NTNU

Norwegian University of
Science and Technology

Variation in chemical composition and genetic differentiation among bilberry (Vaccinium myrtillus L.) populations on a latitudinal gradient

Eva Sofie Dahlø

Biology

Submission date: May 2011

Supervisor: Tor-Henning Iversen, IBI

Co-supervisor: Jens Rohloff, IBI

Henrik Jensen, IBI

Norwegian University of Science and Technology
Department of Biology

Variation in chemical composition and genetic differentiation among bilberry (*Vaccinium myrtillus* L.) populations on a latitudinal gradient

Eva Sofie Dahlø

Master of Science in Cell and Molecular Biology

Department of Biology

Norwegian University of Science and Technology

Trondheim, May 2011

Supervisors:

Dr. Jens Rohloff

Dr. Henrik Jensen

Professor Tor-Henning Iversen

NTNU – Trondheim
Norwegian University of
Science and Technology

Cover photo: Bilberry (*Vaccinium myrtillus* L.). Photo courtesy of Tor Egil Kvalnes.

Acknowledgements

This Master's thesis is a part of my Master's Degree in Cell and Molecular Biology at the Department of Biology at the Norwegian University of Science and Technology (NTNU) in Trondheim. Parts of the work were carried out at the Plant Biocentre (a subdivision at Dep. of Biology) and funding was provided by the Blueberry Project (NFR project no. 184797) and NTNU.

I would like to thank my supervisors, Dr. Jens Rohloff, Dr. Henrik Jensen and Professor Tor-Henning Iversen for all help and guidance. Also, I need to thank Bioforsk for sampling berries in Troms and Hedmark, Grete Rakvaag for answering the many questions during the chemical analyses and the other employees at the Department of Biology that helped in some way or another. In addition, I need to thank Irene B. Ytterdal for making those long days and nights in the laboratory rather amusing. Special thanks go to my parents for always believing in me, and to my fiancé Thomas Kvalnes for help, belief, and love whom no one can replace. You are priceless to me at so many levels.

Trondheim, May 2011

Eva Sofie Dahlø

Table of Contents

Summary

Bilberry (*Vaccinium myrtillus*) is native to Europe and North America and constitutes an important nutritional resource for both humans and animals. Over the years, a series of chemical analyses have revealed several health-beneficial compounds in bilberry, and with the current demand of the berries mainly covered by Sweden and Eastern Europe, there has become an increasing desire to cultivate bilberry in Norway. In order for such cultivation to be successful an increased knowledge about bilberry is seen as essential and thus several studies have investigated the chemical composition of the berry. However, the underlying genetic diversity and the variation between populations in biochemical compounds remain to be thoroughly investigated. Therefore, the aim of present study was to investigate the differences in biochemical composition between populations of bilberry distributed on a latitudinal gradient, and estimate the level of genetic variation within and among the populations. This in order to examine whether biochemical composition was reflected by observed levels of genetic variation.

Bilberries from four Norwegian populations at three regions differing in latitude were analysed for content of total phenolics (TPH), total anthocyanins (ACY) and antioxidant activity (FRAP). Furthermore, metabolic profiling was performed by gas chromatography-mass spectrometry (GC-MS) to reveal biochemical differences between the populations in content of sugars, acids and some simple phenolics. Multivariate statistics were performed and revealed a clustering of samples from the two locations in Mid-Norway, and a clustering of the northern with the southern population. In addition, there were found significant differences in some of the compounds between the populations.

Genetic analyses using four microsatellites were carried out to examine whether metabolic differences between populations were reflected by genetic differentiation. Despite the significant differences between populations in the concentrations of some metabolites no significant genetic differentiation was found. Thus, it seems that the variation in biochemical compounds discovered among populations could be environmentally induced differences on a similar genetic background. However, due to the limited number of working microsatellites and the fact that these molecular markers are neutral, there is still a possibility that the

genetic differences causing compound concentrations to differ could be so minor as to remain undetected. Hence, further studies utilizing more microsatellite markers or new state-of-the-art molecular techniques are needed to determine whether this result holds and is valid also for genetic variation in coding parts of the genome.

1. Introduction

1.1. Bilberry (*Vaccinium myrtillus*)

The diploid (2n = 48) bilberry (*Vaccinium myrtillus*), which is native to Europe and North America, is a deciduous rhizomatous perennial shrub of the *Ericacea* family. In Norway, new leaves sprout in late April or beginning of May and are shed in late September or October. The plants bloom in May, with flowers situated on one-year-old twigs, and during July and August, after pollination, berries develop and ripen to a dark blue colour (Ritchie, 1956). Bilberry is also referred to as European blueberry, while blueberry usually refers to two of its relatives, *Vaccinium corymbosum* and *Vaccinium angustifolium.* These two species have berries that are morphological different from bilberry, with whitish fruit flesh, and are often referred to as highbush and lowbush blueberry, respectively (Camire, 2002; Prior et al., 1998). They are native to Canada and USA where they have been commercially harvested for several decades (Puupponen-Pimiä et al., 2005).

The desire to cultivate bilberry has been steadily increasing over several years (Paassilta et al., 2009; Puupponen-Pimiä et al., 2005). In Norway, in spite of the abundance of wild growing bilberry, the industrial demand has mainly been covered by non-cultivated imports from Sweden and Eastern Europe. With the Norwegian processing industry using several tons of bilberry each year (Paassilta et al., 2009), there is also an increased desire to cultivate this species in Norway. Thus, the biochemical composition of bilberry is of great commercial interests for several reasons. Earlier chemical analyses have revealed high concentrations of several compounds with known health-beneficial properties (Bomser et al., 1996; Cacace and Mazza, 2003; Formica and Regelson, 1995; Nohynek et al., 2006; Serafini et al., 1998). For instance, berries from bilberry have been shown to contain a far higher concentration of antioxidants compared to highbush blueberry (Halvorsen et al., 2002). Furthermore, sugars and organic acids together with multiple aroma compounds have an important role in determining the character of the flavour properties and the quality of the berries (Viljakainen et al., 2002). As the current knowledge about the range of geographical variation among populations in biochemical composition is limited, an increased effort is needed to study this in order to better exploit the many good properties of bilberry.

1.2. Bilberry's secondary metabolites

Secondary metabolites (or compounds) are organic compounds which plants do not utilize in primary metabolism (photosynthesis, respiration and reproduction), but indeed they draw energy and necessary precursors from these processes. Their functions, which for many of the compounds still remain unknown, are very diverse. Several are involved in defence mechanisms, e.g. lignin and tannins, while others act as colorants, attractants or signalling molecules. Secondary metabolites can be classified into three major classes on the basis of their chemical structure; terpenes, nitrogen-containing compounds and phenolics (Taiz and Zeiger, 2006). Phenolic compounds (phenolics) are particularly known for their many health-beneficial qualities and their biosynthesis pathway is quite well known (Jaakola et al., 2002) (Figure 1.1).

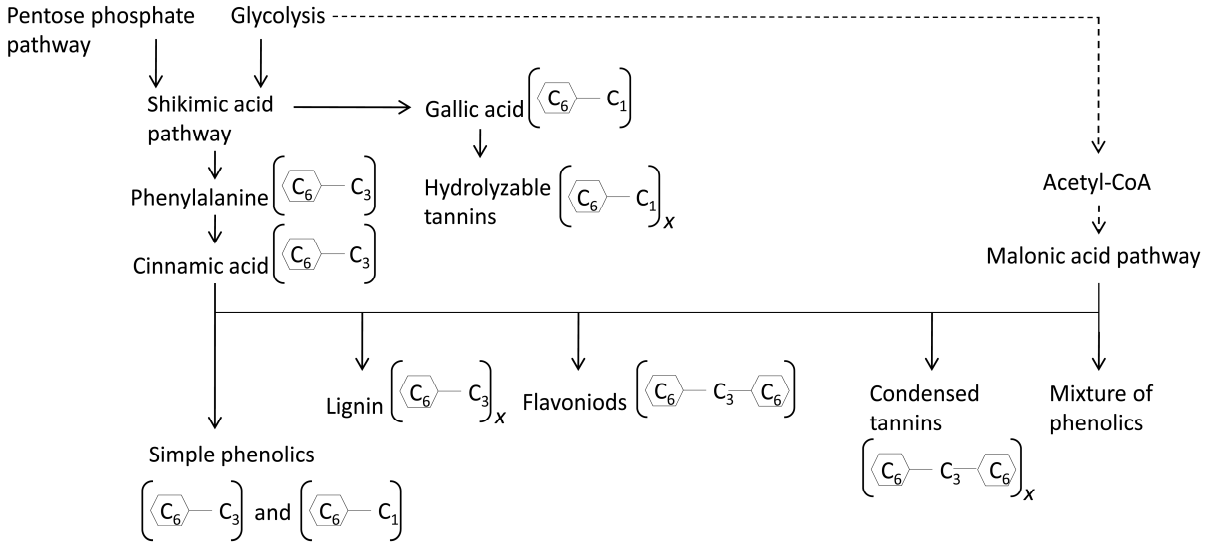

Figure 1.1: Phenolics are synthesized in several ways (Puupponen-Pimiä et al., 2005; Taiz and Zeiger, 2006). Most higher plants derive phenolics from phenylalanine and malonic acid pathway constitutes only a minor role in higher plants. Structures in brackets represent the basic structure of carbon skeletons together with the number of carbons in a carbon chain connected to the benzene ring for the compounds. The *x* subscript outside the brackets of some of the compounds indicates that the compounds can consist of several units of the basic structure inside the brackets.

Bilberry contains several phenolics such as lignin, flavonoids, tannins and a variety of other phenolics (e.g. simple phenolics) (Bravo, 1998; Puupponen-Pimiä et al., 2005; Taiz and Zeiger, 2006) (see Figure 1.1). A major part of flavonoids is constituted by anthocyanins,

which are water-soluble pigments, while simple phenolics, on the other hand, consist of several phenolic acids such as cinnamic acid, gallic acid and caffeic acid. Some of these simple phenolics can form more complex compounds, e.g. an ester between caffeic and quinic acid together form the compound chlorogenic acid (Puupponen-Pimiä et al., 2001; Taiz and Zeiger, 2006). Anthocyanins consist of anthocyanidins connected to sugar moieties and are together with chlorogenic acid two important components responsible for giving bilberries its high antioxidant activity (Puupponen-Pimiä et al., 2005). The berries contain five different anthocyanidins (cyanidin, delphinidin, peonidin, petunidin and malvidin) and three sugar moieties (3-O-arabinoside, 3-O-glucosides and 3-O-galactosides) which combine to form 15 different anthocyanins (Martinelli et al., 1986). Anthocyanins are responsible for the characteristic dark blue colour of bilberries (Prior et al., 1998) and with their antioxidant properties they protect the plant tissue from oxidative stress (Cacace and Mazza, 2003; Serafini et al., 1998). The role of anthocyanins as a protective agent against ultra violet (UV) light (Lois, 1994) has been studied in both bilberry (Taulavuori et al., 1998) and several other plant species (Lois, 1994). For instance, Jaakola et al. (2004) found an increased expression of genes involved in flavonoid biosynthesis, including cyanidin glycosides, in bilberry leaves under increased UV radiation. Furthermore, by using mutant plants, Havaux and Kloppstech (2001) and Li et al. (1993) showed that flavonoid-deficient mutants displayed increased photosystem destruction and sensitivity to UV-B radiation, thus demonstrating the importance of flavonoids in the protection from UV light.

