CRANBERRY METABOLOMICS:

NEW APPROACHES FOR PHYTOCHEMICAL CHARACTERIZATIONS

by

Paula Naomi Brown

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE COLLEGE OF GRADUATE STUDIES

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

December 2011

Abstract

There is a long history of use and modern commercial importance of large (*Vaccinium macrocarpon* Aiton) and small cranberries (*V. oxycoccus* L. and *V. vitis-idaea* L.) in North America. While epidemiological research indicates cranberries have positive health benefits, identifying specific phytochemicals for disease prevention remains elusive. The central objective of this research was to develop phytochemical characterization tools for comparing commercially cultivated cranberries and two wild-harvested *Vaccinium s*pecies. A method was developed and validated to quantify cyanidin-3-O-galactoside (C3Ga), cyanidin-3-O-glucoside (C3Gl), cyanidin-3-O-arabinoside (C3Ar), peonidin-3-O-galactoside (P3Ga) and peonidin-3-O-arabinoside (P3Ar) in cranberry fruit products. The relative standard deviation (%RSDr) of anthocyanins ranged from 1.77% to 3.31% with the method detection limit (MDL) for C3Ga, C3Gl, C3Ar, P3Ga, and P3Ar estimated at 0.018, 0.016, 0.006, 0.013, and 0.011 µg/mL, respectively. In biological replicates of *V. macrocarpon* cultivars, anthocyanin content (mg/g) was determined to be 7.98±5.83 in Ben Lear; 7.02±1.75 in Bergman; 6.05± 2.51 in GH1; 3.28± 1.88 in Pilgrim and 2.81 ±0.81 in Stevens. *V. oxycoccus* contained the five major anthocyanins with the ratio of glycosylated peonidins to cyanidins varying from *V. macrocarpon*. *V. vitis-idaea* contained no measurable glycosylated peonidins. Determination and statistical analysis of serotonin, melatonin and ascorbic acid content in the three *Vaccinium* species, found only vitamin C positively correlated with antioxidant activity. Metabolome comparisons made from untargeted metabolomics by ultra-fast liquid chromatography (UFLC) with time-of-flight (TOF) mass spectrometry (MS), found different pools of common metabolites extracted in methanol (7635), 70% ethanol (5832), and water (5664). Each cranberry matrix contained unique compounds; fruit=3680, product 1=3464,

product 2=3500. Clustering was visualized using principal component and partial least squares discriminate analyses with application of univariate statistics to mitigate false discoveries. A significance algorithm found 1987, 716 and 1556 compounds (p-value<0.05) in methanol, 70% ethanol and water extracts, respectively. In metabolomic profiles of *V. macrocarpon* varietals 6481 compounds were conserved and 136 (Ben Lear), 84 (Bergman), 91 (GH1), 128 (Pilgrim) and 165 (Stevens) unique observations were made. Across the three *Vaccinium* species 8000-10,000 phytochemicals were detected with 4624 compounds conserved between them. Together these data establish targeted and untargeted methods for phytochemical characterization of cranberries, providing foundational chemotaxonomic knowledge and new insights into the maintenance of health in traditional North American diets.

Preface

The work contained in this thesis was designed, conducted, analyzed and interpreted by Paula N. Brown. Collections of *Vaccinium macrocarpon* Aiton were made by Paula N. Brown in collaboration with Ocean Spray Canada Ltd. in the lower Mainland of British Columbia. The collections of *Vaccinium vitis-idaea* L. and *Vaccinium oxycoccus* L. were made at Haida Gwaii, British Columbia, Canada, with assistance from the local community and especially Mr. Tim Wolver. Versions of several of the chapters are either published, submitted to or in preparation to submit to scientific journals for publication as below:

- A version of Chapter 2 has been published as Brown PN, Shipley PR (2011) Single laboratory validation of a method to determine anthocyanins in cranberry fruit and cranberry fruit products by high performance liquid chromatography with ultraviolet detection. *Journal of the AOAC International* 94**(2):** 459-466. Paula Brown designed the method validation study, conducted all statistical analysis of the data and wrote the manuscript with Dr. Paul Shipley.
- Chapter 3 is based on work conducted on *Vaccinium macrocarpon* Aiton samples provided by Ocean Spray Canada Ltd. from their processing station in Richmond, BC. This manuscript has been submitted for publication as follows; Brown PN, Shipley PR, Murch SJ (2011) Anthocyanins diversity within a commercial production of cranberries (*Vaccinium macrocarpon* Aiton) in British Columbia. *Phytochemical Analysis*, submitted on July 18th, 2011.
- The work presented in Chapter 4 has been prepared as a manuscript and submitted for publication. Brown PN, Finley JP, Le CH, Murch SJ (2011) Metabolomic Profiling of Cranberry (*Vaccinium Macrocarpon* Aiton) Fruit and Cranberry Dietary Supplements by

Time-of-Flight Mass Spectrometry with Chemometric Data Analysis. *Analytical and Bioanalytical Chemistr*y, submitted July 5th, 2011.

- For Chapter 5, comparative phytochemistry of five varietals of *Vaccinium macrocarpon* Aiton has been prepared for submission as follows; Brown PN, Murch SJ, Shipley PR (2011) Phytochemical diversity in five varietals of cranberries (*Vaccinium macrocarpon* Aiton) cultivated in British Columbia. *Journal of Agriculture and Food Chemistry*, in preparation.
- A version of Chapter 6 was presented by Paula Brown in the Economic Botany Session of Botany 2011 on July 11th, 2011 in St. Louis, Missouri, USA. This work has been prepared as Brown PN, Turi CE, Shipley PR, Murch SJ (2011) Phytochemical Discovery in Large Cranberry (*Vaccinium macrocarpon* Ait.) and Small Cranberry (*Vaccinium oxycoccus* L. and *Vaccinium vitis-idaea* L.) in British Columbia. *Planta Medica*, submitted August 3rd, 2011.

Table of Contents

Abstract ..... **ii**

Preface ..... **iv**

Table of Contents ..... **vi**

List of Tables ..... **xii**

List of Figures ..... **xv**

Acknowledgements ..... **xix**

Chapter 1: Introduction **1**

Chapter 2: Single Laboratory Validation of a Method to Determine Anthocyanins in Cranberry Fruit Products by High Performance Liquid Chromatography with Ultraviolet Detection **15**

Synopsis 15

Experimental 16

Principle 16

Test Samples 16

Reagents and Supplies 17

Apparatus 19

LC Conditions 20

Calculations 22

Preparation of Test Samples 23

A. Freeze Dried Raw Material and Powdered Extract: 23

B. Juice: 23

C. Juice Cocktail: 24

Preparation of Calibration Solutions 24
Single Laboratory Validation Parameters 25
Selectivity 25
Linearity 26
Method Detection Level and Limit of Quantification 26
Precision (Repeatability) 27
Accuracy 27
Stability Study – Reference Standard and Test Material Solution 28
Youden Ruggedness Trial: A Factorial Study 28
Results and Discussion 29
Method Validation Results – Performance Characteristics 29
Selectivity 29
Linearity and Analytical Range 30
Limits of Detection and Limits of Quantification 30
Precision (Repeatability) 30
Accuracy 32
Stability Study – Reference Standard and Test Material Solution 32
Youden Ruggedness Trial: A Factorial Study 32
Chapter 3: Anthocyanin Diversity within a Commercial Production of Cranberries (*Vaccinium macrocarpon* Aiton) in British Columbia 35
Synopsis 35
Experimental 36
Collection of Cranberry Samples 36

Sample Preparation 36
Chemical Calibration Standards 38
Chromatographic Conditions 39
Anthocyanin Quantification and Statistical Analysis 40
Results 40
Cranberry Production 40
Chemical Reference Materials and Calibration 41
Quantification of Anthocyanins in Cranberry 42
Discussion 47
Chapter 4: Metabolomic Profiling of Cranberry (*Vaccinium macrocarpon* Aiton) Fruit and Cranberry Dietary Supplements by Time-of-Flight Mass Spectrometry with Chemometric Data Analysis 50
Synopsis 50
Experimental 51
Test Materials 51
Metabolite Solvent Extraction Study 51
Ultra-Fast Liquid Chromatography with Time-of-Flight Mass Spectrometry 52
Untargeted Exploratory Data Analysis 53
Metabolite Counts & Subtractive Metabonomics 53
Data Distribution 55
Principal Component Analysis (PCA) 55
Significance Analysis 56
AUC and Kruksal-Wallis P-Value 56

Percent Explained Variance 57
Significance Analysis of Microarrays (SAM) Statistic 57
Results and Discussion 58
Untargeted Exploratory Data Analysis 58
Metabolite Counts & Subtractive Metabonomics 58
Data Distribution 59
Principal Component Analysis (PCA) 59
Significance Analysis 62
AUC and Kruksal-Wallis P-Value 62
Percent Explained Variance 66
SAM Statistic 67
Chapter 5: Phytochemical Diversity of Cranberry (*Vaccinium macrocarpon* Aiton) Cultivars by Anthocyanin Determination and Metabolomic Profiling with Chemometric Analysis 74
Synopsis 74
Materials and Methods 77
Plant Materials 77
Chemicals and Calibration Standards 78
Sample Preparation 78
Anthocyanin Determination by UFLC-DAD 79
Metabolomic Profiling by UPLC-TOF-MS 80
Data Processing and Subtractive Metabonomics 81
Significance Analysis 82

Multivariate Analysis 83
Results and Discussion 84
Determination of Anthocyanins in Cranberry Varietals 84
Metabolomic Profiling by UPLC-TOF-MS 88
Chemometric Analysis of Metabolomic Data 89
Chapter 6: Phytochemical Discovery in Large Cranberry (*Vaccinium macrocarpon* Aiton), Small Cranberry (*Vaccinium oxycoccus*) and Lingonberry (*Vaccinium vitis-idaea*) 99
Synopsis 99
Materials and Methods 100
Plant Materials 100
Reagents and Standards 101
Anthocyanin Determination by HPLC-DAD 102
Melatonin and Serotonin Determination 103
Determination of Free Radical-Scavenging Capacity 104
Ascorbic Acid Determination 105
Statistical Analysis 106
Metabolomic Profiling by UFLC-TOF-MS 106
Exploratory Data Analysis 106
Significance Analysis 107
Results 108
Discussion 122
Chapter 7: Conclusion 126

References **131**

Appendix 1: Supplementary Statistics for Chapter 5 ..... **146**

List of Tables

Table 2-1: Anthocyanin concentrations in mixed standard solution. 18

Table 2-2: Approximate concentrations (µg /mL) of individual anthocyanins at each linearity of the calibration curve 25

Table 2-3: Parameters used in the factorial study. 29

Table 2-4: The method detection limit (MDL) and limit of quantification (LOQ) calculated for each of the analytes used in this study. 30

Table 2-5: Precision results summary of *Vaccinium macrocarpon* test samples. 31

Table 2-6: Negative control spike recovery results. 32

Table 3-1: Calculated concentration (µg/mL) of anthocyanins, corrected for determined purity, at each linearity level of the calibration curve. 38

Table 4-1: Description of cranberry test materials employed in metabolomics study. 52

Table 4-2: Differences in metabolomics profiles of three different solvent extractions of cranberry fruit and two cranberry finished products. 60

Table 4-3: The m/z values of compounds identified as important based on a total AUC $\geq$ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in methanol. 64

Table 4-4: The m/z values of compounds identified as important based on a total AUC $\geq$ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in ethanol. 65

Table 4-5: The m/z values of compounds identified as important based on a total AUC $\geq$ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in water. 66

Table 4-6: The m/z values corresponding to the data with the highest % explained variance in the first two principal components when comparing all product types in each of the extraction solvents and the total AUC and *p*-values associated with each m/z value (variable). 68

Table 4-7: The m/z values identified as significant using the SAM algorithm in R to compare across products types in each extraction solvent with the calculated *p*-value, total AUC and % explained variance of the variable. 70

Table 5-1: Summary of LC-MS-TOF metabolite counts and differences in metabolomics profiles of five cultivars of *Vaccinium macrocarpon* Aiton. 89

Table 5-2: Summary of LC-MS-TOF metabolomic m/z values identified by the SAM statistic in order of ranked significance, *d(i)*, at a false discovery rate of 1.78% as well as the Q-residual values of the compounds in the PLS DA model for up to LV2 and up to LV4. 95

Table 5-3: The observed average abundance (n=5), total AUC and ranking by *d(i)* value of the m/z values identified by the SAM statistic as significant at a false discovery rate of 1.78% comparing across the five *V. macrocarpon* cultivars..... 97

Table 6-1: Summary of LC-MS-TOF metabolite counts and differences in metabolomic profiles of *V. macrocarpon, V. oxycoccus and V. vitis-idaea*..... 112

Table 6-2: The m/z values identified as significant and ranked by *d(i)* value, comparing across all three *Vaccinium* species, using the SAM algorithm in R at a false discover rate of 0.91% with the calculated Kruskal-Wallis *p*-value and total AUC. 119

Table A-1: ANOVA comparing cyanidin-3-*O*-glycosides content among cultivars using biological replicates 146

Table A-2: ANOVA comparing peonidin-3-*O*-glycosides content among cultivars using biological replicates 146

Table A-3: ANOVA comparing cyanidin-3-*O*-glycosides content among cultivars using analytical replicates..... 147

Table A-4: ANOVA comparing peonidin-3-*O*-glycosides content among cultivars using analytical replicates..... 147

Table A-5: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Ben Lear cultivar 148

Table A-6: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Bergman cultivar 148

Table A-7: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in GH1 cultivar 148

Table A-8: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Pilgrim cultivar..... 149

Table A-9: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Stevens cultivar 149

Table A-10: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Ben Lear cultivar..... 149

Table A-11: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Ben Lear cultivar 150

Table A-12: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in Ben Lear cultivar 150

Table A-13: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Bergman cultivar 150

Table A-14: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Bergman cultivar 151

Table A-15: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in Bergman cultivar 151

Table A-16: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in GH1 cultivar 151

Table A-17: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in GH1 cultivar 152

Table A-18: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in GH1 cultivar 152

Table A-19: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Pilgrim cultivar 152

Table A-20: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Pilgrim cultivar 153

Table A-21: Student *t*-tests comparing peonidin-3-*O*-galactoside content between analytical and biological replicates in Pilgrim cultivar 153

Table A-22: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Stevens cultivar 153

Table A-23: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Stevens cultivar 154

Table A-24: Student *t*-tests comparing peonidin-3-*O*-galactoside content between analytical and biological replicates in Stevens cultivar 154

List of Figures

Figure 1-1: Reports of cranberry in the scientific literature as indexed by Web of Science to April 28th, 2011. 2

Figure 1-2: Structures for the 5 anthocyanins peonidin-3-*O*-galactoside (P3Ga), cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-arabinoside (P3Ar), and cyanidin-3-*O*-glucoside (C3Gl). 10

Figure 2-1: The order of anthocyanin elution in a mixed anthocyanin reference standard. 21

Figure 2-2: Calculated effect of factors on total anthocyanins (mg/g) extracted from cranberry extract powder. Factor effects are as follows: J = sample mass; K = sonication time; C = % HCl in solvent; M = shaking time; N = sonicator temperature; O = injection volume; P = centrifugation time. 33

Figure 3-1: Wet-picked harvest of cranberry (*Vaccinium macrocarpon*) in the lower mainland region of British Columbia, Canada. 37

Figure 3-2: Chromatogram showing retention times of anthocyanins at 520 nm for a freeze-dried cranberry fruit sample and a mixed calibration standard. 42

Figure 3-3: Phyto-chemical diversity in samples of *Vaccinium macrocarpon* Aiton collected at mature harvest from farms across British Columbia. A. cyanidin-3-*O*-galactoside, B. cyanidin-3-*O*-arabinoside, C. peonidin-3-*O*-galactoside, D. peonidin-3-*O*-arabinoside. Results are expressed as mg/100mg ± standard error of the mean. 43

Figure 3-4: Variability in the total % of anthocyanins in cranberry (*Vaccinium macrocarpon* Aiton) berries collected in the fall of 2009 in British Columbia. 44

Figure 3-5: Comparison of the individual anthocyanins in (*Vaccinium macrocarpon* Aiton) berries collected in the fall of 2009 in British Columbia, Canada as variability in the concentration of individual anthocyanins in the berries and as variability in the percent of each anthocyanin relative to the total anthocyanin content ± standard error of the mean. 45

Figure 3-6: Relationships in the observed molar ratios of anthocyanins in cranberry fruit. A. molar ratio of cyanidin-containing anthocyanins to peonidin was 88%, B. molar ratio of peonidin glycosides with a galactose moiety compared to arabinose was 217%, C. molar ratio of cyanidin glycoside with a galactose moiety compared to arabinose was 128%. 46

Figure 4-1: Schematic of approach to creating subtractive datasets. 54

Figure 4-2: PCA score plots generated for UFLC-ToF-MS profiles of extracted freeze-dried cranberry fruit. Product 1 and Product 2 in methanol, 70%-aqueous ethanol and water, grouped by (a) product type and (b) extraction solvent. 61

Figure 5-1: Breeding map showing origins of the 5 cultivars of *Vaccinium macrocarpon* Aiton being studied in relation to other common cultivars, along with the estimated production percentages of each cultivar in BC. 76

Figure 5-2: Collection area (Left: orange pin) in Lower Mainland, British Columbia and specific sites of 5 cultivars of *Vaccinium macrocarpon* Aiton collected (R: yellow pins). 77

Figure 5-3: Average anthocyanin content (µg/g) in each cultivar for biological replicates (n=5) and extractions from a composite sample of cranberries (n=5) with error bars representing the standard error of the mean for (A) cyanidin-3-*O*-galactoside, (B) cyanidin-3-*O*-glucoside, (C) cyanidin-3-*O*-arabinoside, (D) peonidin-3-*O*-galactoside, (E) peonidin-3-*O*-arabinoside, and (F) the total anthocyanin content in mg/g of dry weight material (DW)..... 85

Figure 5-4: Comparison of the ratio of glycosylated peonidins to cyanidins in five *Vaccinium macrocarpon* cultivars as a percent of the total anthocyanins ± standard error of the mean in (A) analytical replicates and (B) biological replicates. 87

Figure 5-5: Multivariate analysis of the UFLC-TOF-MS metabolomic profiles of the cranberry varietals. **(A)** The PLS-DA score plot for LV1/LV2 and **(B)** the corresponding loadings plot for the PLS-DA model, data with calculated Kruskal-Wallis *p*-value <0.05 is highlighted in red. **(C)** Shows the loadings plot re-modeled with all values having a *p*-value >0.05 removed with values identified as significant by SAM statistic highlighted in red and **(D)** the corresponding re-modeled PLS-DA score plot for LV1/LV2. 92

Figure 5-6: Multivariate analysis of the UFLC-TOF-MS metabolomic profiles of the cranberry cultivars. (A) The PLS-DA score plot for LV4/LV2 and (B) the corresponding loadings plot for the PLS-DA model, data with calculated Kruskal-Wallis *p*-value <0.05 is highlighted in red. (C) Shows the loadings plot re-modeled with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted in red and (D) the corresponding re-modelled PLS-DA score plot for LV4/LV2. 93

Figure 6-1: Collections of wild *Vaccinium* species in Haida Gwaii, fall 2010. (A) Location of Haida Gwaii, British Columbia. (B) Collection sites in Haida Gwaii. (C) Samples of *V. oxycoccus* (top) and *V. vitis-idaea* (bottom) with specimens for chemical analyses. 102

Figure 6-2: Phytochemical comparisons of *V. macrocarpon, V. oxycoccus, and V. vitis-idaea* as average amount in n=5 biological replicates with standard error of the mean (A) Anthocyanin concentration in dry tissue (B) Melatonin and

seratonin content (C) Ascorbic acid content (D) Amount of cranberry tissue (n=3) for 50% reduction in DPPH response..... 109

Figure 6-3: Comparison of the ratio of glycosylated peonidins to glycosylated cyanidins in the three *Vaccinium* species as a percent of the total anthocyanins ± standard error of the mean..... 110

Figure 6-4: Chemical structures of **(A)** melatonin (N-acetyl-5-hydroxytryptamine) and **(B)** serotonin (5-hydroxytryptamine). 111

Figure 6-5: **(A)** Multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium* species. PCA Score plot of first two principal components and **(B)** the corresponding loadings plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value of >0.05 removed and **(D)** the re-modeled PCA score plot..... 114

Figure 6-6: Curve generated from plotting the averaged area under the univariate ROC curve for comparisons across *Vaccinium* species versus the statistical *p*-value with variables identified as significant by the SAM statistic highlighted in red..... 115

Figure 6-7: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium spp*. **(A)** PLS-DA Score plot of first and third linear variants and **(B)** The corresponding loadings plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. vitis-idaea* & *V. oxycoccus* (green) and all 3 species (grey) and **(D)** The re-modeled PLS-DA plot. 118

Figure 6-8: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium spp*. **(A)** PLS-DA Score plot of first and third linear variants and **(B)** the corresponding plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. oxycoccus* (green) and all 3 species (grey) and **(D)** the re-modeled PLS-DA plot..... 120

Figure 6-9: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of three *Vaccinium spp*. (A) PLS-DA Score plot of second and fourth linear variants and (B) the corresponding loadings plot with values having a

calculated Kruskal-Wallis *p*-value of <0.05 identified in red. (C) The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. vitis-idaea* & *V. oxycoccus* (green) and all 3 species (grey) and the remodeled PLS-DA plot..... 121

Acknowledgements

I would like to express my sincere gratitude to my senior supervisors, Dr. Susan J. Murch and Dr. Paul R. Shipley, for the opportunity to do this research and for their support. My heartfelt thanks go to Dr. Murch, a true mentor, who motivated me to not only think outside the box, but to move the box to new and exciting places and to Dr. Shipley, whose reassurance and understanding made the trip less bumpy.

It is my pleasure to acknowledge my supervisory committee members, Dr. E. Neeland and Dr. R. O'Brien for their encouragement and support throughout the graduate program. I would also like to acknowledge the other graduate and undergraduate students of the "Murchley" lab who made this journey that much more fun, in particular "King Cuong" Le for assistance the LC/TOF/MS and discussions on all things chemometric and Chrissy Turi for DPPH assays and sharing her love of plants. I would also like to thank Brian Mauza (Ocean Spray Canada Ltd.) and Tim Wolver (Haida Gwaii) for their assistance in field collections. Many thanks to the NRG Team at BCIT who "kept it together" in my absence and to Dr. Kim Dotto the Dean of Applied Research at BCIT who trust in means so very much. Special thanks to Michael Chan and Jamie Finley, the only people I know who get as excited about statistics as I do. Without the gentle pushing, well move like a shove, from Dr. James Watzke and some fantastic colleagues and friends, I would never have taken the plunge, thank you Dr. Betz and Dr. Dentali.

Finally, and most importantly, I could not have done this without the support of my Mom and Dad, who are truly the best parents in the world and my amazing girls, Lauren and Maddie, who are my joy and hold my heart.

For my Dear Old Dad from his Darling Daughter

Chapter 1: Introduction

Over the past decade, a large body of epidemiological evidence has emerged that indicates that dietary intake of berries may have a positive effect on human health (Seeram, 2008a). However, the state of the science for most berries and disease endpoints remains in its infancy as investigators try to turn general epidemiological observations about the role of diets "rich in fruits and vegetables" or "rich in antioxidants" (Brunner et al. 2007; Cherubini et al., 2008; Liu & Russell, 2008) into concrete information about the role of specific berries or even specific berry phytochemicals in disease prevention (Kok et al., 2008; Mutanen et al., 2008; Seeram, 2008b). Despite the activity and interest in the role of berries for long-term prevention of chronic disease, Seeram (2008b) noted, during an overview and introduction to research in the field during the International Berry Health Benefits Symposium held in Corvalis, Oregon from June 13–14, 2005, that there have been very few prospective clinical trials. There has, however, been a great deal of research recently into the benefits of cranberry and cranberry preparations and reports have grown exponentially since 2000 (**Figure 1-1**). While research has included studies on potential benefits of cranberries for cardiovascular disease (Neto, 2007) and cancer (Neto et al., 2008) for the most part the focus has been on the role of cranberries in the prevention and treatment of urinary tract infections (UTIs) (Howell, 2007; Jepson & Craig, 2008).

In North America, there are three species generally recognized as cranberries, *Vaccinium macrocarpon* Aiton, *Vaccinium oxycoccus* and *Vaccinium vitis-idaea* L., (McCown & Zeldin, 2005; Suda & Lysák, 2001; Vander Kloet, 1983). *V. marcocarpon*

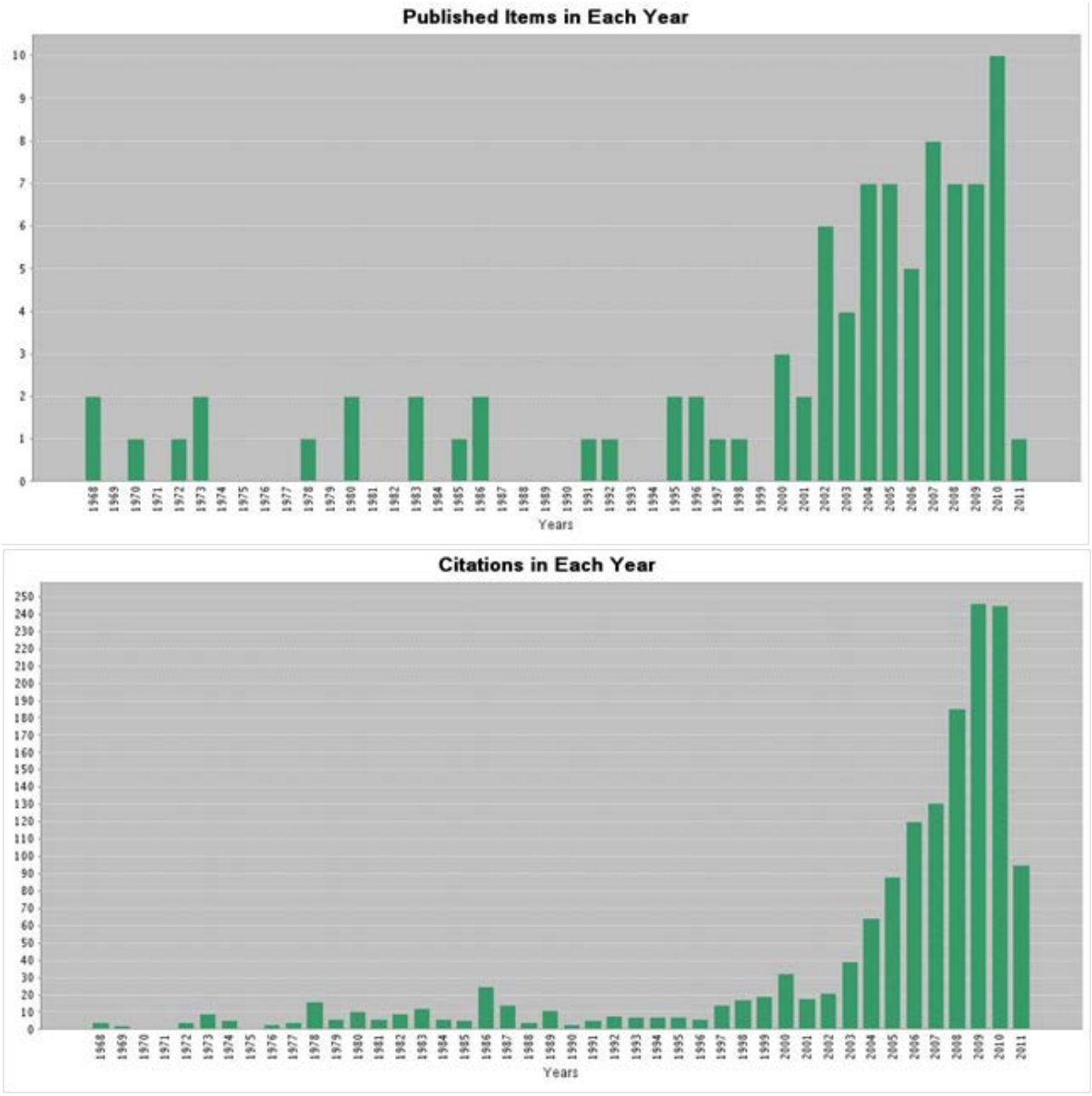

Figure 1-1: Reports of cranberry in the scientific literature as indexed by Web of Science to April 28th, 2011.

is commonly known as 'Large Cranberry' due in part to having significantly larger fruits and leaves than the other two species. This diploid (2n = 24) is native to the Northeastern USA, been bred through agricultural selection and is grown as a commercial crop throughout North America (Hancock et al., 2008; McCown & Zeldin, 2005; Vander Kloet, 1983). *Vaccinium oxycoccus* L, (2n = 24, 48, 72) commonly known as 'Bog

Cranberry', and *Vaccinium vitis-idaea* L., (2n = 24) commonly known as 'Rock Cranberry' in North America, and 'Lingonberry' in Europe are both found in native populations across North America and Europe including widespread natural populations in the coastal and mountainous regions of British Columbia (Douglas et al., 1999; Jacquemart, 1997; Mabberley, 1997; Turner, 2004; Vander Kloet, 1983).

Cranberries have had important roles in traditional health and culture of indigenous people across North America as well as modern uses in the natural health products industry. Several uses of *V. oxycoccus* by the indigenous peoples of northern Canada have been described including consumption of fresh berries, berries stewed with fish, fish eggs or blubber, boiled fruit eaten with meat, fresh or dried fruit stored for winter and fruit preserved in grease and stored in birch bark baskets underground (Moerman, 2004). The Clallam used leaves of the plant to make a tea and the Ojibwa used an infusion of the plant as a treatment of mild nausea (Moerman, 2004). Turner reported that the peoples of Haida Gwaii know *V. oxycoccus* by the name 'dah' and that the name for the species also means "buying" indicating the importance of the crop for trade (Turner, 2004). Patches of berry plants may have been considered the property of a family and passed through generations as fruit was picked in the fall, made into jelly or dried with other fruits (Turner, 2004).

V. vitis-idaea, though morphologically and taxonomically closer to the bilberry, *Vaccinium myritilus* L., than to the cranberry species, was essentially treated and used in the same way as *V. oxycoccus* (Hancock et al., 2008; Kuhnlein & Turner, 1996; Lätti et al., 2011). *V. vitis-idaea* berries, known by the traditional names 'sk'aagii chaay' which translates to "dog-salmon eggs" or 'tllgaa gaanga' which translates to "ground/earth

berries" have been used by the Haida for food, a practice that still continues (Moerman, 2004; Turner, 2004). Moerman reports that the Haida used *V. vitis-idaea* berries for food (Moerman, 2004) and Turner notes that the practice continues and that people now make preserves with them (Turner, 2004). The berries were mixed with boiled fish eggs, livers, air bladders and fat to make a winter meal for the Woodland Cree and that they also used the berries to color porcupine quills or strung on a string to make a necklace (Moerman, 2004). Further, the leaves of the plant were smoked as a tobacco alternative or substitute by the Inuktitut (Moerman, 2004).

Traditionally *V. macrocarpon* berries were gathered from August through the fall, even when still unripe, allowed to ripen and then eaten either fresh or cooked (Kuhnlein & Turner, 1996). Moerman reported the use of wild *V. macrocarpon* by the Algonquin, Chippewa, Ojibwa and Iroquois in baked, dried and raw foods, mixed with corn breads, and sold or traded for other commodities (Moerman, 2004). In the modern marketplace *V. macrocarpon* is one of the significant success stories of the functional foods industry, with by far the most commercial plantings among the three species. Cultivation of native and hybrid varieties of *V. macrocarpon* have been established in low lying bogs across North American and parts of Europe (Suda & Lysák, 2001; Vander Kloet, 1983) with the largest cultivation centers located in Wisconsin, Massachusetts and New Jersey in the United States and in British Columbia, Canada (Kuhnlein & Turner, 1996; McCown & Zeldin, 2005). The American cranberry can be considered a relatively young agricultural crop species, having only been domesticated within the last 160 years (Hancock et al., 2008). Active breeding and selection of cranberry cultivars is also relatively recent and unlike many other staple crop species, the cultivated cranberries are little evolved from

their wild relatives (Hancock et al., 2008). It was not until the 1990's that much of the acreage planted with native selections were replaced with 1st generation hybrids, however, even today many fields still contain native cultivars (Hancock et al., 2008).

The first large-scale cranberry breeding program was initiated by the United States Department of Agriculture (USDA) in 1929 (Dana, 1983; Hancock et al., 2008; Vorsa, 1994). The objective of the program was to develop cranberry cultivars resistant to false blossom disease, a serious disease caused by a phytoplasma carried by the blunt-nosed leafhopper (Dana, 1983; Hancock et al., 2008; Vorsa, 1994). Other selection criteria for the breeding program included yield, fruit rot resistance, keeping quality, fruit appearance, colouring, harvest date, and fruit size (Vorsa, 1994). Evidence from other domesticated crops has shown that breeding can significantly affect secondary metabolite levels and diversity within plants (Fernie & Schauer, 2009; Hopkins et al., 2009; Wink, 1988). These effects are perhaps most prominent when examining secondary metabolites that are associated with chemical defense (Hopkins et al., 2009; Wink, 1988). It has been reported cranberry breeding has had consequences on anti-herbivore defense, as evidenced by an increased susceptibility by two cultivars from a selection and breeding program, as compared to three native varieties, to Gypsy moth caterpillars (Rodriguez-Saona et al., 2011). The same study found that the reduced resistance in the bred varieties correlated to a reduction in the levels of several secondary metabolites in the leaves of the plants (Rodriguez-Saona et al., 2011). Anthocyanins, thought to have several protective biological roles including protection from ultraviolet radiation and solar exposure, cold and drought resistance, and pathogen defense (Chalker-Scott, 1999), are reported to vary between cultivars and time of harvest (Viskelis et al., 2009; Vorsa et al., 2003).

British Columbia is the second largest producer of cranberries in Canada with 88 growers producing 82 million lbs. of fruit in 2009 contributing 12 % to the total North American production (Agriculture & Agrifood Canada, 2010). In 2009 the farm gate sales of cranberry fruit were valued at almost $47 million Canadian (BC Cranberry Growers Association, 2009). Of all the cranberries harvested in Canada every year, about 60% are grown for Ocean Spray which represents as much as 20 per cent of the total North American Ocean Spray cranberry crop (Ocean Spray Canada Ltd., 2011). Almost 96% of all cranberries produced in North America are processed, most commonly for cranberry juice and cranberry juice blends (Knudson, 2008). Since the early 19th century cranberries and cranberry juice have been used for various health effects, most notably the treatment of urinary tract infections (UTIs) (Klein, 2010; Pérez-López et al., 2009). More recently cranberries have become of popular dietary supplement and reportedly led the Dietary Supplements section of the Mainstream Food, Drug, and Mass Market division of the American economy for 2010 with an annual value of $35,806,000 USD (Blumenthal et al., 2011). Health Canada's Natural Health Products Directorate has published monographs for cranberry and those products meeting the compendial standards are permitted to make the traditional use claim, "to help prevent recurrent urinary tract infections" (Government of Canada, 2011a, 2011b).

For preparations made from fruit, Health Canada states the dosage for the product must be equivalent to 10-30 g of fresh fruit per day and for preparations made from juice the dosage for the product must be equivalent to 400-1200 mg of dried juice per day. These dosages are meant to be equivalent to the amounts used in clinical trials that demonstrated positive results for the prevention of UTI (Government of Canada, 2011a,

2011b). However, details about the nature of study materials is lacking in much of the clinical literature and the equivalence of the active constituents or markers among cocktails, concentrates, and capsules/tablets has not been described. As such a comparison among study agents and trial results has not been possible and appropriate cranberry product dose and duration of intervention for prevention of UTI not established (Masson et al., 2009).

While cranberry juice cocktail is the most studied product, concentrates and encapsulated powders have also been used (Jepson & Craig, 2008; Masson et al., 2009). Some sources suggest six capsules of dried cranberry powder are equivalent to 2 oz. of cranberry juice cocktail (Klein, 2010). The single-strength juice is highly acidic and astringent which makes the juice unpalatable at full strength (Hong & Wrolstad, 1986a; Leahy et al., 2001). Accordingly, the juice drink, i.e., cranberry juice cocktail, is a mixture of single-strength cranberry juice, sweetener, water, and vitamin C (Kuzminski, 1996). Juice cocktails can be sweetened with fructose or artificially sweetened and the percent concentrate used in the cocktails ranges from about 25% to 80%, although cocktail with 33% pure juice is most common (Cunningham et al., 2004; Hong & Wrolstad, 1986a; Leahy et al., 2001).

Chemical analyses of cranberries and cranberry products have identified multiple bioactive flavonoids including flavonols, flavan-3-ols (catechins), anthocyanidins, anthocyanins, and proanthocyanidins (Neto, 2007; Nowack & Schmitt, 2008; Pappas & Schaich, 2009). Derived from phenylpropenoid pathways via naringenin, (Ververidis et al., 2007) both anthocyanins and proanthocyanidins (PAC) are found in appreciable quantities in cranberry fruit (Brown & Shipley, 2011; Pappas & Schaich, 2009; Prior et

al., 2001; Vvedenskaya et al., 2004) and have been the focus of many recent studies (Basu et al., 2010; DeFuria et al., 2009; Di Martino et al., 2006; Greenberg et al., 2005; Janecki & Kolodziej, 2010; Johnson-White et al., 2008; Karlsen et al., 2007; Milbury et al., 2010; Szajdek & Borowska, 2008). It has been suggested that the *in vivo* uropathogenic bacterial anti-adhesion properties are due to PAC content (Liu et al., 2006; Foo et al., 2000a; Foo et al., 2000b; Howell et al. 2006) although the concentration necessary to obtain the desired effect remains unknown. Further, the more highly polymerized proanthocyanins are poorly absorbed in the gut (Gonthier et al., 2003; Serrano, Puupponen-Pimiä et al., 2009) and there is conflicting evidence from *in vivo* studies concerning whether oligomeric proanthocyanidins undergo a degree of depolymerisation (Abia & Fry, 2001; Rios et al., 2002; Tsang et al., 2005). Proanthocyanidin dimers and trimers have demonstrated absorption *in vivo* (Baba et al., 2002; Del Bas et al., 2005; Tanaka et al., 2003) and a recent review (Serrano et al., 2009) concludes that while higher order polymers have limited bioavailability they are partially metabolized and available for absorption at different sites in the gastrointestinal tract. However, it has been reported likely that other phytochemicals besides proanthocyanins contribute to the overall biological activity of cranberry, including its anti-oxidant activity (Nowack, 2007). As such, other methods for the evaluation and standardization of cranberry products remain necessary.

Commonly, the quinic acid content and ratio of quinic acid to malic acid are used to calculate percentage of cranberry juice content in juice drinks and to assess cranberry juice authenticity (Kuzminski, 1996; Nagy & Wade, 1995). Unfortunately these acids are not unique or specific to cranberry and as such, employing this measure to determine

quality will not preclude the possibility of adulteration with other juices. Although anthocyanin content of cranberry can change and degrade with processing and storage, the anthocyanin profile is unique to cranberry, and its qualitative pattern characteristic (Nagy & Wade, 1995). According to (Prior et al., 2001), cranberry primarily contains the anthocyanins peonidin-3-*O*-galactoside (P3Ga), cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-arabinoside (P3Ar), and smaller amounts of cyanidin-3-*O*-glucoside (C3Gl) and petunidin-3-*O*-galactoside (**Figure 1-2**).

A broad spectrum of methods have been used for the assay of cranberry phenolic constituents and in a recent review of the analytical literature for berries it was observed these methods are interference-prone and consequently lead to over- or underestimation of the polyphenol content (Krenn et al., 2007). Anthocyanin content of cranberry can degrade with processing and storage and the anthocyanin profile is unique to cranberry, its qualitative pattern characteristic, making it a useful tool for evaluating the identity and quality of cranberry and its products (Hong & Wrolstad, 1986b; Nagy & Wade, 1995; Prior et al., 2001; Woodward et al., 2009). There have been numerous studies showing the bioavailability of anthocyanins (McGhie & Walton, 2007; Wiczkowski et al., 2010) and in particular cranberry anthocyanins post consumption of juice (Milbury et al., 2010). The ability of compounds to be absorbed and available to target tissue is necessary in order to exert biological activity. Thusly, anthocyanin content in cranberry fruit is of particular practical and scientific interest.

