*(2), 140-146.

Weising, K., Nybom, H., Wolff, K., & Meyer, W. (1995). DNA Fingerprinting in Plants and Fungi. CRC, Press. *Inc., Boca Raton, USA*.

Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research, 18*(24), 7213-7218.

Westermeier, R. (2016). *Electrophoresis in practice: a guide to methods and applications of DNA and protein separations*: John Wiley & Sons.

Wexelsen, H. (1933). Linkage of a quantitative and a qualitative character in barley. *Hereditas, 17*(3), 323-341.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research, 18*(22), 6531-6535.

Winder, C. (2001). The toxicology of chlorine. *Environmental Research, 85*(2), 105-114.

Winkel-Shirley, B. (2002). Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology, 5*(3), 218-223.

Witsenboer, H., Michelmore, R., & Vogel, J. (1997). Identification, genetic localization, and allelic diversity of selectively amplified microsatellite polymorphic loci in lettuce and wild relatives (Lactuca spp.). *Genome, 40*(6), 923-936.

Wu, K., Jones, R., Danneberger, L., & Scolnik, P. A. (1994). Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research, 22*(15), 3257.

Wu, W., Bromberg, P. A., & Samet, J. M. (2013). Zinc ions as effectors of environmental oxidative lung injury. *Free Radical Biology and Medicine, 65*, 57-69.

Wu, X., Cao, G., & Prior, R. L. (2002). Absorption and metabolism of anthocyanins in elderly women after consumption of elderberry or blueberry. *The Journal of nutrition, 132*(7), 1865-1871.

Wu, X., Kang, J., Xie, C., Burris, R., Ferguson, M. E., Badger, T. M., & Nagarajan, S. (2010). Dietary Blueberries Attenuate Atherosclerosis in Apolipoprotein E-Deficient Mice by Upregulating Antioxidant Enzyme Expression–4. *The Journal of nutrition, 140*(9), 1628-1632.

Yang, S., Bishop, J., & Webster, M. (2008). Colonization genetics of an animal-dispersed plant (Vaccinium membranaceum) at Mount St Helens, Washington. *Molecular Ecology*, *17*(3), 731-740.

Yu, B. P. (1994). Cellular defenses against damage from reactive oxygen species. *Physiological Reviews*, *74*(1), 139-162.

Yun, J. M., Afaq, F., Khan, N., & Mukhtar, H. (2009). Delphinidin, an anthocyanidin in pigmented fruits and vegetables, induces apoptosis and cell cycle arrest in human colon cancer HCT116 cells. *Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center*, *48*(3), 260-270.

Zhishen, J., Mengcheng, T., & Jianming, W. (1999). The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals. *Food Chemistry*, *64*(4), 555-559.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, *20*(2), 176-183.