1.3. Genetic and environmental effects on secondary metabolites

Genetic and environmental factors interact to determine the amount of total phenolics and total anthocyanins, and thus the antioxidant activity found in berries. Studies have shown that the environmental impact of abiotic factors, which may differ with growing location and season, can be substantial. For instance, Martinussen et al. (2009) found an increasing concentration of total phenolics with decreasing temperatures for bilberry, a similar trend in concentration has been found with increasing maturity of berries (Connor et al., 2002b; Prior et al., 1998). Also, the biosynthesis of anthocyanins is clearly influenced by environmental factors (Connor et al., 2002a). Temperature and water stress have been known to have a significant impact but, in particular, concentration of anthocyanins have been shown to be affected by light, in both the visible and UV-B range of the spectra (Atlegrim and Sjøberg, 1996; Howard et al., 2003; Åkerström et al., 2010). As changes in latitude are always associated with changes in climatic properties, such as temperature, day length and UV radiation, latitude could in many cases be a good indicator for fruit quality (Zheng et al., 2009b). In particular, at northern latitudes, where there is long summer days accompanied with diurnal changes in temperature between day and night, the conditions are expected to be ideal for an increased biosynthesis of anthocyanins in berries of bilberry (Lätti et al., 2008; Åkerström et al., 2010). Accordingly, areas with periods of midnight sun have particularly been found to be good for the development of higher concentrations of anthocyanins in plants (Jaakola et al., 2004; Lätti et al., 2008). However, there are also some contradictory results, for instance Rieger et al. (2008) found a decrease in anthocyanin content with increased altitude, even though solar radiation increases (Blumthaler et al., 1997) and temperatures generally decrease with rising altitude.

Antioxidant activity often closely traces the concentrations of total phenolics and total anthocyanins (Prior et al., 1998) and factors that will affect concentration of phenolics also influence antioxidant activity. These include, for instance, maturity of berries, preharvest environmental conditions, postharvest storage conditions and processing. There could of course also be genetic differences between populations that could lead to differences in antioxidant activity (Connor et al., 2002b; Ehlenfeldt and Prior, 2001; Prior et al., 1998). Specifically, content and composition of total phenolics in plants have earlier been thought

to have a strong genetic component. This is supported by Martinussen et al. (2009) which showed that, when grown under identical climatic conditions, clones of bilberries originating from northern latitudes in Finland still were better adapted to lower temperatures (12 °C) and longer days (24 h light) than clones originating from southern regions. Thus, the fact that clones from different regions displayed their characteristic metabolite concentrations despite growing in the same environment, strongly suggested that the differences in metabolite concentrations had a genetic origin. Furthermore, Howard et al. (2003) investigated antioxidant activity (measured as oxygen radical-absorbing capacity, ORAC), total phenolics and anthocyanin content among genotypes of different *Vaccinium* species and found that the variation among genotypes was much greater than the variation observed between growing seasons. They argued that the concentrations of the different compounds seemed to have a strong genetic basis with only weak influence of environmental conditions. A similar result was obtained by Connor et al. (2002a) for antioxidant activity (ORAC) and phenolic content among *Vaccinium* species. The reported year-to-year fluctuations at each location in their study were considerably smaller than variation among genotypes. However, as mentioned above, many studies have revealed that also environmental factors have an important role in determining metabolic profiles of berries (Martinussen et al., 2009; Åkerström et al., 2010) and more effort should be put into exploring the range of genetic variation within and among geographically separated populations to further expand our knowledge on this topic.

1.4. Genetic variation within and between populations

Genetic diversity of plant species may be utilized, in particular, when selecting genotypes for cultivation. Knowledge of genetic diversity will promote the use of genetic variation in crop improvement and makes it be possible to select superior genotypes for cultivation (Debnath, 2009). There are several elements which may affect the genetic diversity in populations. For bilberry, its growth form and mating system can greatly influence the level of diversity and the spatial distribution of individuals (Albert et al., 2004). For instance, Albert et al. (2004) investigated the level of genetic variation between and within populations of bilberries from different habitats in Belgium by use of random amplified polymorfic DNA markers (RAPD). They found that most of the total genetic variation was located within populations (86.19 %), with only a low proportion of the variation due to differences between populations (13.81 %). These results were probably due to bilberries mixed mating system with some vegetative reproduction, that may lead to low within population variation, and predominant sexual reproduction with long distance dispersal, which may lead to low between population variation (Albert et al., 2004). A similar result was found by Persson and Gustavsson (2001) in their study of lingonberry (*Vaccinium vitis-idea*), which has a similar mating system as bilberry, and by Bartish et al. (1999) which studied woody and herbaceous species, which where long-lived and had outcrossing breeding systems. However, current knowledge of bilberries genetic background is scare and further studies are needed to draw any conclusion regarding the nature of genetic differences among bilberry populations at different geographical scales. Furthermore, there is a need to investigate the correspondence between metabolic profiles and genetic diversity, something which only a selection of studies (e.g. Chan et al., 2010; Mochida et al., 2009) thus far have tried. Of current knowledge, none have previously investigated this in bilberry.

1.5. Aim of the study

The aim of the present study was to address three issues. First, the biochemical and bioactive composition of the bilberries was determined to reveal whether there were any differences between the four bilberry locations (Tromsø, Langvatnet, Lierne and Hedmark) dispersed along on a latitudinal gradient in Norway. Second, the level of genetic variation within and among the bilberry populations was estimated to examine the degree of genetic differentiation among the locations. Finally, based on metabolic clustering and pairwise genetic differentiation it was discussed whether the biochemical composition of the different locations were reflected by observed levels of genetic variation and differentiation between the locations.

This master thesis will contribute to increased knowledge of the quality, flavour and genetics of bilberry across locations separated on a latitudinal gradient. It will provide insight into whether the variation in quality and flavor of berries are mostly a result of genetic differences or are induced by environmental conditions. Thus the information gained from this thesis will be beneficial for several aspects of the potential cultivation of this species in Norway.

2. Materials and methods

2.1. Sampling of berries

Samples of berries and leaves were collected from four locations in three different regions in Norway. Tromsø (T) in Troms county (North Norway), the two locations Lierne (LI) and Langvatnet (LV) in Nord-Trøndelag county (Mid-Norway) and Hedmark (H) in Hedmark county (South Norway) (Figure 2.1).

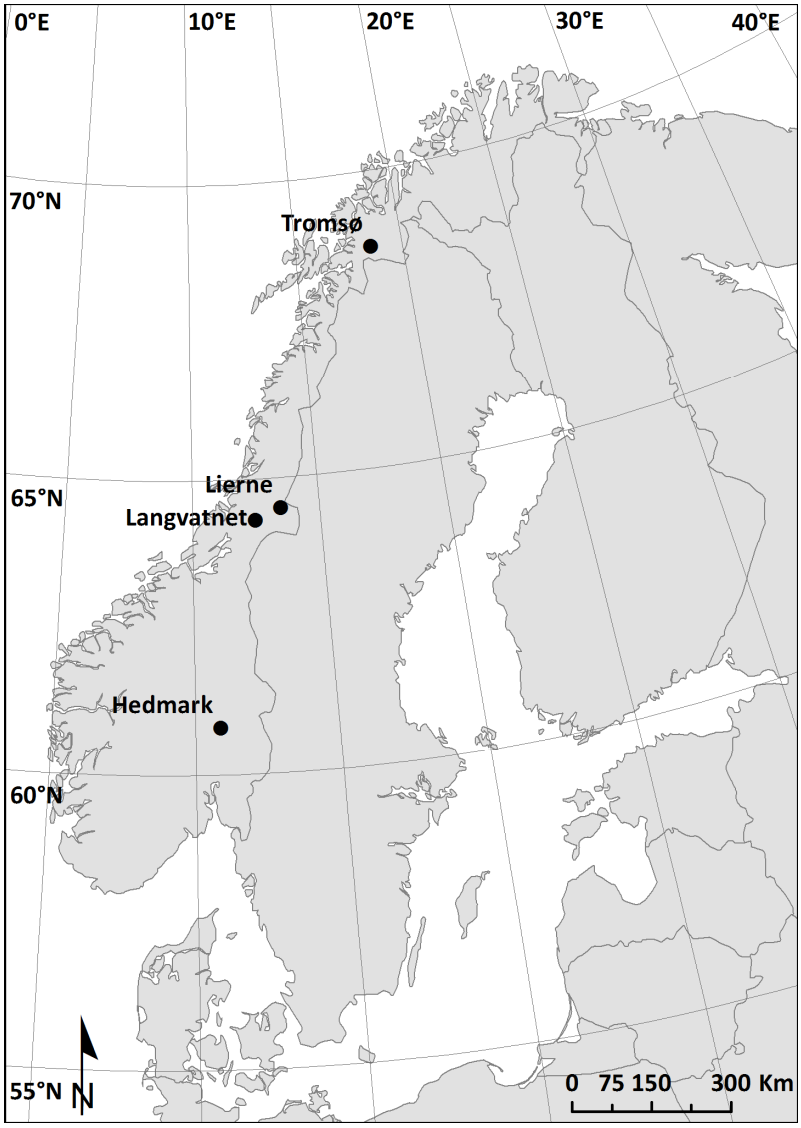

Figure 2.1: An overview of the spatial distribution of the four bilberry locations in present study. Coordinates for Tromsø are 68.6°N and 18.5°E, Lierne 64.5°N and 13.5°E, Langvatnet 64.3°N and 12.5°E and Hedmark 60.8°N and 10.9°E. The map was created in ArcGIS 10 using ArcMap.

The fields at these locations are part of the study area of the on-going (from 2008 throughout 2011) Blueberry project (NFR project no. 184797). Sample areas were all undisturbed and had not been subjected to any unnatural external influence (e.g. treatment of pesticides) preceding collection of samples. Sampling was carried out in summer 2009 during August (early to mid) at all locations. The samples from Lierne, Langvatnet and Hedmark were collected from ten individual plants, while for Tromsø seven individuals were sampled. For Lierne and Langvatnet, berries and leaves from each individual were collected into 50 mL tubes and kept cold (ca 4 °C) for some hours until arrival at the laboratory were all samples were stored at -80 °C prior to further analyses. The same procedure was followed for samples from Tromsø and Hedmark except that these were shipped on dry ice (ca -80 °C) to Trondheim before being stored at -80 °C in the same freezer.