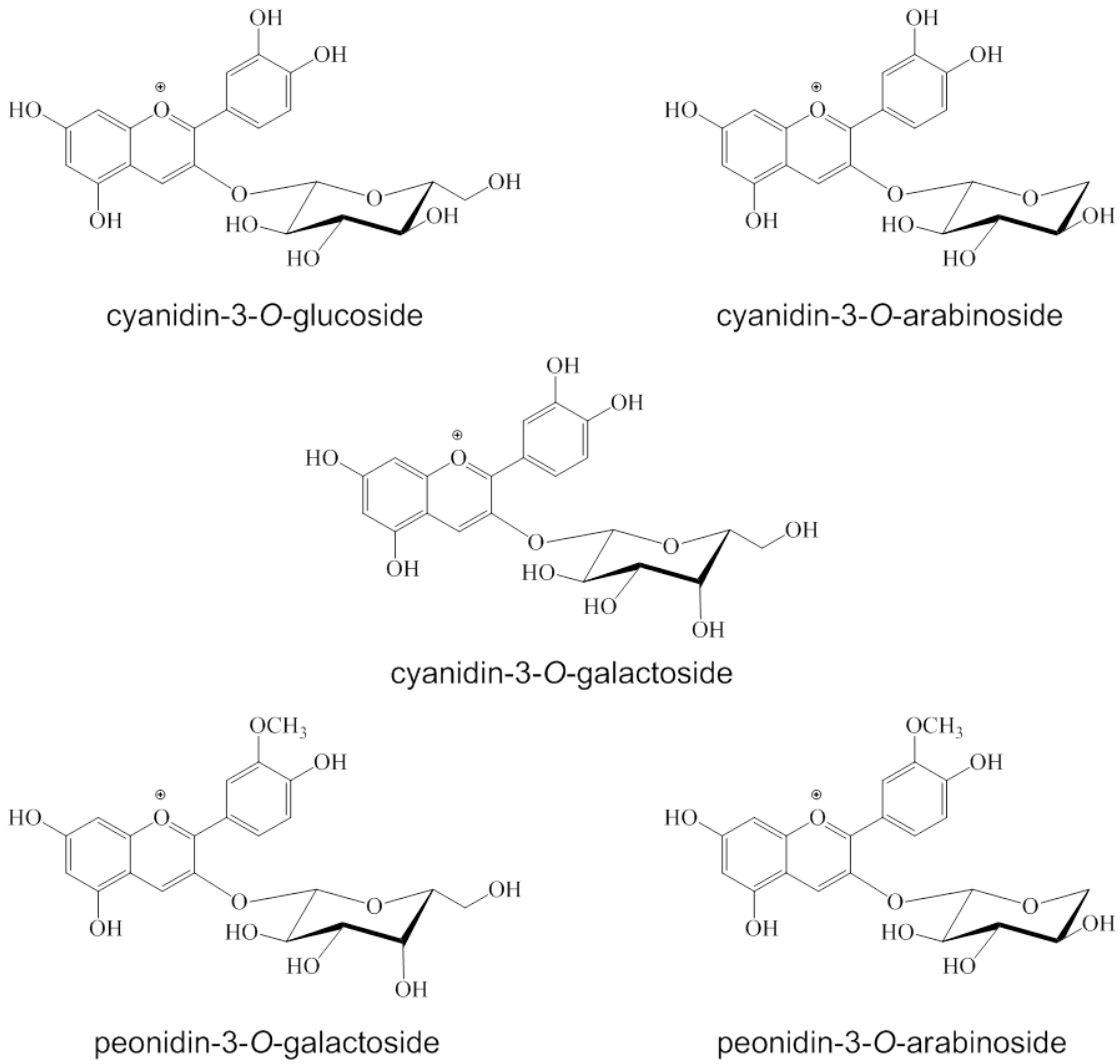

Figure 1-2: Structures for the 5 anthocyanins peonidin-3-*O*-galactoside (P3Ga), cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-arabinoside (P3Ar), and cyanidin-3-*O*-glucoside (C3Gl).

Two common approaches for determining anthocyanin content in cranberry products are spectrophotometry and high performance liquid chromatography with ultraviolet detection (HPLC-UV) HPLC (Chandra et al., 2001; Zhang et al., 2004). Spectrophotometry is intended only to provide an estimation of total anthocyanin content and many HPLC methods force hydrolysis of the anthocyanins prior to chromatographic determination of the anthocyanidin aglycones (Chandra et al., 2001; Krenn et al., 2007).

Both the spectrophotometric and hydrolysis methods can be useful in cases where there is no doubt about the identity and purity of the test materials, however, both are non-specific and susceptible to interferences caused by adulteration with other anthocyanin-rich berries (Chandra et al., 2001). The lack of validated and specific analytical methodology has also lead to an inability to understand or determine the conditions and mechanisms by which cranberry phytochemistry may be varied and optimized. Few methods have been systematically evaluated for accuracy, precision, or reliability and often the analysis of finished products is not within the scope of the method (Betz et al., 2011). With the anthocyanin profile of the large cranberry unique as compared to other *Vaccinium* species commonly used as dietary supplements, a validated analytical method able to provide a positive indicator of identity along with a quantitative measure of unique phytochemicals is highly desirable (Betz et al., 2007; Coates & Meyers, 2011).

Another approach to determine the overall quality of a complex botanical material, such as cranberry, is metabolomics profiling. Described as untargeted quantitative chemical analysis of the whole array of small molecules contained within a cell or tissue (Dunn et al., 2005; Hall et al. 2005; Wishart, 2008), metabolomics analysis is quickly becoming an important tool for the characterization and assessment of commercial herbal products and plant extracts (Chen et al., 2007; Liang et al., 2010; van der Kooy et al., 2009). The collection of data coming from metabolomic analyses is commonly derived from analytical techniques including nuclear magnetic resonance (NMR) (Issaq et al., 2009; Ludwig & Viant, 2010; Schripsema, 2010), liquid chromatography with mass spectrometry (LC-MS) (Allwood & Goodacre, 2010; Issaq et al., 2009), gas chromatography with mass spectrometry (GC-MS) (Allwood et al., 2009; Issaq et al.,

2009) and near-infrared (NIR) analysis (Cozzolino, 2011; Issaq et al., 2009), with NMR and LC-MS being most common (Issaq et al., 2009; Okada et al., 2010; Sumner et al., 2003). Each technique provides unique insights into the metabolome of the material being studied and for this reason methodological limitations can arise in interpreting the generated data including that of sample preparation considerations (Kim & Verpoorte, 2010), detection limit of the analytical technique used (Allwood et al., 2009), and the inherent biological variance in the generated data (Sumner et al., 2003).

The complex metabolic fingerprint generated can be subjected to multivariate statistical data analysis in order to reduce the complexity of the data and detect the pattern changes relating to the environmentally or genetically induced variations in the metabolic composition. The treatment of metabolomic data by multivariate data analyses to develop models for meaningful interpretations is collectively known as “chemometrics”. Making use of various models to interpret data by comparing metrics is referred to as chemoinformatics. This may incorporate classification accuracy (use of permutation techniques), model sensitivity (such as use of the receiver operating characteristic (ROC) curve and similar tools), and derivations of eigenvalues from multivariate models, all of which provide tools for the assessment of the metabolome data under study (Enot & Draper, 2007). Avoiding false discoveries when interpreting metabolomic data usually requires that more than one approach or model be assessed (Broadhurst & Kell, 2006) and ideally multivariate models developed from metabolomics data should be validated (Jansen et al., 2010; Westerhuis et al., 2008).

Focusing on the entire metabolome is a fundamental shift from the characterization of specific classes of plant secondary metabolites and there have been significant advances

made in agriculture and food research to distinguish and characterize related materials through the combination of analytical chemistry and chemometrics (Fernie & Schauer, 2009; Bertram et al., 2010; Cellini et al., 2004; Ku et al., 2010; Louw et al., 2009; Stewart et al., 2011). The following two approaches to phytochemical characterization and differentiation were explored; targeted quantitative determination of known analytes of interest or physical properties and untargeted metabolomics profiling, in order to determine what role chemotype and phenotype may play in potential health benefits attributed to *Vaccinium macrocarpon* or other *Vaccinium* species native to BC**.** The specific objectives of this thesis are to:

I. Develop, optimize and validate a quantitative analytical method for accurate determination of anthocyanins in cranberry (*Vaccinium macrocarpon* Aiton*)* fruit and fruit products.

II. Utilize the validated analytical methodology to conduct a baseline assessment of anthocyanin content and evaluate the diversity within a cranberry crop at harvest.

III. Develop a methodology for the characterization of the metabolome of cranberry fruit and products. Determine and compare the metabolomic profiles of cranberry fruit and demonstrate a unique chemoinformatic approach for the characterization of cranberry fruit and products.

IV. Determine and compare the metabolomic profiles of several cranberry varietals to better understand their diversity and to explore phytochemical relationships between them.

V. Determine and compare the chemical diversity in commercial cranberry and wild populations of native British Columbian cranberries from the Haida Gwaii and

explore the phytochemical relationships between them as well as develop a model for metabolomics analysis to identify commonalities and differences between closely related species.

Together these studies are a comprehensive evaluation of the phytochemistry of *Vaccinium macrocarpon* varietals and other *Vaccinium* species native to BC and provide insight into the variability that exists between commercial producers, the challenges and opportunities afforded by the processes of genetic selections in breeding agricultural crops and the potential of native BC berries species, not currently cultivated, to play a future role in human health and nutrition. The comparative metabolomics described identify the degree of phytochemical commonality and difference among *Vaccinium macrocarpon* varietals and two related *Vaccinium* species. A new approach is described to evaluate the quality of multivariate models generated from metabolomics data. The model developed for experimental design and statistical analysis can be applied to a range of other complex natural products thereby supporting research to establish links between supplementing the diet and the prevention and/or treatment of health conditions.

Chapter 2: Single Laboratory Validation of a Method to Determine Anthocyanins in Cranberry Fruit Products by High Performance Liquid Chromatography with Ultraviolet Detection

Synopsis

Quantitative methods intended for determination of the quality of commercial products bear the weight of regulatory scrutiny and therefore it is essential methods are demonstrated to be accurate, precise, and reproducible. For cranberry fruit and natural health products derived from cranberry fruit anthocyanins are used as a determinant of quality (Krenn et al. 2007; Kuzminski, 1996). Spectrophotometry and HPLC analysis of the anthocyanidins (aglycones) post-hydrolysis can provide an estimate of total anthocyanin content, in cases where the identity and purity of the test article is known. This non-specific approach is susceptible to interferences caused by adulteration with other anthocyanin-rich berries; as such methods to directly measure intact anthocyanins are preferred (Zhang & Zuo, 2004). When measuring intact anthocyanins, it is common industry practice to use a high-purity C3Gl standard to make an external calibration curve and then use total peak areas at 520 nm to calculate total anthocyanin content in mg/g (Andersen, 1989; Cassinese et al., 2007; Chandra et al., 2001; Chen et al., 2007; Deubert, 1978; Fuleki & Francis, 1968; Mazza et al., 2004). This process is called normalization and the validity of the approach is based on the assumption that all anthocyanins have the same UV response to that of C3Gl. With all five major anthocyanins currently commercially available individual reference standards for each analyte of interest was employed rather than using only C3Gl. The HPLC-UV method

described for quantification of the five predominant cranberry anthocyanins was adapted from literature methods (Chandra et al., 2001; de Pascual-Teresa, Santos-Buelga, & Rivas-Gonzalo, 2002; Fuleki & Francis, 1968; Moyer et al., 2002) and validated for use as a quality control tool in analysis of commercial cranberry fruit products. The method described quantifies the five major anthocyanins in raw cranberry, cranberry juice, cranberry juice cocktail, and cranberry extract powder, by employing a seven point standard curve generated using mixed standard solutions containing C3Ga, C3Ar, C3Gl, P3Ga, and P3Ar.

Experimental

Principle

This method is suitable for the determination of cyanidin-3-*O*-galactoside, cyanidin-3-*O*-arabinoside, cyanidin-3-*O*-glucoside, peonidin-3-*O*-galactoside, and peonidin-3-*O*-arabinoside in freeze dried raw cranberry, cranberry juice, cranberry juice cocktail and cranberry extract powder. The extraction methodology uses an acidified 95% alcohol 5% water extraction solvent as described by Fuleki & Francis (1968). The chromatography conditions were slightly modified from methods described by Chandra et al. (2001) and de Pascual et al. (2002). Quantification of the anthocyanins of interested is achieved through the use of standard curves prepared from mixed anthocyanin solutions.

Test Samples

Raw cranberries were obtained from Ocean Spray Canada (Richmond, BC, CAN) and freeze dried upon receipt. Cranberry juice, cranberry juice cocktail and cranberry

powdered extract samples were bought from local stores. All test samples were stored at 4^{o}C. The same lot of each material is used throughout the validation studies.

Reagents and Supplies

Chemicals

HPLC grade methanol, acetonitrile, hydrochloric acid, and phosphoric acid were obtained from VWR International (Edmonton, AB, CAN). Ascorbic acid (purity ≥99.0%) and phosphoric acid, 85% in H_2O (purity ≥99%), were obtained from Sigma-Aldrich Canada Ltd. (Oakville, ON, CAN).

Solutions

The following solutions were prepared for this study:

a) Reference standard diluent – 98:2 (v/v) methanol: conc. HCl (HCl in H_2O, 33-40%)

b) Extraction solvent – 98:2 (v/v) methanol: conc. HCl (HCl in H_2O, 33-40%)

c) HPLC Mobile Phase A (MPA): 99.5:0.5 (v/v) water: conc. phosphoric acid

d) HPLC Mobile Phase B (MPB): 50.0:48.5:1.0:0.5 (v/v) water:acetonitrile:glacial acetic acid: conc. phosphoric acid

Reference Standards

Reference standards for each of the five anthocyanins of interest were obtained from Cerilliant Corp. (Round Rock, Texas). Individual 1000 µg/mL stock solutions of each anthocyanin standard were prepared by weighing 10.0 ±1.0 mg of each standard to separate amber 10-mL volumetric flasks. Approximately 5 mL of a 2% (v/v) conc. HCl in

methanol solution was added to each flask and sonicated until all solid material dissolved. The flasks were then brought to volume with the addition of 2% (v/v) conc. HCl in methanol solution. Solutions were stored at -20 ^{o}C when not in use. A mixed anthocyanin reference standard solution was prepared by transferring 400 µL of the C3Ga stock solution, 100 µL of the C3Gl stock solution, 400 µl of the C3Ar solution, 70 µl of the P3Ga solution and 30 µL of the P3Ar solution and 8100 µl of a 2% (v/v) HCl in methanol solution to a test tube. The tube was vortexed for 30 seconds to produce a mixed standard solution with the concentrations shown in **Table 2-1**. The purity of each individual chemical standard was determined chromatographically using LC-MS analysis. The actual concentration of each anthocyanin in the stock solution was then calculated after purity corrections were applied. Good agreement between the reported and observed purities was observed for most of the standards, but a significant disparity in the reported purity and the measured purity of C3Ar was noted. The mixed anthocyanin reference solution was stored at -20 ^{o}C.

Table 2-1: Anthocyanin concentrations in mixed standard solution.

Anthocyanin	**Theoretical Conc. (µg/ml)**	**Purity (%)**	**Actual Conc. (µg/ml)**	**Chemical formula**
C3Ga	40	91.32	36.53	$C_{21}H_{21}ClO_{11}$
C3Gl	10	98.30	9.83	$C_{21}H_{21}ClO_{11}$
C3Ar	40	44.18	17.67	$C_{20}H_{19}ClO_{10}$
P3Ga	70	92.44	64.71	$C_{22}H_{23}ClO_{11}$
P3Ar	30	90.47	27.14	$C_{21}H_{21}ClO_{10}$

Apparatus

a) *Analytical balance* - Mettler Toledo AE 260 Analytical Delta Range (±0.1 mg) (VWR International, Edmonton, AB., Canada).

b) *Centrifuge* – Eppendorf 5804 Table Top Centrifuge (VWR International, Edmonton, AB, Canada).

c) *Wrist Action Shaker* – Burrell Model BT Wrist Action Shaker (VWR International, Edmonton AB, CAN).

d) *Syringes* – 3 ml Luer-lok® fitted with PTFE filter, 0.45 µm and 0.2 µm pore size, 25 mm diameter (Fisher Scientific, Ottawa, ON, Canada) or equivalents.

e) *Vortex mixer* – Thermo Scientific Maxi Mix 1 (VWR International, Edmonton, AB, Canada).

f) *Micropipettes* – Eppendorf Reference Series, 100, 200 and 1000 µl (VWR International, Edmonton, AB, Canada).

g) *HPLC system* – Agilent 1100 Series liquid chromatograph equipped with quaternary pump and degasser (G1354A), temperature controlled column compartment (G1316A), temperature controlled auto-sampler (G1327A), standard flow-cell 10 mm, 13 µl, 120 bar (G1315-60012), Diode-array detector (G1315B), HPLC 2D ChemStation Software (G2175AA), and online degasser (1322A) (Agilent Technologies, Mississauga, ON, Canada).

h) Ultrasonic water bath – Bransonic 3510 (VWR International, Edmonton, AB, Canada).

i) Coffee grinder – Black and Decker smart grind.

LC Conditions

Instrument: Agilent HP1100 HPLC equipped with diode array detector, or equivalent

Auto-sampler Temperature: 4 ^{o}C

Analytical Column: Cosmosil 5C18-PAQ 4.6 mm x 150 mm, 5 µm (Nacalai USA Inc., San Diego, CA, USA)

Column Temperature: 25 ^{o}C

Detector Conditions (nm): 520 nm (8 nm bandwidth), no reference

Flow Rate: 0.9 mL/min

Run Time: 35 minutes with 5 minutes post time for column equilibration

Injection Volume: 10 µL

Gradient Conditions:

Time (min)	**% Mobile phase B**
0	10
28	50
32	75
32.1	10
35	10

The order of elution is C3Ga (15.8-15.9 min), C3Gl (16.7-16.9 min), C3Ar (17.7-17.9 min), P3Ga (18.3-18.5 min) and P3Ar (20.2-20.4 min), see **Figure 2-2.**

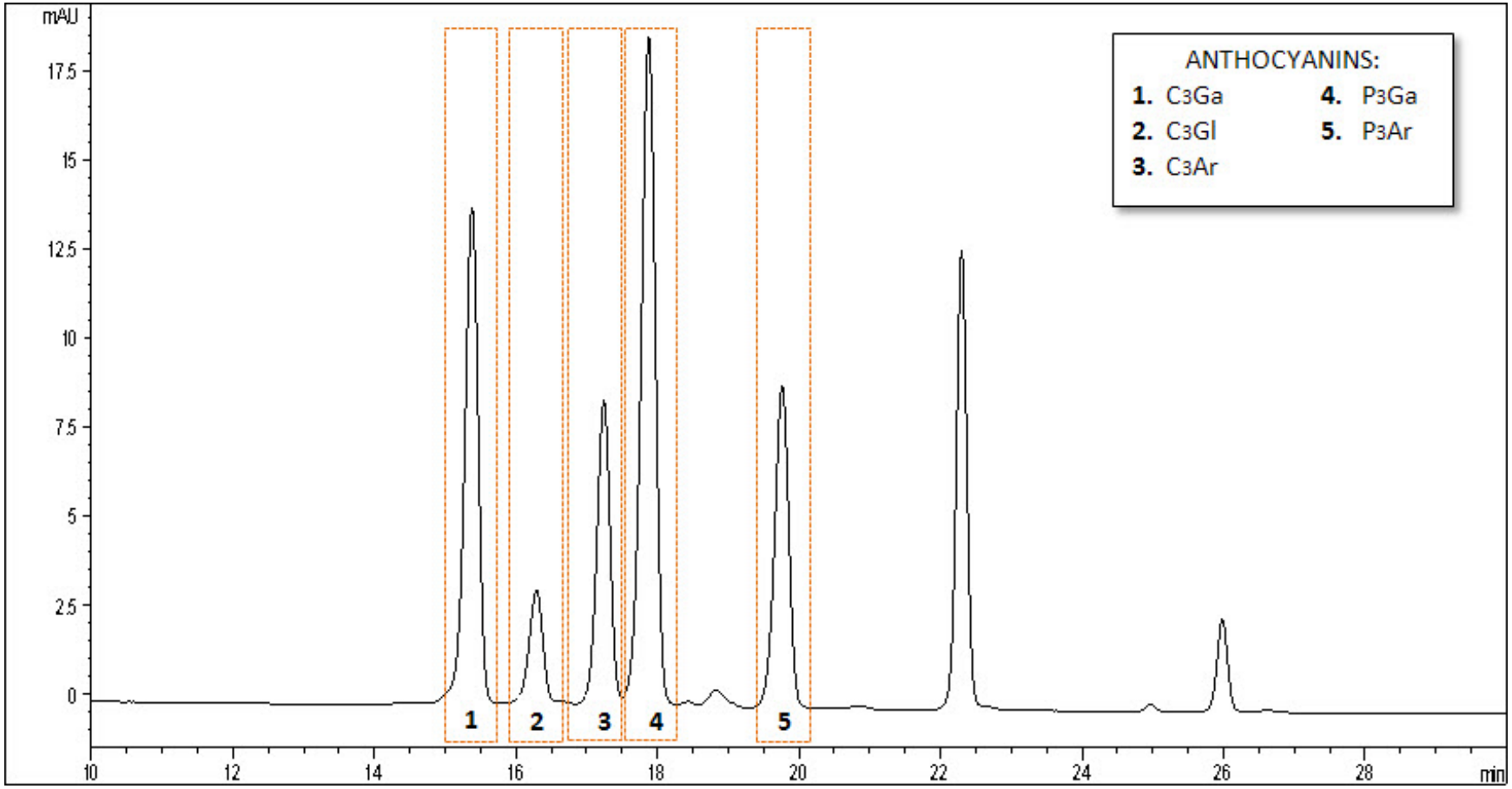

Figure 2-1: The order of anthocyanin elution in a mixed anthocyanin reference standard.

Calculations

Individual anthocyanins from solid samples are quantified in % w/w using the following equation:

$$\frac{P_0 - b_0}{m_0} \times \frac{V}{W} \times \frac{D}{100}$$

P_0 = Peak area of target analyte in sample chromatogram

b_0 = y-intercept of calibration curve for the target analyte

m_0 = Slope of calibration curve for the target analyte

V = Volume of Test Solution, in mL

W = Dry weight of sample, in g

D = Dilution Factor = 1

Individual anthocyanins from liquid samples are quantified in µg/mL using the following equation:

$$\frac{P_0 - b_0}{m_0} \times D$$

P_0 = Peak area of target analyte in sample chromatogram

b_0 = y-intercept of calibration curve for the target analyte

m_0 = Slope of calibration curve for the target analyte

D = Dilution Factor = 1

The calculations used to determine the Horwitz Ratio (HorRat), a normalized performance parameter used to evaluate overall method precision, are provided below.

RSDr (found, %): $$RSDr = \frac{SD(r)}{mean} \times 100$$

Where SD(r) = the population standard deviation.

PRSDr (RSDr calc, %): $$PRSDr = 2C^{-0.15}$$

Where C = the concentration of the analyte expressed as a mass fraction

HorRat value: $$HorRat = \frac{RSDr}{PRSDr}$$

Preparation of Test Samples

A. Freeze Dried Raw Material and Powdered Extract:

Grind freeze dried cranberries and cranberry extract powder to < 60 mesh (250 µm) powder. Weigh 0.250 g (+/-0.025 g) of test article to a 50-mL conical tube, add 20 mL of extraction solvent using a graduated cylinder, mix for 10 seconds with a vortex mixer, sonicate for 15 minutes, and shake on an angle at 180 rpm for 30 minutes. Mix test solutions for 10 seconds with a vortex mixer and centrifuge at 5000 rpm for 5 minutes. Decant the supernatant to a 25-mL volumetric flask and bring to 25 mL volume with extraction solvent. Invert the flask 20 times to mix and filter approximately 1 mL of the solution through a 0.45-µm Teflon filter into an amber HPLC vial for analysis.

B. Juice:

Thoroughly mix the juice by inverting the container approximately 20 times or until no sediment is visible. Centrifuge 10 mL of juice at 5000 rpm for 10 minutes. Dilute 1000

µL with 500 µL extraction solvent and filter approximately 1 mL of the solution through a 0.45-µm Teflon filter into amber HPLC vials.

C. Juice Cocktail:

Thoroughly mix the juice cocktail by inverting the container approximately 20 times or until no sediment is visible. Centrifuge 10 mL of cocktail at 5000 rpm and filter approximately 1 mL of the solution through a 0.45-µm Teflon filter into amber HPLC vials.

Preparation of Calibration Solutions

The mixed anthocyanin reference solution (as described above) is diluted with a solution of 2% (v/v) conc. HCl in methanol, as per the dilution scheme presented below, mixed well and stored at -20 ^{o}C when not in use. Linearity 7—Pipette 500 µL of the mixed anthocyanin reference solution

a) Linearity 6—Dilute 500 µL of linearity 7 solution with 500 µL of reference solution diluent.
b) Linearity 5—Dilute 500 µL of linearity 6 solution with 500 µL of reference solution diluent
c) Linearity 4—Dilute 500 µL of linearity 5 solution with 500 µL of reference solution diluent.
d) Linearity 3—Dilute 500 µL of linearity 4 solution with 500 µL of reference solution diluent.

e) Linearity 2—Dilute 500 µL of linearity 3 solution with 500 µL of reference solution diluent.

f) Linearity 1—Dilute 500 µL of linearity 2 solution with 500 µL of reference solution diluent.

The final concentrations of the calibration standards would produce standard curves that would capture the range of each of the anthocyanins of interest typically found in cranberry fruit and their products. The approximate concentrations for each of the samples are presented in **Table 2-2** below.

Table 2-2: Approximate concentrations (µg /mL) of individual anthocyanins at each linearity of the calibration curve

Linearity	[C3Ga]	[C3Gl]	[C3Ar]	[P3Ga]	[P3Ar]
7	36.53	9.83	17.67	64.71	27.14
6	18.27	4.91	8.84	32.35	13.57
5	9.13	2.46	4.42	16.18	6.78
4	4.57	1.23	2.21	8.09	3.39
3	2.28	0.61	1.10	4.04	1.70
2	1.14	0.31	0.55	2.02	0.85
1	0.57	0.15	0.28	1.01	0.42

Single Laboratory Validation Parameters

This method described was validated according to AOAC International guidelines for conducting a Single Laboratory Validation (SLV) (AOAC International, 2002).

Selectivity

Selectivity was demonstrated by injecting each of the reference analytes and showing that all anthocyanins to be quantified were well resolved from potential matrix interferences at 520 nm and confirmed by a lack of signals from the matrix blanks. Resolution was

considered adequate if baseline separation was observed between the analyte peaks and any other peaks in the chromatogram.

Linearity

The linearity of each analyte was demonstrated for a 7-point calibration standard curve, prepared as described above. The calibration curves for each anthocyanin were plotted and linear regression used to determine the slope and y-intercept. Each calibration curve was visually inspected to confirm linearity. The correlation coefficient (r^2) of the regression lines were calculated using linear regression for each quantified anthocyanin. An r^2 of $\geq$99.5% was considered linear and thus acceptable for quantifying the anthocyanins in the different matrices.

Method Detection Level and Limit of Quantification

The absence of a suitable sample blank precluded the use of the IUPAC method for determination of detection limits for the analytes. As an alternative, the detection limit for each of the analytes was determined using the EPA Method Detection Level (MDL) which is defined as the minimum concentration of substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero (Environmental Protection Agency, 2003). First, a set of nine replicate standard solutions containing a very low concentration of each of the analytes was used to determine the MDL. The calculation of the MDL is as follows:

$$MDL = s \times t_{(0.01,n-1)}$$

where:

s = the sample standard deviation of the replicates

$t_{(0.01,n-1)}$ = the t statistic with α = 0.01 and n - 1 degrees of freedom.

A second set of seven replicate standard solutions containing another very low level of each of the analytes was also prepared and analyzed. Variance ratio tests to between the two replicate sets were used to ensure the variances at low levels were equal. The limit of quantification (LOQ) was calculated as 10 times the sample standard deviation of the results for the replicates used to determine the MDL.

Precision (Repeatability)

Precision was evaluated by analyzing multiple replicates of each test sample. Four replicate preparations of each material were prepared and analyzed on each of three separate days. In total, twelve replicates of each material were prepared. The within-day, between day and total standard deviations were calculated for each of the individual anthocyanins, in each of the four materials. The HorRat value for each analyte in each material was also calculated to assess the overall precision of the method as per the method described by Horwitz (Horwitz, 1982).

Accuracy

Dried elderberry was used as the matrix blank and considered representative of dietary supplement finished products. The matrix blank was spiked with cyanidin-3-O-galactoside at three spiking levels, corresponding to approximately 200%, 100%, and

50% of the expected concentrations in actual sample preparations. Each level was prepared in quadruplet, and analyzed on three separate days along with unfortified controls.

Stability Study – Reference Standard and Test Material Solution

A freeze dried cranberry sample and a mid-range mixed anthocyanin reference standard solution prepared on the first day of the precision analysis were used to assess the stability of the analytes in solution. Both samples were stored in the freezer at -25 ^{o}C for 13 days. Each of the samples was analyzed as per the described method at t = 0, seven, and thirteen days. Total anthocyanin concentration in the test solutions were quantified using standard curves generated from freshly prepared standards and the results obtained at the three different time points compared.

Youden Ruggedness Trial: A Factorial Study

A factorial study was performed to determine whether deviations in selected factors would affect the method performance. The trial used for this test was described by W.J. Youden (Youden & Steiner, 1975) and further detailed by Wernimont (Wernimont, 1987) and examined seven factors at two levels in eight experiments. The seven factors examined were; sample mass, sonication time, extraction solvent concentration, shaking time, sonicator bath temperature, injection volume, and centrifugation time. The parameter levels are designated in **Table 2-3**. Both freeze dried cranberries and a cranberry extract powder were used as sample matrices for the factorial study. Two-tailed variance ratio tests were used to compare the variances for each analyte observed from

the eight experiments. These variances were also compared to the variance obtained for each of the analytes for the cranberry extract powder for the precision analysis describe above.

Table 2-3: Parameters used in the factorial study.

Factor	High Value	Low Value	Factor Effect	Effect
Sample mass	A = 275 mg	a = 175 mg	A and a	J
Sonication time	B = 20 min	b = 10 min	B and b	K
% HCl in solvent	C = 2.5%	c = 1.5%	C and c	L
Shaking time	D = 40 min	d = 20 min	D and d	M
Sonicator temperature	E = 40 °C	e = ice bath	E and e	N
Injection volume	F = 15µL	f = 10 µL	F and f	O
Centrifugation time	G = 10 min	g = 5 min	G and g	P

Results and Discussion

Method Validation Results – Performance Characteristics

Identification of C3Ga, C3Gl, C3Ar, P3Ga and P3Ar in validation test samples was possible by comparing retention times and UV profiles to those of the mixed anthocyanin reference standard. Quantification of analytes was conducted using linear regression analysis from triplicate standard curves prepared on three separate days. Quadruplicate samples were also prepared on each of these days. Test sample matrices included freeze dried cranberry, cranberry extract powder, cranberry juice and cranberry juice cocktail.

Selectivity

C3Ga, C3Gl, C3Ar, P3Ga and P3Ar were well resolved from potential matrix interferences at 520 nm.

Linearity and Analytical Range

All of the standard curves produced over the course of the study appeared linear upon visual inspection with all calculated correlation coefficients above 99.5%. These results confirm the linearity of the analytical range studied. The analytical range was approximately 0.57-36.53 µg/mL for C3Ga, 0.15-9.83 µg/mL for C3Gl, 0.28-17.67 µg/mL for C3Ar, 1.01-64.71 µg/mL for P3Ga, and 0.42-27.14 µg/mL for P3Ar. For solid materials, as prepared by the described method this translates to 0.06-3.65 mg/g for C3Ga, 0.02-0.98 mg/g for C3Gl, 0.03-1.77 mg/g for C3Ar, 0.10-6.47 mg/g for P3Ga, and 0.04-2.71 mg/g for P3Ar.

Limits of Detection and Limits of Quantification

Variance checks showed that the method used was applicable for the analytes. The MDL and LOQ for each of the analytes are reported in **Table 2-4**.

Table 2-4: The method detection limit (MDL) and limit of quantification (LOQ) calculated for each of the analytes used in this study.

Analyte	MDL (µg/mL)	LOQ (µg/mL)
C3Ga	0.02	0.06
C3Gl	0.02	0.05
C3Ar	0.01	0.02
P3Ga	0.01	0.04
P3Ar	0.01	0.03

Precision (Repeatability)

Results from the precision analysis indicated that all analytes had adequate precision in each of the matrices. HorRat values were also acceptable for the anthocyanins ranging

from 0.18 to 1.06. AOAC guidelines describe the acceptable range of HorRat values to be between 0.3 – 1.3. The low HorRat values (<0.3) were considered acceptable as the analysis was performed by a well-trained analyst under tightly controlled situations (AOAC International, 2002). **Table 2-5** summarizes the precision results obtained over the course of the three days.

Table 2-5: Precision results summary of *Vaccinium macrocarpon* test samples.

Matrix	Analyte	Mean mg/g	s(w)	s(b)	s(t)	RSDr	HorRat
Freeze Dried Cranberry	C3Ga	1.11	0.02	0.02	0.03	2.96	0.53
	C3Gl	0.03	0.01	0.01	0.04	12.70	1.32
	C3Ar	0.63	0.01	0.01	0.02	2.65	0.44
	P3Ga	1.82	0.03	0.03	0.05	2.62	0.51
	P3Ar	0.68	0.01	0.01	0.02	2.70	0.45
Cranberry Extract Powder	C3Ga	0.39	0.01	0.01	0.01	2.77	0.42
	C3Gl	0.02	0.01	0.01	0.01	11.20	1.07
	C3Ar	0.17	0.01	0.01	0.01	2.27	0.31
	P3Ga	0.44	0.01	0.01	0.01	2.73	0.43
	P3Ar	0.15	0.01	0.01	0.01	3.31	0.44
Matrix	**Analyte**	**Mean µg/mL**	**s(w)**	**s(b)**	**s(t)**	**RSDr**	**HorRat**
Cranberry Juice	C3Ga	38.10	0.55	0.49	0.74	1.94	0.21
	C3Gl	1.03	0.03	0.05	0.054	5.28	0.33
	C3Ar	23.50	0.54	0.48	0.72	3.08	0.31
	P3Ga	74.10	0.98	0.87	1.31	1.77	0.21
	P3Ar	29.90	0.51	0.39	0.64	2.15	0.22
Cranberry Juice Cocktail	C3Ga	3.76	0.06	0.45	0.07	1.98	0.15
	C3Gl	0.12	0.01	0.02	0.02	19.8	0.89
	C3Ar	2.04	0.04	0.03	0.05	2.61	0.18
	P3Ga	6.49	0.13	0.03	0.01	1.99	0.16
	P3Ar	2.24	0.06	0.03	0.06	2.85	0.28

s(w) = within-day standard deviation
s(b) = between-day standard deviation
s(t) = total standard deviation
RSDr = Total repeatability standard deviation

Accuracy

A spike recovery study spiking C3Ga into negative control material (freeze dried elderberry) at 200%, 100% and 50% of expected levels was completed. The average recovery over these three levels was 99.79% with a percent coefficient of variation (%CV) of 4.59%. Refer to **Table 2-6** for spike recovery results.

Stability Study – Reference Standard and Test Material Solution

No significant deviation in the total anthocyanin content was observed in the mixed standard and extracted solution of freeze dried cranberry fruit when determined at seven and thirteen days, as compared to t = 0. These results suggest the anthocyanins are stable in solution when stored at -25 ^{o}C for at least 13 days.

Table 2-6: Negative control spike recovery results.

C3Ga Spike Level	% Recovery	%CV
Level 1 (50%)	98.24	5.18
Level 2 (100%)	99.43	4.95
Level 3 (200%)	101.69	2.89

Youden Ruggedness Trial: A Factorial Study

The calculated effects of varying the various parameters did not appear to have a significant effect on the total % anthocyanins extracted and determined. F-ratio variance tests indicated that none of the calculated variances for each of the analytes for each of the factors analyzed were significantly different from the variances calculated for the analytes in the cranberry extract powder sample during precision testing. This is illustrated graphically in **Figure 2-3,** whereby each calculated effect (J-P, refer to **Table 2-3**) is plotted with the axis limits set as the overall method precision. Results of the

Youden factorial study indicate that the optimized method is not significantly affected by changes to those specific factors.

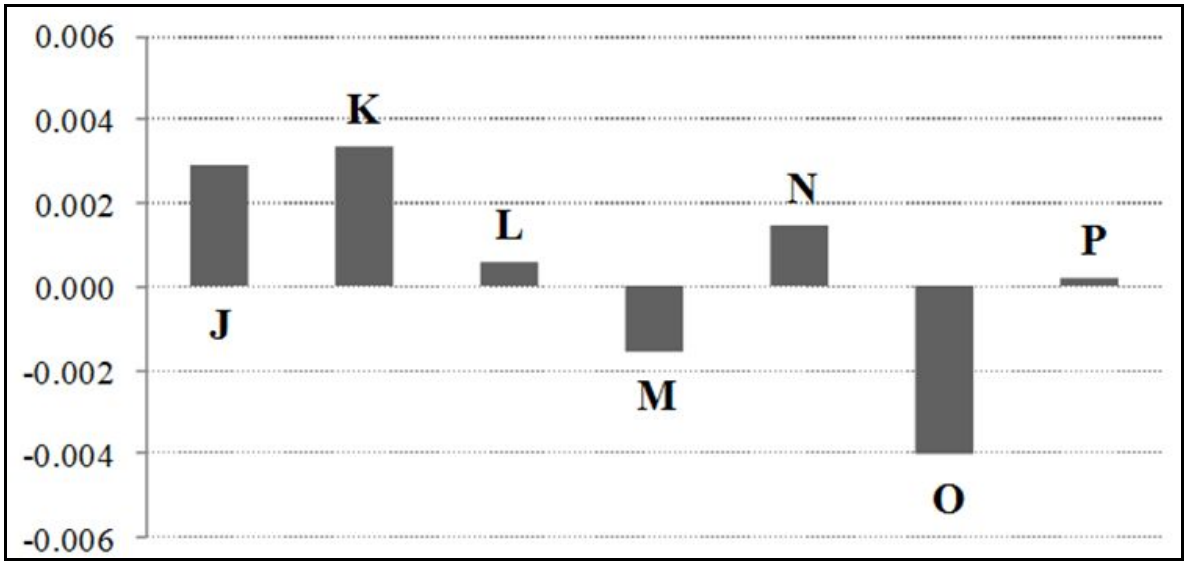

Figure 2-2: Calculated effect of factors on total anthocyanins (mg/g) extracted from cranberry extract powder. Factor effects are as follows: J = sample mass; K = sonication time; C = % HCl in solvent; M = shaking time; N = sonicator temperature; O = injection volume; P = centrifugation time.

With the prevalence of cranberry products in the marketplace and ongoing interest in the scientific community on the potential health benefits of cranberry fruit it is important that analytical methods to access the quality of cranberry fruit and products are validated. The HPLC method described herein for the determination of the five major anthocyanins in raw materials and select finished product containing cranberry (*Vaccinium macrocarpon*) was subjected to a single laboratory validation study as per AOAC guidelines (AOAC International, 2002), a stability study, and Youden ruggedness trial. In summary, the results of these studies demonstrates that the method described is fit for the purpose of determining cyanidin-3-galactoside, cyanidin-3-glucoside, cyanidin-3-arabinoside, peonidin-3-galactoside and peonidin-3-arabinoside in cranberry fruit and

finished products (dried extracts, juice and juice cocktail). It is recommended the method be further evaluated for reproducibility and a composite measure of variation that includes between-laboratory and within-laboratory variation, through a collaborative study. The paucity of information concerning purity and stability of commercially available calibration standards for anthocyanins makes it difficult to determine whether the process of normalization to C3Gl is valid for determination of all anthocyanins. Efforts should be made to independently confirm the purity of chemical reference materials prior to use as calibration standards for quantification.

Chapter 3: Anthocyanin Diversity within a Commercial Production of Cranberries (*Vaccinium macrocarpon* Aiton) in British Columbia

Synopsis

Understanding the medicinal efficacy of cranberry bioactives is complicated by complex human physiology and the microbiome. Compounds must be absorbed and be available to target tissue in order to exert biological activity. There have been numerous studies showing the bioavailability of anthocyanins (McGhie & Walton, 2007; Wiczkowski et al., 2010) and in particular cranberry anthocyanins post consumption of juice (Milbury et al., 2010). Therefore, the anthocyanin content in cranberry fruit is of particular practical and scientific interest. The objective of the current study was to determine the anthocyanin diversity among commercially grown cranberries harvested from production sites across the lower mainland of British Columbia. British Columbia is the second largest producer of cranberries in Canada. As BC's largest berry crop, 88 growers produced 82 million lbs. of fruit in 2009 contributing 12 % to the total North American production (Agriculture & Agrifood Canada, 2010). In 2009 the farm gate sales of cranberry fruit were valued at almost $47 million Canadian (BC Cranberry Growers Association, 2009). Of all the cranberries harvested in Canada every year, about 60% are grown for Ocean Spray which represents as much as 20% of the total North American Ocean Spray cranberry crop (Ocean Spray Canada Ltd., 2011). Herein we report the modification of a previously validated HPLC method for determining cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-glucoside (C3Gl), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-galactoside (P3Ga) and peonidin-3-*O*-arabinoside (P3Ar) in cranberry fruit

(Brown & Shipley, 2011). The optimized method was used to assess the consistency of anthocyanin content across the 2009 British Columbia cranberry crop as a model for phytochemical contents in commercial products.

Experimental

Collection of Cranberry Samples

Twenty samples of cranberry fruit, randomly selected from different commercial trucks originating from 20 different production sites during the 2009 harvest across the lower mainland of British Columbia, were provided by Ocean Spray Canada Ltd. (Richmond, BC, CAN). All cranberries were harvested using a standard "Wet-Pick" method which includes harvest flooding (**Figure 3-1A-C**). Fresh cranberry samples were received directly from Ocean Spray Canada Ltd. before the harvested berries were sent to the processing facility in Richmond, BC. All samples were immediately frozen and stored at -20 °C. A randomized sample selection procedure was followed to subsample three replicates from each of the original 20 samples. For each replicate, berries were randomly chosen as a lot sample from each bulk sample, freeze dried (Modulyo Freeze-Dry System; Fischer Scientific, Ottawa, ON, CAN) and then ground to a <60 mesh (250 μm) powder.

Sample Preparation

Test samples of freeze-dried cranberry fruits, subsampled from the original 20 samples of the 2009 harvest, were prepared for quantitative analysis by HPLC using previously published validated analytical method (Brown & Shipley, 2011). In brief, freeze-dried

cranberries were weighed (0.250 g ±0.025 g) into a 50-mL conical tube, 20 mL of HPLC grade MeOH–concentrated HCl (98:2, v/v) was added and samples were mixed with a vortexer. (Thermolyne Maxi Mix 1; Fisher Scientific) for 10 seconds, sonicated

Figure 3-1: Wet-picked harvest of cranberry (*Vaccinium macrocarpon*) in the lower mainland region of British Columbia, Canada.