2.2. Biochemical analyses

Laboratory work concerning biochemical analyses was performed at the Plant Biocentre, Department of Biology at NTNU. Total phenolics (TPH) of berries was analysed by use of the Folin-Ciocalteu (FC) method, total anthocyanin (ACY) content by pH-differential method and antioxidant activity by use of ferric reducing ability of plasma (FRAP) assay. All analyses were based on spectrophotometric detection. Furthermore, metabolic profiling was done by use of gas chromatography-mass spectrometry (GC-MS) to reveal biochemical differences regarding sugars, acids and simple phenolics.

2.2.1. Extraction of bilberry juice

Prior to analyses, bilberry juice had to be extracted from the samples. Approximately 180 mg of juice was extracted from 2000 mg of homogenized berries, and 1.8 mL 80 % methanol (CH_3OH) (Merck, Darmstadt, Germany) was added to each sample. The methanol contained 25 µg ribitol ($C_5H_{12}O_5$) (Fluka Biochemika, Buchs, Switzerland) per mL methanol which was utilized as an internal standard in the GC-MS analysis. The tubes were first shaken for 15 min on Edmund Büchler Swip shaker (Hechingen, Germany), and then centrifuged for 5 min (13,000 rpm) at 4 °C in IEC micromax RF (Watertown, Massachusetts, USA). The supernatant was pipetted to a new tube and stored in the freezer. This raw extract of the bilberry juice was used in the further analyses.

2.2.2. Folin-Ciocalteu method

Total phenolic content in berries was analyzed by use of Folin-Ciocalteu (FC) method (Singleton, 1965) and expressed as mg gallic acid equivalents ($C_7H_6O_5$) (Fluka Biochemika, Buchs, Switzerland) per 100 g fresh weight. In this procedure, 125 µL of FC reagent (Sigma-Aldrich, St. Louis, Missouri, USA) (diluted 1:10 with distilled water) was added to wells in a 96-well plate. Then, 25 µL gallic acid of different concentrations (0 µg/mL, 20 µg/mL, 40 µg/mL, 60 µg/mL, 80 µg/mL, 100 µg/mL and 120 µg/mL) followed by 100 µL 7.5 % (weight to volume) sodium carbonate (Na_2CO_3) (Merck, Darmstadt, Germany) was added. This was

used to set up a standard curve by linear regression (see Figure A1, Appendix I) from which the equation was solved to give an expression for concentration. Then, absorbance of samples could be entered to estimate concentration of total phenolics. For sample analysis, bilberry raw extracts were further diluted 1:10 with methanol to fit the range of the linear curve, and the same procedure as outlined above, with 25 µL sample extract replacing gallic acid, was followed. Three replicas for each sample were performed (see Table A1, Appendix II for the estimated concentration and standard error (SE) for each sample). The plate was shaken and incubated for two hours prior to spectrophotometric analysis in a 96-well microplate spectrophotometer (Labsystems Multiscan MS, Helsinki, Finland) at 750 nm.

2.2.3. pH-differential method

Total anthocyanin content in the samples was analyzed by use of pH-differential method (Giusti and Wrolstad, 2001) and expressed as mg cyanidin 3-glucoside per 100 g fresh weight. Absorbance of the samples was measured at two different pH values (1 and 4.5) and two different wavelengths (700 nm and 510 nm). The buffers of different pH values were prepared with modification of the protocol used by Giusti and Wrolstad (2001). Buffer of pH 1 (0.025 M) was prepared based on potassium chloride (KCl) and buffer of pH 4.5 (0.4 M) by sodium acetate ($C_2H_3NaO_2$), and both were adjusted to appropriate pH values with hydrogen chloride (HCl). All chemicals were purchased from Merck (Darmstadt, Germany). 20 µl of raw extract of bilberry juice (without further dilution) was added to cuvettes together with 0.5 mL each of the two buffers, and the cuvettes were shaken briefly and incubated 15 min prior to analysis in BioRad Smartspec^{TM} Plus spectrophotometer (Hercules, California, USA). One replica per sample was performed (see Table A1, Appendix II for the estimated concentration for each sample).

Anthocyanin content was calculated as cyanidin 3-glucoside because of its historical usage in similar assays, wide commercial availability, and because this compound is the most common anthocyanin pigment in bilberry (Burdulis et al., 2007; Burdulis et al., 2009; Jaakola et al., 2004). It is customary to calculate total monomeric anthocyanin pigment using the molecular weight and molar extinction coefficient of the major anthocyanin in the sample

matrix (Wrolstad et al., 2005). Total monomeric anthocyanin pigment (*y* mg/L) in the sample was calculated by use of the following equation:

$$y = \frac{A \times MW \times DF \times 1000}{\varepsilon \times l}$$

The difference in absorbance in the two buffers at two different wavelengths is proportional to the anthocyanin content in the samples, thus absorbance (*A*) was calculated as *A* = [(A_{510nm} pH 1.0)-(A_{700nm} pH 1.0)]-[(A_{510nm} pH 4.5)-(A_{700} pH 4.5)]. *MW* represents the molecular weight of 449.2 for cyandin 3-glucoside, *DF* is dilution factor of 26, ε is molar extinction coefficient of 26,900 L x mol^{-1} x cm^{-1} (Jurd and Asen, 1966) and *l* represents the pathlength (1 cm in present calculation, Wrolstad et al., 2005).

2.2.4. Ferric reducing ability of plasma (FRAP)

Antioxidant activity of berries was analyzed by use of ferric reducing ability of plasma (FRAP) method developed by Benzie and Strain (1996), with some modification, and expressed as mmol ferric iron reduced (Fe^{II}) per 100 g fresh weight. FRAP reagent was prepared by mixing 25 mL acetate buffer of pH 3.6 [3.1 g sodium acetate trihydrate ($C_2H_3NaO_2$ x $3H_2O$) (Sigma-Aldrich, St. Louis, Missouri, USA) and 16 mL acetic acid ($C_2H_4O_2$) (Merck, Darmstadt, Germany) per liter buffer solution], 2.5 mL solution of 10 mmol/L tripyridyltriazine ($C_3N_3[C_{15}H_{12}N_3]$) (TPTZ) (Fluka Biochemika, Buchs, Switzerland) mixed in 40 mmol/L hydrogen chloride (HCl) and 2.5 mL 20 mmol/L Iron (III) chloride hexahydrate ($FeCl_3$ x $6H_20$) (Sigma-Aldrich, St. Louis, Missouri, USA). 300 µL of this FRAP reagent was added to wells in a 96-well plate, followed by 10 µL iron (II) sulfate heptahydrate ($FeSO_4$ x $7H_2O$) (Sigma-Aldrich, St. Louis, Missouri, USA) of different concentrations (0 µmol/L, 250 µmol/L, 500 µmol/L, 750 µmol/L, 1000 µmol/L, 1250 µmol/L and 1500 µmol/L). This was used to set up a standard curve by linear regression (see Figure A2, Appendix I) from which the equation was solved to give an expression for concentration. Then, absorbance of samples could be entered to

estimate antioxidant activity. For sample analysis, bilberry raw extracts were further diluted to 1:10 with methanol to fit the range of the linear curve, and the same procedure as outlined above, with 10 µL sample extract replacing iron (II) sulfate heptahydrate, were followed. Three replicas for each sample were performed (see Table A1, Appendix II for the estimated concentration and SE of each sample). The plate was shaken and incubated for precisely four minutes before any measurements in a 96-well microplate spectrophotometer (Labsystems Multiscan MS, Helsinki, Finland) at 595 nm were performed.

2.2.5. Gas chromatography-mass spectrometry (GC-MS)

Gas chromatography-mass spectrometry (GC-MS) was utilized to detect nutritional compounds such sugars, organic acids and simple phenolics. For the analysis, 750 µL of the raw extract of bilberry juice (without further dilution) was pipetted into 2 mL Eppendorf tubes, dried in a Savant SpeedVac Plus (ThermoQuest, San Jose, California, USA) overnight and stored at -80 °C. For the derivatization of extracted sugars, acids and secondary metabolites, the dried residues were re-dissolved in 80 µL of 20 mg/mL methoxyamine hydrochloride (CH_3ONH_2 x HCl) (Fluka Biochemika, Buchs, Switzerland) in pyridine (Fluka Biochemika, Buchs, Switzerland) and left at 30 °C for 90 min in the Sanyo Gallenkamp orbital Incubator (Loughborough, UK) at 200 rpm. Finally, samples were treated with 80 µL of N-Methyl-N-(trimethylsilyl)trifluoroacetamide ($C_6H_{12}F_3NOSi$) (MSTFA) (Fluka Biochemika, Buchs, Switzerland) at 37 °C for 30 min in the same incubator at 200 rpm. Samples were then transferred to 1.5 mL autosampler vials with glass inserts (samples were pipetted into the insert) and stored at -20 °C prior to GC-MS. A high resolution GC-MS (Varian Star 3400 CX, Saturn 3, Agilent technologies, Santa Clara, California, USA) was utilized for all analyses and sample volumes of 1 µL were injected with a split ratio of 25:1. GC separations were carried out on a HP-5MS capillary column (30 m × 0.25 mm inner diameter (i.d.), with film thickness 0.25 µm). The injection temperature was 230 °C and the interface was set to 250 °C. The carrier gas was helium (He) at a constant flow rate of 1 mL min^{-1}. The GC temperature program was held isothermically at 70 °C for 5 min, ramped from 70 °C to 310 °C at a rate of 5 °C min, and finally held at 310 °C for 7 min (analysis time of 60 min). The MS source was adjusted to 230 °C and a mass range of *m/z* 50-550 was recorded. All mass spectra were

acquired in EI mode. Compounds were quantified based on the internal standard ribitol and expressed as µg per 100 g fresh weight. Sample Langvatnet 9 (LV9) had to be excluded from data processing as the GC-MS results were unsatisfying for this sample. See Tables A2 to A5 in Appendix II for the estimated concentration for each sample.

Chromatogram visualization and peak integration of GC-MS data was carried out using the SaturnView™ instrumental software. For mass spectra evaluation and peak identification, the AMDIS software version 2.64 was used in combination with the mass spectral libraries NIST 05 database and a target TMS database (Max-Planck Institute for Molecular Plant Physiology, Golm, Germany).

2.3. Genetic analyses

For estimation of genetic diversity at the inter- and intrapopulation levels, microsatellite markers were utilized. Analyses were carried out at the Biodiversity/Molecular Ecology lab at Department of Biology, NTNU.

2.3.1. Extraction of DNA

Extraction of DNA from bilberry leaves was carried out by use of DNeasy® Plant Mini Kit protocol (QIAGEN Cat. No.: 69104) (Venlo, Netherlands). Approximately 100 mg of frozen leaf tissue was disrupted by use of Retsch QIAGEN TissueLyser (Venlo, Netherlands). The remaining extraction was carried out as described in the QIAGEN DNeasy® Plant Mini Kit protocol.

2.3.2. Microsatellites, PCR and electrophoresis

Potentially suitable microsatellite loci were chosen searching the literature (Bassil et al., 2010; Boches et al., 2006; Boches et al., 2005). Among the 25 published microsatellite loci developed for bilberry or highbush blueberry, 16 promising loci were chosen for testing in eight randomly selected individuals distributed across the four study locations. The potentially suitable loci were chosen for testing based on their variability in bilberry and highbush blueberry as well as in other related species where their amplification success and variability had been determined. In addition, only loci with a simple repeat motif were selected, and the expected size range was used as a criterion for selection to allow multiplexing in two panels. Out of the 16 loci that were originally tested only four microsatellite loci were suitable for further analyses. Table 2.1 provides an overview of the names, amplification success, variability and suitability of the 16 loci tested. The four suitable microsatellite loci were originally included in either of the two multiplex panels and included loci with overlapping size range and the same fluorescence label. Hence,

multiplexing could not be carried out and Polymerase Chain Reaction (PCR) was carried out separately for each locus.

Table 2.1: Locus name, fluorescence label, amplification success, variability and suitability of the 16 microsatellite loci tested in samples from eight randomly selected individuals distributed across the four study locations in present study. Only four loci (in bold) were suitable for further analysis.

Locus	Fluorescence label	Amplification	Variability	Suitability
CA23F	FAM	Yes	Monomorphic	Poor
CA112F	VIC	No	Unknown	Unknown
CA169F	FAM	Yes	Monomorphic	Poor
CA344F	VIC	Yes	Monomorphic	Poor
CA421F	**NED**	**Yes**	**7 alleles**	**Good**
CA483F	**FAM**	**Yes**	**4 alleles**	**Good**
CA787F	FAM	Yes	Monomorphic	Poor
CA794F	VIC	Yes	Unspecific amplification	Poor
CA94F	FAM	Yes	Monomorphic	Poor
NA398	FAM	Yes	Monomorphic	Poor
NA741	**PET**	**Yes**	**2 alleles**	**Good**
NA961	**FAM**	**Yes**	**3 alleles**	**Good**
NA1040	PET	Yes	Unspecific amplification	Poor
VCC_I2	NED	Yes	Monomorphic	Poor
VCC_I8	NED	Partly (5/8)	2 alleles? (Only homozygotes)	Poor[a]
VCC_J5	VIC	Partly (5/8)	3 alleles?	Poor[b]

[a] The marker might contain null alleles (Pemberton et al., 1995), i.e. alleles which cannot be detected by the assay.

[b] Further optimizing of the process in subsequent studies might render the locus suitable.

Prior to PCR a 10 µM primer solution was made for each forward (F) and reverse (R) microsatellite primer. The four microsatellite loci chosen for analysis were CA421F (0.250 µL 10 µM F and R primer solution per sample), CA483F (0.125 µL 10 µM F and R primer solution per sample), NA741 (0.250 µL 10 µM F and R primer solution per sample) and NA961 (0.125 µL 10 µM F and R primer solution per sample). PCR was run with a touchdown PCR program on GeneAmp PCR system 9700 (Applied Biosystems, Foster City, California, USA) with 5 µL QIAGEN multiplex solution and 3 µL extracted DNA (approx. 10 to 20 ng/µL). Ultra pure water, MilliQ (MilliPore) H_2O was added to give a total amount of 10 µL reaction volume for each sample. The PCR profile started with a denaturing step of 15 min at 94.1 °C followed by 12 cycles of 30 s at 94 °C, 90 s initially at a temperature of 62 °C (then decreasing by 1 °C every cycle), and 60 s at 72 °C. This was followed by 23 cycles of 30 s at 94 °C, 90 s at 50 °C

and 60 s at 72 °C. Finally, there was an additional 5 min at 60 °C and an indefinite hold at 4 °C.

After PCR 10 µL highly deionized (Hi-Di) formamide (Applied Biosystems, Foster City, California, USA) was mixed with 0.5 µL size standard (GeneScan 600 LIZ size standard, Applied Biosystems, Foster City, California, USA) and 1 µL of PCR-product per sample. This mix of formamide, size standard and PCR product was applied to a separate well for each sample in an Applied Biosystems 96-well plate designed for use in the ABI 3130xl Genetic Analyser (Applied Biosystems, Foster City, California, USA). Electrophoresis to separate and visualize individual alleles for each microsatellite locus was carried out in a 16 capillary ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, California, USA).

2.3.3. Microsatellite genotyping

Individual genotypes at the four microsatellite loci were scored using GeneMapper Software version 4.0 (Applied Biosystems, Foster City, California, USA).

2.4. Data analysis

Handling of data from both the biochemical and the genetic analyses was performed using R version 2.12.1 (R Development Core Team, 2010). All estimates are given as mean ± 1 standard error (SE). To test whether there were any differences between locations in concentrations of different compounds, analyses of variance (ANOVA) and a multivariate analysis of variance (MANOVA) was performed. For compounds with significant differences between locations a Tukey's honest significant difference (HSD) test was run to find, through pairwise comparison, which means that significantly differed from one another. All these analyses were performed in R software.

Multivariate statistics were utilized for a multiple comparison of all metabolite concentrations including total phenolics (TPH), total anthocyanins (ACY) and antioxidant activity (FRAP) across the different locations. Two multivariate statistics, hierarchical clustering in MultiExperiment Viewer (MeV) version 4.6.2 (Saeed et al., 2003) and principal component analysis (PCA) in R software were utilized to explore the metabolite profile of the samples. The hierarchical clustering creates clusters of samples based on their metabolic similarities while the PCA aims at extracting patterns in the variation of the data by stepwise fitting of orthogonal axes (principal components) in the direction of highest variation.

Genetic differentiation between populations was estimated by F_{ST} using the package HIERFSTAT in R software (Goudet, 2005). Program ISOLDE in GenePop version 4.0 (Raymond and Rousset, 1995b) was used to run a Mantel's test (Mantel, 1967) for the isolation by distance analysis. Here, a transformation of F_{ST} (i.e. $F_{ST}/1\text{-}F_{ST}$) and the natural logarithm of geographic distance were used (Rousset, 1997). Prior to these analyses the four microsatellite loci were tested for deviance from Hardy-Weinberg equilibrium using GenePop version 4.0 (Raymond and Rousset, 1995b). GenePop version 4.0 was also used to estimate observed (H_O) and expected (H_E) heterozygosity. Further estimates of the level of intra-population genetic variation, such as the number of alleles (N_A) and allelic richness (A_R), was calculated using the software FSTAT version 2.9.3.2 (Goudet, 1995).

3. Results

3.1. Biochemical composition

The multivariate analysis of variance (MANOVA) for all compounds combined (total phenolics (TPH), total anthocyanins (ACY), antioxidant activity (FRAP), fructose, glucose, sucrose, arabinose, ribose, xylose, citric acid, malic acid, fumaric acid, quinic acid, chlorogenic acid, shikimic acid, myo-inositol, galactinol, chiro-inositol, not annotated (NA) trisaccharide and not annotated (NA) monosaccharide) indicated that there were significant differences between the four locations in the concentrations of some of these compounds ($F_{3,32}$= 5.132 , $P < 0.001$). Mean concentrations and standard errors (SE) for all compounds in the four locations are found in Tables A1 to A5 in Appendix II.

3.1.1. Concentrations of TPH, ACY and FRAP

Total phenolics (TPH), total anthocyanins (ACY), and antioxidant activity (FRAP) was quantified for the four different locations (see Figure 3.1). Figure 3.1 shows the mean concentrations of total phenolics (TPH) and total anthocyanins (ACY), the ratio between ACY and TPH and the mean antioxidant activity (FRAP) for the different locations. The estimates are based on ten individuals for Lierne, Langvatnet and Hedmark and seven for Tromsø. Table A1 in Appendix II provides the estimated concentrations of each individual sample.

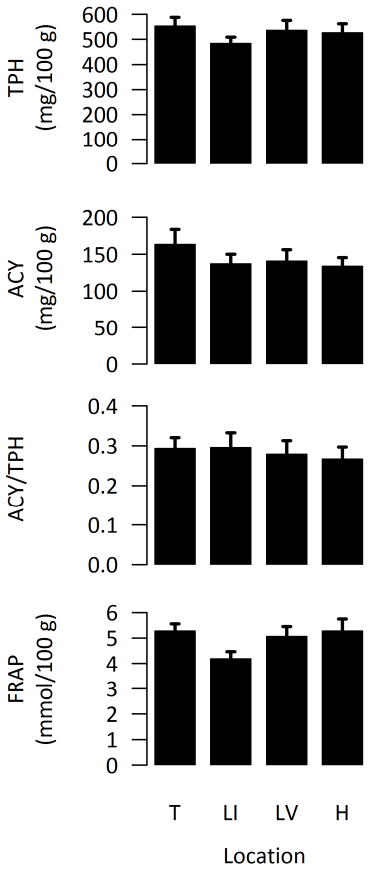

Figure 3.1: Barplots of the concentrations of total phenolics (TPH) (mg/100 g fresh weight) and total anthocyanins (ACY) (mg/100 g fresh weight), the ratio between ACY and TPH, and the antioxidant activity (FRAP) (mmol/100 g fresh weight) for the four locations Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Means for the four locations ± standard errors (SE) are given.

For total phenolics (TPH) Tromsø (T) had the highest concentration, closely followed by Langvatnet (LV) and Hedmark (H), while Lierne (LI) had the lowest of all locations (range of mean TPH: 483.916 (Lierne) - 552.704 (Tromsø) mg/100 g). However, all concentrations are fairly similar, and no significant differences were found between the locations ($F_{3,33}$ = 0.668, *P* = 0.577). For total anthocyanins (ACY) Tromsø again showed the highest concentration, while the three other locations had a noticeable lower but similar result (range of mean ACY: 133.563 (Hedmark) - 163.080 (Tromsø) mg/100 g). Thus no significant differences between the locations was found ($F_{3,33}$ = 0.652, *P* = 0.587). The ratio between total anthocyanins and total phenolics were also very similar for all locations, particularly for Tromsø and Lierne (range of mean ACY/TPH: 0.266 (Hedmark) - 0.295 (Lierne), and no significant difference was

found between the locations ($F_{3,33}$ = 0.162, *P* = 0.921). For antioxidant activity (FRAP) some differences could be observed with Lierne having the lowest estimate (range mean of FRAP: 4.181 (Lierne) - 5.273 (Hedmark) mg/100 g), however, again no significant difference was found between the locations ($F_{3,33}$ = 1.897, *P* = 0.149). Another aspect to notice is the levels of antioxidant activity rather closely trace the concentrations of total phenolics with only a minor dissimilarity in Langvatnet and Hedmark.

3.1.2. Sugars, acids and simple phenolics

Figure 3.2 below shows the mean concentrations (µg/100 g fresh weight ± SE) of the sugars investigated at the four locations and Tables A2 to A5 in Appendix II provides the estimated concentrations of each individual sample.