(Branson Model 3510R-MTH Ultrasonic Cleaner; VWR International, Mississauga, ON, CAN) for 15 minutes, and shaken on an angle at 180 rpm for 30 minutes (Wrist Action Shaker model 57040-82; Burrell Scientific; Pittsburg, PA, USA). The supernatant was decanted into a 25-mL glass volumetric flask and samples were brought to final volume of 25 mL with the extraction solvent. Samples were mixed well by inversion and approximately 1 mL of each sample solution was filtered (0.45-µm Teflon filters; VWR International; Edmonton, AB, CAN) into an amber vial for HPLC analysis.

Chemical Calibration Standards

The individual anthocyanin chemical calibration standards were prepared by Cerilliant Corp. (Round Rock, TX, USA) in Snap-N-Shoot® format as either ~500 or 250 µg/mL solutions of 2% concentrated HCl in methanol (v/v). Mixed chemical calibration standards were prepared as per **Table 3-1** by serial dilution with 2% concentrated HCl in methanol (v/v). The purity of each chemical calibration standard was determined chromatographically, immediately prior to use. The individual Snap-N-Shoot® chemical calibration standards and mixed standard solutions were stored at -20 ^{o}C when not in use.

Table 3-1: Calculated concentration (µg/mL) of anthocyanins, corrected for determined purity, at each linearity level of the calibration curve.

Linearity	[C3Ga]	[C3Gl]	[C3Ar]	[P3Ga]	[P3Ar]
5	62.26	66.34	35.96	53.92	39.01
4	31.13	33.17	17.98	26.96	19.50
3	6.23	6.63	3.60	5.39	3.90
2	0.62	0.66	0.36	0.54	0.39
1	0.31	0.33	0.18	0.27	0.20

Chromatographic Conditions

An 1100 Series Agilent (Mississauga, ON, CAN) liquid chromatography system equipped with vacuum degasser, binary pump, autosampler, temperature controlled column compartment and diode array detector was used for quantification of cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-glucoside (C3Gl), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-galactoside (P3Ga), and peonidin-3-*O*-arabinoside (P3Ar) in cranberry samples. The method was adapted from a previously validated HPLC method (Brown & Shipley, 2011). Method modifications focused on shortening the analysis time while maintaining peak resolution by employing a shorter C18 column with smaller particle size and increasing the flow rate to reduce run time to allow for high-throughput analysis. In brief, anthocyanins were chromatographically separated at ambient temperature using a reverse phase Zorbax SB-C18 Rapid Resolution HT column (4.6 x 50 mm, 1.8 µm) with a mobile composition of (A) 0.5% water–phosphoric acid (99.5:0.5, v/v) and (B) water–acetonitrile–glacial acetic acid–phosphoric acid (50.0:48.5:1.0:0.5, v/v/v/v). The optimized gradient program is 9%-36% B over 8.0 min, 36%-60% B over 0.5 min, 60%-80% B over 0.5 min, and 80%-9% B over 0.1 min, hold at 9% B for 1.4 min. The injection volume and flow-rate were 5 µL and 2.1 mL/min, respectively. The analytes were detected at 520 nm with a diode array detector. Data was collected and analyzed using ChemStation software (Rev. B.03.01) from Agilent Technologies (Mississauga, ON, CAN).

Anthocyanin Quantification and Statistical Analysis

Quantification of anthocyanins was accomplished through the use of standard curves obtained through the analysis of the mixed standard solutions prepared as per the procedure described above. All samples were analyzed in triplicate. Single factor ANOVA followed by post-hoc Tukey's HSD analysis ($p = 0.05$) were used to compare the quantities of the individual and total anthocyanins in the samples.

Results

Cranberry Production

The majority of cranberries in BC are harvested starting the third week of September and throughout till the middle of November using a standard "Wet-Pick" method which includes harvest flooding. The beds are partially flooded and the fruit is "beaten" off the vine using a specialized harvester (**Figure 3-1A**); the fruit floats to the surface (**Figure 3-1B**); the lots are then further flooded and the floating fruit is corralled (**Figure 3-1C**) and loaded onto trucks for delivery to the receiving station. Although the wet harvested fruit is primarily used for processed cranberry products like juice and sauce, new specialty product lines are emerging within the cranberry industry such as natural health products. Once harvested, cranberries are field tested using a standardized Brix assay, a gross measure of total anthocyanin contents and moisture contents and all non-cranberry waste or poor quality fruit (rot or damage) are removed. Random testing for pesticide residue is performed prior to shipment and delivery to the processing facility where fruit is frozen awaiting production.

Chemical Reference Materials and Calibration

The purity and concentration of the anthocyanin calibration standards was determined at the time of manufacture and monitored prior to use, as the long term stability of the solutions was unknown. A significant disparity was observed in the initial reported purity and the measured purity for C3Ar and P3Ar in the anthocyanin Snap-N-Shoot® chemical calibration standards. The rate of degradation of each anthocyanin to its corresponding aglycone and sugar moiety varied with C3Gl being most stable and glycosylated peonidins (P3Ar, P3Ga) exhibiting the highest degree of hydrolysis. Previous stability studies (unpublished results) have shown anthocyanins in acidified solvents do not exhibit hydrolysis for up to 7 hours at room temperature if protected from light and are stable at 4 °C for at least 7 days. It was recommended by the contracted manufacturer, Cerilliant Corp. (Round Rock, TX, USA), that the Snap-N-Shoot® chemical calibration standards be tested prior to use. The actual concentration of each anthocyanin in the stock solution was calculated after purity corrections were applied (**Table 3-2**). The analytes of interest were identified by matching the retention time to chemical reference materials (**Figure 3-2**) and quantification by an external calibration curve constructed from calibration standards. A comparison of standard and sample chromatograms demonstrated excellent reproducibility of the method with minimal variance in elution time (**Figure 3-2**). The order of elution was C3Ga (2.62–2.64 min; % RSD 2.9), C3Gl (3.02–3.04 min; % RSD 2.61), C3Ar (3.21–3.23 min; % RSD 2.61), P3Ga (3.78–3.80 min; % RSD 3.70), and P3Ar (4.43–4.49 min; % RSD 2.10). Method validation parameters were published previously and defined the limit of quantification (LOQ) as 10

times the determined method detection level (MDL) (Environmental Protection Agency, 2003) and reported the LOQ to < 0.06 µg/mL (Brown & Shipley, 2011).

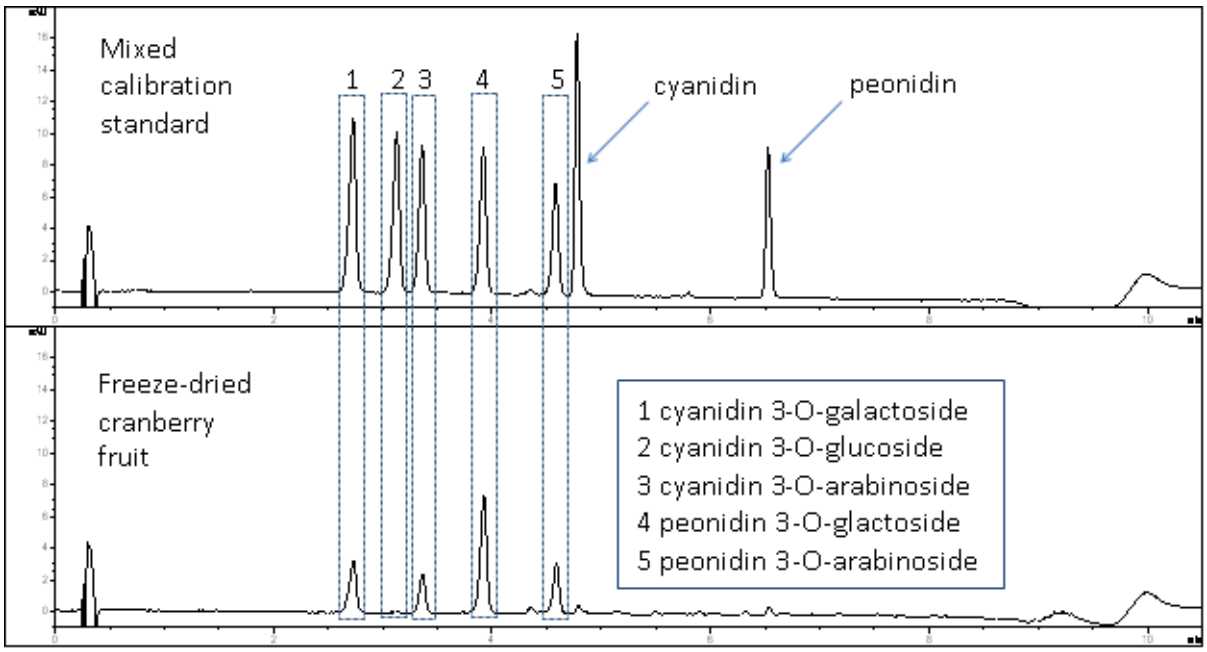

Figure 3-2: Chromatogram showing retention times of anthocyanins at 520 nm for a freeze-dried cranberry fruit sample and a mixed calibration standard.

Quantification of Anthocyanins in Cranberry

Significant differences in anthocyanin content were observed among the 20 samples from the 2009 cranberry harvest. The concentration of C3Ga varied 6 fold from 18.7 ± 1.0 to 3.1 ± 0.23 µmol/g dry weight (DW) (**Figure 3-3A**) while variability in the concentration of C3Ar was slightly less with a 5.1 fold difference in concentrations ranging from 12.7 ± 0.32 to 2.5 ± 0.11 µmol/g DW (**Figure 3-3B**). Similar variability was found in the peonidin derivatives. A nearly 5 fold difference was observed in the P3Ga content ranging from 23.8 ± 0.6 µmol/g DW to 4.8 ± 0.4 µmol/g DW (**Figure 3-3C**) and a 3.7

fold difference was observed in the P3Ar concentrations ranging from 9.6 ± 0.8 µmol/g DW to 2.6 ± 0.12 µmol/g DW (**Figure 3-3D**).

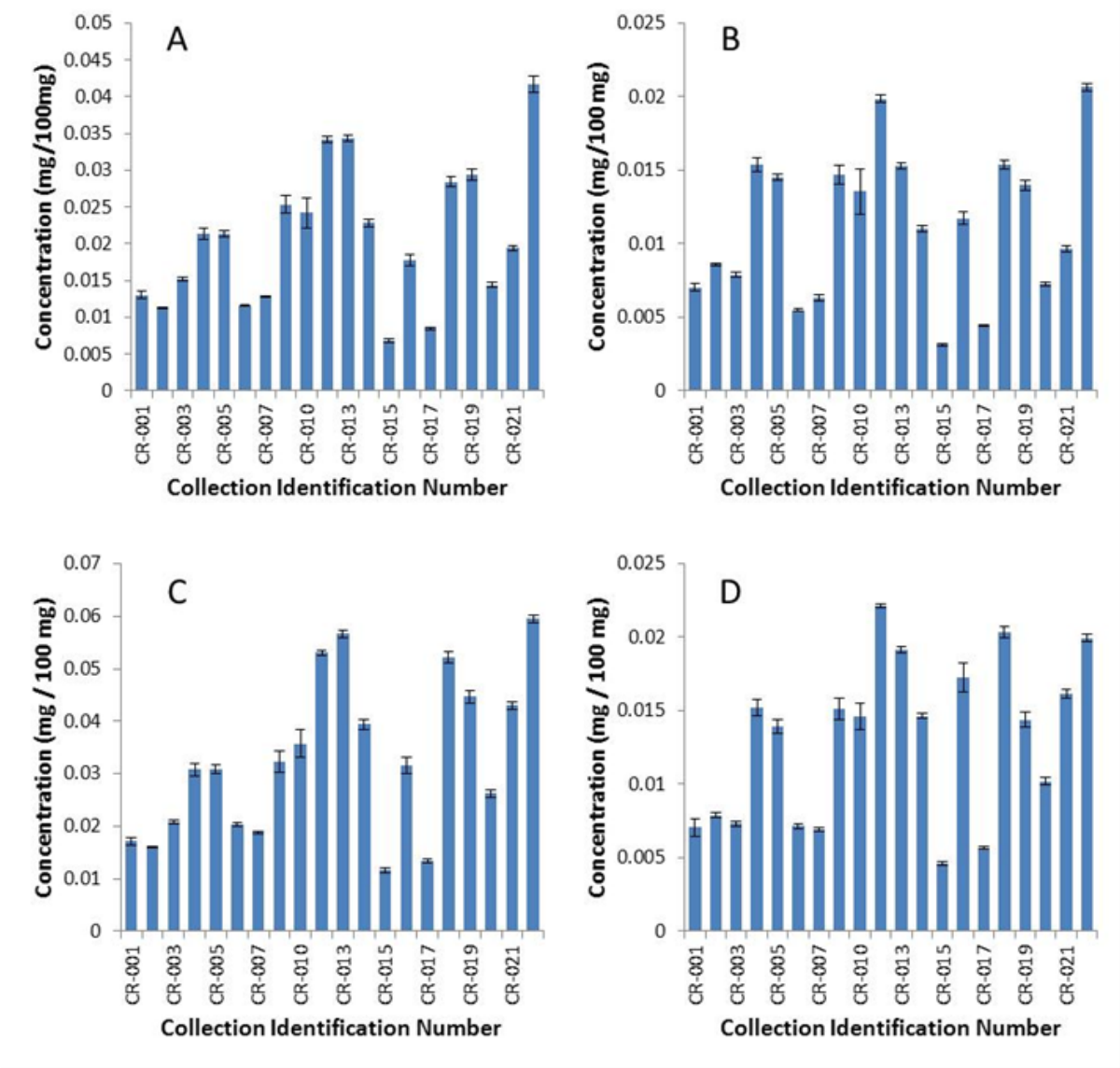

Figure 3-3: Phyto-chemical diversity in samples of *Vaccinium macrocarpon* Aiton collected at mature harvest from farms across British Columbia. **A.** cyanidin-3-*O*-galactoside, **B.** cyanidin-3-*O*-arabinoside, **C.** peonidin-3-*O*-galactoside, **D.** peonidin-3-*O*-arabinoside. Results are expressed as mg/100mg ± standard error of the mean.

When the data for individual collections are expressed as a percentage representing total anthocyanin content, it becomes apparent that individual collection samples vary largely in the sum amount of the measured anthocyanins (**Figure 3-4**).

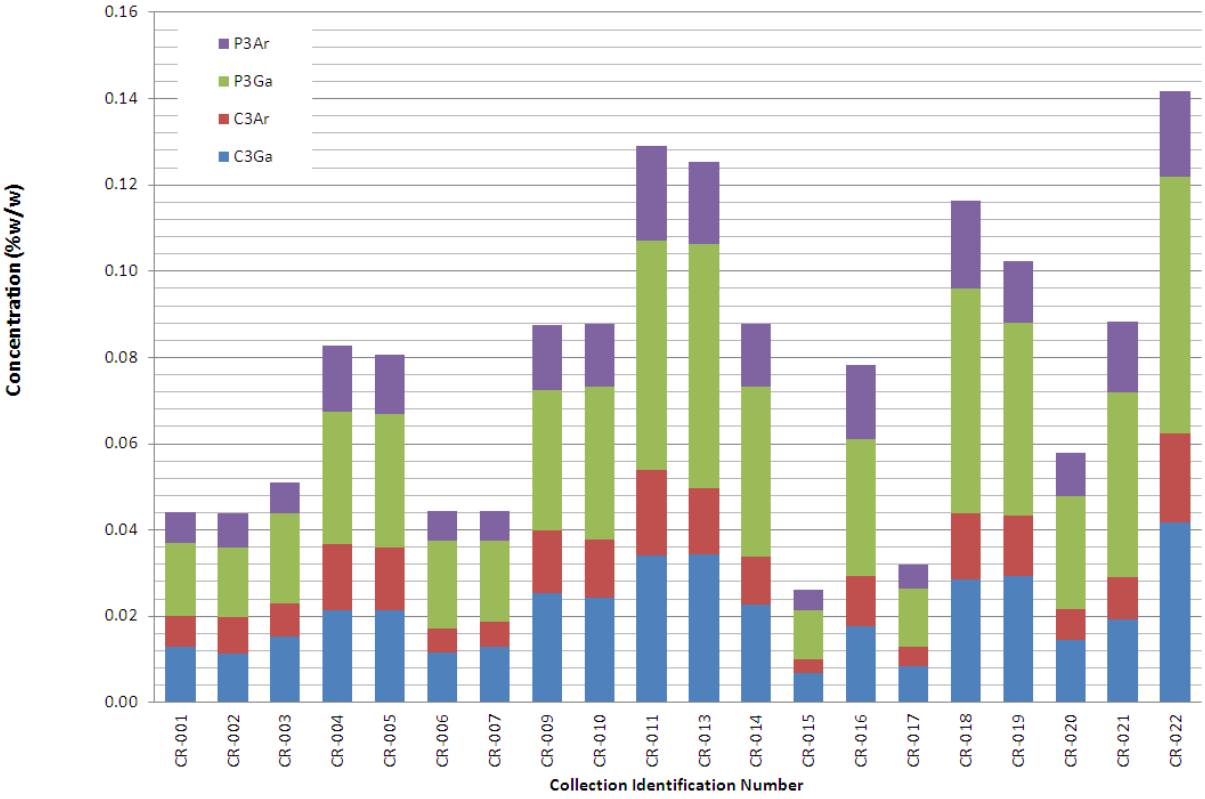

Figure 3-4: Variability in the total % of anthocyanins in cranberry (*Vaccinium macrocarpon* Aiton) berries collected in the fall of 2009 in British Columbia.

Across all of the samples, significantly more P3Ga was quantified than any of the other anthocyanins while the concentration of the two arabinoside-containing anthocyanins was not significantly different (**Figure 3-5**). However, the relative percentages of each of the four measured anthocyanins within each sample were very highly conserved (**Figure 3-5**). A comparison of the molar ratio of the anthocyanins in cranberry provides new insights into the allocation of biochemical resources within the fruit (**Figure 3-6**). The ratio of cyanidin-containing to peonidin-containing anthocyanins was 0.85:1.00,

indicating that peonidin accumulation is metabolically-favoured. Further, the allocation of sugar residues to the anthocyanins favours the galactose moiety over the arabinose with a ratio of 2.17:1.00 for peonidin and 1.19:1.00 for cyanidin (**Figure 3-6**).

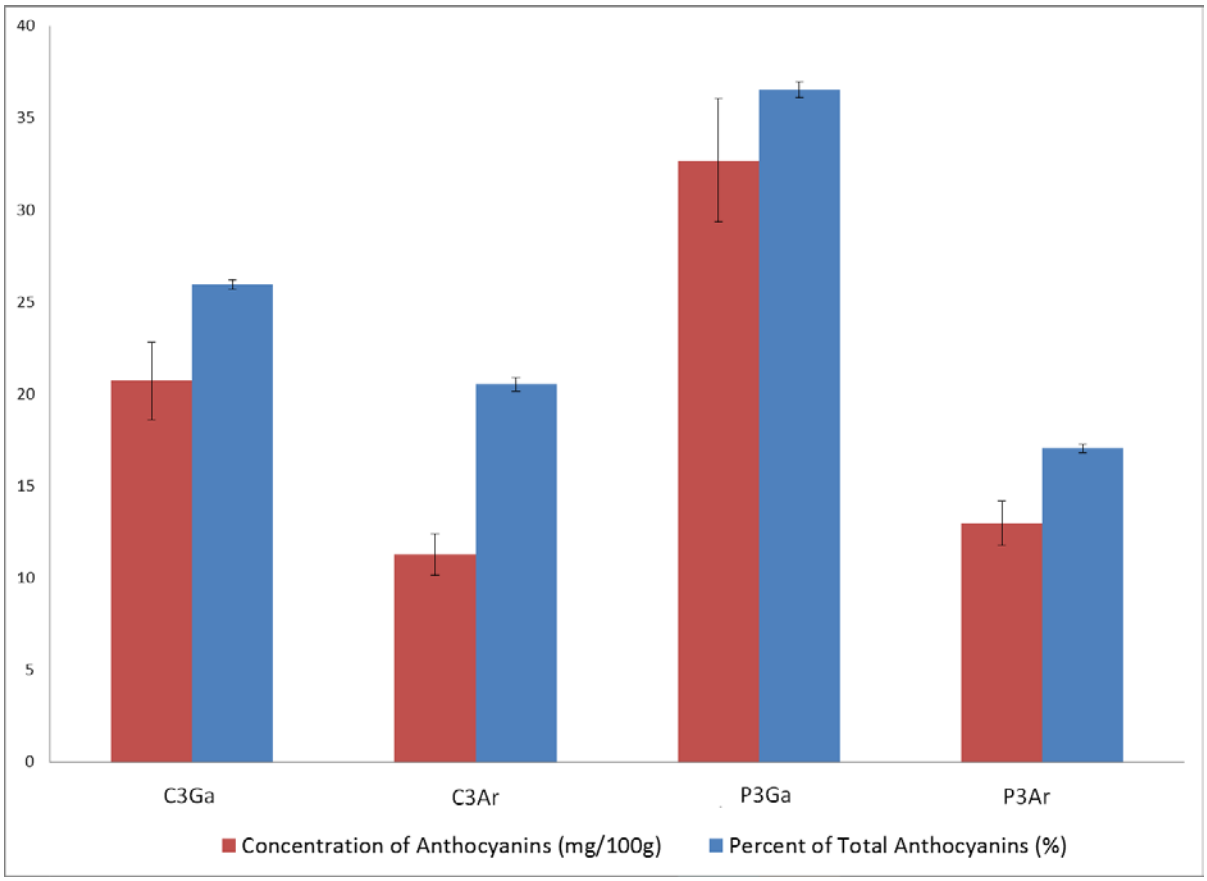

Figure 3-5: Comparison of the individual anthocyanins in (*Vaccinium macrocarpon* Aiton) berries collected in the fall of 2009 in British Columbia, Canada as variability in the concentration of individual anthocyanins in the berries and as variability in the percent of each anthocyanin relative to the total anthocyanin content ± standard error of the mean.

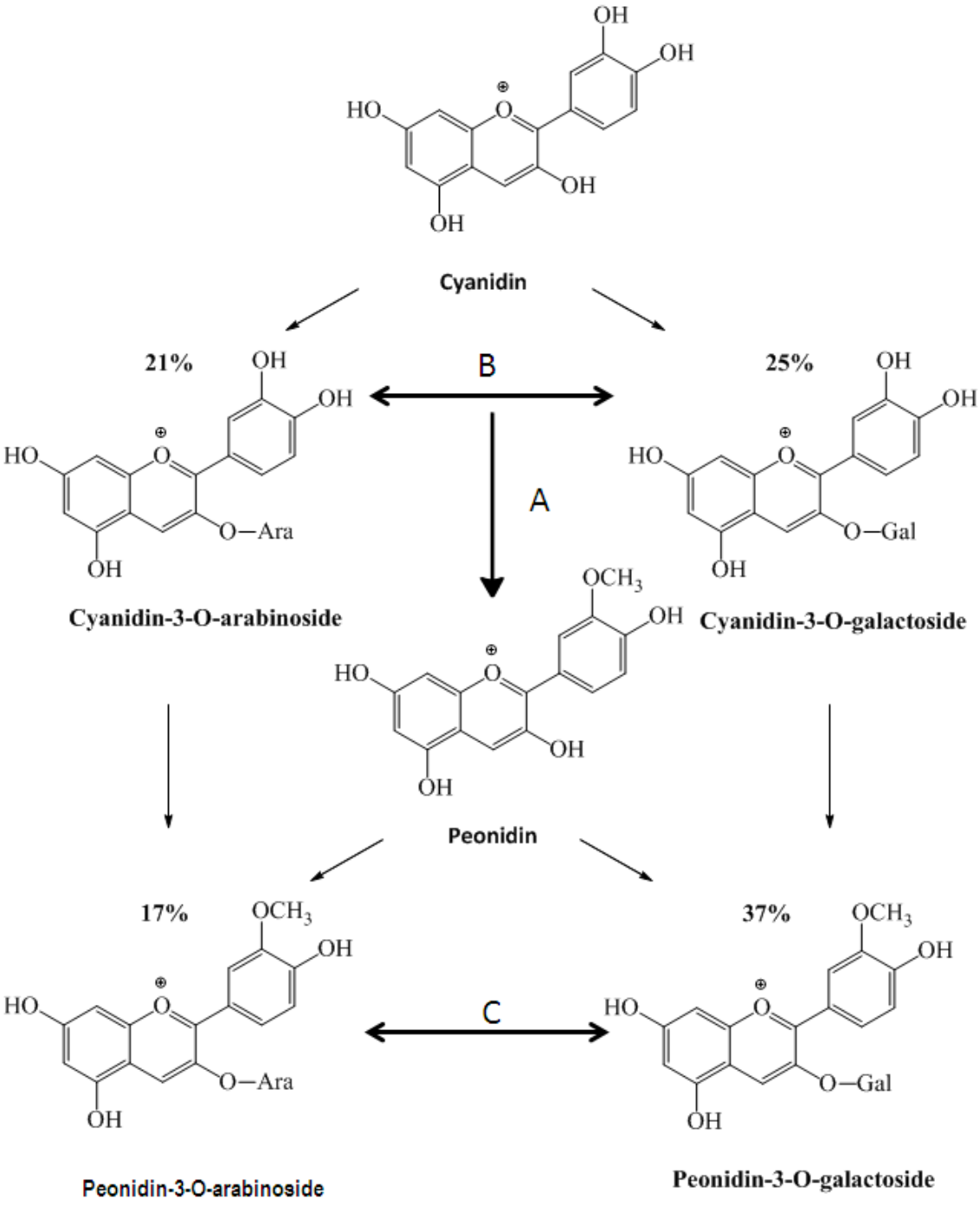

Figure 3-6: Relationships in the observed molar ratios of anthocyanins in cranberry fruit. **A**. molar ratio of cyanidin-containing anthocyanins to peonidin was 88%, **B.** molar ratio of peonidin glycosides with a galactose moiety compared to arabinose was 217%, **C.** molar ratio of cyanidin glycoside with a galactose moiety compared to arabinose was 128%.

Discussion

Cranberries are small, woody vines that stay in commercial fields for more than 25 years and have a rough density of about 400 upright shoots per square ft. (Sandler, 2011). A mature crop (>5 years) of cranberries can yield about 260 100-lb. barrels per acre with minimal fertilizer application under commercial production (Sandler, 2011). Wet-harvested cranberries, 95% of BC's harvest, are primarily used for the processing market, such as juice production (BC Ministry of Agriculture, 2010). Raw cranberries contain approximately 86.5% water and 13.5% solids which include carbohydrates, organic acids, terpenes, sterols and polyphenolic compounds (Howell, 2007; Kuzminski, 1996; Neto, 2007; Pappas & Schaich, 2009). The pressed juice contains from 5-7% solids, the majority of which is sugars and organic acids (BC Ministry of Agriculture, 2010; Cunningham et al., 2004; Ververidis et al., 2007). The optimization of growing and harvesting techniques is essential for the production of a high value crop with adequate profiles of bioactive metabolites.

Anthocyanins of cranberries are important determinants of quality in production and highly investigated bioactive constituents. Recently, there has been growing interest and reports of cranberry anthocyanin content, medicinal potential and metabolism. The most significant findings of the current research are: (a) the variability in total anthocyanin content between different production sites in a cranberry harvest and (b) the highly conserved relative distribution of anthocyanidin skeletons and sugar moieties. A number of different factors affect the anthocyanin content of berries in general and cranberries in particular. Anthocyanin accumulation and the identity of the anthocyanins that are found in specific fruits is highly correlated to exposure of the fruit to UV light, the

microenvironment of fruit production, availability of nitrogen and phosphorous, or water stress (Martens et al., 2010). It is interesting to note that individual collections (CR-011, 13, 18 and 22) had significantly higher concentrations of all 4 anthocyanins in the harvested fruit (**Figure 3-4**). These data may reflect extrinsic differences such as farm location, management practices, and degree of ripeness at harvest.

Anthocyanins are among the estimated 10,000 different flavonoids found in plants. It is thought that while in plant tissues, anthocyanins form a tertiary structure which protects them from hydrolysis and produces colour variations (Clifford, 2000; Yoshida et al., 2009). The main anthocyanin constituents of cranberry have been found to be C3Ga, C3Ar, P3Ga and P3Ar. Previous researchers have shown patterns of distribution and concentrations of cyanidins and peonidins in cranberries that are not significantly different from the current data (Leusink et al., 2010) but the specific biochemical pathways of production of individual anthocyanins remain elusive. There are at least 37 genes that have been identified as playing a role in anthocyanin biosynthesis and metabolism in plants (Martens et al., 2010). In grapes, Arabidopsis and other plants, anthocyanidin synthase (also known as leucoanthocyanidin dioxygenase or LDOX) is considered to be the pivotal enzyme for biosynthesis of anthocyanins from flavonoid skeletons (He et al., 2010; Martens et al., 2010). Our data demonstrate a preferential allocation of galactose as the sugar moiety and peonidin as the anthocyanidin skeleton over cyanidin for accumulation in the cranberry fruit.

In conclusion, these data show the utility of rigorous, validated analytical methods for quantification of the anthocyanin contents of the commercial cranberry crop as well as the variability that can exist between individual growers and producers. The data also

provide an interesting possible avenue for research to determine whether there is preferential biochemical synthesis in the pathway or whether the transport of specific anthocyanins to the developing fruit is favoured in cranberry. The universal distribution of anthocyanins across the cranberry fruit could make it an ideal model system for studies of source-sink relationships in flavonoid metabolism and biochemical studies of plant metabolism.

Chapter 4: Metabolomic Profiling of Cranberry (*Vaccinium macrocarpon* Aiton) Fruit and Cranberry Dietary Supplements by Time-of-Flight Mass Spectrometry with Chemometric Data Analysis

Synopsis

As attempts are being made to understand the relationship between supplementing the diet with cranberry and the prevention and/or treatment of conditions, such as UTI, it is essential to develop a better understanding of the phytochemical determinants of quality. The use of metabolomic profiling combined with multivariate data analysis offers an innovative way to assess cranberry and products derived from cranberry that reflect of the whole biochemical composition. Described as the study of "unique chemical fingerprints" left behind by cellular processes (Daviss, 2005) or untargeted quantitative chemical analysis of the whole array of small molecules contained within a cell or tissue (Dunn et al., 2005; Hall et al., 2005; Wishart, 2008), metabolomic analysis is quickly becoming an important tool for detection of adulterants and the characterization and assessment of commercial herbal products and plant extracts. The objectives of this study were to develop a statistical approach for characterization of cranberry fruit and two cranberry products by analysis of metabolic fingerprints generated from extraction by three different solvents (methanol, 70% aqueous ethanol and water) commonly used in production of cranberry food and supplement products. Characterization of the metabolomic data using comparative and subtractive statistical analyses was used to identify unique compounds within each extract. These techniques compared and

contrasted the metabolomes of the different products and allowed for the determination of their degrees of consistency and difference. These chemometric techniques used to identify unique compounds within the different extracts and subsequently rank these compounds demonstrate a unique chemoinformatic approach for the characterization of cranberry fruit and products derived from cranberry fruit. The use of this approach serves as a unique model for evaluating natural health products.

Experimental

Test Materials

Cranberry fruit, obtained post-farm gate from Ocean Spray Canada Ltd. (Richmond, BC), were grown in peat, and harvested in 2007 using a standard "Wet-Pick" method. The samples were received frozen and maintained at -20° C, then freeze dried (Thermo Scientific Super Modulyo Freeze Dryer; Fischer Scientific) and ground to a <60 mesh (250 µm) powder. Cranberry fruit products were purchased from commercial vendors in Burnaby, BC (see **Table 4-1**).

Metabolite Solvent Extraction Study

The freeze dried, ground cranberry fruit from the 2007 harvest, cranberry finished product 1 and cranberry finished product 2 were weighed n=9 each, 0.259 ± 0.025 g into 15.0-mL conical centrifuge tubes for a total of 27 test samples. Each test material type was extracted in triplicate with 10 mL of the solvent, (methanol, 70% aqueous ethanol, or water) by vortex mixing for 1 min (Thermolyne Maxi Mix 1 Vortex Mixer; Fisher Scientific), sonicating for 20 minutes (Branson Model 3510R-MTH Ultrasonic Cleaner;

VWR), and centrifuging at 5000 rpm (4500 x *g*) for 5 min (Eppendorf Tabletop Centrifuge 15804R, VWR). The supernatant was collected and the solvent evaporated to dryness under nitrogen. The residue was then re-dissolved in methanol:water (80:20), vortexed, and filtered through a 0.45-µm PVDF syringe filter to HPLC vials for analysis.

Table 4-1: Description of cranberry test materials employed in metabolomics study.

Product	**Manufacturer/ Source**	**Product Description Dosage/Serving**	**Ingredients**
Cranberry Fruit	Ocean Spray Canada Ltd.	2007 Harvest, Lower Mainland of British Columbia	Whole cranberry fruit (*Vaccinium macrocarpon* Aiton), freeze-dried, ground to 60 mesh (<250 µm)
Product 1 (Capsules)	Cranberry; Swiss Natural Sources	Cran-Max® cranberry 34:1 fruit extract; Cranberry juice powder 20:1 fruit extract Recommended dose: 1 500 mg capsule daily	Contains: Cran-Max® cranberry (*Vaccinium macrocarpon* Aiton) 34:1 fruit extract (250 mg/capsule); Cranberry juice powder (*Vaccinium macrocarpon* Aiton) 20:1 fruit extract (250mg/capsule)
Product 2 (Caplets)	CranRich Super strength Cranberry Concentrate, Natural Factors	Cranberry concentrate 36:1 (*Vaccinium macrocarpon* fruit juice) Recommended dose: 1-2 500 mg softgels daily	500 mg CranRich cranberry concentrate (*Vaccinium macrocarpon*, fruit juice) 36:1; Softgel (gelatin, glycerin, purified water), rice starch, silica, magnesium stearate

Ultra-Fast Liquid Chromatography with Time-of-Flight Mass Spectrometry

Experiments were performed with an ACQUITY^{TM} series Ultra Performance Liquid Chromatography System (Waters Inc., Mississauga, ON) coupled with a Micromass LCT Premier^{TM} series ToF-MS (Waters Inc., Mississauga, ON) and controlled with a MassLynx V4.1 Data Analysis System (Waters). Chromatographic separation was achieved with a Waters BEH Acquity C_{18} column, 2.1 x 150 mm, 1.7 µm with the following mobile phase conditions: 0.1% aqueous formic acid:acetonitrile (0.0 - 10.0

min, 95:5-5:95 v/v, 10.0-15.0 min, 5:95 v/v, 15.0-20.0 min, 5:95-95:5 v/v, 20.0-25.0 min, 95:5 v/v).

Untargeted Exploratory Data Analysis

Microsoft Excel (Microsoft Corp., Redmond, WA), R (R Foundation, GNU) and Solo+MIA (Eigenvector Research Inc., Wenatchee, WA) were used for data processing and principal component analyses (PCA). The metabolomic data for each product, with blanks removed (subtracted), were assessed without scaling or further data transformation. The data were compiled to an ASCII text file with identifiers of samples in rows (objects) and retention time, m/z ratio and abundance as columns (variables).

Metabolite Counts & Subtractive Metabonomics

A series of automated functions to create subtractive datasets, according to the schematic in **Figure 4-1,** were created in Microsoft Excel using sequential algorithms and functions designated "Subtractive Metabonomics". Although initially labour intensive, the custom macros designed for this process provided detailed information from any metabolite dataset in a relatively short period and could be adapted to a variety of different experimental designs.

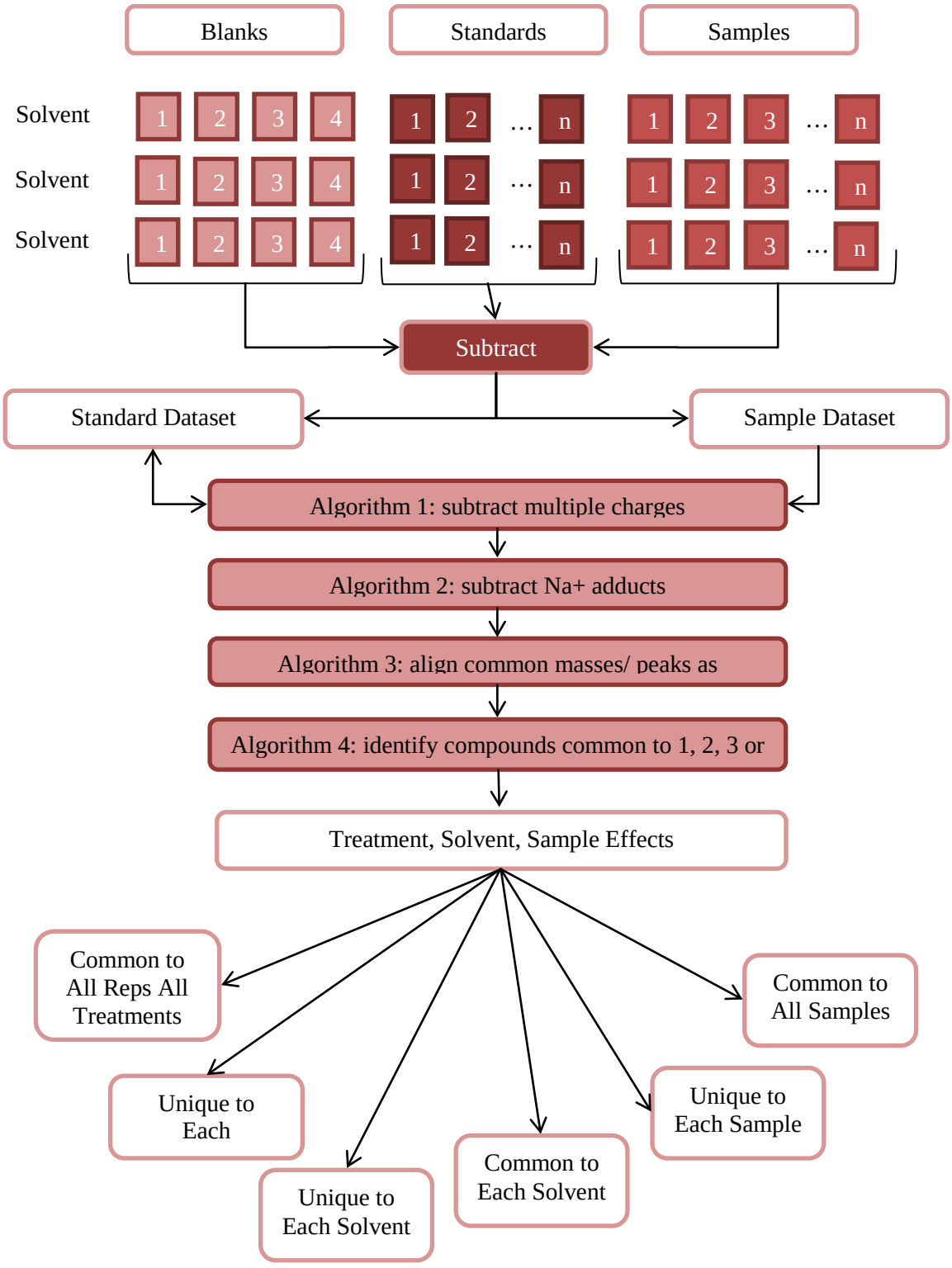

Figure 4-1: Schematic of approach to creating subtractive datasets.

Data Distribution

Probability plots (quantile-quantile or Q-Q) were generated for mean-centered experimental datasets from the stats package available in R using the qqnorm algorithm, which plots calculated quantiles for observed data versus theoretical quantiles for normal distribution of the data (R Development Core Team, 2010). Q-Q plots were generated to compare the mean-centered experimental data distribution to a theoretical normal distribution of data for individual experimental conditions, i.e. by product and solvent. Comparative Q-Q plots were generated to compare directly the probability distribution between mean-centered sets of products within the same solvent system. This was done using the qqplot algorithm in R, which plots calculated quantiles for each dataset being compared.

Principal Component Analysis (PCA)

An unsupervised analysis, PCA, was applied to the data as a means to observe variance for the multivariate dataset by creating covariance matrices with transformation into a coordinate system. The resulting PCA score plot has principal components on each axis that represent a reduction of the number of data dimensions to those, which preserve most of the variance seen in the measured dataset. The associated loadings plot reflects the influence of the variables in the PCA model generated. In all cases, auto scaling was selected for pre-processing before applying the PCA algorithm. PCA score and loading plots were generated for the entire dataset to determine clustering by product type and by extraction solvent employed.

Significance Analysis

Three different approaches were taken to determine the potential significance of data within the entire dataset and subsets of data grouped by extraction solvent. The % explained variance was calculated from the error matrix of up to four principal components from the PCA model, this is akin to interpreting the loadings plot of the PCA analysis directly. While values having the highest % explained variance do contribute maximally to variance and differentiation, it has been demonstrated that seemingly important variables identified by this technique are often found not to be influential when the variable is isolated and the data remodelled (Breiman, 2001; Broadhurst & Kell, 2006). To mitigate the potential of falsely identifying values as significant additional univariate approaches were undertaken to evaluate data significance: the area under the ROC Curve (AUC) with a calculated Kruksal-Wallis *p*-value (Tuszynski, 2010) and the significance analysis of microarrays (SAM) statistic (Tusher et al., 2001).