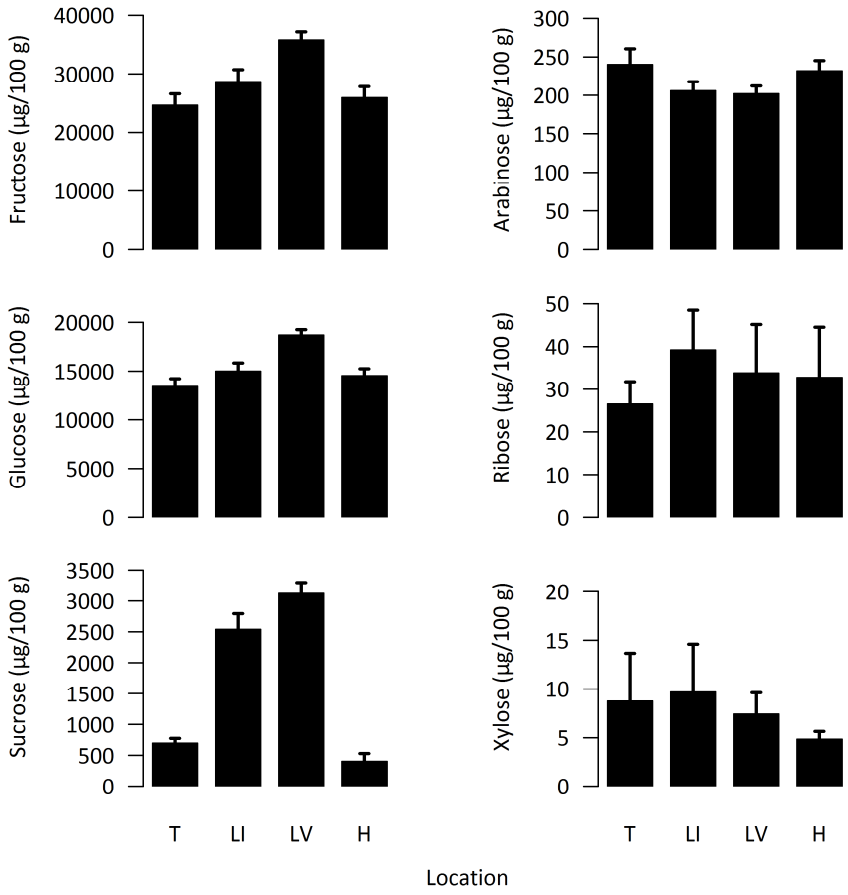

Figure 3.2: Barplots of the concentrations (µg/100 g fresh weight ± SE) of the sugars, fructose, glucose, sucrose, arabinose, ribose and xylose for the four different locations Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Means for the four locations ± standard errors (SE) are given.

On the left in Figure 3.2 the sugars of highest concentrations are listed. For fructose, samples from Langvatnet (LV) and Lierne (LI) had the highest concentrations, while Tromsø (T) and Hedmark (H) showed lower but fairly similar results. There was significant differences in the concentration of fructose between the locations ($F_{3,32}$ = 6.571, P = 0.002). The pairwise comparisons indicated significant differences between Langvatnet and Hedmark (P = 0.004), Langvatnet and Lierne (P = 0.049) and Tromsø and Langvatnet (P = 0.003). The remaining pairwise comparisons were not significantly different (P > 0.05). The same pattern was maintained for glucose ($F_{3,32}$ = 9.094, P < 0.001) and even more pronounced, for sucrose ($F_{3,32}$ = 53.669, P < 0.001). Pairwise comparisons revealed significant differences in glucose concentration between Langvatnet and Hedmark (P = 0.001), Langvatnet and Lierne (P = 0.005) and Tromsø and Langvatnet (P < 0.001), while the remaining locations were not significantly different (P > 0.05). For sucrose there were significantly differences between

Lierne and Hedmark ($P < 0.001$), Langvatnet and Hedmark ($P < 0.001$), Tromsø and Lierne ($P < 0.001$) and Tromsø and Langvatnet ($P < 0.001$), with the remaining comparisons being non-significantly different ($P > 0.05$). Regarding 5-carbon sugars depicted on the right hand side of Figure 3.2, none of the sugars differed significantly in concentration between the locations (arabinose: $F_{3,32} = 1.788$, $P = 0.169$; ribose: $F_{3,32} = 0.235$, $P = 0.871$; xylose: $F_{3,32} = 0.383$, $P = 0.765$). However, arabinose levels tended to be lower in Lierne and Langvatnet, ribose showed a noticeable difference between Tromsø and Lierne, and xylose concentrations were lowest in samples from Hedmark.

Mean concentrations (µg/100 g fresh weight ± SE) of the organic acids investigated at the four locations are shown below in Figure 3.3. The most abundant of these were quinic acid, citric acid and malic acid. Tables A2 to A5 in Appendix II provides the estimated concentrations of each individual sample.

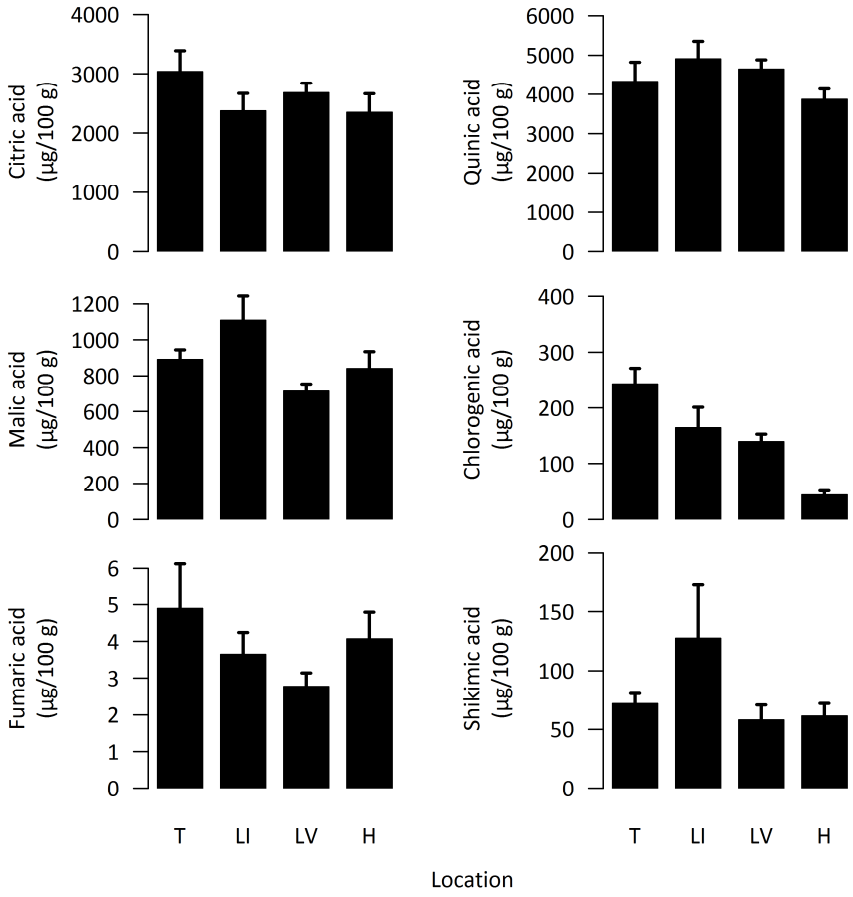

Figure 3.3: Barplots of the concentrations (µg/100 g fresh weight) of the organic acids, citric acid, malic acid, fumaric acid, quinic acid, chlorogenic acid and shikimic acid for the four different locations Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Means for the four locations ± standard errors (SE) are given.

The results for citric, malic and fumaric acid are fairly different, however, only for malic acid were there marginally significant differences between the locations ($F_{3,32}$ = 3.111, *P* = 0.039). Pairwise comparisons indicated that the only significant difference were between Langvatnet and Lierne (*P* = 0.027), the remaining locations had similar concentrations (*P* > 0.05). Malic acid concentrations were highest in bilberries from Lierne and Tromsø. Even though there were no significant differences between the locations (citric acid: $F_{3,32}$ = 1.122, *P* = 0.355; fumaric acid: $F_{3,32}$ = 1.325, *P* = 0.283), citric acid content was highest in Tromsø and Langvatnet, while fumaric acid showed highest concentrations in samples from Tromsø and Hedmark and lowest levels in the samples from Langvatnet. On the right hand side of Figure 3.3, quinic acid had by far the highest concentration when compared to the phenolic structures chlorogenic acid and shikimic acid. The concentrations of quinic acid were very similar for all locations, thus no significant differences were found ($F_{3,32}$ = 1.601, *P* = 0.208).

On the other hand, chlorogenic acid showed a clear decrease in concentration from Tromsø to Hedmark populations ($F_{3,32}$ = 10.488, P < 0.001). The pairwise comparisons revealed significant differences between Lierne and Hedmark (P = 0.005), Langvatnet and Hedmark (P = 0.039), Tromsø and Hedmark (P < 0.001) and Tromsø and Langvatnet (P = 0.043), but not between the two remaining comparisons (P > 0.05). Bilberries from Lierne showed a noticeable higher concentration of shikimic acid compared to the other locations, however, the variation between individuals within the locations was also high and no significant differences were detected ($F_{3,32}$ = 1.517, P = 0.229).

Among the five remaining compounds investigated, galactinol and NA monosaccharide showed significant differences between the locations (galactinol: $F_{3,32}$ = 6.671, P = 0.002; NA monosaccharide: $F_{3,32}$ = 4.794, P = 0.007; see also Figure A3, Appendix II). Pairwise comparisons revealed that differences in galactinol concentration were found between Langvatnet and Hedmark (P < 0.001), Tromsø and Hedmark (P = 0.042) and Langvatnet and Lierne (P = 0.047), the remaining locations were not significantly different (P > 0.05). The concentration of NA monosaccharide displayed significant differences between Tromsø and Langvatnet (P = 0.012), but there was no differences between the remaining locations (P > 0.05). The three final compounds investigated displayed no significant differences between the locations (myo-inositol: $F_{3,32}$ = 1.520, P = 0.228; chiro-inositol: $F_{3,32}$ = 1.526, P = 0.226; NA trisaccharide: $F_{3,32}$ = 0.293, P = 0.830; see Figure A3, Appendix II).

3.1.3. Multivariate statistics

3.1.3.1. Hierarchical clustering

The hierarchical clustering grouped samples from the four different locations based on the metabolic similarities of the 20 compounds, including total phenolics (TPH), total anthocyanins (ACY) and antioxidant activity (FRAP) (Figure 3.4).

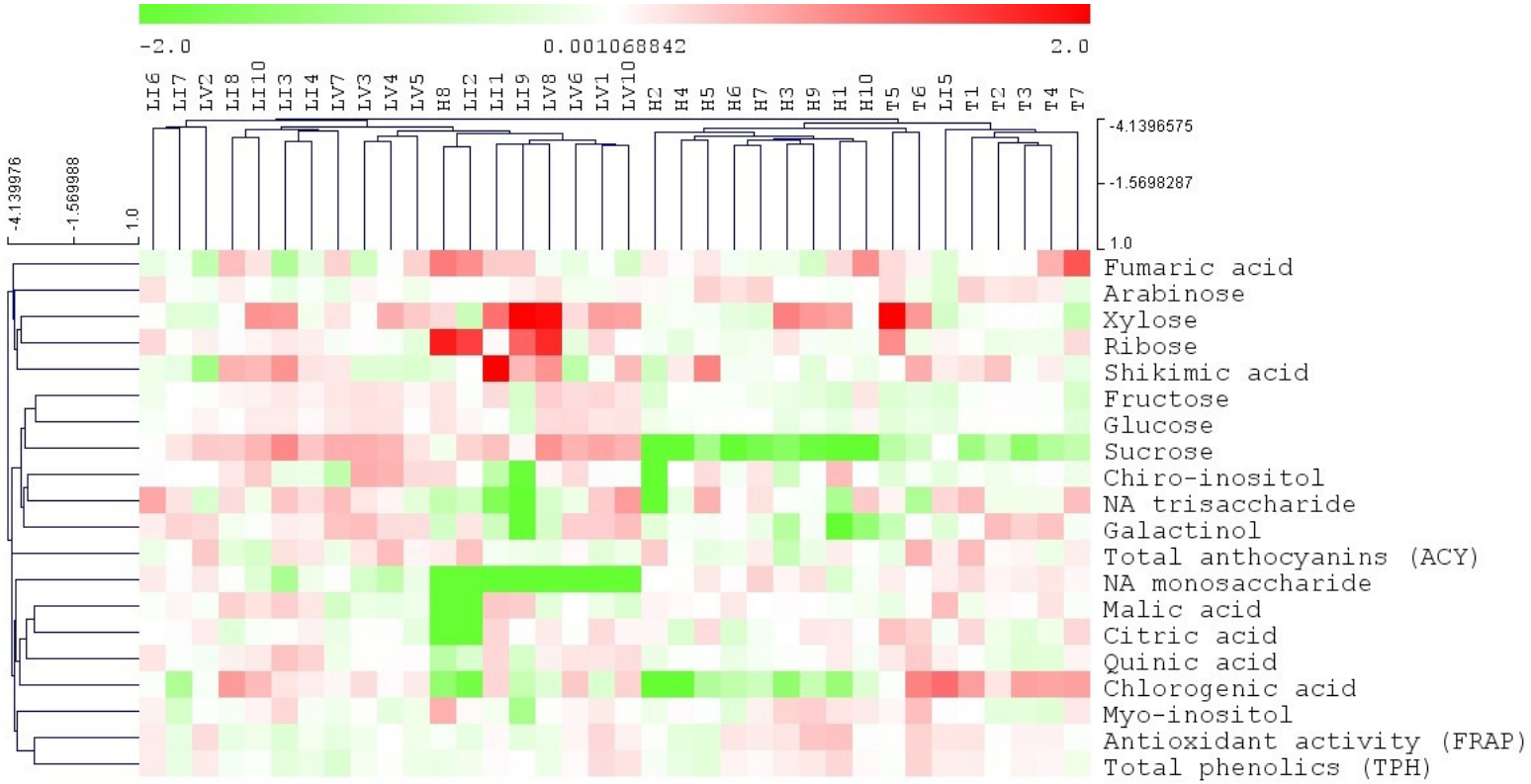

Figure 3.4: Heat map of clustering of metabolite data for all compounds in all the samples from the different locations, Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Green colour represents a lower concentration of a compound compared to the median and red a higher concentration. On the right hand side of the map, compound names are indicated, the left side shows the compound clustering, whilst the clustering at the top of the map shows the names and grouping of all samples. Sample LV9 was excluded from the analysis.

The heat map in Figure 3.4 shows that samples from same location to a large extent group together. On top of the map two major clusters are seen, with samples from Lierne and Langvatnet at the left and Hedmark and Tromsø on the right. Lierne and Langvatnet forms a cluster where the samples are very much mixed. Samples from Hedmark and Tromsø, however, do form a cluster but their samples still remains somewhat separated from each other within the cluster. In addition, two samples (LI5 and H8) are exceptions of this location clustering.

From Figure 3.4 it is evident that different compounds are clustered together, as seen at the left side of the map, according to their levels of concentration in the different bilberry samples. Arabinose, xylose, ribose and shikimic acid had elevated levels of concentration in some samples from Lierne and Langvatnet and thus clustered together. Furthermore, the cluster of fructose, glucose and sucrose together displayed an elevated concentration in samples from Lierne and Langvatnet and low concentration in Hedmark and Tromsø. NA monosaccharide, malic acid and citric acid, which also clustered, displayed a lowered concentration in some samples from Hedmark, Lierne and Langvatnet.

3.1.3.2. Principal Component Analysis (PCA)

A principal component analysis (PCA) was performed to further reveal the internal structures in the variation of the compound concentrations among the different locations. The first two principal components (PCs) were used to visualize the structure in the data in a PCA biplot and accounted for 37 % of the total variance (Figure 3.5).

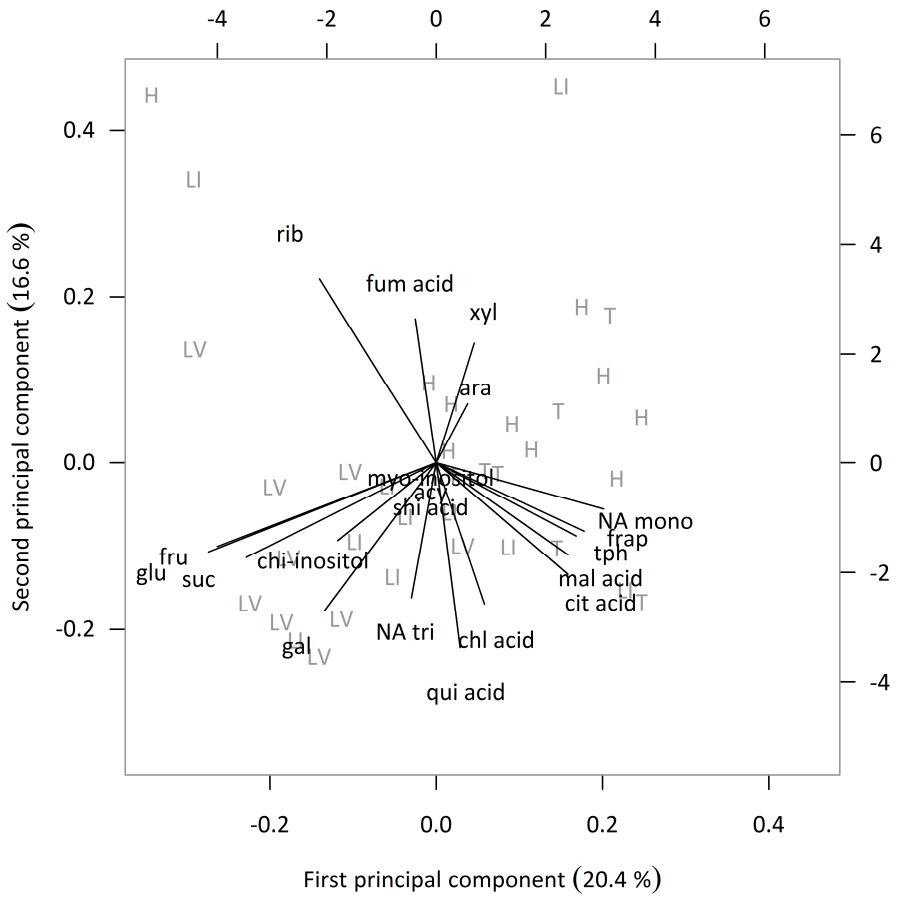

Figure 3.5: Biplot from a principal component analysis (PCA) of metabolite concentrations for all compounds from the four locations, Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Position of PC scores for 36 samples (sample LV9 had to be excluded) is shown in grey letters, and the vectors (arrows), one for each of 20 metabolite compounds, point in the direction of increasing concentration. Arrows that point in the same direction indicate that concentrations of compounds are correlated. Abbreviated names of compounds are shown at the end of the arrows. Tph = total phenolics, acy = total anthocyanins , frap = antioxidant activity, fru = fructose, glu = glucose, suc = sucrose, ara = arabinose, xyl = xylose, rib = ribose, cit acid = citric acid, mal acid = malic acid, fum acid = fumaric acid, qui acid = quinic acid, chl acid = chlorogenic acid, shi acid = shikimic acid, gal = galactinol, chi inositol = chiro-inositol, NA tri = not annotated (NA) trisaccharide, NA mono = not annotated (NA) monosaccharide. Myo-inositol was not abbreviated.

The figure shows that the variation explained by the first principal component is 20.4 %, and the second explains 16.6 %. The samples, as indicated in gray in the biplot, reveal a grouping of Tromsø (T) and Hedmark (H) at the upper right side of the plot and a grouping of Lierne (LI) and Langvatnet (LV) at the lower left side of the figure, thus samples from each locations associated similarly with the two PCs. The compounds fructose, glucose, sucrose, chiro-inositol and galactinol have a strong negative loading on principal component 1 (PC1). On the other side, NA monosaccharide, malic acid, citric acid, antioxidant activity (FRAP) and

total phenolics (TPH) have positive loadings on PC1. Regarding principal component 2 (PC2), fumaric acid, ribose, xylose, and arabinose have a positive loading along this axis. In contrast, NA trisaccharide, chlorogenic acid, quinic acid and galactinol have a strong negative loading on PC2. Furthermore, it is also seen in Figure 3.5 that several other compounds which are associated with PC1 have slight negative loadings on PC2. The biplot also reveals which concentrations of compounds that are correlated and responded similar to the PCs. NA monosaccharide, antioxidant activity (FRAP), total phenolics (TPH), malic acid and citric acid are closely associated at the right hand side. Another clear association is fructose, glucose and sucrose at the left.

From the biplot some patterns in the distribution of compound concentrations among the locations are visible. Lierne (LI) and Langvatnet are associated with higher concentrations of fructose, glucose, sucrose, together with chiro-inositol, galactinol and to some extent NA trisaccharide, quinic acid and chlorogenic acid. Samples from Hedmark and Tromsø were associated with the 5-carbon sugars arabinose and xylose, and fumaric acid, NA monosaccharide, malic acid, citric acid, antioxidant activity (FRAP) and total phenolics (TPH) .

3.2. Genetic analyses

3.2.1. Genetic differentiation between locations

The genetic differentiation estimated as F_{ST} across all four locations was 0.026 (95% CI: -0.015, 0.050). Thus, the level of genetic differentiation between the locations was low and not significantly different from zero. The level of genetic differentiation between pairs of locations varied between -0.036 and 0.055 (Table 3.1). However, none of the pairwise F_{ST} estimates were significantly different from zero.

Table 3.1: Level of genetic differentiation (F_{ST}) between pairs of locations. 95% confidence intervals are given in parentheses.

	Hedmark (H)	Lierne (LI)	Langvatnet (LV)
Lierne (LI)	0.038 (-0.052, 0.106)	-	-
Langvatnet (LV)	-0.036 (-0.046, -0.026)	0.047 (-0.034, 0.122)	-
Tromsø (T)	0.010 (-0.036, 0.047)	0.055 (-0.021, 0.128)	0.038 (-0.008, 0.097)

There was also no significant relationship between the level of genetic differentiation and geographic distance between pairs of locations (Mantel test: *P* = 0.650). This is clearly seen in Figure 3.6.