AUC and Kruksal-Wallis P-Value

Within a binary comparison of products within a solvent category, ROC curves for each model were generated and the AUC for each model computed. The AUC value of each metabolite was obtained from the colAUC algorithm within the caTools package in R (Tuszynski, 2010). The mean of the three calculated binary AUC values per metabolite was obtained and designated as "total AUC" within each solvent category (Hand & Till, 2001). Next, using the tertiary comparison of products of the solvent category, the p-value was derived from the non-parametric Kruskal-Wallis test. Each metabolite was assigned a p-value of the tertiary comparison using the kruskal.test algorithm within R.

Values for which the AUC was ≥ 0.8333 and the associated p-value < 0.05 were identified as significant.

Percent Explained Variance

To determine how well the individual metabolites are described by the PCA model, the % total variance for up to four principal components were calculated from the error matrix derived from the PCA model using R. The derivation of the % variance per metabolite was obtained via the pcaVarexpl algorithm within the chemometrics package in R (Filzmoser & Varmuza, 2010). The calculated % variance was used to rank the measured metabolites and used with PCA to observe the contributed effect of individual metabolites.

Significance Analysis of Microarrays (SAM) Statistic

The SAM statistic (Tusher et al., 2001) and associated false discovery rates (FDR) were calculated using the SAM algorithm from the siggenes package in R for subsets of data grouped by extraction solvent. Due to the limited number of replicates, n=3, permutations were used to set the expected *d(i)* as the null level of abundance which allows for the comparison of the observed *d(i)* based on actual data and the expected *d(i)*. To consider which metabolites differ significantly in terms of abundance, an artificially selected threshold (Δ) was applied to flag metabolites beyond this threshold boundary. The associated false discovery rate (FDR) was determined by starting with the smallest *d(i)* among the metabolites with largest abundance and the least negative *d(i)* among the metabolites with the smallest abundance and counting the number of metabolites that

exceeded the horizontal cut-offs. The count is averaged from all permutations and then divided by the total number of metabolites called significant.

Results and Discussion

Untargeted Exploratory Data Analysis

Metabolite Counts & Subtractive Metabonomics

Evaluation of the metabolomics fingerprint allowed for identification of several interesting patterns within the dataset (**Table 4-2**). Each compound was identified by a unique elution time and m/z signature as well as a relative intensity value. Baseline noise and compounds common to solvent blanks were eliminated from the dataset prior to evaluation (**Figure 4-2**). Just slightly greater than 16,000 distinct compounds were found in the metabolomic profiles of the cranberry and cranberry products with about 7,000 compounds found in all replicates of the fruit regardless of solvent used. About 39% - 50% of the phytochemistry described by the metabolomic fingerprinting was consistent between the cranberry fruit and the two commercial products giving an indication of the amount of plant-based compounds that are conserved in the manufacturing process (**Table 4-2**, lines M4, Et4, W4). In the methanol extracts of cranberry fruit, product 1 and product 2, 2570, 3627, and 3222 compounds respectively, were found only in methanol extracts and not in the other solvents (**Table 4-2**, line M5) and 3,025 of the compounds identified in methanol extracts of cranberry fruit were not found in the methanol extracts of either of the products (**Table 4-2**, line M6). Similarly, chemical fingerprinting of ethanol extracts of cranberry fruit identified 3636 compounds that did not appear in the fingerprints of the products (**Table 4-2**, line Et6) and water extracts

were found to contain 3213 compounds that were not found in either product (**Table 4-2**, line W6). Together, these data help to identify phytochemical constituents of fruit that are retained or lost in the processing and can be used for improved product quality. Further, the identification of a few, high concentration individual compounds that are unique to the cranberry fruit or products provide potential biomarkers for product quality and efficacy (**Table 4-2**, lines M7, Et7, and W7).

Data Distribution

When the distribution of experimental data for each product type, for each extraction solvent, was compared directly to the theoretical distribution, the distribution did not follow a normal distribution pattern, as expected (data not shown). The non-parametric distribution of the data for each experimental condition was found to be unique. All statistical operations for evaluating data variance and significance were selected for suitability of use with non-parametric datasets.

Principal Component Analysis (PCA)

The score plots of the first and second principal component from the PCA model allow visualization of differentiation in the data for all solvents by the product type (**Figure 4-2a**) and for all products by the extraction solvent employed (**Figure 4-2b**). Despite the differences noted from the metabolite counts for each solvent, overall the dataset is more clearly differentiated within the 95% confidence boundaries by product type than by the varied extraction solvents. It is interesting to note that while almost all replicate extractions of product 1, regardless of solvent, fall within the 95 % confidence boundary

Table 4-2: Differences in metabolomics profiles of three different solvent extractions of cranberry fruit and two cranberry finished products.

		Cranberry	Product 1	Product 2
A1	Common to Replicates for All Solvent	7080	6311	6157
Methanol Extracts				
M1	Total # compounds in each methanol extract	9826 ± 473	8593 ± 831	8508 ± 816
M2	Total # compounds extracted by methanol	16475	16005	15148
M3	Total common to all methanol extracts		7635	
M4	% of common chemistry	46	48	50
M5	Only extracted in methanol	2570	3627	3222
M6	Total unique compounds	3025	2028	2034
M7	% Abundance of unique compounds (% of TIC)	2.9	5.0	3.7
70% Ethanol Extracts				
Et1	Total # compounds in each ethanol extract	7445 ± 3446	7019 ± 453	6864 ± 610
Et2	Total # compounds extracted by ethanol	14944	13072	13384
Et3	Total common to all ethanol extracts		5832	
Et4	% of common chemistry	39	45	44
Et5	Only extracted in ethanol	2035	1974	2420
Et6	Total unique compounds	3636	2220	2604
Et7	% Abundance of unique compounds (% of TIC)	5.9	3.7	15.2
Water Extracts				
W1	Total # compounds in each water extract	7959 ± 630	7061 ± 254	6704 ± 296
W2	Total # compounds extracted by water	14710	13403	12746
W3	Total common to all water extracts		5664	
W4	% of common chemistry	39	42	44
W5	Only extracted in water	2513	2444	2231
W6	Total unique compounds	3213	2194	2698
W7	% Abundance of unique compounds (% of TIC)	4.2	4.1	4.6

around the cranberry fruit extracts, there is significantly less overlap between the cranberry fruit and product 2 (**Figure 4-2a**).

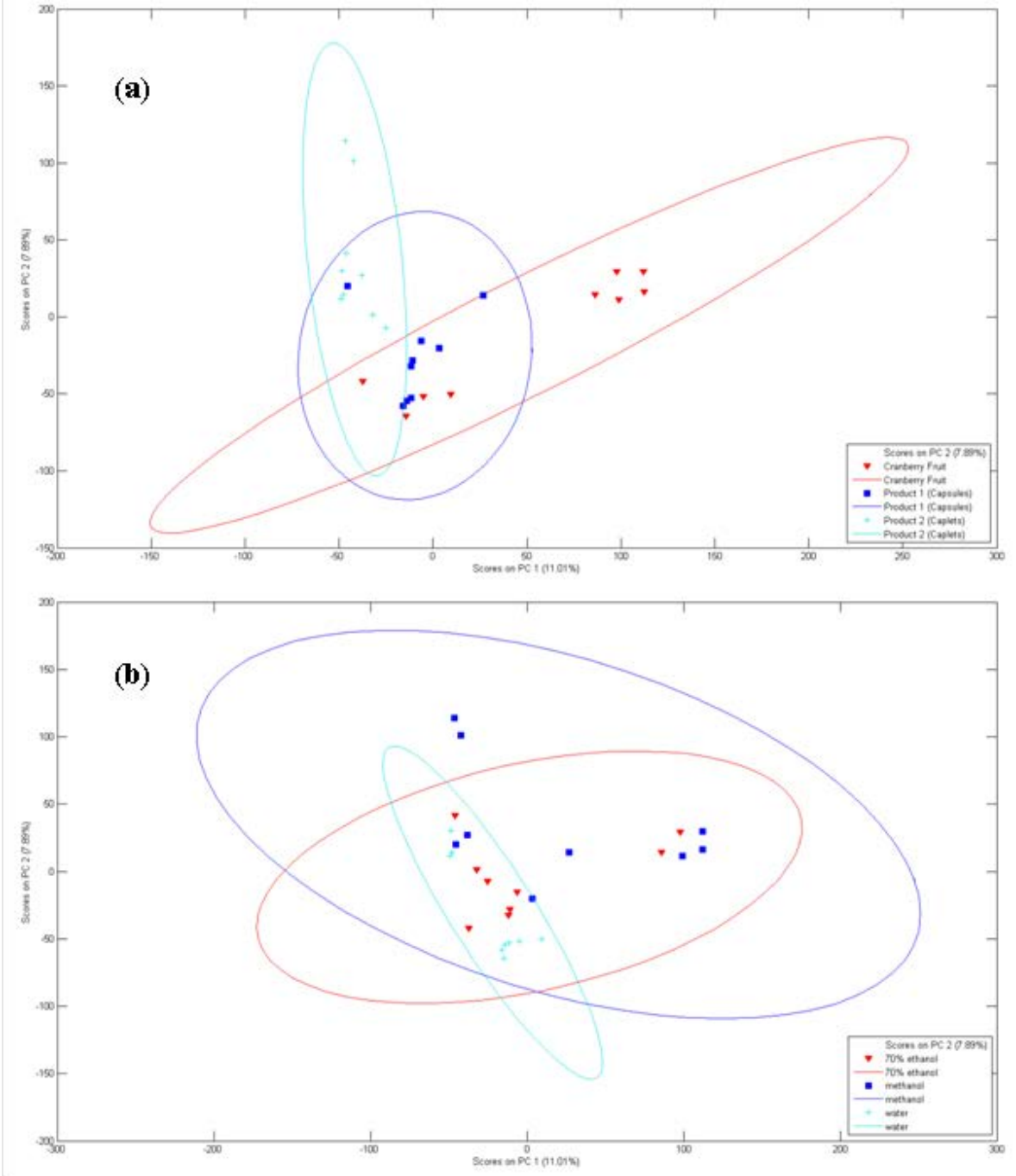

Figure 4-2: PCA score plots generated for UFLC-ToF-MS profiles of extracted freeze-dried cranberry fruit. Product 1 and Product 2 in methanol, 70%-aqueous ethanol and water, grouped by **(a)** product type and **(b)** extraction solvent.

Significance Analysis

AUC and Kruksal-Wallis P-Value

The importance of compounds in discriminating cranberry fruit from the two-finished product forms was evaluated by comparing the total AUC and the associated statistical *p*-value of the data sets grouped by extraction solvent. The m/z value for compounds with *p*-values <0.05 and total AUC ≥ 0.8333, which is a reflection of distribution across the product types and replicates within a product, are presented in **Tables 4-3** to **4-5.** When comparing data from a given extraction solvent from the fruit and two finished products a total AUC of 1 reflects two distribution scenarios; either the compound exists in all product types with up to two absent replicates in a product type or the compounds exists in two product types but with differentiation in abundance. For the other AUC values represented, with a *p*-value up to 0.0496, the compound is present in all or up to two product types with the AUC value being dependent on absence in replicates within product types and/or similarity in abundance levels between product types.

For all product types extracted with methanol (**Table 4-3**) 688 compounds were found to have a total AUC of 0.8333 and *p*-value of 0.0221, which represented the case, where the compound was detected in only one of the 3 product types. There were three sets of p-values associated with m/z values (144 compounds) that had a total AUC of 1. At a *p*-value of 0.0273, there were 57 compounds of which 35 were found in all replicates of all products at abundance levels distinguishable per product type. The remaining 22 compounds with the same *p*-value and total AUC also existed in all products, but were found to be undetectable in a single replicate of one of the products. The 38 compounds found to be in all product types, but undetected in 2 replicates for one of the products

resulted in a *p*-value of 0.0265. Compounds found to be present in 2 of the 3 product types, of which there were 49, had a *p*-value of 0.0241 and total AUC of 1. In the case where the compound was found only in the finished products, as opposed to the cranberry fruit it is possible the compounds are either an artefact of the processing or are concentrated during the manufacturing process and are actually present in the cranberry fruit but below the detection limit.

The products extracted with 70% aqueous ethanol, shown in **Table 4-4**, contained 254 compounds with a total AUC of 0.8333 and *p*-value of 0.0221. In every case, these compounds were found to present in only one of the three product types. The 56 m/z values that had an associated total AUC of 1 are separated into 3 groupings by p-value. The 20 compounds present in products 1 & 2, but absent in the cranberry fruit, and one compound (m/z 533.0886) found only in cranberry fruit and product 1 all had a *p*-value of 0.241. The 17 m/z values with a p-value of 0.0265 were all compounds that were found in all product types but absent in 2 replicates of a given product type; specifically 2 of 17 were found absent in 2 replicates of product 1 and 15 of 17 were absent in 2 replicates of cranberry fruit. Those values with a total AUC of 1 and *p*-value of 0.0273 are of particular interest as those 5 compounds (m/z values: 841.7087, 796.728, 523.3578, 501.3721, 439.0192) were present in all replicates of product types at abundance levels distinguishable for each product type which merits further characterization. The remaining 14 of 19 compounds with a p-value of 0.0273 existing in all replicates of products 1 and 2, but were absent in one replicate of the cranberry fruit.

Table 4-3: The m/z values of compounds identified as important based on a total AUC ≥ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in methanol.

Methanol extracts			
Kruskal-Wallis p-value	Total AUC	# of metabolites	M/Z Value
0.0221	0.8333	688	-
0.0241	1.0000	49	-
0.0265	1.0000	38	-
0.0273	1.0000	57	-
0.0329	0.9444	1	566.361
0.0329	0.9444	1	568.35
0.0340	0.9444	235	-
0.0347	0.8889	516	-
0.0349	0.9630	102	-
0.0379	0.9630	43	-
0.0390	0.9630	56	-
0.0394	0.9444	1	859.251
0.0439	0.9444	1	667.399
0.0447	0.9259	1	763.558
0.0459	0.9259	163	-
0.0484	0.9259	1	772.31
0.0496	0.9259	34	-
	Total #	**1987**	

Evaluating the data from the water-extracted products (**Table 4-5**) resulted in 488 compounds at a *p*-value 0.0221and a single compound (m/z value 863.3893) at a *p*-value of 0.0211 all with a total AUC of 0.8333 that reflects the presence of these compounds in only one of the three product types. There were four groupings of compounds found to have a total AUC of 1. At a *p*-value of 0.0234 was a single compound, m/z 100.0756, present at distinguishable levels in only cranberry fruit (0.11±0.3) and product 2 (0.25±0.08). The group of compounds at AUC 1 had a p-value of 0.0241 and represented

the presence of each compound in 2 of 3 product types. With a *p*-value of 0.0265, there were 100 compounds in the third grouping where each compound was found in all products but in one of the product types was only found in 1 of the replicates. The last compound grouping, 120 compounds in total with a *p*-value of 0.0273, reflects two scenarios of distribution; either present in all replicates of all product types (78 compounds) or found in all products but not detected in a single replicates of the one of the product types (35 compounds).

Table 4-4: The m/z values of compounds identified as important based on a total AUC ≥ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in ethanol.

70% aqueous ethanol extracts			
Kruskal-Wallis p-value	Total AUC	# of metabolites	M/Z Value
0.0221	0.8333	254	-
0.0241	1.0000	20	-
0.0265	1.0000	17	-
0.0273	1.0000	19	-
0.0340	0.9444	97	-
0.0347	0.8889	198	-
0.0349	0.9630	6	-
0.0379	0.9630	14	-
0.0390	0.9630	16	-
0.0447	0.9259	1	512.371
0.0458	0.8704	1	369.228
0.0458	0.8704	1	293.021
0.0459	0.9259	57	-
0.0496	0.9074	1	463.024
0.0496	0.9259	14	-
	Total #	**716**	

Table 4-5: The m/z values of compounds identified as important based on a total AUC ≥ 0.8333 and with a $p < 0.05$ when comparing across all product types extracted in water.

Water extracts			
Kruskal-Wallis p-value	Total AUC	# of metabolites	M/Z Value
0.0211	0.8333	1	863.3893
0.0221	0.8333	488	-
0.0234	1.0000	1	100.0756
0.0241	1.0000	100	-
0.0265	1.0000	71	-
0.0273	1.0000	120	-
0.0336	0.8889	1	437.1010
0.0340	0.9444	130	-
0.0347	0.8889	376	-
0.0349	0.9630	37	-
0.0379	0.9630	29	-
0.0390	0.9630	61	-
0.0447	0.9259	1	367.2658
0.0458	0.8704	1	409.1991
0.0459	0.9259	104	-
0.0496	0.9259	35	-
	Total #	**1556**	

Percent Explained Variance

Similar to the practice of pulling data off a PCA loading plot, the total % explained variance could be calculated to identify those data points (observations) which contributed maximally to the overall variance within the dataset. Across the entire dataset of all product types and solvent, 103 compounds contributed to over 90% explained variance in the first and second principal component, while there were a total of 12697 compounds that had less than 10% explained variance. This indicates that 49.4 % of the compounds found in cranberry fruit metabolome are conserved for products manufactured into dosage forms. While often only those metabolites identified as

"unique" and contributing maximally to the variance within datasets are paid much interest, but as noted by Broadhurst and Kell (2006) there must be some consideration taken that false discoveries can be made. It is possible that a compound found only in one replicate within an experiment could be attributed an artificially high explained variance based on its absence from other datasets. It is for this reason other treatments to prioritize the metabolomic data were sought.

The % explained variance in the first and second principal components when comparing all product types was calculated for subset of data grouped by extraction solvents. In **Table 4-6** are listed the m/z values that exhibit the highest % explained variance in each data subset along with the associated total AUC and *p*-value. The m/z values with *p*-values >0.05, suggesting they are not significant, are highlighted in **Table 4-6** and account for half of the identified values in methanol and 8 of 10 m/z values with high % explained variance in 70% aqueous ethanol. This suggests that relying solely on % explained variance may not be the best approach for identifying significant compounds in the metabolomics data. Interestingly, the m/z values identified as having the highest % explained variance in the water-extracted products were all were found to be significant with respect to total AUC and *p*-value.

SAM Statistic

When using permutations, it is important estimate the percentage of metabolites identified by chance, the false discovery rate (FDR). As the threshold Δ decreases, the number of significant metabolites increases (closer to the value of $d(i)_{actual}=d(i)_{expected}$), but this is at the cost of increasing FDR. For every comparison of the 3 product types by

Table 4-6: The m/z values corresponding to the data with the highest % explained variance in the first two principal components when comparing all product types in each of the extraction solvents and the total AUC and *p*-values associated with each m/z value (variable).

Description of Occurance	M/Z Value	explained variance	p-value	Total AUC
Methanol extracts				
Observed in fruit	845.7227	99.33%	0.0221	0.8333
Observed in fruit & Product 2	248.9951	99.36%	0.0241	1.000
Observed in Product 1 & Product 2	707.1908	99.32%	0.0625	0.8519
	392.2108	99.51%	0.0625	0.8519
	279.029	99.43%	0.0347	0.8889
	251.0727	99.55%	0.0625	0.8519
Observed in all test materials	517.3313	99.32%	0.0509	0.9259
	443.3862	99.44%	0.0273	1.0000
	346.0445	99.42%	0.0265	1.0000
	273.037	99.51%	0.0728	0.8889
70% aqueous ethanol extracts				
Observed in fruit & Product 1	757.1484	99.33%	0.1258	0.8333
	190.0868	99.20%	0.1988	0.7593
Observed in fruit & Product 2	412.1059	99.16%	0.0347	0.8889
Observed in Product 1 & Product 2	290.1213	99.30%	0.0347	0.8889
Observed in all test materials	613.0269	99.26%	0.1825	0.8148
	551.0988	99.21%	0.4128	0.7222
	508.2533	99.24%	0.0608	0.8889
	433.1109	99.18%	0.3319	0.7593
	308.1311	99.22%	0.0605	0.8519
	269.2235	99.23%	0.5760	0.6481
water extracts				
Observed in fruit & Product 1	697.0822	99.44%	0.0241	1.0000
Observed in Product 1 & Product 2	769.2545	99.50%	0.0241	1.0000
	294.1154	99.56%	0.0340	0.9444
Observed in Product 2	933.2276	99.46%	0.0221	0.8333
	819.1052	99.48%	0.0221	0.8333
	799.2535	99.48%	0.0221	0.8333
	788.1461	99.46%	0.0221	0.8333
	748.613	99.54%	0.0221	0.8333
	318.9778	99.52%	0.0221	0.8333
	310.1404	99.62%	0.0221	0.8333

a common extraction solvent, a list of significant m/z values is given with an associated FDR in **Table 4-7**. In comparing the m/z values identified as significant by the SAM statistic (FDR 10%) to the associated % explained variance of those variables for the first two principal components, only m/z 845.7227 found in the methanol extracted cranberry fruit was consistently identified as significant. The other seven m/z values from the methanol extracts, when comparing all three product types, all had explained variance of >98.9%. Further, all of the compounds in the methanol dataset identified as significant by the SAM statistics had total AUC values ranging from 0.8333-1.0000 and associated *p*-values of <0.05.

With a calculated false discovery rate of 15%, there were 17 m/z values identified as significant in the metabolomics profiles from the 70% aqueous ethanol extracts. While the majority of these identified compounds has an explained variance ranging from 48.8-58.5%, three m/z values found only in product 1 (885.5636, 881.1926, 841.2387) and one found in all three test materials (326.3762) had only about 29% explained variance in the first two principal components and as such would not be considered important or influential with respect to covariance in the PCA model. The total AUC and p-values for each of the m/z values identified as significant by the SAM statistic are also found in **Table 4-7**. Unlike with the methanol dataset there were two m/z values (326.3762, 387.1789) identified as significant by SAM in the 70% aqueous ethanol dataset that had p-values of >0.05, which suggests data is not statistically significant with respect to distribution across replicates.

Table 4-7: The m/z values identified as significant using the SAM algorithm in R to compare across products types in each extraction solvent with the calculated *p*-value, total AUC and % explained variance of the variable.

Description of Occurance	M/Z Value	*d(i)*	p-value	Total AUC	% explained variance
Methanol extracts: calculated false discovery rate = 10%					
Observed only in cranberry fruit	845.7227	1772	0.0221	0.8333	99.33%
	779.435	1697	0.0221	0.8333	98.94%
	343.0365	2357	0.0221	0.8333	99.14%
	261.1252	3824	0.0221	0.8333	99.26%
	222.1465	6082	0.0221	0.8333	98.92%
Observed in all test materials	954.6111	2131	0.0273	1.0000	99.27%
	247.1691	3755	0.0459	0.9259	99.03%
70% aqueous ethanol extracts: calculated false discovery rate = 15%					
Observed only in cranberry fruit	318.3207	328	0.0221	0.8333	51.81%
Observed only in Product 1	885.5636	326	0.0221	0.8333	29.14%
	881.1926	638	0.0221	0.8333	29.05%
	841.2387	303	0.0221	0.8333	29.47%
Observed only in Product 2	966.5259	275	0.0221	0.8333	57.77%
	864.771	351	0.0221	0.8333	46.73%
	449.8076	353	0.0221	0.8333	58.54%
	428.2075	271	0.0221	0.8333	47.07%
	420.268	1786	0.0221	0.8333	48.80%
	266.1803	432	0.0221	0.8333	55.26%
Observed in all test materials	841.7087	1307	0.0273	1.0000	44.75%
	387.1789	2730	0.0549	0.9074	50.58%
	326.3762	274	0.0605	0.8519	28.61%
	261.2195	1101	0.0459	0.9259	49.36%
Observed only in Products (1 & 2)	417.2127	400	0.0347	0.8889	49.34%
	219.2105	311	0.0347	0.8889	47.85%
Observed only in cranberry fruit and Product 2	475.3222	308	0.0347	0.8889	55.17%
Water extracts: calculated false discovery rate = 7%					
Observed only in Cranberry fruit	739.5641	1795	0.0221	0.8333	76.19%
	613.3321	7122	0.0221	0.8333	74.85%
	200.187	3797	0.0221	0.8333	71.52%
	194.1179	2066	0.0221	0.8333	71.48%
Observed only in Product 1	549.1552	3807	0.0221	0.8333	72.04%
Observed only in Product 2	955.2333	1895	0.0221	0.8333	98.09%
	329.9232	4716	0.0221	0.8333	98.31%
Observed in all	276.1068	4076	0.0552	0.8889	97.89%
Observed only in Products (1 & 2)	559.1321	4232	0.0347	0.8889	98.08%

From the water extracted metabolomic profiles of all products types, the SAM statistic identified nine variables as significant at a false discovery rate of 7%. Four of these variables were found exclusively in the cranberry fruit and had m/z values with 71.5-76.2% explained variance. Only one significant variable was found in product 1 only, m/z 549.1552, which had an explained variance of 72.0%. The remaining four significant variables (m/z values), with about 98% explained variance, were found in product 2 only (955.2333, 329.9232), in both finished products (559.1321) and in all test materials (276.1068). However, the compound identified as significant and present in all materials (m/z 276.1068) had an associated *p*-value > 0.05 and the calculated total AUC was 0.8889 which is reflective of the compound not being found consistently in the replicates of each products type. Specifically the abundance, while present in consistent levels for all replicates of product 2, was found at a much lower abundance in only 1 of 3 replicates for the fruit and 2 of 3 replicates for product 1.

Applying the SAM statistic to the metabolomic data provides a way of determining significance with respect to the abundance of each compound detected when comparing across products. Unlike the % explained variance, which is based on the total covariance of each compound in the PCA model, the SAM statistic incorporates a measure of deviation within replicates when comparing the abundance across the products in the dataset. In this way, this significance test acknowledges the potential for false discovery of significant components and the threshold can be set to reduce the false discovery rate making this a very useful technique for prioritizing metabolites in a metabolomics data set for further study.

The 2008 Cochrane Review on the effectiveness of cranberry products in preventing UTIs considered 10 clinical studies; 6 focused on cranberry juice, 3 on cranberry finished products and 1 included both product types (Jepson & Craig, 2008). In the study conducted by Stothers *et al.* (Stothers, 2002) both cranberry juice and cranberry tablets were found effective, as compared to placebo, in decreasing the number of patients experiencing at least one symptomatic UTI per year. Although the Cochrane Review concluded there is some evidence that cranberry juice decreases the number of symptomatic UTIs, the authors further concluded that the optimum dosage and method of administration (e.g. juice, tablets, capsules) was not clear (Jepson & Craig, 2008). With the wide variety of products in the marketplace and limited amount of product information provided in clinical studies (Gagnier et al., 2006; Wolsko et al., 2005), it is important that tools are developed to better characterize these complex products, ensuring their quality and efficacy. In this study, metabolomic profiles were generated by UFLC-ToF-MS for cranberry fruit and two commercial cranberry-based products as a model system to explore the potential of metabolomics profiling with chemometric analysis to evaluate the complex phytochemistry of natural products.

The use of chemometric and statistical approaches for prioritizing data or determining the significance of compounds in the metabolomics profiles may facilitate the determination of compounds which are important to the quality and efficacy for products. Comparing the variables identified in the PCA model, which is designed to find patterns in large multivariate data sets, with the highest % explained variance and being the most influential in the model with variables identified in other approaches that determine significance can help mitigate the occurrence of identifying false positives. Examination

of the total AUC and calculated Kruksal-Wallis *p*-value provides information on the distribution of the data across products and between replicates, while the SAM statistic ranks significance based on the abundance of metabolites across a data set with consideration given to the deviation within the replicates. Thus, finding significant compounds that also exhibit high % explained variance in PCA model would increase confidence in such compounds being important in a given product or plant material. Additionally, the use of total AUC values measures the ability of the UFLC-ToF-MS metabolomic profile to classify correctly a compound to the products under study. This interpretation, along with the p-value reflect the significance of the abundance distribution across products and helps establish a starting point to understand how the metabolome of each product is related to each other and create a pathway to finding compounds associated with product efficacy.

Our objective was to develop and evaluate a set of statistical tools for analysis of mass spectrometry-based metabolomics of natural products using cranberry as a model. Overall, these data and analyses demonstrate that natural products metabolomics should carefully consider the chemodiversity of products as well as the potential for false discoveries when interpreting models generated from multivariate data. Utilizing a combination of approaches to identifying compounds of significance allows for the development of a more robust approach to interpreting and utilizing metabolomic profiling as a quality control tool and for identifying compounds potentially important to efficacy.

Chapter 5: Phytochemical Diversity of Cranberry (*Vaccinium macrocarpon* Aiton) Cultivars by Anthocyanin Determination and Metabolomic Profiling with Chemometric Analysis

Synopsis

The current study was designed to characterize and quantify the phytochemical diversity among five cranberry cultivars commonly grown in British Columbia, Canada: Ben Lear, Stevens, Pilgrim, Bergman and GH1. Ben Lear, a native landrace selection from Wisconsin domesticated around 1900, ripens early and has very good productivity but is susceptible to rot and other diseases (Agriculture Research Service, 2009). The Pilgrim, Stevens and Bergman cultivars are all the result of a USDA cranberry breeding program started in 1929 (Dana, 1983; Hancock et al., 2008). The cultivated culivar 'Stevens' is a cross between the native McFarlin and Potter cultivars and has become one of the most widely used cranberry varieties in cultivation (Hancock et al., 2008; N. Vorsa, 1994). Pilgrim, a cross of McFarlin and Prolific, and Bergman, a cross of Early Black and Searles, are also being grown throughout North America (Agriculture Research Service, 2009; Vorsa, 1994). GH1 is a highly productive cross of Rezin Mcfarlin and Searles that was developed by Ed Grygleski and released in 2004 (Agriculture Research Service, 2009). The breeding origins and relationship between the 5 cultivars collected, along with an estimate of production percentages in BC are depicted in **Figure 5-1**. The specific objectives of the study were to quantify the biological variation of anthocyanin content within and across the cultivars by a previously validated high performance liquid chromatography with ultraviolet (HPLC-UV) detection method (Brown & Shipley,

2011) adapted to rapid resolution chromatography (Brown, Shipley et al., n.d.) and to develop a model approach that describes phytochemical diversity and relationships between cultivars through Ultra Performance Liquid Chromatography with Time of Flight Mass Spectrometry (UPLC-TOF-MS) metabolomic profiling. A systematic examination of UPLC-TOF-MS metabolomic profiles from commonly produced cultivars of *Vaccinium macrocarpon* was undertaken in a process referred to as "subtractive metabonomics" (Brown, Finley et al., n.d.; Brown et al., 2011). This comparison of metabolites from multiple replicates and related biological systems or populations, moves beyond characterization at a singular timepoint to support identification of relationships across population metabolomes, in this study, cultivars of *Vaccinium macrocarpon*. Multivariate data analysis was employed to assist in evaluation of relationships identified by comparing the metabolomes and visualization of clustering across the cultivars. Different statistical tools were applied to the data set to evaluate data significance and the robustness of the multivairate models developed. This "chemoinformatic" approach introduces a process by which incorrect designation of metabolites as significant and potential misclassifications of relationships can be minimized when considering metabolomic data and offers a new model for interpreting metabolomics data and describing metabonomic changes in the overall phytochemistry of cranberry cultivars (Broadhurst & Kell, 2006; Enot & Draper, 2007).

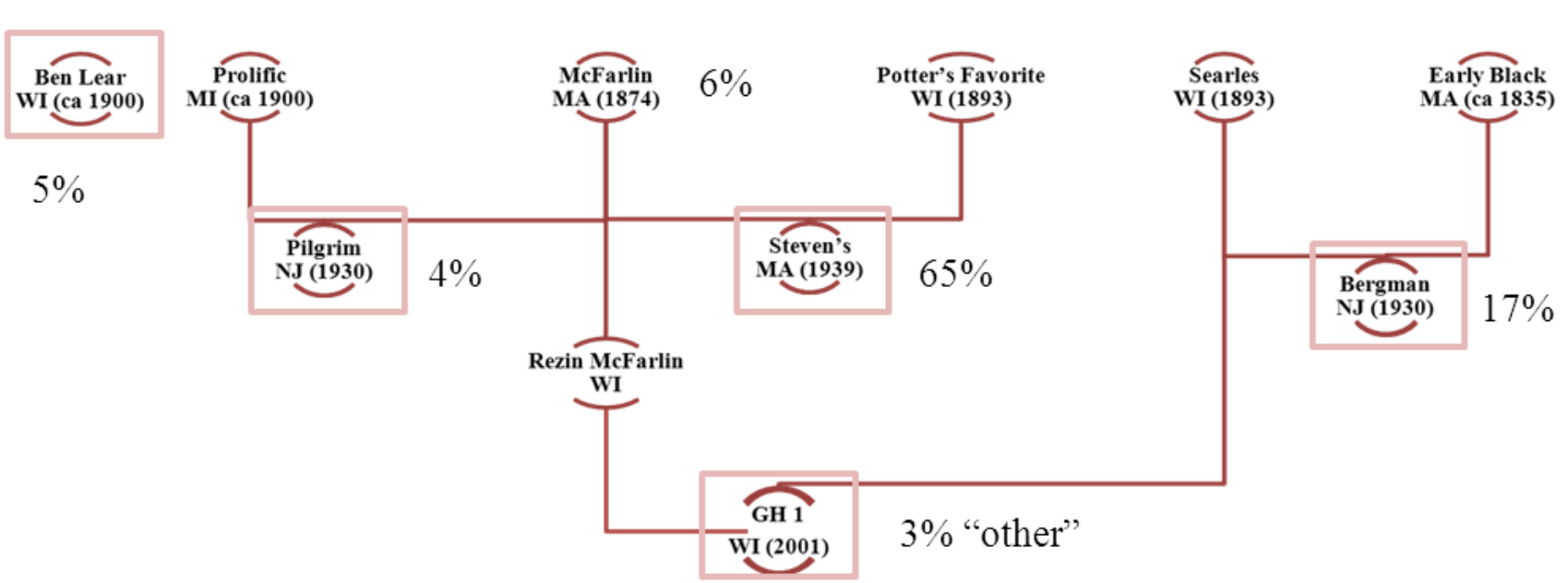

Figure 5-1: Breeding map showing origins of the 5 cultivars of *Vaccinium macrocarpon* Aiton being studied in relation to other common cultivars, along with the estimated production percentages of each cultivar in BC.

Materials and Methods

Plant Materials

Vaccinium macrocarpon Aiton samples were collected from the Lower Mainland of British Columbia from five sites (**Figure 2**) on Oct. 16 of 2010 : Ben Lear (N49°11.861', W123°02.557'), Pilgrim (N49°10.944', W123°00.759'), Stevens (N49°11.065', W123°00.523'), Bergman (N49°11.760', W123°02.109'), and GH1 (N49°11.869, W123°02.564'). Cranberry production at all sites was under similar growing conditions (peat) and utilized the same watershed. For all cultivars, the berries were collected immediately prior to the scheduled wet harvesting of the site. Berries were frozen within 4 hours of collection and stored at -20 ºC until use in chemical and metabolomic analyses.

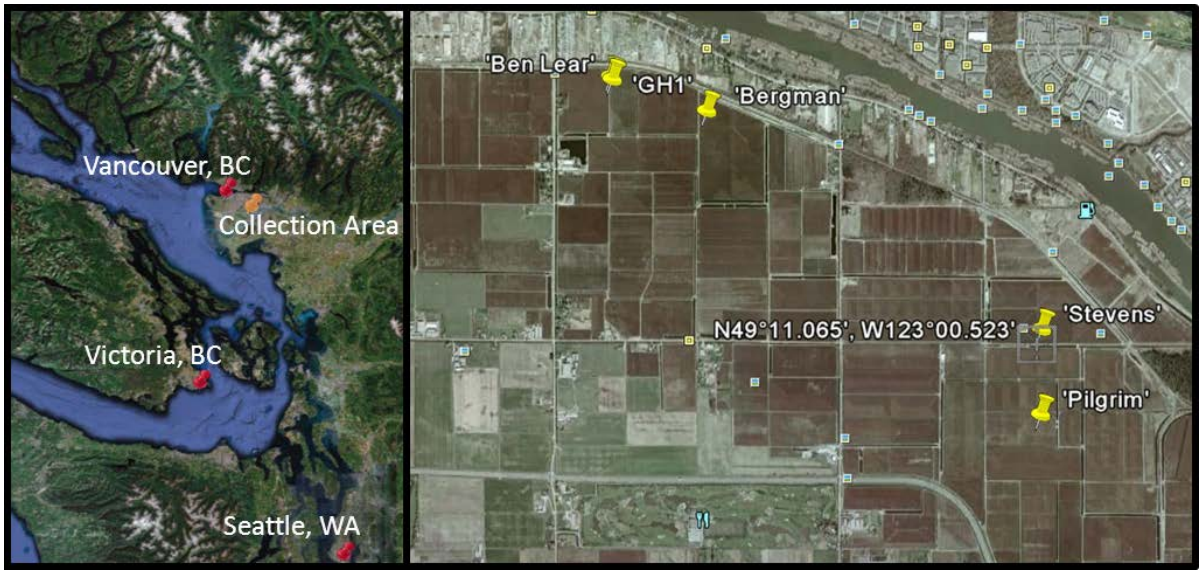

Figure 5-2: Collection area (Left: orange pin) in Lower Mainland, British Columbia and specific sites of 5 cultivars of *Vaccinium macrocarpon* Aiton collected (R: yellow pins).

Chemicals and Calibration Standards

The chemical reagents (hydrochloric acid, glacial acetic acid, phosphoric acid, formic acid) used in this study were of analytical grade and obtained from Sigma-Aldrich Canada (Oakville, ON, CAN). Chromatography and extraction solvents (acetonitrile, methanol, water) were of HPLC grade or equivalent and obtained from VWR International (Edmonton, AB, CAN). Individual anthocyanin chemical calibration standards (250-500 µg/ml) were prepared in 2% concentrated HCl in methanol (v/v) by Cerilliant Corp. (Round Rock, Texas, USA) in Snap-N-Shoot® format. Mixed chemical calibration standards were prepared for n=5 levels by serial dilution with 2% concentrated HCl in methanol (v/v). The individual Snap-N-Shoot® chemical calibration standards were stored at -20 ^{o}C when not in use and the purity was determined chromatographically immediately prior to use.

Sample Preparation

The experiment was designed to compare two different sampling methods, biological and analytical. *Sampling for biological variability*: A randomized sample selection procedure was followed to subsample 5 replicate berries from each cultivar collection, each individual berry was weighed, freeze-dried (Modulo Freeze-Dry System; Fisher Scientific, Ottawa, ON, CAN), reweighed to determine % dry matter, and prepared as individual berries for analysis using previously published validated analytical method (*16*). *Sampling for analytical variability*: Twenty individual cranberries were selected at random from the field collections, combined into a bulk sample, weighed, freeze-dried, reweighed to determine % dry matter, ground to a <60 mesh powder and subsampled in 5

replicates. All cranberry samples were weighed (0.250 g ±0.025 g) into 50-mL conical tubes, 20 mL of HPLC grade MeOH–concentrated HCl (98:2, v/v) was added and samples were mixed with a vortex mixer (Thermolyne Maxi Mix 1; Fisher Scientific; Ottawa, ON, CAN) for 10 seconds, sonicated (Branson Model 3510R-MTH Ultrasonic Cleaner; VWR International, Mississauga, ON, CAN) for 15 min, and shaken on an angle at 180 rpm for 30 minutes (Wrist Action Shaker model 57040-82; Burrell Scientific, Pittsburgh, PA, USA). The supernatant was decanted into a 25-mL glass volumetric flask and samples were brought to final volume of 25 mL with the extraction solvent. Samples were mixed well by inversion and approximately 1 mL of each sample solution was filtered (0.45-µm Teflon filters; VWR International; Edmonton, AB, CAN) into an amber vial for HPLC analysis.

Anthocyanin Determination by UFLC-DAD

The content of cyanidin-3-*O*-galactoside, cyanidin-3-*O*-arabinoside, cyanidin-3-*O*-glucoside, peonidin-3-*O*-galactoside, and peonidin-3-*O*-arabinoside was determined in the samples as per a published analytical method (Brown & Shipley, 2011; Brown, Shipley, et al., n.d.). In brief, anthocyanins were chromatographically separated using an 1100 Series Agilent (Mississauga, ON, CAN) liquid chromatography system at ambient temperature using a reverse phase Zorbax SB-C18 Rapid Resolution HT column (4.6 x 50 mm, 1.8 µm) with a mobile composition of (A) 0.5% water–phosphoric acid (99.5:0.5, v/v) and (B) water–acetonitrile–glacial acetic acid–phosphoric acid (50.0:48.5:1.0:0.5, v/v/v/v). The optimized gradient program is 9%-36% B over 8.0 min, 36%-60% B over 0.5 min, 60%-80% B over 0.5 min, and 80%-9% B over 0.1 min, hold at 9% B for 1.4

min. The injection volume and flow-rate were 5 µl and 2.1 mL/min, respectively. The analytes were detected at 520 nm, data was collected and analyzed using ChemStation software (Rev. B.03.01) from Agilent Technologies (Mississauga, ON, CAN). Quantification of anthocyanins was accomplished through the use of standard curves obtained through the analysis of the mixed standard solutions prepared as per the procedure described above. Single factor ANOVA was used to compare the quantities of the individual and total anthocyanins levels as well as anthocyanin ratios in the biological and analytical replicates between the cultivars. Student *t*-tests were used to compare the anthocyanin levels and ratios between the biological and analytical replicates for each cultivar.