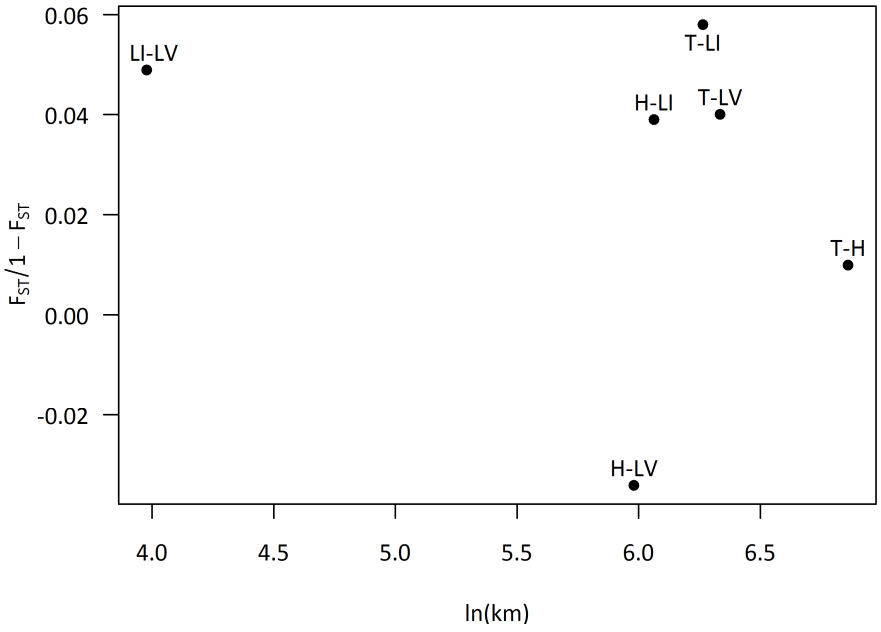

Figure 3.6: Tests of isolation by distance. The relationship between genetic differentiation (given as $F_{ST}/1-F_{ST}$) and the geographic distance (i.e. natural logarithm of number of km) separating each location, Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). The number of km (measured as linear distance) separating the locations was; T-H: 953.47 km, T-LV: 563.64 km, T-LI: 525.41 km, H-LI: 429.63 km, H-LV: 395.91 km and LI-LV: 53.42 km.

3.2.2. Genetic variability within locations

Using four (polymorphic) microsatellite loci a fairly low overall genetic diversity was found with two to eight alleles per locus (Table A6, Appendix III). The mean allelic richness (A_R) of the different locations was Tromsø: 3.745, Lierne: 2.818, Langvatnet: 3.524, Hedmark: 2.966, and mean expected heterozygosity (H_E) of the locations was Tromsø: 0.514, Lierne: 0.538, Langvatnet: 0.532, Hedmark: 0.519. There was no difference between locations in either mean allelic richness (A_R) ($F_{3,12} = 0.211$, $P = 0.887$) or mean expected heterozygosity (H_E) ($F_{3,12} = 0.010$, $P = 0.998$). Thus, there seemed to be similar levels of genetic variation within all the four locations.

4. Discussion

This study showed that there were significant differences in metabolite concentrations between the four studied bilberry locations, but there was little evidence of a latitudinal gradient in the metabolite concentrations. However, using both the compounds that significantly differed between the locations (e.g. fructose, glucose, sucrose, malic acid and chlorogenic acid) and the remaining compounds that make out the metabolic profile, the locations could be grouped according to their metabolic similarities and inference could be drawn regarding the association of different compounds (Figure 3.4 and 3.5). Despite the shown differences in the metabolic compounds of the locations there were no genetic differentiations between the locations with the microsatellites utilized in the study (Table 3.1), in addition there was found similar levels of genetic variation within all the four locations.

4.1. Biochemical composition

4.1.1. Levels of metabolite concentration and latitudinal differences

Concentrations of total phenolics (TPH) from the present study were in the same order as those found in earlier studies of *Vaccinium* species (Connor et al., 2002a; Giovanelli and Buratti, 2009; Ochmian et al., 2009; Pop et al., 2008; Prior et al., 1998). For instance, Giovanelli and Buratti (2009) found mean total phenolics for bilberry in Italy to range from 577 to 614 mg/100 g between two locations, which is comparable to the present study were mean concentrations ranged from 484 to 553 mg/100 g. Furthermore, the present study confirms earlier findings of higher concentrations of total phenolics in bilberry than in highbush and lowbush blueberry (e.g. Connor et al., 2002a; Giovanelli and Buratti, 2009; Prior et al., 1998). For total phenolics within the present study, samples from Tromsø had the highest concentrations closely followed by Langvatnet, Hedmark and Lierne. However, despite earlier evidence for a positive relationship between lower temperatures and total phenolic concentrations in berries of bilberry, with northern populations generally

containing more phenolics (Martinussen et al., 2009), there was no significant difference detected between the locations in the present study.

Earlier studies on total anthocyanin (ACY) concentration have reported concentrations within the range of the results from the present study (Lätti et al., 2008; Åkerström et al., 2010) however, also concentrations twice the order of the present results have been reported (Giovanelli and Buratti, 2009; Prior et al., 1998). The means of total anthocyanin concentration in the present study ranged from 134 to 163 mg/100 g, with a slightly higher concentration found in the northernmost population, Tromsø. This is in accordance with other studies (Jaakola et al., 2004; Lätti et al., 2008; Åkerström et al., 2010), where an accumulation of anthocyanins often is a result of prolonged days and lower temperatures, as most often found at northern latitudes (Lätti et al., 2010; Åkerström et al., 2010). However, in present study this is only evident as a trend since no significant differences between the locations were detected.

Anthocyanins comprise a major fraction of the total phenolics in *Vaccinium* species (Prior et al., 1998). For instance, Prior et al. (1998) found the ratio between total anthocyanins and total phenolics to be 0.57 for bilberry and equally Giovanelli and Buratti (2009) in their investigations found ratios of 0.57 and 0.56 for two different locations. Such a relationship was also found in the present study, however, with mean ratios ranging from 0.27 to 0.29 between locations, the fraction is slightly lower than reported by Prior et al. (1998) and Giovanelli and Buratti (2009). Since concentrations of total phenolics are in accordance with these studies, the lower ratios are due the lower concentrations of anthocyanins found in present study. Thus, anthocyanins accounts of less than 50 % of total phenolics in present study and this might affect antioxidant activity in these berries. However, earlier studies of other *Vaccinium* species have found rather high potential for variation in the anthocyanin/total phenolics ratio, for instance Moyer et al. (2002) found ratios in cultivars of highbush blueberry to range from 0.28 to as high as 0.63. Further, the present results showed no significant differences in the ratio between the locations, the northern location was equal to both the Mid-Norway and southern locations. Thus, anthocyanins seem to constitute a similar portion of total phenolics irrespective of changes in latitude (Figure 3.1).

Mean estimates of antioxidant activity (FRAP) found within the present study (4.18 to 5.27 mmol/100 g) were within the range of estimates previously found for bilberry in some studies (Giovanelli and Buratti, 2009) but lower than the estimates of 7.57 to 8.86 mmol/100 g reported by Halvorsen et al. (2002) when analysing bilberry from Poland, Norway and Sweden. Furthermore, the results confirm that bilberry has a higher antioxidant activity than what is usually found in highbush and lowbush blueberry (Giovanelli and Buratti, 2009; Halvorsen et al., 2002; Prior et al., 1998; but see Moyer et al. 2002). Looking at Figure 3.1 one aspect that is noteworthy is that the relative differences in antioxidant activity (FRAP) between locations is similar to those found for the concentrations of total phenolics (TPH) with only a minor change in Langvatnet and Hedmark. Total phenolics and therein anthocyanins are large contributors to the high antioxidant activity in bilberries (Camire, 2002; Prior et al., 1998; Rice-Evans et al., 1996) and it is well established that there is a strong positive relationship between total phenolics, anthocyanin content and antioxidant activity (Connor et al., 2002a; Ehlenfeldt and Prior, 2001; Moyer et al., 2002; Prior et al., 1998). Furthermore, antioxidant activity usually coincides better with total phenolics than anthocyanins in the investigated *Vaccinium* species (Howard et al., 2003; Moyer et al., 2002; Prior et al., 1998) something which is reflected by the present results (see Figure 3.4 and 3.5). As the concentrations of total phenolics are within previously found range (Prior et al., 1998), the lower estimate of antioxidant activity compared to Halvorsen et al. (2002) seems not to be a reflection of the amount of total phenolics but rather the lower concentration of total anthocyanins. These concentrations have been known to be affected by factors such as sampling date, genetic differences, preharvest environmental conditions, postharvest storage conditions and processing (Connor et al., 2002b; Ehlenfeldt and Prior, 2001; Pop et al., 2008) and consequently affects antioxidant activity in bilberry samples.

As mentioned above, sampling date may affect the amount of total phenolics and total anthocyanins, and thus also the antioxidant activity (Burdulis et al., 2007; Connor et al., 2002b; Prior et al., 1998). This is due to the positive relationship between increased maturity at harvest and the concentrations (Prior et al., 1998). Thus, there is a possibility for an effect of sampling date on the results of the present study, since sampling in Mid- and South Norway was carried out in mid August (16^{th} and 20^{th}) and sampling in Tromsø took place at

the beginning of August (6th of August). However, this does not seem to be the case as Tromsø both had highest concentrations for total phenolics and high antioxidant activity despite having the earliest sampling date. Thus, an effect of sampling date can be ruled out and other factors such as genetic differences, preharvest environmental conditions, and postharvest storage conditions and processing should be considered. The only difference in the postharvest storage conditions in the present study was that samples from Tromsø and Hedmark were shipped on dry ice to Trondheim. However, this should not have had any effect on the results and it is assumed that postharvest storage conditions were similar for all samples. Furthermore, all processing of berries in the laboratory were performed in a uniform manner. Thus, processing should not have affected samples differently and preharvest environmental conditions and genetic differences are most likely the major factors that have contributed to differences between populations in present study.

The concentrations of metabolites analyzed by GC-MS in the present study (fructose, glucose, sucrose, arabinose, ribose, xylose, citric acid, malic acid, fumaric acid, quinic acid, chlorogenic acid, shikimic acid, myo-inositol, galactinol, chiro-inositol, NA trisaccharide and NA monosaccharide) will never be an exact reflection of their real values because only one standard with known concentration was utilized, and variation in the detector response for different compounds (chemical structures) was not taken into consideration. Hence, the method used in present study was an approximation to quantify differences between samples (semiquantitative), and not a method to calculate exact concentrations for each compound. Thus, absolute concentrations of compounds measured by GC-MS will not necessarily be comparable with concentrations in published literature. However, a comparison of amounts of the different compounds between samples will be unbiased and reliable.