Metabolomic Profiling by UPLC-TOF-MS

Cranberry metabolomes were analyzed as per a previously established protocol (Brown, Finley, et al., n.d.) with an ACQUITY^{TM} series Ultra Performance Liquid Chromatography System (Waters Inc., Mississauga, ON) coupled with a Micromass LCT Premier^{TM} series TOF-MS (Waters Inc., Mississauga, ON) and controlled with a MassLynx V4.1 Data Analysis System (Waters). Chromatographic separation was achieved with a Waters BEH Acquity C_{18}, 2.1 x 150 mm, 1.7 µm column with mobile phase conditions: 0.1% aqueous formic acid:acetonitrile (0.0 - 10.0 min, 95:5-5:95 v/v, 10.0-15.0 min, 5:95 v/v, 15.0-20.0 min, 5:95-95:5 v/v, 20.0-25.0 min, 95:5 v/v). A 25 min run time was used with a flow rate of 0.25 mL/min and column temperature of 30 °C. The autosampler was at 4 °C with an injection volume of 5 µL. A Waters 1525

HPLC Binary Solvent manager provided a steady flow of 2ng/mL leucine enkephalin at 10 µL/min. as the internal calibrant for flight tube length in mass spectrometer.

Data Processing and Subtractive Metabonomics

The metabolomics data of the five *Vaccinium macrocarpon* cultivars, as taken from UPLC-TOF-MS experiments, was compiled with blanks summed and then subtracted with any resulting negative values reset to zero. After blank subtraction, the data was set up in an ASCII text file where rows consist of sample identifiers as objects and columns consist of retention time, m/z ratio and abundance as variables. Compiled data was analysed in Microsoft Excel (Redmond, WA, USA), with R (The R Foundation for Statistical Computing, Vienna, AUT) and Solo+MIA (Eigenvector Research Inc., Wenatchee, WA, USA) used for statistical analysis including multivariate modelling. A series of automated functions to create subtractive datasets were created in Microsoft Excel using sequential algorithms and functions designated "Subtractive Metabonomics" (Brown, Finley, et al., n.d.; Brown et al., 2011). The custom macros designed for this process serve to locate standards, align retention times, remove multiply charged ions, etc. and can provided detailed information from any metabolite dataset. After blank subtraction, the data was also set up in an ASCII text file where rows consist of sample identifiers as objects and columns consist of retention time, m/z ratio and abundance as variables for import into R (R Foundation, GNU) and Solo+MIA (Eigenvector Research Inc., Wenatchee, WA) for further statistical analysis including multivariate modelling.

Significance Analysis

To assess the degree of significance of metabolites/compounds in comparisons of the metabolomics profiles of the 5 cranberry cultivars, different univariate statistical tools were employed. First, Receiver Operator Characteristic (ROC) curves were generated and the Area Under the ROC curve (AUC) were computed for each binary comparison between the 5 cultivars using the "*colAUC*" algorithm in the "*caTools*" package (Tuszynski, 2010). The ROC curve is a plot of the sensitivity for a binary classification system (Broadhurst & Kell, 2006). The accuracy of this plot is determined by assessing the AUC, whereby an AUC of 1 would indicate 100% sensitivity at a 0% false positive rate. To generate a "total" AUC value that captures all possible binary comparisons between the cultivars the mean of all 10 binary AUC values at each m/z value was determined (Hand & Till, 2001). In this way the AUC can be used to capture the distribution of abundance across the cultivars and between replicates.

The analysis can become more meaningful when results are supplemented with univariate variance analysis such as the non-parametric Kruskal-Wallis one-way analysis of variance (Kruskal & Wallis, 1952). At each m/z value the Kruskal-Wallis *p*-value was calculated for the replicate abundance data (n=5 per cultivar) using the "*kruskal.test*" algorithm derived from the "*stats"* package available in R. The *p*-values and total AUC of the metabolomic data are plotted in Excel to enabled identification of those variables that exhibited high variances in abundance across cultivars and small deviations within replicates. This process was employed as a tool to assist in identification of components (variables) in the metabolomics profiles that may be significant in terms of contributing to cultivar differentiation.

As the AUC is a binary comparison, it is unable to capture interactions or relationships that may exist across the metabolites in the metabolomics profiles, The Significance of Microarrays (SAM) analysis described by Tusher et al.(Tusher et al., 2001) is another approach for identifying metabolites (variables) of significance that ia based on the distribution of abundance in the metabolomics profiles of each cultivar. The SAM statistic was applied to identify m/z values of significance when comparing across the 5 cultivars with respect to abundance; the false discovery rate (FDR) was calculated at each threshold selected (Tusher et al., 2001). The ranked groups of values and associated FDRs were compiled using the "*siggenes"* package done in R (Schwender, 2009). The m/z values identified by the SAM statistic were distinguished in the plot of *p*-value against total AUC of each metabolite.

Multivariate Analysis

Both unsupervised (PCA) and supervised (PLS-DA) algorithms were selected to observe variance exhibited by the metabolomic data and relate categorical information (sample identifiers) to the abundance data. The only preprocessing of the data prior to applying PCA and PLS-DA algorithms was autoscaling performed in Solo+MIA. Score plots were generated to visualize clustering by varietal and loadings plots for examining the distribution of data values. Within the loading plots for the PLS-DA models, the calculated Kruskal-Wallis *p*-value for abundances at each m/z value were labeled as two groups; values with a calculated *p*-value >0.05 and values with *p*-value of < 0.05. Those data with *p*-values greater than 0.05 were removed from the data set, and the remaining data remodeled. The Q-residual value of each m/z value (variable) was taken from the

PLS-DA loading plots (Eissa, Batson, Wise, & Gallagher, 1996). Each Q-residual was transformed into % Q-residual by the following formula;

$$\left[1 - \frac{Q_{res} - Q_{res}^{min}}{Q_{res}^{max}}\right] \times 100 = \% \, Q_{res}$$

where Q_{res}^{min} is the minimal Q-residual of the entire metabolomic dataset and Q_{res}^{max} is the maximal Q-residual of the entire metabolomic dataset. The % Q-residual was used to observe the ability of the chosen model to capture the contributed information existing in the data at each m/z value.

Results and Discussion

Determination of Anthocyanins in Cranberry Varietals

Ben Lear, Bergman and GH1 had significantly higher anthocyanin levels than Pilgrim and Stevens (**Figure 5-3**). These trends are in agreement with other studies that reported higher anthocyanin levels in Ben Lear than in Stevens et al., 2009; N. Vorsa et al., 2003). As noted by Vorsa et al. (2003) this variation in the anthocyanin levels can be partially explained by the differences in the size of the fruit, as a negative correlation has been observed between fruit productivity and anthocyanin content (Vorsa et al., 2003).

Fruit size differs across the varietals with Ben Lear and Bergman having the smallest fruit with cup counts of 70-90 and 65-80, respectively, and Pilgrim and Stevens having the largest fruit with cup counts of 46-66 and 50-60, respectively. These observations are consistent with anthocyanin concentration being higher in the skin thereby affecting the surface area to volume ratio and the resultant measured values. With

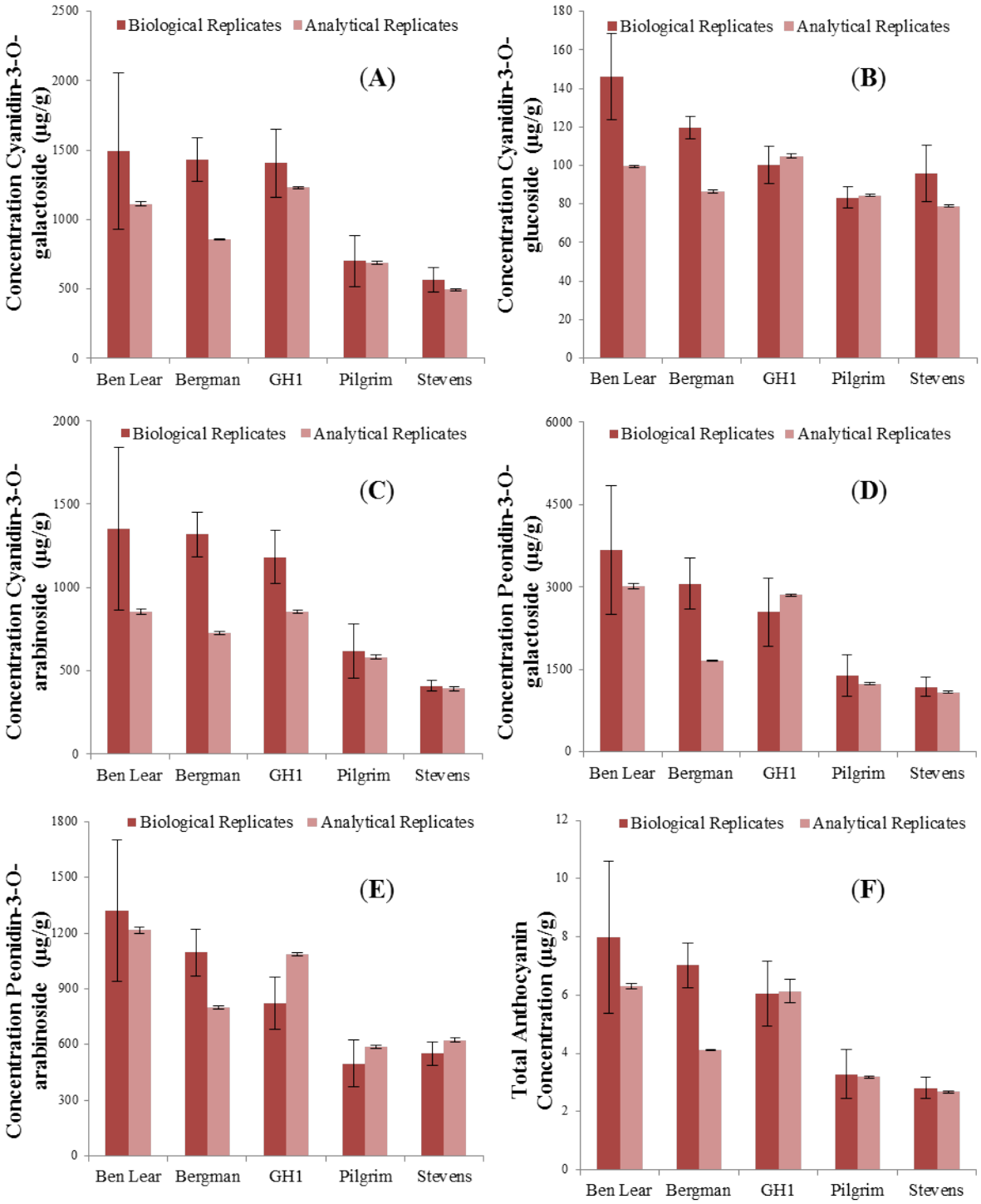

Figure 5-3: Average anthocyanin content (µg/g) in each cultivar for biological replicates (n=5) and extractions from a composite sample of cranberries (n=5) with error bars representing the standard error of the mean for **(A)** cyanidin-3-*O*-galactoside, **(B)** cyanidin-3-*O*-glucoside, **(C)** cyanidin-3-*O*-arabinoside, **(D)** peonidin-3-*O*-galactoside, **(E)** peonidin-3-*O*-arabinoside, and **(F)** the total anthocyanin content in mg/g of dry weight material (DW).

other crops it has been observed that cultivars bred for commercial production have reduced levels of certain of secondary metabolites and chemical diversity as compared to wild type (Frost et al., 2008; Rodriguez-Saona et al., 2011; Wink, 1988). To allow for high productivity and large fruit size, some trade-offs are expected as additional energy and nutrients are devoted to fruit production (Frost et al., 2008; Rodriguez-Saona et al., 2011; Wink, 1988).

Both Pilgrim and Steven's were selected in part due to their large fruit size and high productivity, however this selection may have translated to the inhibition of anthocyanin producing pathways. Careful consideration and monitoring of several variables is needed when breeding cultivars to ensure desired properties are retained and/or improved upon. For example the GH1, a relatively new cultivar, exhibits high productivity without a decrease in anthocyanin production. This cultivar had relatively high levels of anthocyanins, yet is of similar size and is more productive than Stevens and Pilgrim. There are efforts being made to develop cultivars that are productive, have large fruit and high anthocyanin levels (N. Vorsa, 1994; N. Vorsa & Polashock, 2005).

Figure 5-4 shows the comparison between the biological and analytical replicates. Although at first glance the biological replicates appears to show higher anthocyanin values, Student t-tests indicated that only in Bergman was a statistically significant difference ($p<0.05$) detected. The biological replicates had a large variance associated with it whereas the analytical replicates showed a very low level of variance. The results herein presented provide an interesting representation of the trade-offs inherent in composite sampling. The main advantage of composite sampling is a reduction in measurement costs. Analysis of n=20 can thus be accomplished in one

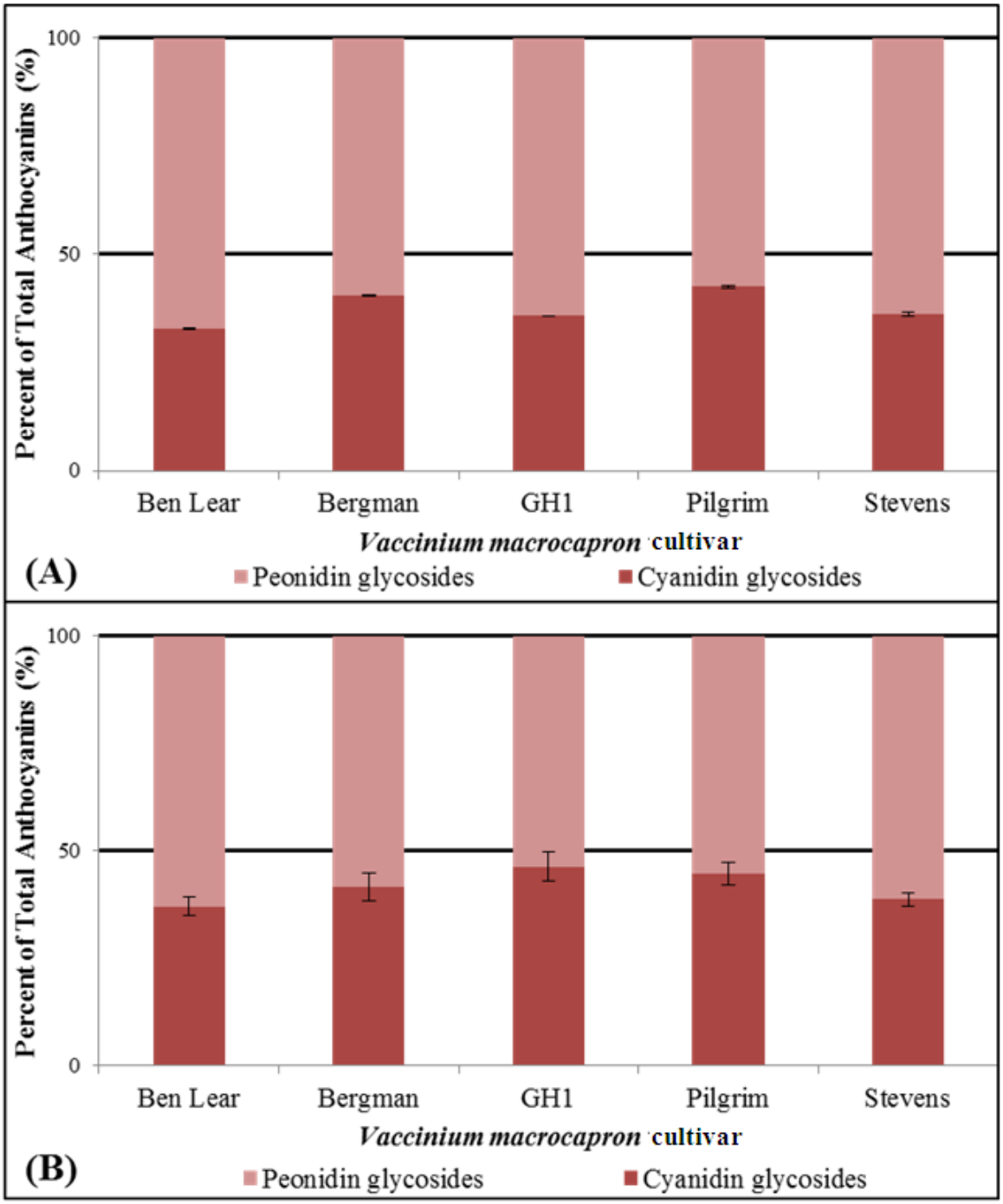

Figure 5-4: Comparison of the ratio of glycosylated peonidins to cyanidins in five *Vaccinium macrocarpon* cultivars as a percent of the total anthocyanins ± standard error of the mean in **(A)** analytical replicates and **(B)** biological replicates.
Ben Lear, GH1 and Stevens were found to have lower cyanidin glycoside to peonidins.

measurement using a validated method and the result provides a good estimate of the population mean. The observed variance from the analysis of the 5 analytical replicates is predominately due to the inherent precision of the analytical method. The trade-off to this reduced testing approach is in a loss of information concerning the inherent variance of the sample population. As shown in the biological replicates, there is a large variance in the anthocyanin levels between the individual berries within a given cultivar population. One implication of the lost information can be seen when comparing the cyanidin glycoside to peonidin glycoside ratios (**Figure 5-4**). The analytical replicates of glycoside ratios than Bergman and Pilgrim, however no significant differences in the ratios were detected when comparing the biological replicates. As the variance in the composite sample is dominated by the method precision and not the variability of the berries, a proper comparison of the ratios cannot be accomplished. This is of particular interest as it is often necessary to pool biological replicates to provide enough material for the analytical technique being utilized. This would have, in this case, potentially led to incorrect conclusions as to differences between the cultivars that actually lie within the range of biological variation. Multiple biological replicates are required to provide the necessary variance information and thus allow for a proper comparison of the populations.

Metabolomic Profiling by UPLC-TOF-MS

The metabolite counts and results of subtractive metabonomics are found in **Table 5-1** which makes direct comparisons across the metabolomics fingerprints of five *V. macrocarpon* cultivars. Using subtractive metabonomic algorithms described previously

(Brown, Finley, et al., n.d.) 6481 compounds were found conserved across all varietals with 136 (Ben Lear), 84 (Bergman), 91 (GH1), 128 (Pilgrim) and 165 (Stevens) unique compounds observed. About 55-57% of the phytochemistry described in the metabolomic fingerprints was common across all cultivars, with the exception of the metabolome of Stevens which shared 65% of the observed phytochemistry with all other cultivars. From the binary comparisons of the metabolomics profiles, we observe Stevens and Bergman and Stevens and Ben Lear have the fewest metabolites in common with 83 and 84, respectively, while Stevens and GH1 share the most in common with 214.

Table 5-1: Summary of LC-MS-TOF metabolite counts and differences in metabolomics profiles of five cultivars of *Vaccinium macrocarpon* Aiton.

***Vaccinium macrocapron* cultivar**	**Ben Lear**	**Bergman**	**GH1**	**Pilgrim**	**Stevens**
Total # compounds observed	11544	11395	11322	11736	10038
Average # compounds per cultivar	5971 ± 195	5944 ± 131	5848 ± 147	6330 ± 222	4477 ± 945
Compounds in all replicates	1717	1828	1616	2108	252
Compounds in at least 50% of replicates	5254	5262	5225	5674	3765
Unique to each cultivar	136	84	91	128	165
Common between cultivar	6481				
Common between cultivars	**Ben Lear**	**Bergman**	**GH1**	**Pilgrim**	**Stevens**
Ben Lear	-	152	100	171	85
Bergman	152	-	96	140	83
GH1	100	96	-	132	214
Pilgrim	171	140	132	-	118
Stevens	85	83	214	118	-

Chemometric Analysis of Metabolomic Data

In the PCA score plot of the first and second principal component, significant differentiation amongst the cranberry cultivars from metabolomic profiles was not

observed, despite the differences noted in the metabolite counts for each cultivar (data not shown). To observe how the metabolomic profiles of the cultivars may be differentiated, PLS-DA was selected as the next modeling tool (**Figure 5-5A, B**). In the score plot modeling linear variates one (LV1) and two (LV2) differentiation of Stevens and GH1 from the other cultivars is observed with the 95% confidence boundaries around Bergman, Ben Lear, and Pilgrim overlapping (**Figure 5-5A**). The linked loadings plot (**Figure 5-5B**) indicates minor clustering of the data by cultivar; Stevens (upper right corner), GH1 (bottom center) and the intense cluster of metabolites where Bergman, Ben Lear, and Pilgrim are located (upper left corner). This clustering is also evident in the values with a Kruskal-Wallis *p*-value of <0.05 (highlighted in red), which indicates the data contributing to the clustering pattern is statistically significant. To further evaluate the relationships between the cultivars and robustness of the data in LV1/LV2, metabolites with *p*-value of > 0.05 were peeled off the PLS-DA model and the score and loading plots remodeled (**Figure 5-5C, D**). For the metabolite total of 15064 values, 1749 had a calculated $p<0.05$, resulting in a loss of 88.4% of the data.

With a limited number of samples (n=5) evaluated per *Vaccinium* species the suggested way of validating or testing the performance of the multivariate model by randomly assigning 25% of the samples as "unknown" to test against a model built of the remaining data (Varmuza & Filzmoser, 2009), is not possible so another approach must be taken. Just as an analytical method can be orthogonally validated by employing an independent method, the results of the PLS-DA model can be compared to a univariate statistical analysis, the significance analysis of microarray (SAM) statistic. The values identified as significant by the SAM statistic are identified (in red) in the LV1/LV2 loadings plot

(**Figure 5-5C**) the loadings to observe whether they fall within the clusters that define the positioning of the five cultivars in the remodeled LV1/LV2 score plot (**Figure 5-5D**). With a false discovery rate of 1.78%, the 20 values identified by the SAM statistic are unlikely to be false positives (**Table 5-2**). The SAM values positioned within the clustered areas and at their outermost edges where the contribution of the values to the score plot clustering is greatest (**Figure 5-5C**), is supportive of the PLS-DA model. For example, the SAM algorithm identified m/z value of 276.3168, only observed in the Stevens cultivar, as significant and this value appears on the edge of the cluster in the loading plot that represents the Stevens cultivar in the PLS-DA model. Similarly, m/z values of compounds observed only in the metabolomics profiles of the GH1 cultivar should be in the lower bottom corner of the loadings plot (Figure 5C) near the LV1 value of zero and in this region two m/z values, 986.9273 and 398.9882, were identified as significant by the SAM statistic. Although m/z 986.9273 is observed only in GH1, m/z 398.9882 was observed in both GH1 and Ben Lear.

Examining the score plot **(Figure 5D**) we see that Ben Lear is positioned closest to GH1 and the PLS-DA model for LV1/LV2 is again consistent with the SAM statistic. Often differentiation is not observed in score plot of the first two linear variates and other combinations can be considered. **Figure 5-6A,** the score plot of LV4 and LV2, shows a very different clustering pattern than **Figure 5-5A**. Ben Lear, GH1 and Bergman are somewhat differentiated and the 95% confidence intervals around Stevens and Pilgrim cultivars are overlapping, however, the linked loadings plot (**Figure 5-6B**) shows that values with $p<0.05$ are not concentrated in the same way but rather are spread out in the LV4/LV2 model. When the values with a calculated $p>0.05$ are removed the remodeled

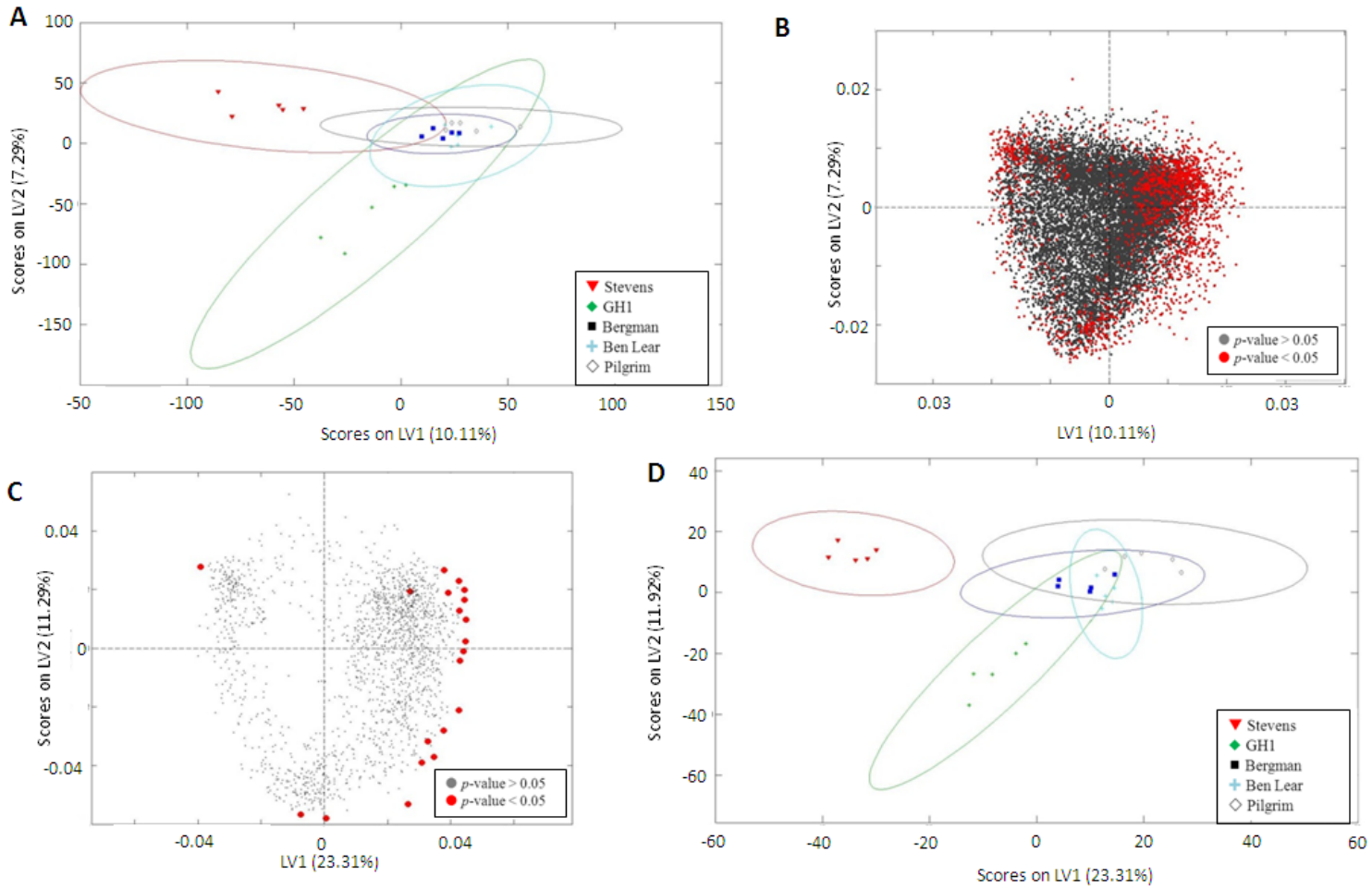

Figure 5-5: Multivariate analysis of the UFLC-TOF-MS metabolomic profiles of the cranberry varietals. **(A)** The PLS-DA score plot for LV1/LV2 and **(B)** the corresponding loadings plot for the PLS-DA model, data with calculated Kruskal-Wallis *p*-value <0.05 is highlighted in red. **(C)** Shows the loadings plot re-modeled with all values having a *p*-value >0.05 removed with values identified as significant by SAM statistic highlighted in red and **(D)** the corresponding re-modeled PLS-DA score plot for LV1/LV2.

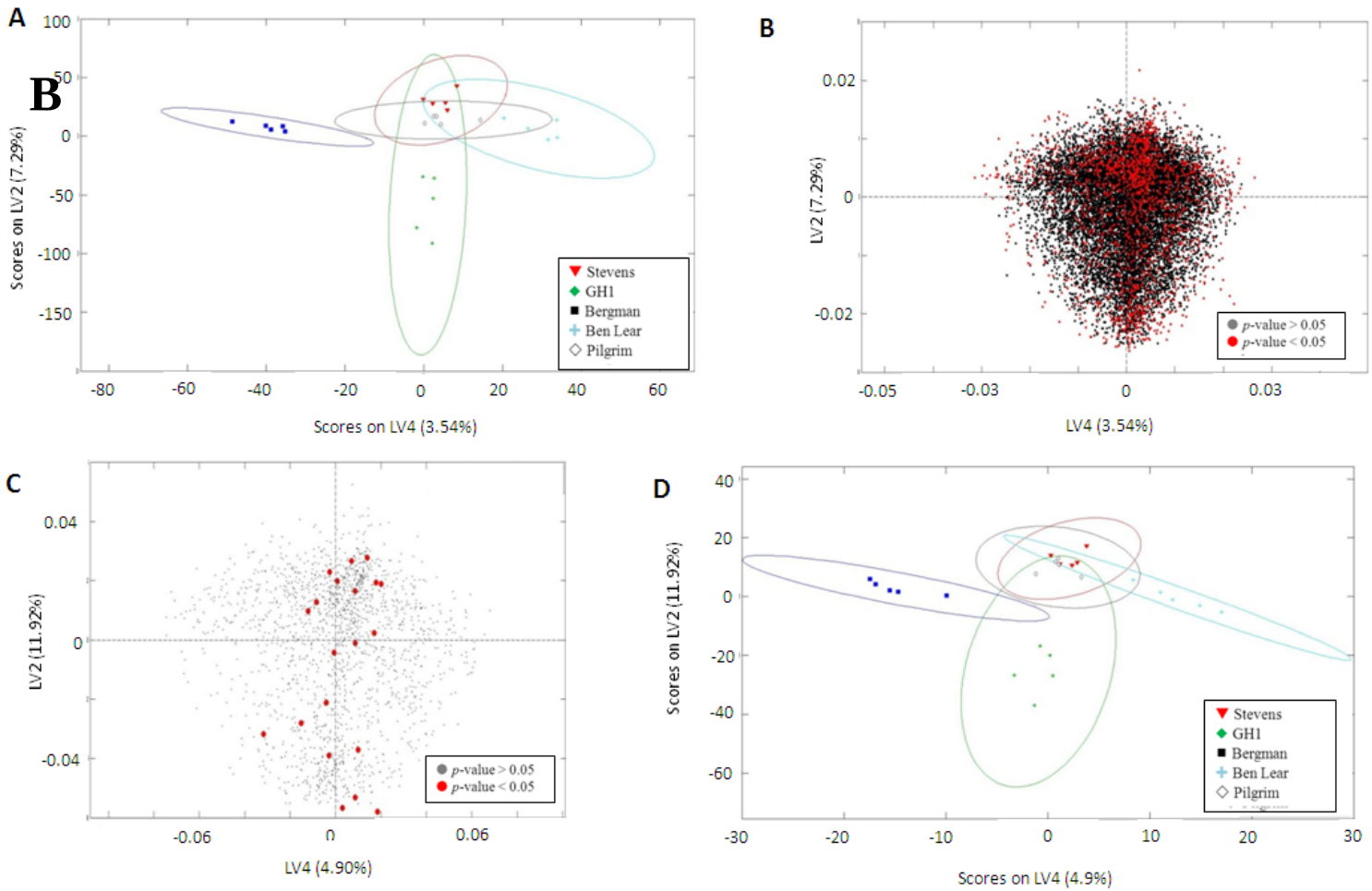

Figure 5-6: Multivariate analysis of the UFLC-TOF-MS metabolomic profiles of the cranberry cultivars. **(A)** The PLS-DA score plot for LV4/LV2 and **(B)** the corresponding loadings plot for the PLS-DA model, data with calculated Kruskal-Wallis *p*-value <0.05 is highlighted in red. **(C)** Shows the loadings plot re-modeled with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted in red and **(D)** the corresponding re-modelled PLS-DA score plot for LV4/LV2.

values, highlighted in red, cluster in a fashion that loosely reflects the differentiation observed in score plot (**Figure 5-6D**) but are not concentrated at the outermost edges of the LV4/LV2 loadings plot.

Based on the combination of orthogonal statistical approaches for evaluating the metabolomics profiles of the five cultivars, we can have more confidence in PLS-DA model of LV1/LV2 which indicates the metabolome of the Stevens cultivar is distinct from Ben Lear, Bergman, GH1 and Pilgrim, although more similar to GH1 which also exhibits a phytochemical signature distinct from Ben Lear, Bergman, and Pilgrim. Although only the LV1/LV2 plots (score and loading) agree with the orthogonal univariate statistics, some merit is seen in the LV4/LV2 plots (**Figure 5-6C**, **D**) where a cluster of the SAM identified metabolites are localized at the outermost edges of a cluster in the loadings plot that reflects where the Stevens and Pilgrim cultivars are positioned in the score plot. This indicates those metabolite values are truly more significant than other metabolites identified by the SAM statistic with respect to the LV4/LV2 PLS-DA model. While generally it is assumed with the inclusion of more latent variables in a PLS-DA model the more variance the model exhibits per metabolite value, the comparison of the Q-residuals for up to LV2 and LV4 shows nearly 50% of the SAM identified values (9 of 20 total) reflect a decreasing % Q-residual with increasing number of latent variables (**Table 5-2**), thereby providing further confidence in the LV2/LV4 PLS-DA model.

The majority of the SAM identified metabolite values are present in at least three of the cultivars (**Table 5-3**). It is interesting to note that where the values are observed in only 4, 3 or 2 cultivars, they were not observed in Stevens. In fact, outside of those

Table 5-2: Summary of LC-MS-TOF metabolomic m/z values identified by the SAM statistic in order of ranked significance, *d(i)*, at a false discovery rate of 1.78% as well as the Q-residual values of the compounds in the PLS DA model for up to LV2 and up to LV4.

m/z Value	d(i) value	Q-residuals (%)	
		up to LV2	up to LV4
296.385	42.9	98.7	97.7
285.291	26.2	42.4	98.4
305.296	25.8	97.1	100.0
294.354	23.9	95.0	96.5
792.901	23.6	93.8	93.9
482.553	23.4	97.9	97.4
584.593	22.7	83.7	96.3
631.546	22.1	99.0	97.9
398.494	21.9	75.2	80.9
280.337	21.9	80.3	89.6
546.612	20.1	98.5	95.9
585.555	19.4	90.1	89.7
986.927	19.0	85.3	83.5
719.697	18.5	83.7	86.8
398.988	18.4	80.1	87.1
233.316	17.0	91.0	88.8
245.276	16.7	96.6	95.3
276.317	16.6	94.6	93.4
736.798	15.8	89.7	88.6
553.652	15.6	87.9	87.3

values found in all 5 cultivars, only 1 m/z value of significance was identified by SAM, m/z 276.317 and observed in the Stevens cultivar (**Table 5-3**). Further investigations on those values not conserved in the Steven's cultivar could lead to a better understanding differences and relationships between the native and hybrid cultivars, such as Stevens and GH1.

Applying univariate statistics as a quality assessment of the PLS-DA models developed from the metabolomics data indicates that LV1/LV2 model is better fit than LV2/LV4 where the cultivars are further differentiated. Having the majority of the % Q-residual

per metabolite value not increasing with an increased number of latent variables, further suggests the PLS-DA models are not entirely reflective of the cranberry varietal metabolomes. However, the metabolomic data and PLS-DA models should be carefully considered as it is expected that the differences and similarities of varietals from a single species would be subtle. This is further illustrated in **Table 5-3** where the majority of the SAM identified metabolite values exist in all or the majority of cultivars indicating that significant values existing only in one cultivar are difficult to find. Strategies exist to better improve the "detection" of important metabolites, although for a cultivar study it may prove more useful to increase the sample size to better illustrate the true picture, decrease the value of the associated FDR with statistical models and cross-validate the multivariate models developed.

Both the targeted and untargeted analyses found significant phytochemical differences among the varietals. If the goal of the commercial production is yield of anthocyanins, then mass plantings of high anthocyanin cultivars such as Ben Lear, Bergman or GH1 would be recommended, however, other agronomic considerations such as yield and disease resistance may be reflected in the large scale production of 'Stevens' by commercial farmers. It is interesting that GH1, a modern cultivar which exhibits both high productivity and anthocyanin content, is reasonably differentiated by both PLSDA-DA models and has metabolites of significance identified in **Table 5-3** both in independent of the other cultivars as well as in common with Ben Lear, Bergman and Pilgrim. Overall the significance analyses identified individual metabolites characteristic of each of the cultivars and a new approach to quality assessment of multivariate models of metabolomics data has been demonstrated. One of the key difficulties with

Table 5-3: The observed average abundance (n=5), total AUC and ranking by *d(i)* value of the m/z values identified by the SAM statistic as significant at a false discovery rate of 1.78% comparing across the five *V. macrocarpon* cultivars.

m/z Value	average abundance for n=5 replicates					total AUC	Rank, d(i) value
	Stevens	GH1	Bergman	Ben Lear	Pilgrim		
	Observed in all cultivar						
792.901	0.074	0.270	0.360	0.483	0.494	0.920	5th, 23.6
296.385	0.107	1.112	0.627	0.806	0.678	0.904	1st, 42.9
305.296	0.446	0.984	1.381	1.637	1.322	0.896	3rd, 25.8
482.553	0.347	0.482	0.908	0.958	1.121	0.896	6th, 23.4
280.337	0.223	1.205	1.516	1.256	0.950	0.880	10th, 21.9
736.798	1.787	7.160	5.645	6.995	6.331	0.852	19th, 15.8
719.697	0.208	1.235	1.092	1.239	0.892	0.836	14th, 18.5
585.555	0.277	1.494	1.693	1.668	1.472	0.788	12th, 19.4
546.612	0.208	1.995	2.134	2.316	2.273	0.782	11th, 20.1
	Observed in 4 cultivars						
584.593	-	0.038	0.077	0.071	0.123	0.928	7th, 22.7
631.546	-	0.026	0.061	0.077	0.093	0.912	8th, 22.1
245.276	-	0.138	0.278	0.264	0.317	0.844	17th, 16.7
233.316	-	0.383	0.546	0.602	0.605	0.796	16th, 17.0
	Observed in 3 cultivars						
398.494	-	-	0.052	0.061	0.084	0.884	9th, 21.9
294.354	-	-	0.043	0.042	0.058	0.866	4th, 23.9
553.652	-	-	0.042	0.060	0.061	0.834	20th, 15.6
	Observed in 2 cultivars						
398.988	-	0.036	-	0.020	-	0.844	15th, 18.4
285.291	-	-	-	0.035	0.150	0.784	2nd, 26.2
	Observed in only one cultivar						
986.927	-	0.029	-	-	-	0.736	13th, 19.0
276.317	0.115	-	-	-	-	0.700	18th, 16.6

metabolomics datasets is the need for methods and statistical approaches to ensure the quality of data sets. The combined application of the univariate approaches, statistical Kruskal-Wallis *p*-values, area under the ROC curve and SAM statistic provide the basis

for establishing quality evaluations of multivariate models as a new tool for metabolomics research with broad future applications.

Chapter 6: Phytochemical Discovery in Large Cranberry (*Vaccinium macrocarpon* Aiton), Small Cranberry (*Vaccinium oxycoccus*) and Lingonberry (*Vaccinium vitis-idaea*)

Synopsis

Cranberries have had important roles in traditional health and culture of indigenous people across North America as well as modern uses in the natural health products industry. In North America, there are three species generally recognized as cranberries, *Vaccinium macrocarpon* Aiton, *Vaccinium oxycoccus* and *Vaccinium vitis-idaea* L., (McCown & Zeldin, 2005; Suda & Lysák, 2001; Vander Kloet, 1983). Many of the medicinal compounds are common among these three *Vaccinium* species (Borowksa et al., 2009; N. Vorsa & Polashock, 2005). For *V. vitis-idaea* Lehtonen et al., found that the species is a rich source of anthocyanins that are eliminated via methylation, oxidation and excretion through urine, a common mechanism to the one proposed for *V. macrocarpon* (Lehtonenet al., 2010; Lehtonen et al., 2009). Polymeric proanthocyanin extracts of both *V. macrocarpon* and *V. vitis-idaea* have also shown antimicrobial activity against *Staphylococcus aureus* and inhibited hemagglutination of *E. coli* in *in vitro* bioassays (Kylli et al., 2011). Berries of all 3 of the species have been found to have a strong potential for detoxification of reactive oxygen species or "antioxidant potential" (Borowksa et al., 2009; Yao & Vieira, 2007). The high antioxidant activity observed in cranberry fruit and extracts is in part attributed to high levels of phenolic compounds (Cherubini et al., 2008; Seeram, 2008a; Viskelis et al., 2009; Zheng & Wang, 2003). It

has been shown however, that these compounds do not explain all of the antioxidant activity observed in extracts (Kähkönen, et al., 2001). Studies in grape and other berries have shown that melatonin and its metabolites can act as signaling molecules that induce antioxidant responses and also directly act as reactive oxygen species (ROS) scavengers (Boccalandroet al., 2011; Murch et al., 2010; Vitalini et al., 2011). Synergism, from compounds such as vitamin C, has also been shown to have a significant effect on antioxidant response (Kähkönen et al., 2001). Unfortunately, the full nutritional and medicinal potential and activity of the berries as well as their importance for maintaining health in traditional and modern diets is not fully understood. With the goal of gaining a better phytochemical understanding of *Vaccinium* species in British Columbia, the specific research objectives of this work were: (a) to collect and investigate wild populations of native BC cranberries from the Haida Gwaii, (b) to compare the relative abundance of known cranberry phytochemicals such as anthocyanins and vitamin C in the species of large and small cranberries, (c) to discover previously un-described phytochemicals in large and small cranberries such as indoleamine neuromodulators and (d) to develop a model for metabolomics analysis to identify commonalities and differences between closely related species.

Materials and Methods

Plant Materials

Vaccinium oxycoccus L. and *Vaccinium vitis-idaea* L. fruit were collected from wild populations in Haida Gwaii in September 2010. Berries, leaves and stems for analytical samples and herbarium vouchers were collected for *Vaccinium oxycoccus* L. in the

muskeg just off the Yellowhead Highway approximately 7 kms south of Masset ($53^0$55'33" N, $132^0$6'25" W) and for *Vaccinium vitis-idaea* L. a short distance south on the same highway just off the pavement where the vegetation is brushed by the highway crews (53^0 51'20" N, $132^0$5'20"W**)** (**Figure 6-1A**, **B**). Species were identified based on local knowledge (Tim Wolthers) and comparison to published flora (Douglas et al. 1999). Berries were immediately frozen and shipped to the Okanagan campus of the University of British Columbia. Leaf and stem collections were pressed and dried with a field press, assembled and deposited as vouchers (2558SJM and 2559SJM) at the herbarium of the Beaty Biodiversity Centre, University of British Columbia and the herbarium at the National Tropical Botanical Garden, Kauai, Hawaii. Fresh fruit of *V. macrocarpon* Aiton was grown in prepared peat beds and harvested by the standard wet-pick method in the Vancouver Greater Regional District under the standard commercial production conditions approved by Ocean Spray Canada Ltd. (Richmond, BC, CAN), and immediately frozen prior to shipping.

Reagents and Standards

HPLC grade methanol, acetonitrile, hydrochloric acid, and phosphoric acid were obtained from VWR International ((Edmonton, AB, CAN). Ascorbic acid (purity ≥99.0%), trolox ((±)-6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) (purity >97%), melatonin (N-acetyl-5-hydroxytryptamine) (purity ≥ 99.5%), serotonin (5-hydroxytryptamine) (purity ≥ 98%) and DPPH (1,1-diphenyl-1-picryhydrazyl) were obtained from Sigma-Aldrich (Oakville, ON, CAN). A mixed reference standard containing cyanidin-3-*O*-galactoside (C3Ga), cyanidin-3-*O*-glucoside (C3Gl), cyanidin-3-*O*-arabinoside (C3Ar), peonidin-3-*O*-galactoside (P3Ga), and peonidin-3-*O*-

arabinoside (P3Ar) at 250-500 µg/mL in methanol were obtained from Cerilliant Corp. (Round Rock, TX, USA).

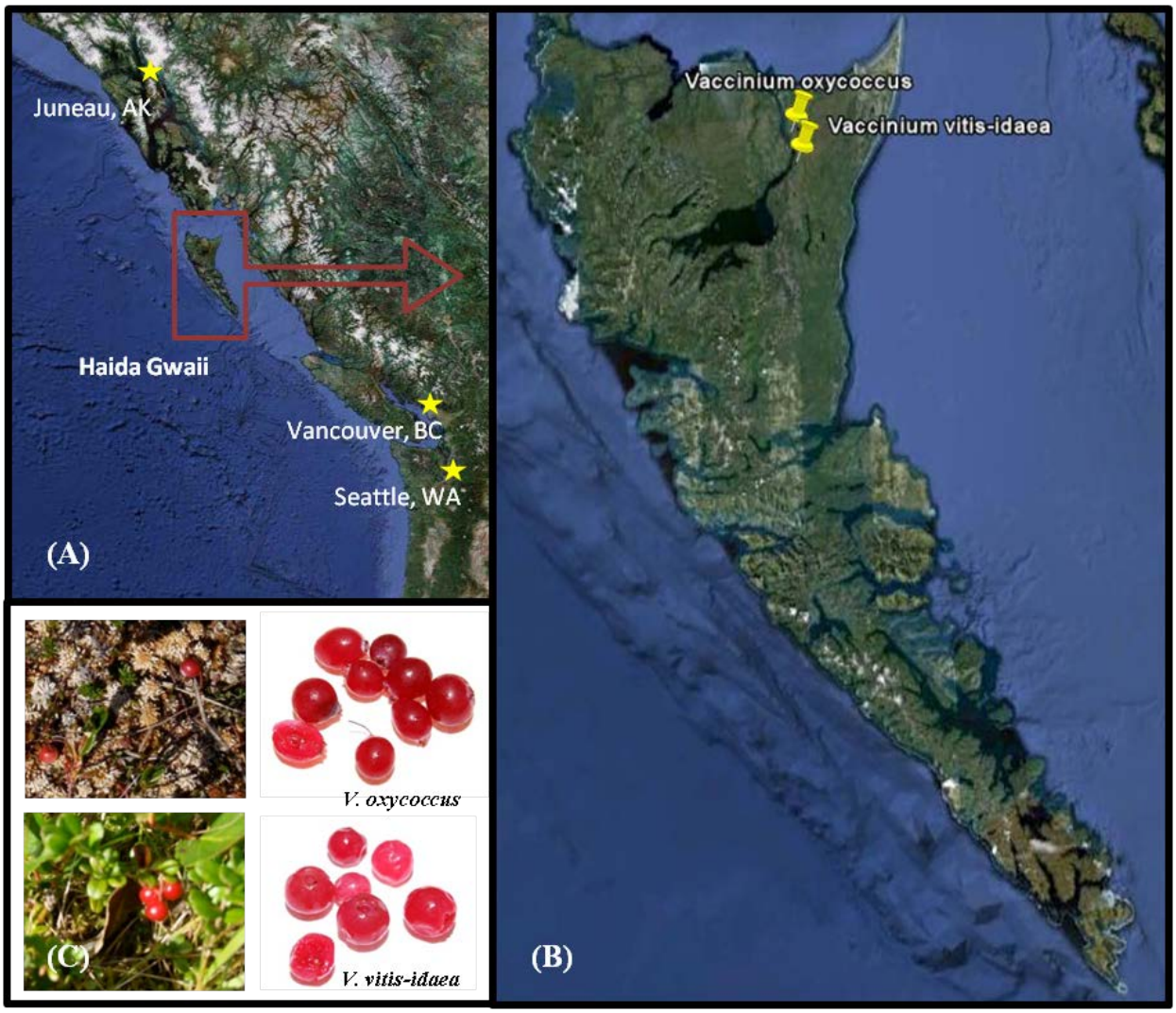

Figure 6-1: Collections of wild *Vaccinium* species in Haida Gwaii, fall 2010. **(A)** Location of Haida Gwaii, British Columbia. **(B)** Collection sites in Haida Gwaii. **(C)** Samples of *V. oxycoccus* (top) and *V. vitis-idaea* (bottom) with specimens for chemical analyses.

Anthocyanin Determination by HPLC-DAD

The anthocyanin content was determined using a previously validated method (Brown & Shipley, 2011) optimized for rapid resolution (Brown, Shipley, et al., n.d.). In brief, freeze dried cranberries were weighed (0.250±0.025g) to 50-mL conical tubes and extracted with 20 mL of MeOH-concentrated HCl (98:2, v/v) by vortex mixing (Thermolyne Maxi Mix 1; Fisher Scientific Company, Ottawa, ON, CAN), sonication for 15 min. (Branson Model 3510R-MTH Ultrasonic Cleaner; VWR International, Edmonton, AB, CAN) and shaking (Burrell Scientific model 57040-82 Wrist Action Shaker, Pittsburgh, PA, USA) at an angle at 180 rpm for 30 min. The supernatant was decanted to a 25-mL glass volumetric flask, brought to volume with extraction solvent, and mixed well. Approximately 1 mL of each sample solution was filtered (0.45-µm PTFE) to amber HPLC vials. Chromatographic separation was achieved on an Agilent 1100 Series HPLC System (ON, CAN) using a Zorbax SB C_{18} Rapid Resolution HT column (4.6 x 50 mm, 1.8 µm), mobile phase A: 0.5% water-phosphoric acid (99.5:0.5, v/v) and mobile phase B: water-acetonitrile-glacial acetic acid-phosphoric acid (50.0:45.5:1:0:0.5, v/v/v/v) with gradient elution: (a) 9%-36% B over 8.0 min. (b) 36%-60% B over 0.5 min. (c) 60%-80% B over 0.5 min. (d) 80%-9% over 0.1 min. (e) 9% B held at 1.4 min. The detection wavelength was 520 nm and all data was collected and analyzed using Agilent ChemStation software, Rev. B.03.01 (Mississauga, ON, CAN).

Melatonin and Serotonin Determination

Melatonin and serotonin were analyzed following a previously published method (Murch et al., 2010). Briefly, frozen berries were sectioned, weighed and homogenized in complete darkness with the extraction completed in less than 15 minutes for each

individual sample. Cranberries, including seed flesh and skin (200 mg), were homogenized for 3 min in 200 µL methanol:water: formic acid (80:20:1 v/v) using a cordless motor Kontes Pellet Pestle™ grinder (NJ, USA) and disposable pestles (Kontes). Samples were centrifuged for 3 min at 16,000 g (Galaxy 16DH Microcentrifuge, VWR International , Arlington Heights, IL, USA) and the resulting supernatant was filtered (0.2 µm, Ultrafree-MC filtered centrifuge tubes; Millipore, Billerica, MA, USA). Samples were kept at 4 °C in total darkness until analysis on a Waters Acquity UPLC system with separation achieved at 30^0C on a Waters BEH C18 UPLC column, (2.1 x 150 mm, 1.7 µm; Waters Corporation, Milford, MA, USA) with gradient elution using 1% aqueous formic acid:acetonitrile (0.0-4.0 min, 95:5-5:95 v/v, 4.0-4.5 min, 95:5—95:5 v/v, 4.5-5.0min, 95:5 v/v) at 0.25 mL/min. Melatonin and serotonin were quantified by electrospray ionization in the positive mode with Time of Flight mass spectrometry (LCT Premier Micromass MS) using the "W" configuration and by comparison to authenticated standards and absolute mass as per the parameters for mass spectrometry (Murch et al., 2010).

Determination of Free Radical-Scavenging Capacity

The 1,1-diphenyl-2-picrylhydrazyl (DPPH) assay employed to assess the free radical-scavenging capacity of berries from the three *Vaccinium* species was adapted from published methods (J. Liu, Hu, Dong, & Hu, 2004; Sánchez-Moreno, 2002; Takahata, Ohnishi-Kameyama, Furuta, Takahashi, & Suda, 2001). Briefly, approximately 0.3 g sections of fruit samples from each berry, prepared in triplicate, were homogenized in 500µL of methanol (30 sec) and centrifuged (Galaxy 16DH Microcentrifuge; VWR

International, AB, CAN) for 3min at 10,000 rpm. The supernatant was collected and diluted to the following concentrations: 0.03, 0.02, 0.015, 0.01, 0.003, 0.0003, 0.00003 (g/ml) and assayed immediately. Trolox ((±)-6-hydroxy-2,5,7, 8 tetramethylchromane-2-carboxylic acid) standard solutions were prepared in methanol at 0, 1, 8, 16, 24, 32, 40, and 50 µM. For each 96 well plate, 100 µL of standards and diluted extracts were placed alongside 200 µL of methanol and extract blanks; all samples were randomly assigned to wells. To each standard and sample 100 µL of 0.0994 mM DPPH was added, the plate placed into a BioTek Synergy H multi-detection microplate reader (Winooski, VT, USA), and slowly shaken for 2 sec prior to acquiring absorbance at 520 nm every 60 sec at 25 ºC. The DPPH radical-scavenging activity was calculated at t=15 mins and expressed as fresh berry weight required to reduce free radicals (DPPH response) by 50%. Three replicates for each of the *Vaccinium* species were used in the determination.

Ascorbic Acid Determination

The ascorbic acid (Vitamin C) assay was modified from previously reported methods with slight modifications (Asami, Hong, Barrett, & Mitchell, 2003). Frozen cranberries were thawed, individually weighed to 15-mL conical tubes in triplicate and extracted in 1 mL of 5% aqueous o-phosphoric acid by mashing with a glass rod followed by sonication for 30 min in an ice bath. The test solutions were filtered (0.45 µm PTFE) to HPLC vials. Analysis was performed on an Agilent UFLC 1200 system ((Mississauga, ON, CAN) with separation achieved with an Inertsil ODS-3 RP C_{18} column (4.6 x 250 mm, particle diameter 5 µm) (GL Science Inc., Torrance, CA, USA) using a gradient elution of mobile phase A: 0.1% aqueous o-phosphoric acid and mobile phase B: 100%

methanol; (a) 0-14 min: 0% B (0.5 mL/min), (b) 15-17: 100% B (1.5 mL/min), (c) 17.1-20 min: 0% B (1.5 mL/min), (d) 21.1-21.5 min: 0% B (0.5 mL/min). The analyte of interest was detected at 245 nm and the data collected and analyzed using Agilent ChemStation software, Rev. B.03.01.

Statistical Analysis

All of the above determinations were performed using n=5 for each *Vaccinium spp.* Correlations between the relative antioxidant potential of each species and the other variables analyzed above were determined by calculation of a Pearson product-moment correlation coefficient (*r*).

Metabolomic Profiling by UFLC-TOF-MS

Experiments were performed as per a previously established protocol (Brown, Finley, et al., n.d.) with an ACQUITY^{TM} series Ultra Performance Liquid Chromatography System (Waters Corporation, Mississauga,ON) coupled with a Micromass LCT Premier^{TM} series ToF-MS (Waters Corporation, Mississauga, ON) and controlled with a MassLynx V4.1 Data Analysis System (Waters Corporation). Phytochemicals were chromatographically separated at 30 °C on a Waters BEH Acquity C_{18} column (2.1 x 150 mm, 1.7 µm) and eluted with a gradient of 1% aqueous formic acid:acetonitrile (0.0 - 10.0 min, 95:5-5:95 v/v, 10.0-15.0 min, 5:95 v/v, 15.0-20.0 min, 5:95-95:5 v/v, 20.0-25.0 min, 95:5 v/v) over 25 min. at 0.25 mL/min. A Waters 1525 HPLC Binary Solvent Manager provided a steady flow of 2 ng/mL leucine enkephalin at 10 µL/min.

Exploratory Data Analysis

For the metabolomic data for each *Vaccinium* species the blanks were summed, subtracted against each sample, and any resultant negative values reset to zero. The data was then assessed without scaling or further data transformation. A series of automated functions to create subtractive datasets were created in Excel (Microsoft Corp., Redmond, WA, USA), as previously described (Brown, Finley, et al., n.d.) using sequential algorithms and functions designated "Subtractive Metabonomics". The data were then compiled to an ASCII text file with identifiers of samples in rows (objects) and retention time, m/z ratio and abundance as columns (variables). Further statistical analyses were accomplished using Solo+MIA (Eigenvector Research Inc., Wenatchee, WA, USA). Principal component analysis (PCA) was applied to the data, creating covariance matrices and transforming them into a coordinate system, as a means to observe variance for the multivariate dataset. Auto scaling was selected for pre-processing before applying the PCA algorithm. Both PCA score and loading plots were generated for the entire dataset to visualize clustering by *Vaccinium* species. The data was also modelled by Partial Least Squares Discriminant Analysis (PLS-DA) as a supervised approach to exploring clustering relationships.

Significance Analysis

Within a binary comparison between each pair of *Vaccinium* species, receiver operating characteristic (ROC) curves for each model were generated using R statistical software (The R Foundation for Statistical Computing, Vienna, Austria). The area under the ROC curve (AUC) value for each metabolite in the binary model was computed using the "colAUC" algorithm within the caTools package in R (Tuszynski, 2010). The mean of the three calculated binary AUC values per metabolite (m/z value) was obtained and

designated as “total AUC” (Hand & Till, 2001). For each metabolite the p-value was derived from the non-parametric Kruskal-Wallis test using the “kruskal.test” algorithm in R.

The SAM statistic (Tusher et al., 2001) and associated false discovery rates (FDR) were calculated using the SAM algorithm from the siggenes package in R comparing across all three *Vaccinium* species being studied. Due to the limited number of replicates, n=5, per species, permutations were used to set the expected *d(i)* as the null level of abundance which allows for the comparison of the observed *d(i)* based on actual data and the expected *d(i)*. To consider which metabolites differ significantly in terms of abundance, an artificially selected threshold (Δ) was applied to flag metabolites beyond this threshold boundary and the associated false discovery rate (FDR) determined. The metabolites as identified by the SAM statistic were distinguished in the plot of p-value vs. total AUC of each metabolite and examined for further interpretation.

Results

While the phytochemical composition of common cranberry (*Vaccinium macrocarpon* Aiton) has been relatively extensively studied, there have been very few studies of the native BC *Vaccinium spp.* including *V. oxycoccus* and *V. vitis-idaea*. Our objective was to develop a qualitative and quantitative understanding of the phytochemical consistency and diversity among wild and cultivated species of *Vaccinium*. To investigate the potential for the berries of the *Vaccinium spp*. to withstand environmental exposures that generate radical oxygen species, known metabolites with strong antioxidant potential

were determined (**Figure 6-2**), including anthocyanins, vitamin C (ascorbic acid), melatonin and serotonin.

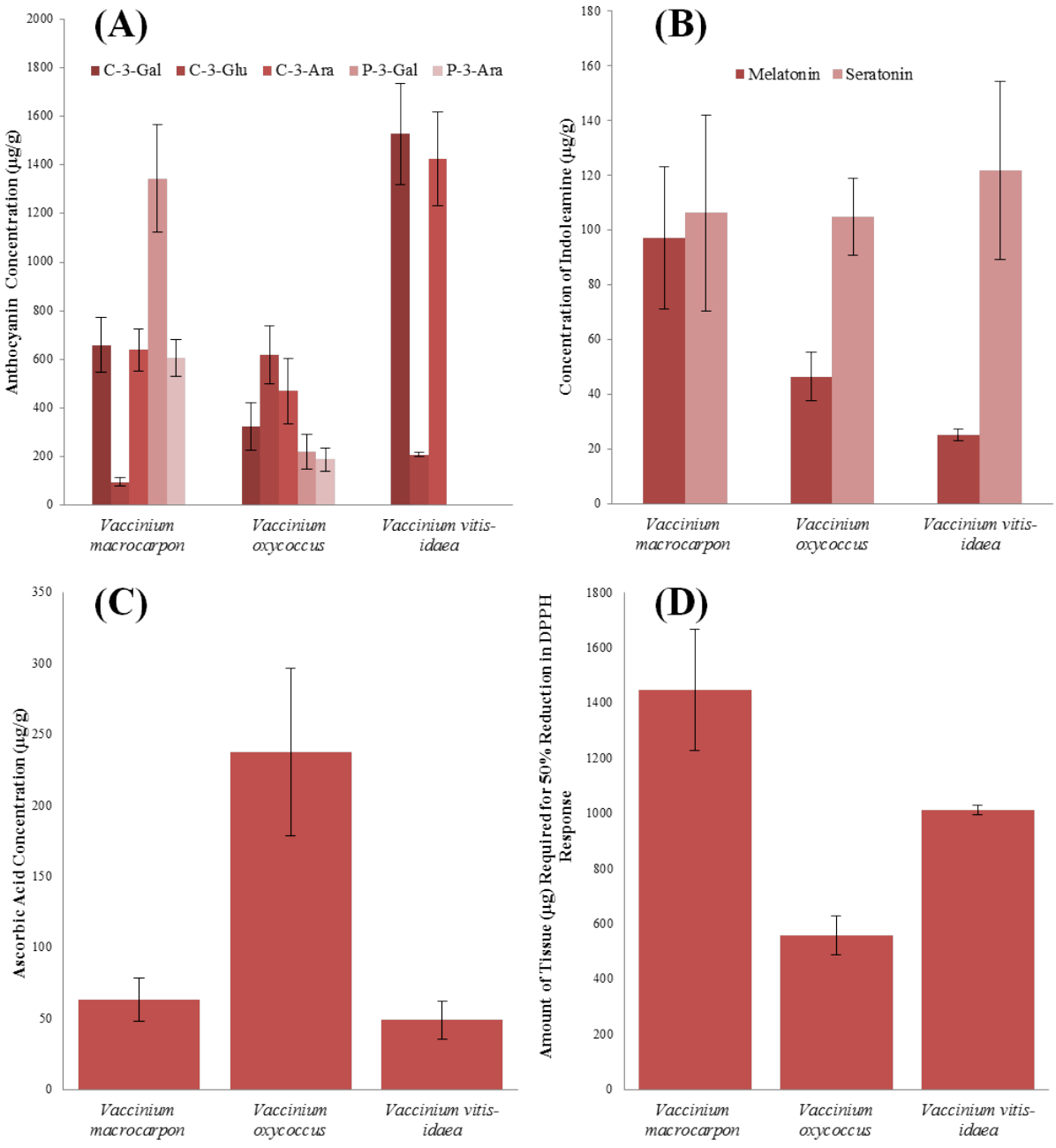

Figure 6-2: Phytochemical comparisons of *V. macrocarpon*, *V. oxycoccus, and V. vitis-idaea* as average amount in n=5 biological replicates with standard error of the mean **(A)** Anthocyanin concentration in dry tissue **(B)** Melatonin and seratonin content **(C)** Ascorbic acid content **(D)** Amount of cranberry tissue (n=3) for 50% reduction in DPPH response.

V. macrocarpon had the highest total anthocyanin content and all five of the major anthocyanins common to cranberry were quantified in the berries (**Figure 6-2A**), with peonidin-3-O-galactoside present at significantly greater concentrations than other anthocyanins. The *Vaccinium oxycoccus* fruit also contained all five of the major anthocyanins but the ratio of glycosylated peonidins to cyanidins was about 20:80, as compared to 60:40 in *V. macrocarpon* (**Figure 6-3**). The *V. vitis-idaea* berries contained measureable amounts of only the glycosylated cyanidin anthocyanins and did not contain measurable amounts of glycosylated peonidins (**Figure 6-2A**). Interestingly, there was a strong negative correlation (r = -0.92) between the anthocyanin content and the relative antioxidant potential (**Figure 6-2A**, **D**).

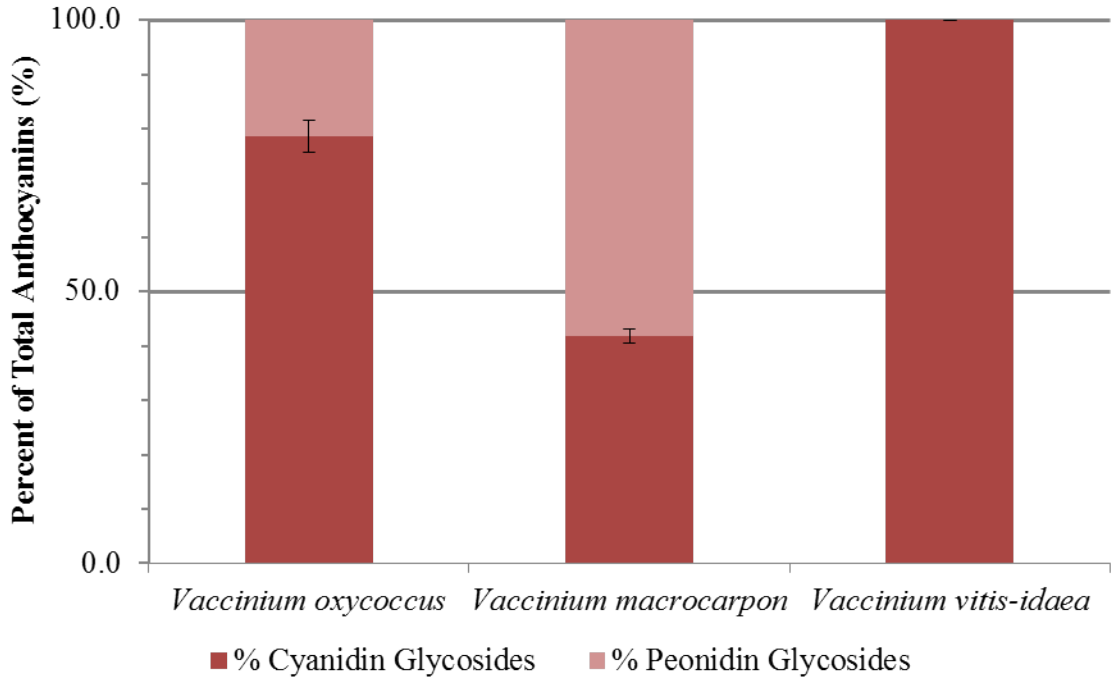

Figure 6-3: Comparison of the ratio of glycosylated peonidins to glycosylated cyanidins in the three *Vaccinium* species as a percent of the total anthocyanins ± standard error of the mean.

Melatonin and serotonin (**Figure 6-4**) are indoleamine neurohormones found in plants, bacteria, fungi and animals (Cole, Cao, Alan, Saxena, & Murch, 2008; Paredes, Korkmaz, Manchester, Tan, & Reiter, 2009) but have not previously been described in these *Vaccinium spp.* We here report the first evidence of these indoleamine antioxidants in *V. macrocarpon, V. oxycoccus* and *V. vitis-idaea* and further find that there was no significant difference in serotonin content between the species but the melatonin content was significantly higher in the commercially cultivated cranberry (*V. macrocarpon*) as compared to the two species native to BC (**Figure 6-2B**). Neither melatonin nor serotonin were strongly correlated with the relative antioxidant potential of the 3 *Vaccinium spp.* with correlation coefficients of r = -0.67 and r = -0.09, respectively (**Fiure 6-2B**, **D**). There were significant differences in the ascorbic acid contents of the 3 species (**Figure 6-2C**) and a strong positive correlation (r = 0.84) between ascorbic acid content and the relative antioxidant potential (**Figure 6-2C**, **D**).

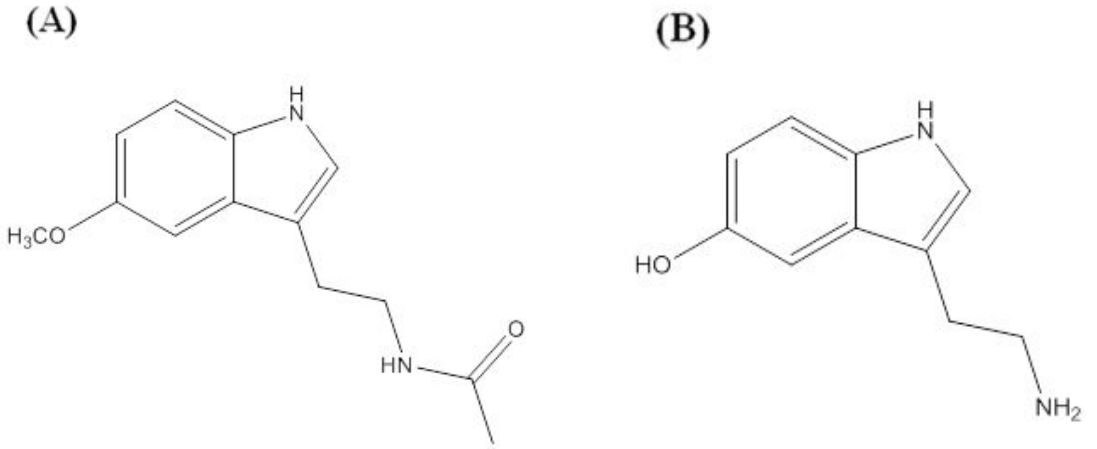

Figure 6-4: Chemical structures of **(A)** melatonin (N-acetyl-5-hydroxytryptamine) and **(B)** serotonin (5-hydroxytryptamine).

The metabolite counts and results of subtractive metabolomics shown in **Table 6-1** make direct comparisons across the three *Vaccinium* species. The total number of compounds observed in the metabolomics profiles was 10038 (*V. macrocarpon*), 8035 (*V. oxycoccus*) and 9285 (*V. vitis-idaea*). A total of 4626 compounds were conserved across all species and in binary comparisons 2257 compounds were conserved between *V. macrocarpon* and *V. vitis-idaea*, significantly more than *V. macrocarpon* and *V. oxycoccus* (1289) or *V. vitis-idaea* and *V. oxycoccus* (1391). Of all compounds observed in the metabolome of *V. macrocarpon*, 18.6% were unique to that species, while only about 10% of the compounds for *V. oxycoccus* and *V. vitis-idaea* were unique to those species.

Table 6-1: Summary of LC-MS-TOF metabolite counts and differences in metabolomic profiles of *V. macrocarpon, V. oxycoccus and V. vitis-idaea*

Description of occurance in metabolomic profiles	*Vaccinium macrocarpon*	*Vaccinium oxycoccus*	*Vaccinium vitis-idaea*
Total number of compounds detected	10038	8035	9285
Average number of compounds per biological replicate	4477±945	3415±600	4090±506
# Compounds found in all replicates	252	296	646
# Compounds in ? 50% of the replicates	3765	2656	3148
Unique to each species	1868	731	1013
V. macrocarpon & *V. vitis-idaea* only	2257		
V. macrocarpon & *V. oxycoccus* only	1289		
V. oxycoccus & *V. vitis-idaea* only	1391		
V. macrocarpon, *V. oxycoccus*, *V. vitis-idaea*	4624		

Commonalities and differences among the berries were determined by multivariate analysis using a PCA model. The score plot of the first (20.24%) and second (10.25%) principal component from the PCA model exhibited interesting clustering (**Figure 6-5A**) with almost complete overlap of the 95% confidence boundaries around *V. oxycoccus* and *V. vitis-idaea* and with the 95% confidence boundary for *V. macrocarpon* encompassing both of the other species data sets. With only 30.49% of the total variance (Q^2) explained

in the model of PC1 & PC2, a Kruskal-Wallis *p*-value was applied to the data and all values associated with a *p*-value <0.05 identified (**Figure 6-5B**). The majority of the values in this loadings plot are clustered on the left side indicating the importance of those metabolite values to *V. oxycoccus* and *V. vitis-idaea* in the score plot (**Figure 6-5A**) rather than for that of *V. macrocarpon.* After applying an inclusivity limit of *p*-value <0.05, the data was re-modeled and a loadings plot (**Figure 6-5C**) and score plot (**Figure 6-5D**) were generated. In the score plot some differentiation of the *Vaccinium* species is now observed with 37.51% explained variance in PC1 and 18.84% in PC2, suggesting this approach results in models that better represent the data in unsupervised analysis. Observing the metabolite value distribution in **Figure 6-5C** shows 3 clearly visible clusters and is an indication of the importance of those values to the positioning of the 3 species in the score plot (**Figure 6-5D**).

The Kruskal-Wallis *p*-value is often plotted against the area under the ROC curve or "AUC", which has been designated as total AUC for our multi-class comparison, as shown in **Figure 6-6**. Values that are plotted with a *p*-value of <0.05 and a total AUC > 0.5 are assessed for their importance in the differentiation in a multi-class comparison, such as the metabolomics profiles of the three *Vaccinium* species. The SAM statistic not only accounts for distribution across products but also the standard deviation of abundance in replicates and for the metabolomics data. For the comparison of the three species a list of significant m/z values, as identified by the SAM statistic at an FDR of 0.91%, is presented in **Table 6-2** and are also highlighted in red in **Figure 6-6** so it can be seen how this univariate significance test compares to AUC and *p*-value. The m/z values

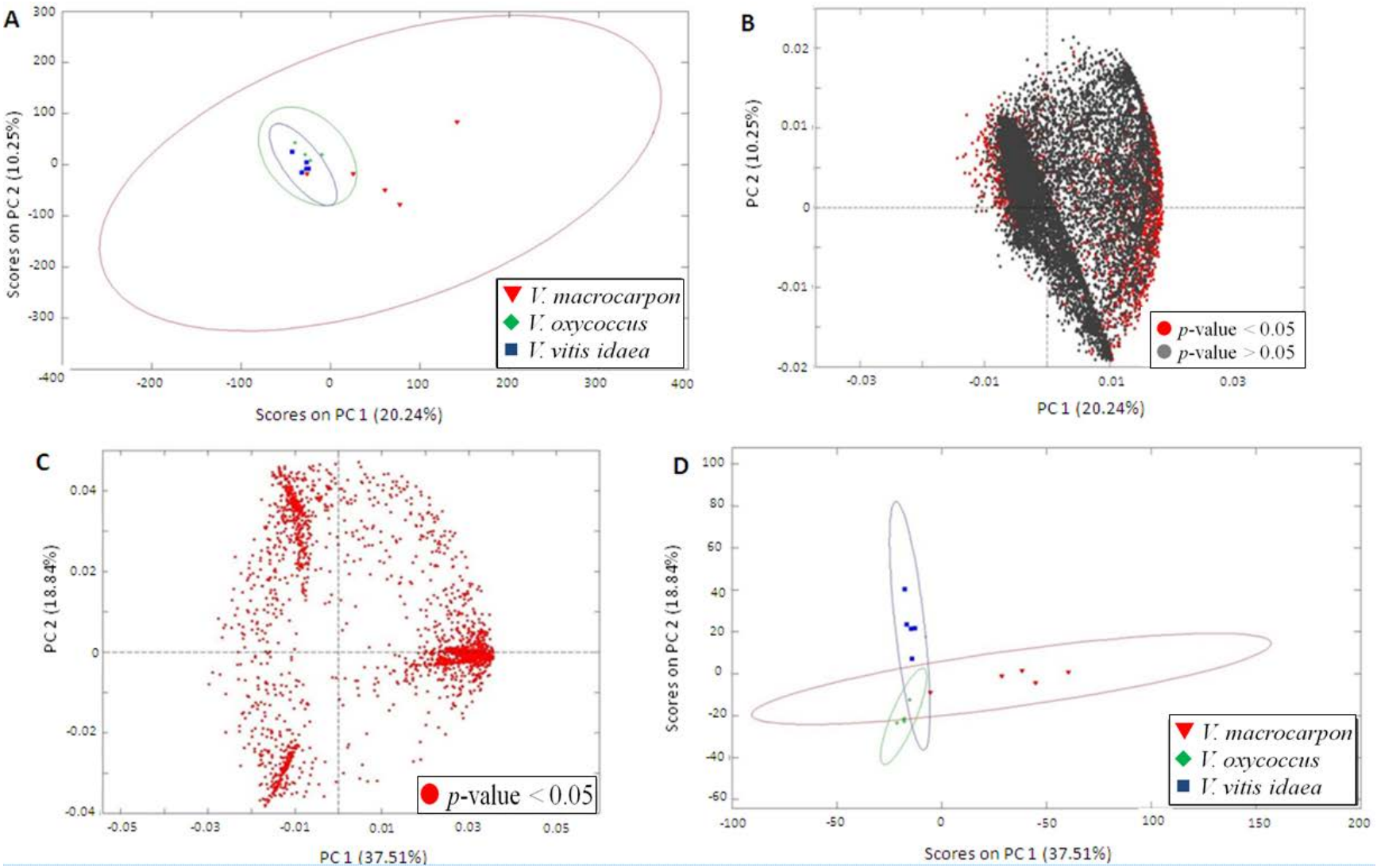

Figure 6-5: (A) Multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium* species. PCA Score plot of first two principal components and **(B)** the corresponding loadings plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value of >0.05 removed and **(D)** the re-modeled PCA score plot.

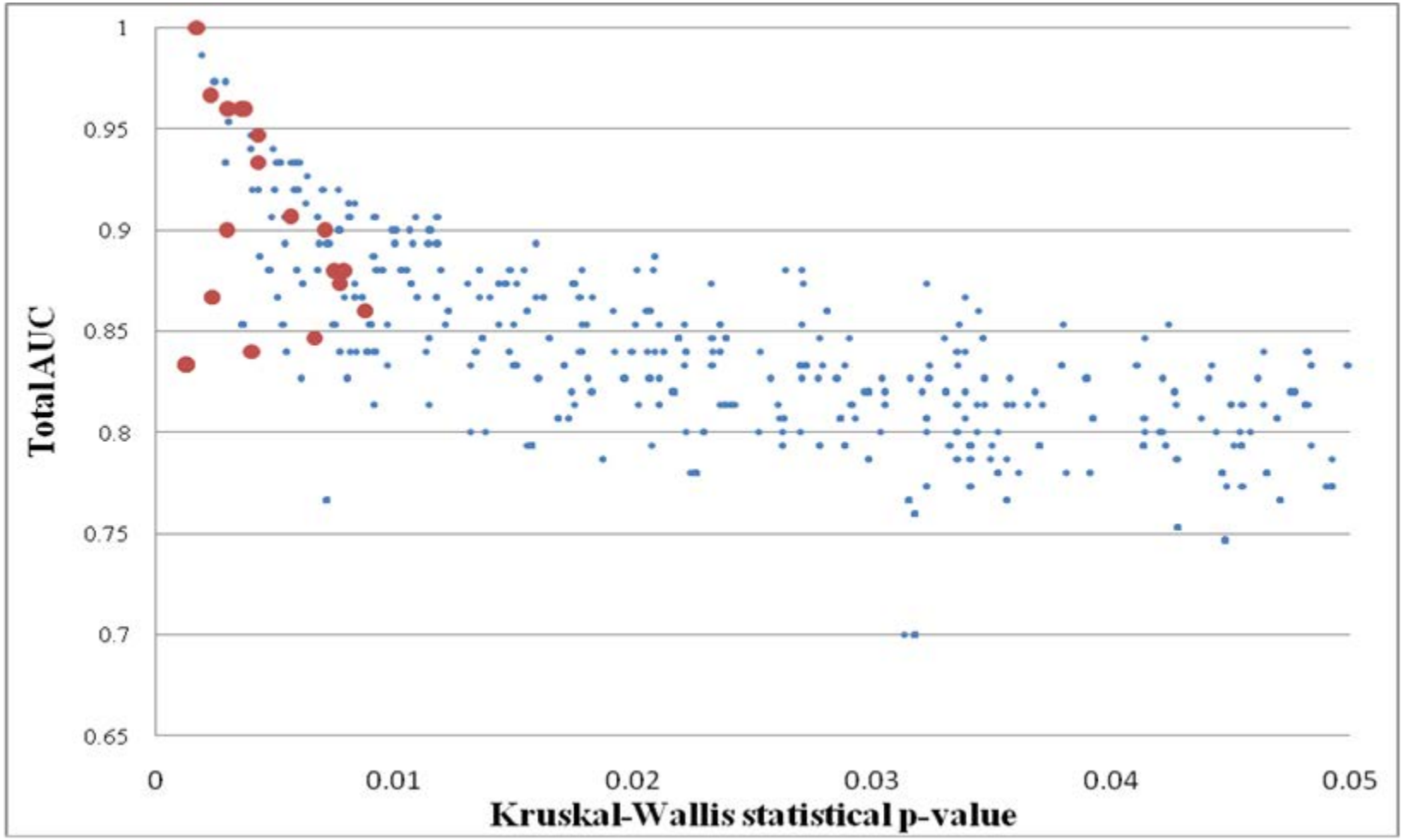

Figure 6-6: Curve generated from plotting the averaged area under the univariate ROC curve for comparisons across *Vaccinium* species versus the statistical *p*-value with variables identified as significant by the SAM statistic highlighted in red.

identified by the SAM statistic as significant had total AUC values ranging from 0.83 to 1.00 and although the SAM statistic ranks the values differently than the total AUC (**Table 6-2**) all values and are concentrated in the region of the plot of AUC against *p*-value that indicates significant data (**Figure 6-6**). The Kruksal-Wallis *p*-values ranged from 0.0013-0.0079 for the values identified by the SAM statistic, which were well below the limit for statistical significance of $p > 0.05$.

Partial least squares discriminant analysis (PLS-DA), a supervised multivariate approach, was applied to the metabolomics data and resulted in a clustering pattern in LV1 and LV2 (**Figure 6-7A**) very similar to the remodeled PCA (**Figure 6-5D**). The same process of identifying values with a *p*-value <0.05 (**Figure 6-7B**), removing the values with *p*-value > 0.05 and then remodeling LV 1 and LV2 to generate a loadings plot (**Figure 6-7C**) and score plot (**Figure 6-7D**) resulted in reduced overlap between the 95% confidenceboundaries around each *Vaccinium sp*. was observed. The distribution in the loadings plot (**Figure 6-7B**) and the number of values with $p<0.05$ (**Figure 6-7C**) should be related to the SAM statistic which contrasts the data sets by metabolite abundance but incorporates a measure of deviation across replicates within each data set. In **Figure 6-7C** the values identified by the SAM algorithm as significant do follow a pattern consistent with the PLS-DA model. For example, the 5 metabolite values identified as significant by SAM statistic and observed only in *V. vitis idaea* (**Table 6-2**) are positioned on the edge in the upper left corner of the *V. vitis idaea* cluster in the loadings plot (**Figure 6-7C**) which reflects the importance of these values in the PLS-DA score plot (**Figure 6-7D**) and indicates the PLS-DA model and results of the SAM algorithm are consistent.

For LV 1 and LV3 the PLS-DA score plot (**Figure 6-8A**) shows complete overlap of the 95% confidence boundaries of *V. vitis-idaea* and *V. oxycoccus* and somewhat with *V. macrocarpon*. Identifying the values with $p<0.05$ (**Figure 6-8B**) and remodeling the data after applying a cut off value of $p<0.05$ (**Figure 6-8C**) results in observed differentiation of *V. vitis-idaea* and *V. oxycoccus* from *V. macrocarpon* in the score plot (**Figure 6-8D**). Just as in the re-modeled PLS-DA of LV1 and LV2 (**Figure 6-7C, 6-7D**), the 5 metabolite values from *V. vitis-idaea* identified as significant in the SAM algorithm are within the clustered area of the loadings plot (**Figure 6-8C**) and from the score plot (**Figure 6-8D**) we can see these values are important to the positioning of *V. vitis-idaea* and for this PLS-DA model, also influence the position of *V. oxycoccus*. For the PLS-DA model of LV2 and LV4, which shows complete differentiation of the three *Vaccinium* species with a confidence boundary of 95% in the score plot (**Figure 6-9A**), an increase in differentiation is not observed in the remodeled score plot (**Figure 6-9D**). Consistent with the remodeled loadings plots for LV1/LV2 (**Figure 6-7C**) and LV1/ LV3 (**Figure 6-8C**), the 5 metabolite values identified as significant by the SAM algorithm are found to be positioned at the edge of the cluster associated with *V. vitis-idaea* in loading plot (**Figure 6-9C**). Again, this is an indication of the importance of these values to the positioning of *V. vitis-idaea* in the score plot (**Figure 6-9D**) and how the three species are differentiated in in this PLS-DA model.

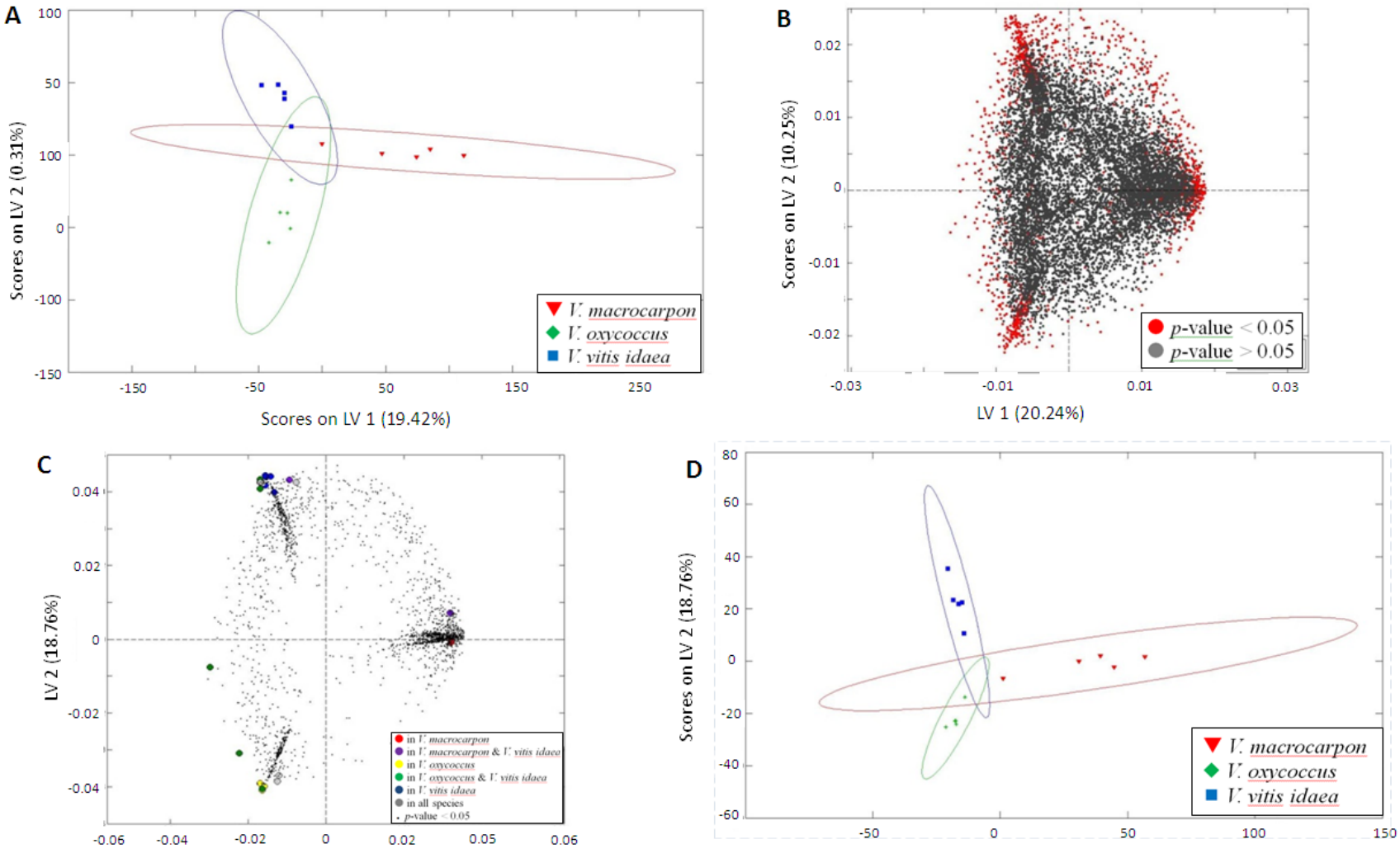

Figure 6-7: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium spp*. **(A)** PLS-DA Score plot of first and third linear variants and **(B)** The corresponding loadings plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. vitis-idaea* & *V. oxycoccus* (green) and all 3 species (grey) and **(D)** The re-modeled PLS-DA plot.

Table 6-2: The m/z values identified as significant and ranked by *d(i)* value, comparing across all three *Vaccinium* species, using the SAM algorithm in R at a false discover rate of 0.91% with the calculated Kruskal-Wallis *p*-value and total AUC.

m/z Value	Ret. Time	*p* -value	Rank, d(i)value	total AUC
Observed in all 3 species				
478.5575	5.19	0.0037	12th, 40.2	0.96
413.4574	5.68	0.0043	5th, 51.1	0.95
747.7100	5.75	0.0071	24th, 33.4	0.90
374.4192	5.47	0.0075	8th, 48.5	0.88
346.3744	5.46	0.0079	17th, 36.3	0.88
396.4475	4.05	0.0088	9th, 43.3	0.86
396.4536	4.37	0.0088	11th, 41.4	0.86
Observed in *V. macrocarpon* & *V. vitis-idaea* only				
413.4493	5.93	0.0030	10th, 42.6	0.96
348.4073	4.36	0.0057	14th, 37.9	0.91
Observed in *V. oxycoccus* and *V. vitis-idaea* only				
368.3629	5.68	0.0017	7th, 49.9	1.00
328.7924	5.90	0.0023	19th, 36.1	0.97
580.4006	5.72	0.0036	4th, 53.2	0.96
667.6591	3.98	0.0043	22nd, 33.8	0.93
575.6078	5.14	0.0077	1st, 105.9	0.87
Observed in *V. macrocapron* only				
739.7589	4.83	0.0024	16th, 36.4	0.87
492.4115	3.93	0.0024	18th, 36.2	0.87
Observed in *V. oxycoccus* only				
669.5615	1.45	0.0067	3rd, 70.8	0.85
333.7948	5.01	0.0013	2nd, 91.5	0.83
206.2153	5.13	0.0013	15th, 36.7	0.83
Observed in *V. vitis idaea* only				
390.3956	5.91	0.0030	6th, 50.9	0.90
395.4561	1.82	0.0024	20th, 34.6	0.87
354.3597	4.15	0.0040	13th, 39.2	0.84
350.3512	5.68	0.0013	21st, 33.8	0.83
390.4157	5.46	0.0013	23rd, 33.7	0.83

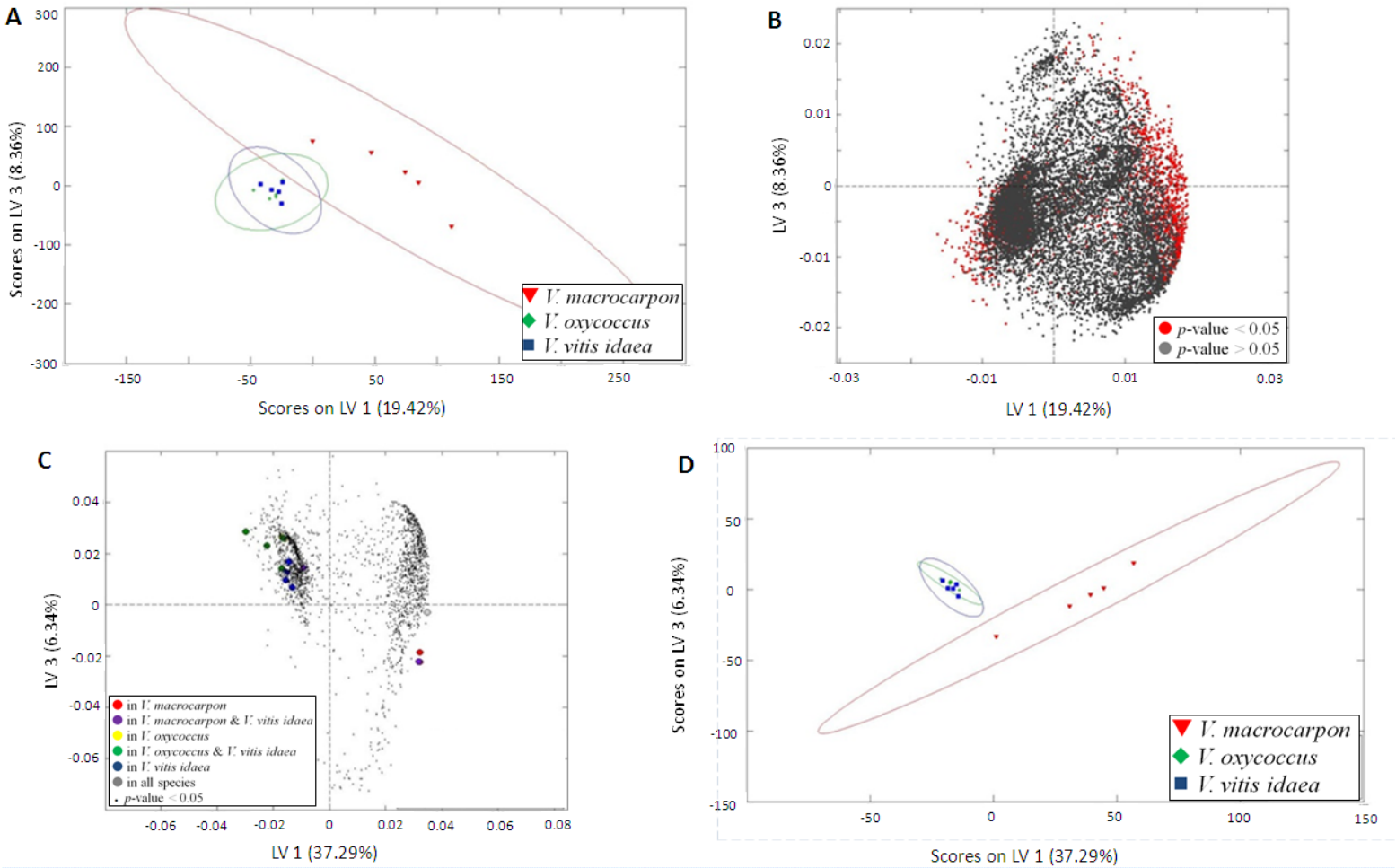

Figure 6-8: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of 3 *Vaccinium spp*. **(A)** PLS-DA Score plot of first and third linear variants and **(B)** the corresponding plot with values having a calculated Kruskal-Wallis *p*-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. oxycoccus* (green) and all 3 species (grey) and **(D)** the re-modeled PLS-DA plot.

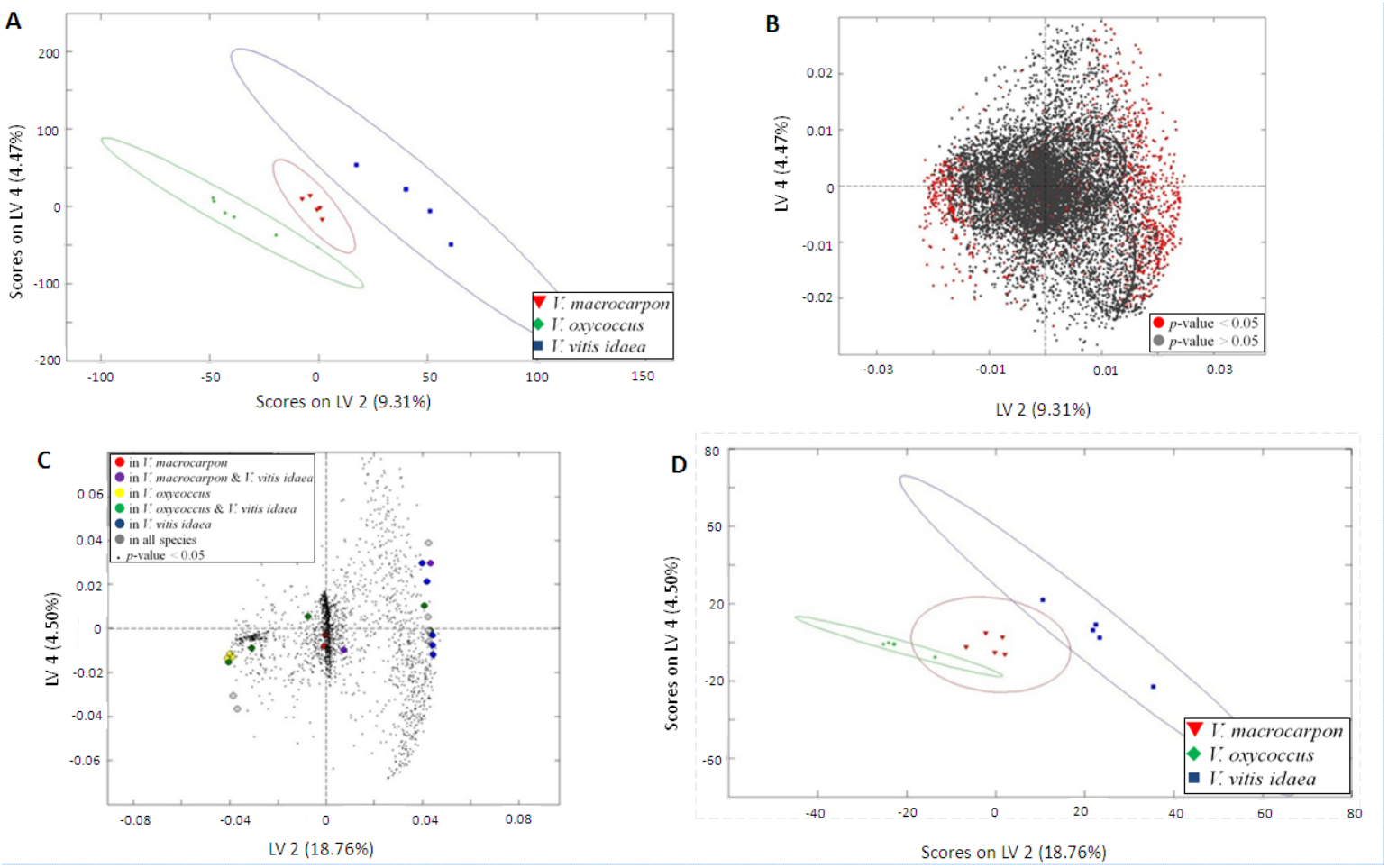

Figure 6-9: Supervised multivariate analysis of the LC-MS-TOF metabolomic profiles of three *Vaccinium spp*. **(A)** PLS-DA Score plot of second and fourth linear variants and **(B)** the corresponding loadings plot with values having a calculated Kruskal-Wallis p-value of <0.05 identified in red. **(C)** The re-modeled loadings plot with all values having a *p*-value >0.05 removed and with values identified as significant by SAM statistic highlighted according to their distribution across the species; observed only in *V. macrocarpon* (red), *V. vitis-idaea* (blue), *V. oxycoccus* (yellow), *V. macrocarpon* & *V. vitis-idaea* (purple), *V. vitis-idaea* & *V. oxycoccus* (green) and all 3 species (grey) and the remodeled PLS-DA plot.

Discussion

Two approaches to phytochemical characterization and differentiation between three *Vaccinium spp.* were employed in this study; targeted quantitative determination of known analytes of interest or physical properties (Cole et al., 2008) and untargeted metabolomics profiling (Murch et al., 2009). One of the most important aspects of the phytochemistry of plant tissues is the relative composition of phytochemicals with the potential to detoxify radical oxygen species, commonly known as antioxidant potential. For cranberry and other berries, it has been hypothesized that abundant anthocyanin composition provides evidence of strong antioxidant activity and has implications for human health (Moyer et al., 2002; Seeram, 2008a; Szajdek & Borowska, 2008; Zheng & Wang, 2003). Further, it has been also hypothesized that indoleamines such as melatonin are present in berries and other fruits to protect the genetic materials from oxidative damage due to environmental stresses (Manchester et al., 2000; Murch et al., 2009, 2010). Our data indicates that neither of these hypotheses can fully account for the phytochemical mechanisms in cranberry as the potential for detoxification of oxygen free radicals was more significantly correlated to ascorbic acid content than to the other antioxidants in the tissues. Previous researchers have also investigated the relationship between vitamin C and the total anthocyanin content of the commercial cranberry and found that the total anthocyanins but not ascorbic acid were significantly correlated with cytochrome c modulated oxidation of 6-hydroxydopamine, an important pathway for neurological health (Yao & Vieira, 2007). These results together with our results indicate two possibilities: (1) the total anthocyanin contents include antioxidant phytochemicals not detected with our targeted analysis of specific anthocyanins or (2)

specific physiological mechanisms affected by individual dietary antioxidants may not be elucidated in a measure of the total antioxidant potential.

The untargeted metabolomics approach to comparative phytochemistry is growing in popularity and application but the size and complexity of data sets can make experimental design and data interpretation difficult. The objectives of the comparative metabolomics described in this study were (1) to identify the degree of phytochemical commonality and difference among *Vaccinium* species and (2) to develop a model for experimental design and statistical analysis that could be applied to a range of other metabolomics profiling. One of the more interesting results of our metabolomics study is the discovery that 4624 compounds identified in the dataset are common to all three of the *Vaccinium* species. Given that we are comparing different species within a single genus and that one of the species was cultivated under commercial conditions while the other two were harvested from wild populations, the degree of phytochemical conservation was 46% of the total chemistry detected in *V. macrocarpon*, 57% of the total chemistry detected in *V. oxycoccus* and 49% of the total chemistry detected in *V. vitis-idaea*. This measure of the conservation of chemistries across species indicates the importance of primary metabolism in the growth and survival of plants under diverse microenvironments.

It has become most common to analyze metabolomics datasets by multivariate analysis using untargeted algorithms like Principle Component Analysis (PCA) (Jansen et al., 2010; Barry M. Wise & Gallagher, 1996) but our data indicate that these simplified analyses may lead to false discovery of phytochemical differences. Typically the influence of the metabolites (variables) in the loadings plots of multivariate analysis such

as PCA and PLS-DA are within the edge regions, furthest away from the origin. While this is true, the observed clustering in the modeling of the *Vaccinium spp.* metabolomics indicates that each species has a grouping of metabolites critical to the posting of the species in the score plot. Hence using the $p<0.05$ cut off ensures the majority of the metabolites kept are significant in the model which exhibits species differentiation in the scores plot, although some that were excluded (with *p*-value of >0.05) were important for the original model in each case (**Figure 7-9**). This indicates using a p-value cut off can have an important impact on how metabolomics data is interpreted when modeled in multivariate analysis.

The two algorithms, AUC and SAM, both capture the distribution of abundance in metabolomics profiles; however the approach is different. The ROC curve is a plot of sensitivity, defined as the true positive rate, and 1-specificty which corresponds to the false positive rate. The accuracy of this plot is determined by assessing the AUC, whereby an AUC of 1 would indicate 100% sensitivity at 0% false positive rate. This analysis becomes more meaningful when taken together with p-values for significance, as in **Figure 6- 6,** and is a way to prioritize metabolomics data and assess the quality of our regression models (PCA, PLS-DA). However, AUC is a binary comparison directly comparing only two species, so for the three species comparison all possible binary comparisons are made and the AUC values are averaged to a "total" AUC to capture all possible comparisons (Hand & Till, 2001). In this way AUC is unable to capture the interactions or relationships that may exist across the metabolites in the metabolomics profiles. The SAM statistic is a step towards capturing the more complex relationships as the degree of significance is compared to the null distribution of the dataset by using

permutations of all metabolite abundances and calculating an associated FDR which is not captured in AUC algorithms.

With the use of the SAM statistic it becomes possible to identify the most significant metabolites per species, as shown in **Table 6-2**. When comparing the SAM values directly to the loadings plot generated in multivariate analysis (see **Figures 6-7C**,**6- 8C** and **6-9C**) the distribution of the values and the number of significant values should be consistent if the MVA model is a good fit for the metabolomics data. The PLS-DA loadings plot of LV2 and LV4, after the p-value <0.05 limit was applied (**Figure 6- 9C**), shows the tightest grouping of SAM significant values originate from *V. vitis-idaea*, whereas the SAM significant values from *V. macrocarpon* and *V. oxycoccus* are more evenly distributed. The results of the SAM algorithm support the positioning of the three species in the LV2/LV4 score plot (**Figure 6- 8D**) and indicates the observed differentiation of the species is a reasonable interpretation of the metabolomics data. Differences are evident between the ranking list from the SAM statistic and the degree of the AUC value (**Table 6-2**), for example an AUC value of 1 is reached with m/z 638.3629, yet is not the topped ranked metabolite value at a *d(i)* of 49.9. Although ranking order in the two approaches is different, all values identified by the SAM algorithm as significant have high AUC values and both operations support the regression models of the metabolomics data. Ideally PLS-DA models of metabolomics data should be validated (Westerhuis et al., 2008) but in cases where the sample size is insufficient to support cross-validation, the combined approach described in this study can provide for a quality assessment based on significance, sensitivity and specificity.

Chapter 7: Conclusion

With the prevalence of cranberry products in the marketplace and ongoing interest in the scientific community on the potential health benefits of cranberry fruit, it is important that tools are developed to better characterize these complex products, ensuring their quality and efficacy. In this thesis, I describe five bodies of work to develop and utilize new tools for characterizing cranberry fruit.

The critical first objective of the thesis was to develop a validated analytical method to quantify the five major anthocyanins in cranberries (*Vaccinium macrocarpon*). An HPLC method was optimized and subjected to a single laboratory validation study as per AOAC guidelines (AOAC International, 2002), a stability study, and Youden ruggedness trial. The results of these studies demonstrated that the method described is fit for the purpose of accurate determination of cyanidin-3-galactoside, cyaniding-3-glucoside, cyanidin-3-arabinoside, peonidin-3-galactoside and peonidin-3-arabinoside in cranberry fruit and finished products (dried extracts, juice and juice cocktail). This work has been published in the *Journal of the AOAC International*.

The second objective of the thesis was the large-scale investigation of the chemical diversity in the anthocyanin content of commercial cranberries at harvest. Samples from 20 different growers were compared to determine the range of different anthocyanin contents in the cranberry crop. The validated method was adapted to ultra-fast liquid chromatography, thereby reducing the run time from 30 to 10 minutes without loss of resolution, allowing for rapid high throughput analysis. Our results show significant variance between grower/producer operations and an interesting conservation of the

relative amounts of cyanidin and peonidin in the berries regardless of the total anthocyanin contents. These data provide an interesting possible avenue for future basic research to determine whether there is preferential biochemical synthesis in the pathway or whether the transport of specific anthocyanins to the developing fruit is favoured in cranberry. The universal distribution of anthocyanins across the cranberry fruit could make it an ideal model system for studies of source-sink relationships in flavonoid metabolism and biochemical studies of plant metabolism. This work has been submitted as an original article to the journal *Phytochemical Analysis*.

The third objective of the thesis was to develop a comprehensive model for metabolomic analyses in cranberry that can be used for phytochemical discovery and assuring fruit identity and quality. Metabolomics is a rapidly growing field and Principle Component Analysis or other types of multivariate analysis have been described as standard approaches to analyzing the large data sets that are generated in untargeted analyses. The cranberry metabolome was characterized by a standardized UFLC reverse phase separation with time of flight mass spectrometry for compound identification. New statistical tools were designed for prioritizing and evaluating significance thereby considering the robustness of the models developed. This “chemoinformatic” approach examines the potential of false discoveries or false positives when interpreting metabolomic data and offers a new model for interpreting metabolomics data that could be applied to investigations of other plant species. The key and novelty of this approach is that it overcomes the limitations of existing methods for describing phytochemical composition and provides an excellent and reproducible tool for describing composition of the entire plant. This tool will be invaluable in basic research on the chemical features

of medicinal and health-promoting plants that describe “quality” in the sense of efficacy and safety. Once the metabolome is known, the techniques can be used for basic quality assurance by discovering the optimum metabolic “fingerprint” and then providing a reliable method for determining whether or not each batch of raw material or finished product match the optimum.

The fourth objective of the research was to use the targeted and untargeted analytical tools that have been developed to investigate the similarities and differences between varietals of *Vaccinium macrocarpon* Aiton grown under standardized commercial conditions in a single watershed in the lower mainland of British Columbia. Five cultivars of commercial cranberries, viz. Ben Lear, Bergman, GH1, Pilgrim and Stevens, were collected from production facilities in the lower mainland of British Columbia for phytochemical analysis of anthocyanin and metabolome contents. Significant differences were found in the anthocyanin contents of the cultivarss with the most commonly planted cultivar, Stevens, showing the lowest anthocyanin contents. If the goal of the commercial production is yield of anthocyanins, then mass plantings of high anthocyanin cultivars such as Ben Lear, Bergman or GH1 would be recommended. However, other agronomic considerations such as yield and disease resistance may be reflected in the large scale production of ‘Stevens’ by commercial farmers. Further, the metabolomics analysis identified individual metabolites that were characteristic of each of the varietals. Using the chemoinformatics approaches, the combined application of the univariate approaches, statistical Kruskal-Wallis *p*-values, area under the ROC curve and SAM statistic provide the basis for establishing quality evaluations of multivariate models as a new tool for metabolomics research with broad future applications.

The final objective of the thesis was to characterize the phytochemical diversity among two species of traditionally used, west coast small cranberry (*Vaccinium oxycoccus* L. and *Vaccinium vitis-idaea* L.) with the large commercial cranberry *(Vaccinium macrocarpon* Ait.) indigenous to the east coast of North America and cultivated in British Columbia. *V. oxycoccus* and *V. macrocarpon* contained the five major anthocyanins known in cranberry but the ratio of glycosylated peonidins to cyanidins varied between them. *V. vitis-idaea* did not contain measurable amounts of glycosylated peonidins. Extracts of all three berries were found to contain serotonin, melatonin and ascorbic acid as well as 8,000 – 10,000 unidentified phytochemicals. Antioxidant potential was not correlated with anthocyanin or indoleamine content but was positively correlated with vitamin C contents. From significance analysis and statistical modeling it was found 2 compounds in *V. macrocarpon,* 3 in *V. oxyc*occus and 5 in *V. vitis-idaea* were both unique to each species and key to the characterization and differentiation of these cranberry metabolomes. Through multivariate modeling differentiation of the species was observed and univariate statistical analysis was employed to provide a quality assessment of the models developed for the metabolomics data. Together, these data provide foundational knowledge about the chemotaxonomy of the *Vaccinium* genus and provide new insights into the maintenance of health in the traditional diets of North Americans.

In summary, the research described in this thesis includes the creation of a toolbox of selected and nonselective chemometric, analytical and statistical tools have been developed and utilized to describe relationships between commercially harvested

cranberries, varietals of the American large cranberry, and highly related North American cranberry species.

References

Abia, R., & Fry, S. C. (2001). Degradation and metabolism of 14C-labelled proanthocyanidins from carob (*Ceratonia siliqua*) pods in the gastrointestinal tract of the rat. *Journal of the Science of Food and Agriculture*, *81*(12), 1156-1165.

Agriculture & Agrifood Canada. (2010). What's new in BC - Spotlight on BC : berries - Summer 2010 (blueberries and cranberries). Retrieved June 13, 2011, from http://www.ats.agr.gc.ca/reg/5552-eng.htm

Agriculture Research Service. (2009). NCGR-Corvallis: *Vaccinium* catalog cranberry cultivars and selections. Retrieved June 24, 2011, from http://www.ars-grin.gov/cor/catalogs/vaccran.html

Allwood, J. W., & Goodacre, R. (2010). An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical Analysis*, *21*(1), 33-47.

Allwood, W. J., Erban, A., de Koning, S., Dunn, W. B., Luedemann, A., Lommen, A., Kay, L., et al. (2009). Inter-laboratory reproducibility of fast gas chromatography–electron impact–time of flight mass spectrometry (GC–EI–TOF/MS) based plant metabolomics, *5*(4), 479-496.

Andersen, Ø. M. (1989). Anthocyanins in fruits of *Vaccinium oxycoccus* L. (small cranberry). *Journal of Food Science*, *54*(2), 383-384.

AOAC International. (2002). *Guidelines for single laboratory validation of chemical methods for dietary supplements and botanicals*. Baltimore, MD, USA: AOAC International.

Asami, D. K., Hong, Y.-J., Barrett, D. M., & Mitchell, A. E. (2003). Comparison of the total phenolic and ascorbic acid content of freeze-dried and air-dried marionberry, strawberry, and corn grown using conventional, organic, and sustainable agricultural practices. *Journal of Agricultural and Food Chemistry*, *51*(5), 1237-1241. doi:10.1021/jf020635c

Baba, S., Osakabe, N., Natsume, M., & Terao, J. (2002). Absorption and urinary excretion of procyanidin B2 [epicatechin-(4[beta]-8)-epicatechin] in rats. *Free Radical Biology and Medicine*, *33*(1), 142-148. doi:16/S0891-5849(02)00871-7

Basu, A., Rhone, M., & Lyons, T. J. (2010). Berries: emerging impact on cardiovascular health. *Nutrition Reviews*, *68*(3), 168-177. doi:10.1111/j.1753-4887.2010.00273.x

BC Cranberry Growers Association. (2009). Industry : British Columbia cranberry growers association. Retrieved June 13, 2011, from http://www.bccranberrygrowers.com/industry/

BC Ministry of Agriculture. (2010). BC farm products A-Z - Cranberry. Retrieved June 13, 2011, from http://www.agf.gov.bc.ca/aboutind/products/plant/cranberry.htm

Bertram, H. C., Weisbjerg, M. R., Jensen, C. S., Pedersen, M. G., Didion, T., Petersen, B. O., Duus, J. Ø., et al. (2010). Seasonal changes in the metabolic fingerprint of 21 grass and legume cultivars studied by nuclear magnetic resonance-based metabolomics. *Journal of Agricultural and Food Chemistry*, *58*(7), 4336-4341. doi:10.1021/jf904321p

Betz, J. M., Brown, P. N., & Roman, M. C. (2011). Accuracy, precision, and reliability of chemical measurements in natural products research. *Fitoterapia*, *82*(1), 44-52. doi:16/j.fitote.2010.09.011

Betz, J. M., Fisher, K. D., Saldanha, L. G., & Coates, P. M. (2007). The NIH analytical methods and reference materials program for dietary supplements. *Analytical and Bioanalytical Chemistry*, *389*(1), 19-25. doi:10.1007/s00216-007-1342-8

Blumenthal, M., Linstrom, A., Lynch, M. E., & Rea, P. (2011). Herb sales continue growth - Up 3.3% in 2010. *HerbalGram*, *90*, 64-67.

Boccalandro, H. E., González, C. V., Wunderlin, D. A., & Silva, M. F. (2011). Melatonin levels, determined by LC-ESI-MS/MS, fluctuate during the day/night cycle in *Vitis vinifera* cv Malbec: evidence of its antioxidant role in fruits. *Journal of Pineal Research*, *51*(2), 226-232. doi:10.1111/j.1600-079X.2011.00884.x

Borowksa, E. J., Mazur, B., Kopciuch, R. G., & Buszewki, B. (2009). Polyphenol, anthocyanin and resveratrol mass fractions and antioxidant properties of cranberry cultivars. *Food Technology and Biotechnology*, *47*(1), 56-61.

Breiman, L. (2001). Statistical modeling: The two cultures. *STATISTICAL SCIENCE*, *16*(3), 199-231.

Broadhurst, D. I., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, *2*(4), 171-196. doi:10.1007/s11306-006-0037-z

Brown, P. N., & Shipley, P. R. (2011). Determination of anthocyanins in cranberry fruit and cranberry fruit products by high-performance liquid chromatography with ultraviolet detection: Single-laboratory validation. *Journal of AOAC International*, *94*(2), 459-466.

Brown, P. N., Finley, J. P., Le, C. H., & Murch, S. J. (n.d.). Metabolomic profiling of cranberry (*Vaccinium macrocarpon* Aiton) fruit and cranberry dietary supplements by time-of-flight mass spectrometry with chemometric analysis. *Analytical and Bioanalytical Chemistry*.

Brown, P. N., Shipley, P. R., & Murch, S. J. (n.d.). Anthocyanin diversity within a commercial production of cranberries (*Vaccinium macrocarpon* Aiton) in British Columbia. *Phytochemical Analysis*.

Brown, P. N., Turi, C. E., Shipley, P. R., & Murch, S. J. (2011). Phytochemical discovery in large cranberry (*Vaccinium macrocarpon* Aiton.), small cranberry (*Vaccinium oxycoccus*) and Lingonberry (*Vaccinium vitis-idaea*). *Planta Medica*.

Brunner, E., Rees, K., Ward, K., Burke, M., & Thorogood, M. (2007). Dietary advice for reducing cardiovascular risk. *Cochrane Database of Systematic Reviews*, *17*(4), CD002128. doi:10.1002/14651858.CD002128.pub3

Cassinese, C., de Combarieu, E., Falzoni, M., Fuzzati, N., Pace, R., & Sardone, N. (2007). New liquid chromatography method with ultraviolet detection for analysis of anthocyanins and anthocyanidins in *Vaccinium myrtillus* fruit dry extracts and commercial preparations. *Journal of AOAC International*, *90*(4), 911-919.

Cellini, F., Chesson, A., Colquhoun, I., Constable, A., Davies, H. V., Engel, K. H., Gatehouse, A. M. R., et al. (2004). Unintended effects and their detection in genetically modified crops. *Food and Chemical Toxicology*, *42*(7), 1089-1125. doi:16/j.fct.2004.02.003

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochemistry and Photobiology*, *70*(1), 1-9. doi:10.1111/j.1751-1097.1999.tb01944.x

Chandra, A., Rana, J., & Li, Y. (2001). Separation, identification, quantification, and method validation of anthocyanins in botanical supplement raw materials by HPLC and HPLC−MS. *Journal of Agricultural and Food Chemistry*, *49*(8), 3515-3521. doi:10.1021/jf010389p

Chen, F., Sun, Y., Zhao, G., Liao, X., Hu, X., Wu, J., & Wang, Z. (2007). Optimization of ultrasound-assisted extraction of anthocyanins in red raspberries and identification of anthocyanins in extract using high-performance liquid chromatography-mass spectrometry. *Ultrasonics Sonochemistry*, *14*(6), 767-778. doi:16/j.ultsonch.2006.12.011

Chen, P., Ozcan, M., & Harnly, J. (2007). Chromatographic fingerprint analysis for evaluation of *Ginkgo biloba* products. *Analytical and Bioanalytical Chemistry*, *389*(1), 251-261. doi:10.1007/s00216-007-1386-9

Cherubini, A., Ruggiero, C., Morand, C., Lattanzio, F., Dell'Aquila, G., Zuliani, G., Di Iorio, A., et al. (2008). Dietary antioxidants as potential pharmacological agents for ischemic stroke. *Current Medicinal Chemistry*, *15*(12), 1236-1248.

Clifford, M. N. (2000). Anthocyanins – nature, occurrence and dietary burden. *Journal of the Science of Food and Agriculture*, *80*(7), 1063-1072. doi:10.1002/(SICI)1097-0010(20000515)80:7<1063::AID-JSFA605>3.0.CO;2-Q

Coates, P. M., & Meyers, C. M. (2011). The National Institutes of Health investment in research on botanicals. *Fitoterapia*, *82*(1), 11-13. doi:16/j.fitote.2010.11.013

Cole, I., Cao, J., Alan, A., Saxena, P., & Murch, S. J. (2008). Comparisons of *Scutellaria baicalensis, Scutellaria lateriflora* and *Scutellaria racemosa:* Genome Size, Antioxidant Potential and Phytochemistry. *Planta Medica*, *74*(4), 474-481. doi:10.1055/s-2008-1034358

Cozzolino, D. (2011). Infrared methods for high throughput screening of metabolites: food and medical applications. *Combinatorial Chemistry & High Throughput Screening*, *14*(2), 125-131.

Cunningham, D. G., Vannozzi, S. A., Turk, T., Roderick, R., O'Shea, E., & Brilliant, K. (2004). Cranberry phytochemicals and their health benefits. In F. Shahidi & D. K. Weerasinghe (Eds.), *Nutraceutical Beverages: Chemistry, Nutrition and Health*. Washington, DC, USA: American Chemical Society.

Dana, M. N. (1983). Cranberry cultivar list (*Vaccinium macrocarpon*). *Fruit Varieties Journal*, *37*, 88-95.

Daviss, B. (2005). Growing pains for metabolomics. *The Scientist*, *19*(5), 25.

de Pascual-Teresa, S., Santos-Buelga, C., & Rivas-Gonzalo, J. C. (2002). LC–MS analysis of anthocyanins from purple corn cob. *Journal of the Science of Food and Agriculture*, *82*(9), 1003-1006. doi:10.1002/jsfa.1143

DeFuria, J., Bennett, G., Strissel, K. J., Perfield, J. W., Milbury, P. E., Greenberg, A. S., & Obin, M. S. (2009). Dietary blueberry attenuates whole-body insulin resistance in high fat-fed mice by reducing adipocyte death and its inflammatory sequelae. *The Journal of Nutrition*, *139*(8), 1510-1516. doi:10.3945/jn.109.105155

Del Bas, J. M., Fernández-Larrea, J., Blay, M., Ardèvol, A., Salvadó, M. J., Arola, L., & Bladé, C. (2005). Grape seed procyanidins improve atherosclerotic risk index and induce liver CYP7A1 and SHP expression in healthy rats. *The FASEB Journal*, *19*(3), 479-481. doi:10.1096/fj.04-3095fje

Deubert, K. H. (1978). A rapid method for the extraction and quantitation of total anthocyanin of cranberry fruit. *Journal of Agricultural and Food Chemistry*, *26*(6), 1452-1453. doi:10.1021/jf60220a041

Di Martino, P., Agniel, R., David, K., Templer, C., Gaillard, J. L., Denys, P., & Botto, H. (2006). Reduction of Escherichia coli adherence to uroepithelial bladder cells after consumption of cranberry juice: a double-blind randomized placebo-controlled cross-over trial. *World Journal of Urology*, *24*(1), 21-27. doi:10.1007/s00345-005-0045-z

Douglas, G. W., Meidinger, D. V., & Pojar, J. (1999). *Illustrated flora of British Columbia, volume 3: Dicotyledons (Diapensiaceae through Onagraceae).*

Victoria, BC, Canada: B.C. Ministry of Environment, Lands and Parks and B.C. Ministry of Forests.

Dunn, W. B., Bailey, N. J. C., & Johnson, H. E. (2005). Measuring the metabolome: Current analytical technologies. *The Analyst, 130*(5), 606-625. doi:10.1039/b418288j

Eissa, A. B., Batson, G., Wise, B. M., & Gallagher, N. B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control, 6*, 329-348. doi:10.1016/0959-1524(96)00009-1

Enot, D. P., & Draper, J. (2007). Statistical measures for validating plant genotype similarity assessments following multivariate analysis of metabolome fingerprint data. *Metabolomics, 3*(3), 349-355. doi:10.1007/s11306-007-0066-2

Environmental Protection Agency. (2003). Guidelines establishing test procedures for the analysis of pollutants; Procedures for detection and quantification, 40 CFR pt. 146, Appendix B rev. 1.11.

Fernie, A. R., & Schauer, N. (2009). Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics, 25*(1), 39-48. doi:16/j.tig.2008.10.010

Filzmoser, P., & Varmuza, K. (2010). Chemometrics: Multivariate statistical analysis in chemometrics. 2010, . Retrieved May 27, 2011, from http://cran.r-project.org/web/packages/chemometrics/index.html

Frost, C. J., Mescher, M. C., Carlson, J. E., & De Moraes, C. M. (2008). Plant defense priming against herbivores: Getting ready for a different battle. *Plant Physiology, 146*(3), 818-824. doi:10.1104/pp.107.113027

Fuleki, T., & Francis, F. J. (1968). Quantitative methods for anthocyanins. 1. Extraction and determination of total anthocyanin in cranberries. *Journal of Food Science, 33*(1), 72-78.

Gagnier, J., Demelo, J., Boon, H., Rochon, P., & Bombardier, C. (2006). Quality of reporting of randomized controlled trials of herbal medicine interventions. *The American Journal of Medicine, 119*(9), 800.e1-800.e11. doi:10.1016/j.amjmed.2006.02.006

Gonthier, M.-P., Donovan, J. L., Texier, O., Felgines, C., Remesy, C., & Scalbert, A. (2003). Metabolism of dietary procyanidins in rats. *Free Radical Biology and Medicine, 35*(8), 837-844. doi:16/S0891-5849(03)00394-0

Government of Canada. (2011a). Monograph: Cranberry Juice, Dried. Retrieved June 13, 2011, a from http://webprod.hc-sc.gc.ca/nhpid-bdipsn/monoReq.do?id=292&lang=eng

Government of Canada. (2011b). Monograph: Cranberry. Retrieved June 13, 2011, b from http://webprod.hc-sc.gc.ca/nhpid-bdipsn/monoReq.do?id=71&lang=eng

Greenberg, J. A., Newmann, S. J., & Howell, A. B. (2005). Consumption of sweetened dried cranberries versus unsweetened raisins for Inhibition of uropathogenic *Escherichia coli* adhesion in human urine: A pilot study. *Journal of Alternative & Complementary Medicine, 11*(5), 875-878. doi:10.1089/acm.2005.11.875

Hall, R. D., Vos, C. H. R., Verhoven, H. A., & Bino, R. J. (2005). Metabolomics for the assessment of functional diversity and quality traits in plants. In S. Vaidyanathan, G. G. Harrigan, & R. Goodacre (Eds.), *Metabolomic analyses: Strategies for systems biology* (pp. 31-45). New York, NY, USA: Springer.

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). East Lansing, MI, USA: Springer. Retrieved from http://www.springer.com/life+sciences/book/978-1-4020-6906-2

Hand, D. J., & Till, R., J. (2001). A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning, 45*(2), 171-186.

He, F., Mu, L., Yan, G.-L., Liang, N.-N., Pan, Q.-H., Wang, J., Reeves, M. J., et al. (2010). Biosynthesis of anthocyanins and their regulation in colored grapes. *Molecules, 15*(12), 9057-9091. doi:10.3390/molecules15129057

Hong, V., & Wrolstad, R. E. (1986a). Cranberry juice composition. *Journal of Association of Official Analytical Chemist, 69*(2), 199-207.

Hong, V., & Wrolstad, R. E. (1986b). Detection of adulteration in commercial cranberry juice drinks and concentrates. *Journal of Association of Official Analytical Chemist, 69*(2), 208-213.

Hopkins, R. J., van Dam, N. M., & van Loon, J. J. A. (2009). Role of glucosinolates in insect-plant relationships and multitrophic interactions. *Annual Review of Entomology, 54*(1), 57-83. doi:10.1146/annurev.ento.54.110807.090623

Horwitz, W. (1982). Evaluation of analytical methods used for regulation of foods and drugs. *Analytical Chemistry, 54*(1), 67A-76A. doi:10.1021/ac00238a002

Howell, A. B. (2007). Bioactive compounds in cranberries and their role in prevention of urinary tract infections. *Molecular Nutrition & Food Research, 51*(6), 732-737.

Issaq, H. J., Van, Q. N., Waybright, T. J., Muschik, G. M., & Veenstra, T. D. (2009). Analytical and statistical approaches to metabolomics research. *Journal of Separation Science, 32*(13), 2183-2199. doi:10.1002/jssc.200900152

Jacquemart, A.-L. (1997). Vaccinium oxycoccos L. (*Oxycoccus palustris* Pers.) and *Vaccinium microcarpum* (Turcz. ex Rupr.) Schmalh. (*Oxycoccus microcarpus* Turcz. ex Rupr.). *Journal of Ecology*, *85*(3), 381-396. doi:10.2307/2960511

Janecki, A., & Kolodziej, H. (2010). Anti-adhesive activities of flavan-3-ols and proanthocyanidins in the interaction of group A-Streptococci and human epithelial cells. *Molecules*, *15*(10), 7139-7152. doi:10.3390/molecules15107139

Jansen, J. J., Smit, S., Hoefsloot, H. C. J., & Smilde, A. K. (2010). The photographer and the greenhouse: how to analyse plant metabolomics data. *Phytochemical Analysis*, *21*(1), 48-60. doi:10.1002/pca.1181

Jepson, R. G., & Craig, J. C. (2008). Cranberries for preventing urinary tract infections (Review). *Cochrane Database Systematic Reviews*, *23*(1), CD001321. doi:10.1002/14651858.CD001321.pub4

Johnson-White, B., Buquo, L., Zeinali, M., & Ligler, F. S. (2006). Prevention of nonspecific bacterial cell adhesion in immunoassays by use of cranberry juice. *Analytical Chemistry*, *78*(3), 853-857. doi:10.1021/ac051700v

Kähkönen, M. P., Hopia, A. I., & Heinonen, M. (2001). Berry Phenolics and Their Antioxidant Activity. *J. Agric. Food Chem.*, *49*(8), 4076-4082. doi:10.1021/jf010152t

Kalt, W., Blumberg, J. B., McDonald, J. E., Vinqvist-Tymchuk, M. R., Fillmore, S. A. E., Graf, B. A., O'Leary, J. M., et al. (2008). Identification of anthocyanins in the liver, eye, and brain of blueberry-fed pigs. *Journal of Agricultural and Food Chemistry*, *56*(3), 705-712. doi:10.1021/jf071998l

Karlsen, A., Retterstøl, L., Laake, P., Paur, I., Kjølsrud-Bøhn, S., Sandvik, L., & Blomhoff, R. (2007). Anthocyanins inhibit nuclear factor-kappaB activation in monocytes and reduce plasma concentrations of pro-inflammatory mediators in healthy adults. *The Journal of Nutrition*, *137*(8), 1951-1954.

Kim, H. K., & Verpoorte, R. (2010). Sample preparation for plant metabolomics. *Phytochemical Analysis*, *21*(1), 4-13. doi:10.1002/pca.1188

Klein, M. A. (2010). Cranberry. In P. M. Coates, J. M. Betz, M. R. Blackman, G. M. Cragg, M. Levine, J. Moss, & J. D. White (Eds.), *Encyclopedia of dietary supplements* (pp. 193-201). London, UK: Informa Healthcare.

Knudson, W. A. (2008). The strategic marketing institute working paper - The economic impact of expanded cranberry production. Retrieved August 3, 2011, from http://productcenter.msu.edu/documents/cranberries.pdf

Kok, T. M., Breda, S. G., & Manson, M. M. (2008). Mechanisms of combined action of different chemopreventive dietary compounds. *European Journal of Nutrition*, *47*(S2), 51-59. doi:10.1007/s00394-008-2006-y

Krenn, L., Steitz, M., Schlicht, C., Kurth, H., & Gaedcke, F. (2007). Anthocyanin- and proanthocyanidin-rich extracts of berries in food supplements--analysis with problems. *Pharmazie, 62*(11), 803-812.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47,* 583-621.

Ku, K. M., Choi, J. N., Kim, J., Kim, J. K., Yoo, L. G., Lee, S. J., Hong, Y.-S., et al. (2010). Metabolomics analysis reveals the compositional differences of shade grown tea (Camellia sinensis L.). *Journal of Agricultural and Food Chemistry, 58*(1), 418-426. doi:10.1021/jf902929h

Kuhnlein, H. V., & Turner, N. J. (1996). Descriptions and uses of plant foods by indigenous peoples. *Traditional plant foods of Canadian indigenous peoples: nutrition, botany, and use* (pp. 173-174). Amsterdam, the Netherlands: Gordon and Breach Publishers.

Kuzminski, L. N. (1996). Cranberry juice and urinary tract infections: Is there a beneficial relationship? *Nutrition Reviews, 54*(11), S87-S90.

Kylli, P., Nohynek, L., Puupponen-Pimiä, R., Westerlund-Wikström, B., Leppänen, T., Welling, J., Moilanen, E., et al. (2011). Lingonberry (*Vaccinium vitis-idaea*) and European cranberry (*Vaccinium microcarpon*) proanthocyanidins: Isolation, identification, and bioactivities. *Journal of Agricultural and Food Chemistry, 59*(7), 3373-3384. doi:10.1021/jf104621e

Lätti, A. K., Riihinen, K. R., & Jaakola, L. (2011). Phenolic compounds in berries and flowers of a natural hybrid between bilberry and lingonberry (*Vaccinium* × *intermedium* Ruthe). *Phytochemistry, 72*(8), 810-815. doi:16/j.phytochem.2011.02.015

Leahy, M., Roderick, R., & Brilliant, K. (2001). The cranberry - promising health benefits, old and new. *Nutrition Today, 36*(5), 254-265.

Lehtonen, H.-M., Lehtinen, O., Suomela, J.-P., Viitanen, M., & Kallio, H. (2010). Flavonol glycosides of sea buckthorn (*Hippophaë rhamnoides* ssp. *sinensis*) and lingonberry (*Vaccinium vitis-idaea*) are bioavailable in humans and monoglucuronidated for excretion. *Journal of Agricultural and Food Chemistry, 58*(1), 620-627. doi:10.1021/jf9029942

Lehtonen, H.-M., Rantala, M., Suomela, J.-P., Viitanen, M., & Kallio, H. (2009). Urinary excretion of the main anthocyanin in lingonberry (*Vaccinium vitis-idaea*), cyanidin 3-*O*-galactoside, and its metabolites. *Journal of Agricultural and Food Chemistry, 57*(10), 4447-4451. doi:10.1021/jf900894k

Leusink, G. J., Kitts, D. D., Yaghmaee, P., & Durance, T. (2010). Retention of antioxidant capacity of vacuum microwave dried cranberry. *Journal of Food Science, 75*(3), C311-C316. doi:10.1111/j.1750-3841.2010.01563.x

Liang, Y., Xie, P., & Chau, F. (2010). Chromatographic fingerprinting and related chemometric techniques for quality control of traditional Chinese medicines. *Journal of Separation Science, 33*(3), 410-421. doi:10.1002/jssc.200900653

Liu, C., & Russell, R. M. (2008). Nutrition and gastric cancer risk: an update. *Nutrition Reviews, 66*(5), 237-249. doi:10.1111/j.1753-4887.2008.00029.x

Liu, J., Hu, L., Dong, Z., & Hu, Q. (2004). DPPH radical scavenging activity of ten natural p-terphenyl derivatives obtained from three edible mushrooms indigenous to China. *Chemistry & Biodiversity, 1*(4), 601-605. doi:10.1002/cbdv.200490050

Louw, L., Roux, K., Tredoux, A., Tomic, O., Naes, T., Nieuwoudt, H. H., & van Rensburg, P. (2009). Characterization of selected South African young cultivar wines using FTMIR spectroscopy, gas chromatography, and multivariate data analysis. *Journal of Agricultural and Food Chemistry, 57*(7), 2623-2632. doi:10.1021/jf8037456

Ludwig, C., & Viant, M. R. (2010). Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochemical Analysis, 21*(1), 22-32. doi:10.1002/pca.1186

Mabberley, D. J. (1997). *The plant-book: A portable dictionary of the vascular plants 2nd Edition*. Cambridge, UK: Cambridge University Press.

Manchester, L. C., Tan, D.-X., Reiter, R. J., Park, W., Monis, K., & Qi, W. (2000). High levels of melatonin in the seeds of edible plants: Possible function in germ tissue protection. *Life Sciences, 67*(25), 3023-3029. doi:16/S0024-3205(00)00896-1

Martens, S., Preuß, A., & Matern, U. (2010). Multifunctional flavonoid dioxygenases: Flavonol and anthocyanin biosynthesis in *Arabidopsis thaliana* L. *Phytochemistry, 71*(10), 1040-1049. doi:16/j.phytochem.2010.04.016

Masson, P., Matheson, S., Webster, A. C., & Craig, J. C. (2009). Meta-analyses in prevention and treatment of urinary tract infections. *Infectious Disease Clinics of North America, 23*(2), 355-385. doi:10.1016/j.idc.2009.01.001

Mazza, G., Cacace, J. E., & Kay, C. D. (2004). Methods of analysis for anthocyanins in plants and biological fluids. *Journal of AOAC International, 87*(1), 129-145.

McCown, B. H., & Zeldin, E. L. (2005). *Vaccinium spp*. Cranberry. In R. E. Litz (Ed.), *Biotechnology of fruit and nut crops, biotechnology in agriculture series* (pp. 247 - 261). Oxfordshire, United Kingdom: CABI Publishing.

McGhie, T. K., & Walton, M. C. (2007). The bioavailability and absorption of anthocyanins: Towards a better understanding. *Molecular Nutrition & Food Research, 51*(6), 702-713. doi:10.1002/mnfr.200700092

Milbury, P. E., Vita, J. A., & Blumberg, J. B. (2010). Anthocyanins are bioavailable in humans following an acute dose of cranberry juice. *The Journal of Nutrition, 140*(6), 1099-1104. doi:10.3945/jn.109.117168

Moerman, D. E. (2004). *Native American ethnobotany*. Portland, OR, USA: Timber Press Inc.

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. doi:10.1021/jf011062r

Murch, S. J., Alan, A. R., Cao, J., & Saxena, P. K. (2009). Melatonin and serotonin in flowers and fruits of *Datura metel* L. *Journal of Pineal Research, 47*(3), 277-283. doi:10.1111/j.1600-079X.2009.00711.x

Murch, S. J., Hall, B. A., Le, C. H., & Saxena, P. K. (2010). Changes in the levels of indoleamine phytochemicals during véraison and ripening of wine grapes. *Journal of Pineal Research, 49*(1), 95-100. doi:10.1111/j.1600-079X.2010.00774.x

Mutanen, M., Pajari, A. M., Päivärinta, E., Misikangas, M., & Oikarinen, S. (2008). Berries as chemopreventive dietary constituents--a mechanistic approach with the ApcMin/+ mouse. *Asia Pacific Journal of Clinical Nutrition, 17*(S1), 123-125.

Nagy, S., & Wade, R. L. (1995). *Methods to detect adulteration of fruit juice beverages vol. 1*. Auburndale, FL, USA: Agscience Inc.

Neto, C. C. (2007). Cranberry and blueberry: Evidence for protective effects against cancer and vascular diseases. *Molecular Nutrition & Food Research, 51*(6), 652-664. doi:10.1002/mnfr.200600279

Neto, C. C., Amoroso, J. W., & Liberty, A. M. (2008). Anticancer activities of cranberry phytochemicals: An update. *Molecular Nutrition & Food Research, 52*(S1), S18-S27.

Nowack, R. (2007). Cranberry juice – a well-characterized folk-remedy against bacterial urinary tract infection. *Wiener Medizinische Wochenschrift, 157*(13-14), 325-330.

Nowack, R., & Schmitt, W. (2008). Cranberry juice for prophylaxis of urinary tract infections - conclusions from clinical experience and research. *Phytomedicine, 15*(9), 653- 667.

Ocean Spray Canada Ltd. (2011). The cranberry harvest. Retrieved June 13, 2011, from http://www.oceanspray.ca/Heritage/The-Beauty-of--the-Cranberry-Harvest.aspx

Okada, T., Afendi, F. M., Altaf-Ul-Amin, M., Takahashi, H., Nakamura, K., & Kanaya, S. (2010). Metabolomics of medicinal plants: the importance of multivariate

analysis of analytical chemistry data. *Current Computer-Aided Drug Design*, *6*(3), 179-196.

Pappas, E., & Schaich, K. M. (2009). Phytochemicals of cranberries and cranberry products: Characterization, potential health effects, and processing stability. *Critical Reviews in Food Science and Nutrition*, *49*(9), 741-781. doi:10.1080/10408390802145377

Paredes, S. D., Korkmaz, A., Manchester, L. C., Tan, D.-X., & Reiter, R. J. (2009). Phytomelatonin: a review. *Journal of Experimental Botany*, *60*(1), 57-69. doi:10.1093/jxb/ern284

Pérez-López, F. R., Haya, J., & Chedraui, P. (2009). *Vaccinium macrocarpon*: An interesting option for women with recurrent urinary tract infections and other health benefits. *Journal of Obstetrics and Gynaecology Research*, *35*(4), 630-639. doi:10.1111/j.1447-0756.2009.01026.x

Prior, R. L., Lazarus, S. A., Cao, G., Muccitelli, H., & Hammerstone, J. F. (2001). Identification of procyanidins and anthocyanins in blueberries and cranberries (*Vaccinium spp*.) using high-performance liquid chromatography/mass spectrometry. *Journal of Agricultural and Food Chemistry*, *49*(3), 1270-1276. doi:10.1021/jf001211q

R Development Core Team. (2010). The R project for statistical computing. Retrieved May 26, 2011, from http://www.r-project.org/

Rios, L. Y., Bennett, R. N., Lazarus, S. A., Rémésy, C., Scalbert, A., & Williamson, G. (2002). Cocoa procyanidins are stable during gastric transit in humans. *The American Journal of Clinical Nutrition*, *76*(5), 1106-1110.

Rodriguez-Saona, C., Vorsa, N., Singh, A. P., Johnson-Cicalese, J., Szendrei, Z., Mescher, M. C., & Frost, C. J. (2011). Tracing the history of plant traits under domestication in cranberries: potential consequences on anti-herbivore defences. *Journal of Experimental Botany*, *62*(8), 2633-2644. doi:10.1093/jxb/erq466

Sánchez-Moreno, C. (2002). Review: Methods used to evaluate the free radical scavenging activity in foods and biological systems. *Food Science and Technology International*, *8*(3), 121-137. doi:10.1106/108201302026770

Sandler, H. A. (2011). Nitrogen and vine-harvest method affect cranberry vine production and yield. *HortTechnology*, *21*(1), 87-97.

Schripsema, J. (2010). Application of NMR in plant metabolomics: techniques, problems and prospects. *Phytochemical Analysis*, *21*(1), 14-21. doi:10.1002/pca.1185

Schwender, H. (2009). siggenes. Retrieved June 3, 2011, from http://www.bioconductor.org/packages/2.4/bioc/html/siggenes.html

Seeram, N. P. (2008a). Berry fruits: Compositional elements, biochemical activities, and the impact of their intake on human health, performance, and disease. *Journal of Agricultural and Food Chemistry*, *56*(3), 627-629. doi:10.1021/jf071988k

Seeram, N. P. (2008b). Berry fruits for cancer prevention: current status and future prospects. *Journal of Agricultural and Food Chemistry*, *56*(3), 630-635. doi:10.1021/jf072504n

Serrano, J., Puupponen-Pimiä, R., Dauer, A., Aura, A.-M., & Saura-Calixto, F. (2009). Tannins: Current knowledge of food sources, intake, bioavailability and biological effects. *Molecular Nutrition & Food Research*, *53*(S2), S310-S329. doi:10.1002/mnfr.200900039

Stewart, D., Shepherd, L. V. T., Hall, R. D., & Fraser, P. D. (2011). Crops and tasty, nutritious food – How can metabolomics help? In R. D. Hall (Ed.), *Annual plant reviews volume 43: Biology of plant metabolomics* (pp. 181-217). Oxford, United Kingdom: Wiley-Blackwell. Retrieved from http://onlinelibrary.wiley.com.ezproxy.library.ubc.ca/doi/10.1002/9781444339956.ch7/summary

Stothers, L. (2002). A randomized trial to evaluate effectiveness and cost effectiveness of naturopathic cranberry products as prophylaxis against urinary tract infection. *Canadian Journal of Urology*, *9*(3), 1558-1562.

Suda, J., & Lysák, M. A. (2001). A taxonomic study of the *Vaccinium* Sect. *Oxycoccus* (Hill) W. D. J. Koch (Ericaceae) in the Czech Republic and adjacent territories. *Folia Geobotanica*, *36*(3), 303-320.

Sumner, L. W., Mendes, P., & Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry*, *62*(6), 817-836. doi:16/S0031-9422(02)00708-2

Szajdek, A., & Borowska, E. J. (2008). Bioactive compounds and health-promoting properties of berry fruits: A review. *Plant Foods for Human Nutrition*, *63*(4), 147-156. doi:10.1007/s11130-008-0097-5

Takahata, Y., Ohnishi-Kameyama, M., Furuta, S., Takahashi, M., & Suda, I. (2001). Highly polymerized procyanidins in brown soybean seed coat with a high radical-scavenging activity. *Journal of Agricultural and Food Chemistry*, *49*(12), 5843-5847. doi:10.1021/jf010307x

Tanaka, N., Sekiya, N., Hattori, M., Goto, H., Shibahara, N., Shimada, Y., & Terasawa, K. (2003). Measurement of plasma procyanidin B-2 and procyanidin B-3 levels after oral administration in rat. *Phytomedicine*, *10*(2-3), 122-126. doi:78/094471103321659816

Tsang, C., Auger, C., Mullen, W., Bornet, A., Rouanet, J.-M., Crozier, A., & Teissedre, P.-L. (2005). The absorption, metabolism and excretion of flavan-3-ols and

procyanidins following the ingestion of a grape seed extract by rats. *British Journal of Nutrition, 94*(02), 170-181. doi:10.1079/BJN20051480

Turner, N. J. (2004). Bog cranberry. *Plants of Haida Gwaii* (p. 124). Winlaw, BC, Canada: Sono Nis Press.

Tusher, V. G., Tibshirani, R., & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America, 98*(9), 5116-5121. doi:10.1073/pnas.091062498

Tuszynski, J. (2010). CRAN - Package caTools. Retrieved June 3, 2011, from http://cran.r-project.org/web/packages/caTools/index.html

van der Kooy, F., Maltese, F., Hae Choi, Y., Kyong Kim, H., & Verpoorte, R. (2009). Quality control of herbal material and phytopharmaceuticals with MS and NMR based metabolic fingerprinting. *Planta Medica, 75*(07), 763-775. doi:10.1055/s-0029-1185450

Vander Kloet, S. P. (1983). The Taxonomy of *Vaccinium Oxycoccus*. *Rhodora, 85*(841), 1-44.

Varmuza, K., & Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. Boca Raton, FL, USA: CRC Press.

Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., & Panopoulos, N. (2007). Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. *Biotechnology Journal, 2*(10), 1214-1234. doi:10.1002/biot.200700084

Viskelis, P., Rubinskienė, M., Jasutienė, I., Šarkinas, A., Daubaras, R., & Česonienė, L. (2009). Anthocyanins, antioxidative, and antimicrobial properties of American cranberry (*Vaccinium macrocarpon* Ait.) and their press cakes. *Journal of Food Science, 74*(2), C157-C161. doi:10.1111/j.1750-3841.2009.01066.x

Vitalini, S., Gardana, C., Zanzotto, A., Fico, G., Faoro, F., Simonetti, P., & Iriti, M. (2011). From vineyard to glass: agrochemicals enhance the melatonin and total polyphenol contents and antiradical activity of red wines. *Journal of Pineal Research, 51*(3), 278-285. doi:10.1111/j.1600-079X.2011.00887.x

Vorsa, N. (1994). Breeding the American cranberry. *Wisconsin Cranberry School 1994 proceedings: Breeding the American cranberry* (pp. 1 - 4). Madison, WI, USA: Board of Regents of the University of Wisconsin System. Retrieved from http://digicoll.library.wisc.edu/cgi-bin/WI/WI-idx?type=turn&entity=WI.cranproc1994.p0001&id=WI.cranproc1994&isize=text

Vorsa, N., & Polashock, J. J. (2005). Alteration of anthocyanin glycosylation in cranberry through interspecific hybridization. *Journal of the American Society for Horticultural Science, 130*(5), 711-715.

Vorsa, N., Polashock, J., Cunningham, D., & Roderick, R. (2003). Genetic inferences and breeding implications from analysis of cranberry germplasm anthocyanin profiles. *J. Amer. Soc. Hort. Sci., 128*(5), 691-697. doi:

Vvedenskaya, I. O., Rosen, R. T., Guido, J. E., Russell, D. J., Mills, K. A., & Vorsa, N. (2004). Characterization of flavonols in cranberry (*Vaccinium macrocarpon*) powder. *Journal of Agricultural and Food Chemistry, 52*(2), 188-195. doi:10.1021/jf034970s

Wernimont, G. T. (1987). *Use of statistics to develop and evaluate analytical methods*. (W. Spendley, Ed.). AOAC International.

Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., Vis, D. J., Smilde, A. K., Velzen, E. J. J., Duijnhoven, J. P. M., et al. (2008). Assessment of PLSDA cross validation. *Metabolomics*, *4*(1), 81-89. doi:10.1007/s11306-007-0099-6

Wiczkowski, W., Romaszko, E., & Piskula, M. K. (2010). Bioavailability of cyanidin glycosides from natural chokeberry (*Aronia melanocarpa*) juice with dietary-relevant dose of anthocyanins in humans. *Journal of Agricultural and Food Chemistry, 58*(23), 12130-12136. doi:10.1021/jf102979z

Wink, M. (1988). Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theoretical and Applied Genetics, 75*(2), 225-233. doi:10.1007/BF00303957

Wise, Barry M., & Gallagher, N. B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control, 6*(6), 329-348. doi:16/0959-1524(96)00009-1

Wishart, D. S. (2008). Applications of metabolomics in drug discovery and development. *Drugs in R&D, 9*(5), 307-322.

Wolsko, P. M., Solondz, D. K., Phillips, R. S., Schachter, S. C., & Eisenberg, D. M. (2005). Lack of herbal supplement characterization in published randomized controlled trials. *The American Journal of Medicine, 118*(10), 1087-1093. doi:16/j.amjmed.2005.01.076

Woodward, G., Kroon, P., Cassidy, A., & Kay, C. (2009). Anthocyanin Stability and Recovery: Implications for the analysis of clinical and experimental samples. *Journal of Agricultural and Food Chemistry, 57*(12), 5271-5278. doi:10.1021/jf900602b

Yao, Y., & Vieira, A. (2007). Protective activities of *Vaccinium* antioxidants with potential relevance to mitochondrial dysfunction and neurotoxicity. *NeuroToxicology*, *28*(1), 93-100. doi:16/j.neuro.2006.07.015

Yoshida, K., Mori, M., & Kondo, T. (2009). Blue flower color development by anthocyanins: from chemical structure to cell physiology. *Natural Product Reports*, *26*(7), 884-915. doi:10.1039/b800165k

Youden, W. J., & Steiner, E. H. (1975). *Statistical manual of the AOAC*. Baltimore, MD, USA: AOAC International.

Zhang, K., & Zuo, Y. (2004). GC-MS determination of flavonoids and phenolic and benzoic acids in human plasma after consumption of cranberry juice. *Journal of Agricultural and Food Chemistry*, *52*(2), 222-227. doi:10.1021/jf035073r

Zhang, Z., Kou, X., Fugal, K., & McLaughlin, J. (2004). Comparison of HPLC methods for determination of anthocyanins and anthocyanidins in bilberry extracts. *Journal of Agricultural and Food Chemistry*, *52*(4), 688-691. doi:10.1021/jf034596w

Zheng, W., & Wang, S. Y. (2003). Oxygen radical absorbing capacity of phenolics in blueberries, cranberries, chokeberries, and lingonberries. *Journal of Agricultural and Food Chemistry*, *51*(2), 502-509. doi:10.1021/jf020728u

Appendix 1: Supplementary Statistics for Chapter 5

Table A-1: ANOVA comparing cyanidin-3-*O*-glycosides content among cultivars using biological replicates

Cyanidin Glycosides Biological Replicates

Summary

Groups	*Count*	*Sum*	*Average*	*Variance*
Ben Lear	5	185.22709	37.04542	21.86338
Bergman	5	208.440122	41.68802	51.90985
GH1	5	231.949652	46.38993	62.21672
Pilgrim	5	223.075364	44.61507	34.50214
Stevens	5	193.376862	38.67537	12.53473

ANOVA: Single Factor

Source	*SS*	*df*	*MS*	*F*	P-*value*	*F crit*
Between	306.5263	4	76.6315757	2.093452	0.119635	2.866081
Within	732.1073	20	36.6053651			
Total	1038.634	24				

Table A-2: ANOVA comparing peonidin-3-*O*-glycosides content among cultivars using biological replicates

Peonidin Glycosides Biological Replicates

Summary

Groups	*Count*	*Sum*	*Average*	*Variance*
Ben Lear	5	314.7729	62.95458	21.86338
Bergman	5	291.5599	58.31198	51.90985
GH1	5	268.0503	53.61007	62.21672
Pilgrim	5	276.9246	55.38493	34.50214
Stevens	5	306.6231	61.32463	12.53473

ANOVA: Single Factor

Source	*SS*	*df*	*MS*	*F*	P-*value*	*F crit*
Between	306.5263	4	76.63158	2.093452	0.119635	2.866081
Within	732.1073	20	36.60537			
Total	1038.634	24				

Table A-3: ANOVA comparing cyanidin-3-*O*-glycosides content among cultivars using analytical replicates

Cyanidin Glycosides Analytical Replicates

Summary

Groups	*Count*	*Sum*	*Average*	*Variance*
Ben Lear	5	164.1151	32.82302	0.037914
Bergman	5	202.1582	40.43165	0.101793
GH1	5	178.6607	35.73215	0.031402
Pilgrim	5	212.564	42.51281	0.590033
Stevens	5	180.7385	36.1477	0.546715

ANOVA: Single Factor

Source	*SS*	*df*	*MS*	*F*	P-*value*	*F crit*
Between	302.733	4	75.68324	289.3407	2.17E-17	2.866081
Within	5.231427	20	0.261571			
Total	307.9644	24				

Table A-4: ANOVA comparing peonidin-3-*O*-glycosides content among cultivars using analytical replicates

Peonidin Glycosides Analytical Replicates

Summary

Groups	*Count*	*Sum*	*Average*	*Variance*
Ben Lear	5	335.8849	67.17698	0.037914
Bergman	5	297.8418	59.56835	0.101793
GH1	5	321.3393	64.26785	0.031402
Pilgrim	5	287.436	57.48719	0.590033
Stevens	5	319.2615	63.8523	0.546715

ANOVA: Single Factor

Source	*SS*	*df*	*MS*	*F*	P-*value*	*F crit*
Between	302.733	4	75.68324	289.3407	2.17E-17	2.866081
Within	5.231427	20	0.261571			
Total	307.9644	24				

Table A-5: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Ben Lear cultivar

t-Test: Ben Lear Cyanidin Glycosides

	Biological	*Analytical*
Mean	37.05	32.82
Variance	21.86	0.04
Observations	5.00	5.00
df	4.00	
t Stat	2.02	
p(T<=t) one-tail	0.06	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.11	
t Critical two-tail	2.78	

t-Test: Ben Lear Peonidin Glycosides

	Biological	*Analytical*
Mean	62.95	67.18
Variance	21.86	0.04
Observations	5.00	5.00
df	4.00	
t Stat	-2.02	
p(T<=t) one-tail	0.06	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.11	
t Critical two-tail	2.78	

Table A-6: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Bergman cultivar

t-Test: Bergman Cyanidin Glycosides

	Biological	*Analytical*
Mean	41.69	40.43
Variance	51.91	0.10
Observations	5.00	5.00
df	4.00	
t Stat	0.39	
p(T<=t) one-tail	0.36	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.72	
t Critical two-tail	2.78	

t-Test: Bergman Peonidin Glycosides

	Biological	*Analytical*
Mean	58.31	59.57
Variance	51.91	0.10
Observations	5.00	5.00
df	4.00	
t Stat	-0.39	
p(T<=t) one-tail	0.36	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.72	
t Critical two-tail	2.78	

Table A-7: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in GH1 cultivar

t-Test: GH1 Cyanidin Glycosides

	Biological	*Analytical*
Mean	46.39	35.73
Variance	62.22	0.03
Observations	5.00	5.00
df	4.00	
t Stat	3.02	
p(T<=t) one-tail	0.02	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.04	
t Critical two-tail	2.78	

t-Test: GH1 Peonidin Glycosides

	Biological	*Analytical*
Mean	53.61	64.27
Variance	62.22	0.03
Observations	5.00	5.00
df	4.00	
t Stat	-3.02	
p(T<=t) one-tail	0.02	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.04	
t Critical two-tail	2.78	

Table A-8: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Pilgrim cultivar

t-Test: Pilgrim Cyanidin Glycosides

	Biological	*Analytical*
Mean	44.62	42.51
Variance	34.50	0.59
Observations	5.00	5.00
df	4.00	
t Stat	0.79	
p(T<=t) one-tail	0.24	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.47	
t Critical two-tail	2.78	

t-Test:Pilgrim Peonidin Glycosides

	Biological	*Analytical*
Mean	55.38	57.49
Variance	34.50	0.59
Observations	5.00	5.00
df	4.00	
t Stat	-0.79	
p(T<=t) one-tail	0.24	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.47	
t Critical two-tail	2.78	

Table A-9: Student *t*-tests comparing anthocyanin content between analytical and biological replicates in Stevens cultivar

t-Test: Stevens Cyanidin Glycosides

	Biological	*Analytical*
Mean	38.68	36.15
Variance	12.53	0.55
Observations	5.00	5.00
df	4.00	
t Stat	1.56	
p(T<=t) one-tail	0.10	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.19	
t Critical two-tail	2.78	

t-Test: Stevens Peonidin Glycosides

	Biological	*Analytical*
Mean	61.32	63.85
Variance	12.53	0.55
Observations	5.00	5.00
df	4.00	
t Stat	-1.56	
p(T<=t) one-tail	0.10	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.19	
t Critical two-tail	2.78	

Table A-10: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Ben Lear cultivar

t-Test: Ben Lear Cyandin-3-O-Galactoside

	Biological	*Analytical*
Mean	1493.64	1112.30
Variance	1586596.02	913.13
Observations	5.00	5.00
df	4.00	
t Stat	0.68	
p(T<=t) one-tail	0.27	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.54	
t Critical two-tail	2.78	

t-Test: Ben Lear Cyandin-3-O-Glucoside

	Biological	*Analytical*
Mean	146.12	99.55
Variance	2544.76	1.20
Observations	5.00	5.00
df	4.00	
t Stat	2.06	
p(T<=t) one-tail	0.05	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.11	
t Critical two-tail	2.78	

Table A-11: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Ben Lear cultivar

t-Test: Ben Lear Cyandin-3-O-Arabinoside

	Biological	*Analytical*
Mean	1351.54	853.53
Variance	1202927	1097.88
Observations	5.00	5.00
df	4.00	
t Stat	1.01	
p(T<=t) one-tail	0.18	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.37	
t Critical two-tail	2.78	

t-Test: Ben Lear Peonidin-3-O-Galactoside

	Biological	*Analytical*
Mean	3671.77	3012.80
Variance	6794080	12212.52
Observations	5.00	5.00
df	4.00	
t Stat	0.56	
p(T<=t) one-tail	0.30	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.60	
t Critical two-tail	2.78	

Table A-12: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in Ben Lear cultivar

t-Test: Ben Lear Peonidin-3-O-Arabinoside

	Biological	*Analytical*
Mean	1319.35	1214.90
Variance	727631	1689.09
Observations	5.00	5.00
df	4.00	
t Stat	0.27	
p(T<=t) one-tail	0.40	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.80	
t Critical two-tail	2.78	

t-Test: Ben Lear Total Anthocyanins

	Biological	*Analytical*
Mean	7982.42	6293.08
Variance	33940875	45748.85
Observations	5.00	5.00
df	4.00	
t Stat	0.65	
p(T<=t) one-tail	0.28	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.55	
t Critical two-tail	2.78	

Table A-13: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Bergman cultivar

t-Test: Bergman Cyandin-3-O-Galactoside

	Biological	*Analytical*
Mean	1430.63	855.40
Variance	120081	80.83
Observations	5.00	5.00
df	4.00	
t Stat	3.71	
p(T<=t) one-tail	0.01	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.02	
t Critical two-tail	2.78	

t-Test: Bergman Lear Cyandin-3-O-Glucoside

	Biological	*Analytical*
Mean	119.81	86.30
Variance	167.85	4.61
Observations	5.00	5.00
df	4.00	
t Stat	5.71	
p(T<=t) one-tail	0.00	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.00	
t Critical two-tail	2.78	

Table A-14: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Bergman cultivar

t-Test: Bergman Cyandin-3-O-Arabinoside

	Biological	*Analytical*
Mean	1318.29	724.88
Variance	91608.63	362.23
Observations	5.00	5.00
df	4.00	
t Stat	4.38	
p(T<=t) one-tail	0.01	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.01	
t Critical two-tail	2.78	

t-Test: Bergman Peonidin-3-O-Galactoside

	Biological	*Analytical*
Mean	3059.28	1656.91
Variance	1090642	277.83
Observations	5.00	5.00
df	4.00	
t Stat	3.00	
p(T<=t) one-tail	0.02	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.04	
t Critical two-tail	2.78	

Table A-15: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in Bergman cultivar

t-Test: Bergman Peonidin-3-O-Arabinoside

	Biological	*Analytical*
Mean	1094.11	798.43
Variance	80602	338.17
Observations	5.00	5.00
df	4.00	
t Stat	2.32	
p(T<=t) one-tail	0.04	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.08	
t Critical two-tail	2.78	

t-Test: Bergman Total Anthocyanins

	Biological	*Analytical*
Mean	7022.13	4121.92
Variance	3078940	1454.83
Observations	5.00	5.00
df	4.00	
t Stat	3.69	
p(T<=t) one-tail	0.01	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.02	
t Critical two-tail	2.78	

Table A-16: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in GH1 cultivar

t-Test: GH1 Cyandin-3-O-Galactoside

	Biological	*Analytical*
Mean	1405.95	1230.08
Variance	297264	277.51
Observations	5.00	5.00
df	4.00	
t Stat	0.72	
p(T<=t) one-tail	0.26	
t Critical one-tail	2.13	
P(T<=t) two-tail	0.51	
t Critical two-tail	2.78	

t-Test: GH1 Cyandin-3-O-Glucoside

	Biological	*Analytical*
Mean	100.24	104.90
Variance	457.30	8.78
Observations	5.00	5.00
df	4.00	
t Stat	-0.48	
p(T<=t) one-tail	0.33	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.65	
t Critical two-tail	2.78	

Table A-17: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in GH1 cultivar

t-Test: GH1 Cyandin-3-O-Arabinoside		
	Biological	*Analytical*
Mean	1181.52	854.53
Variance	129565	599.08
Observations	5.00	5.00
df	4.00	
t Stat	2.03	
p(T<=t) one-tail	0.06	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.11	
t Critical two-tail	2.78	

t-Test: GH1 Peonidin-3-O-Galactoside		
	Biological	*Analytical*
Mean	2541.11	2852.56
Variance	1956676	1408.64
Observations	5.00	5.00
df	4.00	
t Stat	-0.50	
p(T<=t) one-tail	0.32	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.64	
t Critical two-tail	2.78	

Table A-18: Student *t*-tests comparing peonidin-3-*O*-arabinoside content between analytical and biological replicates in GH1 cultivar

t-Test: GH1 Peonidin-3-O-Arabinoside		
	Biological	*Analytical*
Mean	820.98	1085.18
Variance	101070	229.81
Observations	5.00	5.00
df	4.00	
t Stat	-1.86	
p(T<=t) one-tail	0.07	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.14	
t Critical two-tail	2.78	

t-Test: GH1 Total Anthocyanins		
	Biological	*Analytical*
Mean	6049.79	6127.25
Variance	6282951	7315.53
Observations	5.00	5.00
df	4.00	
t Stat	-0.07	
p(T<=t) one-tail	0.47	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.95	
t Critical two-tail	2.78	

Table A-19: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Pilgrim cultivar

t-Test: Pilgrim Cyandin-3-O-Galactoside		
	Biological	*Analytical*
Mean	700.43	685.17
Variance	169775	585.46
Observations	5.00	5.00
df	4.00	
t Stat	0.08	
p(T<=t) one-tail	0.47	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.94	
t Critical two-tail	2.78	

t-Test: Pilgrim Cyandin-3-O-Glucoside		
	Biological	*Analytical*
Mean	83.36	84.31
Variance	163.62	1.78
Observations	5.00	5.00
df	4.00	
t Stat	-0.16	
p(T<=t) one-tail	0.44	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.88	
t Critical two-tail	2.78	

Table A-20: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Pilgrim cultivar

t-Test: Pilgrim Cyandin-3-O-Arabinoside

	Biological	*Analytical*
Mean	616.69	582.24
Variance	129975	1021.24
Observations	5.00	5.00
df	4.00	
t Stat	0.21	
p(T<=t) one-tail	0.42	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.84	
t Critical two-tail	2.78	

t-Test: Pilgrim Peonidin-3-O-Galactoside

	Biological	*Analytical*
Mean	1385.14	1240.57
Variance	704523	2853.06
Observations	5.00	5.00
df	4.00	
t Stat	0.38	
p(T<=t) one-tail	0.36	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.72	
t Critical two-tail	2.78	

Table A-21: Student *t*-tests comparing peonidin-3-*O*-galactoside content between analytical and biological replicates in Pilgrim cultivar

t-Test: Pilgrim Peonidin-3-O-Arabinoside

	Biological	*Analytical*
Mean	495.83	587.23
Variance	79345	322.24
Observations	5.00	5.00
df	4.00	
t Stat	-0.72	
p(T<=t) one-tail	0.25	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.51	
t Critical two-tail	2.78	

t-Test: Pilgrim Total Anthocyanins

	Biological	*Analytical*
Mean	3281.45	3179.51
Variance	3530433	13466.19
Observations	5.00	5.00
df	4.00	
t Stat	0.12	
p(T<=t) one-tail	0.45	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.91	
t Critical two-tail	2.78	

Table A-22: Student *t*-tests comparing cyanidin-3-*O*-galactoside content between analytical and biological replicates in Stevens cultivar

t-Test: Stevens Cyandin-3-O-Galactoside

	Biological	*Analytical*
Mean	566.23	494.14
Variance	36686	350.87
Observations	5.00	5.00
df	4.00	
t Stat	0.84	
p(T<=t) one-tail	0.22	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.45	
t Critical two-tail	2.78	

t-Test: Stevens Cyandin-3-O-Glucoside

	Biological	*Analytical*
Mean	95.65	78.83
Variance	1061	1.29
Observations	5.00	5.00
df	4.00	
t Stat	1.15	
p(T<=t) one-tail	0.16	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.31	
t Critical two-tail	2.78	

Table A-23: Student *t*-tests comparing cyanidin-3-*O*-arabinoside content between analytical and biological replicates in Stevens cultivar

t-Test: Stevens Cyandin-3-O-Arabinoside

	Biological	*Analytical*
Mean	409.22	391.12
Variance	5161	509.02
Observations	5.00	5.00
df	5.00	
t Stat	0.54	
p(T<=t) one-tail	0.31	
t Critical one-tail	2.02	
p(T<=t) two-tail	0.61	
t Critical two-tail	2.57	

t-Test: Stevens Peonidin-3-O-Galactoside

	Biological	*Analytical*
Mean	1182.68	1080.16
Variance	163651	2450.19
Observations	5.00	5.00
df	4.00	
t Stat	0.56	
p(T<=t) one-tail	0.30	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.60	
t Critical two-tail	2.78	

Table A-24: Student *t*-tests comparing peonidin-3-*O*-galactoside content between analytical and biological replicates in Stevens cultivar

t-Test: Stevens Peonidin-3-O-Arabinoside

	Biological	*Analytical*
Mean	552.18	622.65
Variance	19740	570.95
Observations	5.00	5.00
df	4.00	
t Stat	-1.11	
p(T<=t) one-tail	0.17	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.33	
t Critical two-tail	2.78	

t-Test: Stevens Total Anthocyanins

	Biological	*Analytical*
Mean	2805.95	2666.90
Variance	650868	8159.65
Observations	5.00	5.00
df	4.00	
t Stat	0.38	
p(T<=t) one-tail	0.36	
t Critical one-tail	2.13	
p(T<=t) two-tail	0.72	
t Critical two-tail	2.78	