The predominant sugars in bilberry samples from the present study were fructose and glucose, something which was also found by Viljakainen et al. (2002) for bilberry collected in Finland and by Kader et al. (1993) in their study on highbush blueberry. Both studies agree with the present findings that fructose is predominant, glucose closely follow and sucrose is present with a lesser amount. Furthermore, the same pattern is repeated in several other species of berries (Viljakainen et al., 2002). The proportions of fructose and glucose are

similar across the locations as both are major substrates in the hexose phosphate pool and are therefore converted into each other during carbon metabolism (Buchanan et al., 2000). From the present study it is also clear that the main sugars (fructose and glucose) are present in several times higher concentrations than the 5-carbon sugars (arabinose, ribose and xylose) (Figure 3.2). Based on the present results arabinose is the predominant sugar in the juicy parts of berries, while a study by Gross and Sams (1984) on highbush blueberry showed that xylose is the predominant sugar in cell walls of berries and that the content of arabinose was found to be fairly low. This may mainly be due that the cell wall of berries is high in hemicelluose with xylose as the sugar present in largest amount (Sullivan et al., 1960). These results can also be taken as validation for the methods used in present study, in that it has succeeded in separating supernatant from cell walls and other residues.

The two main sugars were, as mentioned, present in much higher concentrations than the 5-carbon sugars, a finding which also holds when comparing to any of the organic acids in the present study. This pattern has also been shown in other studies (Kader et al., 1994; Viljakainen et al., 2002). Among the organic acids, Ayaz et al. (2001) found quinic and citric acid to be present in highest concentrations in bilberries from Turkey, a result which is in accordance with results of the present study (Figure 3.3). On the other hand, Kader et al. (1994) found citric acid to be present in highest concentrations in highbush blueberry, however concentrations of quinic acid was not estimated in this study.

On a latitudinal gradient, Martinussen et al. (2009) found higher sweetness (measured as fructose, glucose and sucrose) in samples of bilberries from northern latitudes in Finland compared to southern samples. As berries are an annual part of the plants, a cryoprotective role of this accumulation of sugars (Taiz and Zeiger, 2006) seems unlikely. On the contrary, analysing samples of black currant (*Ribes nigrum*), Zheng et al. (2009b) found that samples from southern regions in Finland had higher concentrations of fructose, glucose and sucrose. In addition, the present study found no evidence for a latitudinal trend, rather Lierne and Langvatnet possessed significantly higher concentrations of sugars, especially sucrose, than the other two locations. This may potentially be due favourable environmental conditions in Mid-Norway this year or in these locations, alternatively there could be genetic differences between the berries in the locations. In any case, it is not possible to generalize any patterns

in concentrations of sugars based on results from this and previous studies. With regard to acids, Zheng et al. (2009b) found a higher concentration of citric acid in southern regions of Finland compared to the northern but a lower content of malic acid and quinic acid. However, in the present study, there was a trend of lower concentrations of both quinic and citric acid towards south, but the pattern was irregular and not significant. Malic acid did display significant differences between locations, but this was due to a high concentration in one of the Mid-Norway locations and there was no latitudinal trend here either (see Figure 3.3). On the other hand, for chlorogenic acid within present study, the northern location had significantly higher concentrations than the locations to the south, and the concentrations were clearly decreasing with lower latitudes (Figure 3.3). This may reflect an adaption to growing at northern locations, because at northern latitudes, berries will be continuously subjected to abiotic stresses such as frost, almost continuous light during the growing season and often nutrient-poor soils (Martz et al., 2010). Thus, the accumulation of chlorogenic acid may be one way to protect berries from oxidative damage and stress. Indeed, Martinussen et al. (2009) also found higher concentrations of phenolic acids in northern clones in their study on bilberries.

As seen above, few clear patterns in the distribution of sugars or acids are evident. However, climatic conditions have earlier been shown to affect the level of some of the investigated compounds (Zheng et al., 2009a). In their study on *Ribes* species, Zheng et al. (2009a) found an effect of latitude, humidity, temperature and radiation level on the amount of sugars and acids within their cultivars. Thus, it is reasonable to believe that the environment at the growing sites of bilberry have an influence on the biochemical composition and contributes to locations with characteristic chemical profiles.

The concentrations of sugars and acids in various berry species can, like the concentrations of phenolics, be affected by the level of maturation at the date of sampling (Zhang et al., 2011). For instance, Zhang et al. (2011) found an increase in fructose, sucrose, citric and malic acid with increasing stage of development in strawberries (*Fragaria x ananassa*), and the same result has been found in bilberries for fructose, glucose, sucrose and malic acid (Ayaz et al., 2001). However, due to cell degradation occurring in overripe berries Zhang et al. (2011) also found, as expected, a decrease in some of the sugars and acids with additional

development. Thus, it is important to sample berries at the same stage of development when comparing biochemical composition of different locations. In present study, this was taken into consideration and care was made to sample bilberries from all locations within a reasonable short time period after ripening in August.

4.1.2. Metabolic associations among the locations

The grouping of samples in the hierarchical clustering heat map and in the PCA biplot indicates similarities of the metabolic profiles of the samples. Both the heat map in Figure 3.4 and the biplot in Figure 3.5 clearly show that the samples from the different locations to a large degree group together with other samples from the same location. The regional grouping of the samples indicates that samples from each region can be distinguished statistically from samples from other regions based on their metabolic profile. As expected, Lierne and Langvatnet clustered tightly together with both approaches. Their grouping was of no surprise as the locations were within close geographical range and probably had been subjected to similar environmental conditions. In addition, based on their metabolic similarity they could have been closely related due to a common origin after immigration to Norway at the end of the last glacial period (Wille, 1915). The heat map and the PCA biplot both also present a clustering of samples from Hedmark with samples from Tromsø. This was much more surprising as it indicates a metabolic similarity between the bilberries in the far North with those to the far South. Since they are geographically separated by a long distance they most likely have not been subjected to same environmental conditions (e.g. temperature, rainfall, radiation). Thus, other factors must have contributed to their metabolic similarities, e.g. a common genetic background.

A closer look at the biplot shows that some compounds are correlated, such as the sugars to the left (fructose, glucose and sucrose), the acids in the middle (shikimic, chlorogenic and quinic acid) and a mix of compounds to the right (NA monosaccharide, antioxidant activity (FRAP), total phenolics (TPH), malic acid and citric acid). With this it is possible to quickly identify some of the previously mentioned compound associations, such as, the close association between fructose and glucose and the correlations of total phenolics (TPH) and antioxidant activity (FRAP). Furthermore, citric and malic acid are shown to correlate, and

this may be due that both are involved in the citric acid cycle (Buchanan et al., 2000). With this in mind, it can be seen that the grouping of Langvatnet and Lierne is associated with a similar increase in the concentrations of sugars (fructose, glucose and sucrose) together with chiro-inositol and galactinol. While, increases in ribose, fumaric acid, xylose and arabinose together with NA monosaccharide, antioxidant activity (FRAP), total phenolics (TPH), and malic and citric acid group to some extent samples from Tromsø with samples from Hedmark.

The two primary components (PCs) in the biplot accounted for 37 % of the total variance, thus they explain some of the segregation and cluster patterns, but the pattern was clearly more complex than what could be entirely explained by using only two PCs. However, comparing the grouping of samples and compounds in the biplot with the heat map, there is clear similarities that provide confidence in the results. The small discrepancies, which are due to differences in the procedures of hierarchical clustering and PCA (Sumner et al., 2003), do not change the conclusions. Thus, by computing both multivariate analyses, trends that are supported by both analyses will be more reliable and more robust. In particular, both the heat map and the PCA biplot provided evidence for grouping of samples from Lierne with Langvatnet, and Tromsø with Hedmark. The PCA biplot showed a lower concentration of sugars for Hedmark and Tromsø and higher for Lierne and Langvatnet, which is also evident in the heat map. Furthermore, NA monosaccharide showed a higher concentration in Tromsø and Hedmark and a lower in the locations in Mid-Norway, which can be observed by both approaches and suggest that conclusions are robust. Another line of evidence follows from comparison of the concentration of different compounds in barplots in Figure 3.2 and 3.3 and A3 in Appendix II with both the heat map (Figure 3.4) and the PCA biplot (Figure 3.5). For instance, the barplots in Figure 3.2 shows an increase in fructose, glucose and sucrose in Lierne and Langvatnet, which is evident in both the heat map and in the biplot. Thus, there should be no reason to doubt the conclusions reached at this stage.

4.1.3. Effect of sample size on metabolite concentration

The low number of locations and low sample size in each may have had an effect on the robustness of the results in the present study. The effect of low sample size is that each

sample will have a larger effect on the estimates, while this does not cause bias in the results, it will cause relatively large standard errors. This can be seen for some of the metabolites in Figures 3.2 and 3.3 (see also Figure A3, Appendix II), such as the large error bar for shikimic acid in Lierne due to two samples deviating from the rest of the samples at this location. A number of other studies have, as in the present study, based their analyses on samples from only one or a few locations and only one sampling year (Martinussen et al., 2009; Määttä-Riihinen et al., 2004). Increasing the number of locations within years should not generally change conclusions, but will allow higher resolution of patterns, in the same way that increasing the numbers of years should increase the knowledge regarding the nature of variation in the patterns. Thus, when possible, it is always preferable to increase the number of samples per location and the number of sampling years in order to clarify any patterns and make estimates more precise, such that conclusions will be more robust.

4.2. Genetic variation

4.2.1. Microsatellite markers for genetic differentiation

Genetic diversity and estimation of the genetic relationship between individuals and populations may be explored by use of chloroplast or mitochondrial DNA sequences (Freeland, 2005; Renner and Zhang, 2004), random amplified polymorfic DNA markers (RAPD), expressed sequence tag-polymerase chain reaction (EST-PCR) (Albert et al., 2004; Bell et al., 2008), intersimple sequence repeat markers (ISSR) (Debnath, 2009) or microsatellites (Boches et al., 2005). Microsatellite loci are more variable than for example organelle DNA sequences, allozymes or dominant markers (Freeland, 2005; Todokoro et al., 1995). Thus, variable polymorphic microsatellite loci are suitable for revealing intraspecific evolutionary relationships (Bowcock et al., 1994; Todokoro et al., 1995). In addition, they are suitable for investigating the degree of genetic variability within and between locations of the same species at a regional scale. Therefore, microsatellites were chosen for the inter- and intrapopulation study within the present thesis.

Of the 16 microsatellites originally tested, only four were possible to utilize for the genetic differentiation analyses. The four microsatellites that were suitable for analysis in the present study did not correspond entirely with the ones found to be suitable for bilberry in other studies. None of the selected primers (except NA741) have previously been successfully tested in bilberry (Bassil et al., 2010), but all four have been applied in other *Vaccinium* species (Boches et al., 2006). However, several of the other 16 microsatellites which proved unsuitable for present study have been successfully used in previous studies with bilberry (Bassil et al., 2010). Thus, some of the unsuitable 12 microsatellite loci might have turned out to be suitable if it had been possible to put more time and effort into improving the quality of DNA, testing other PCR profiles, or altering salt and primer concentrations. The QIAGEN protocol for DNA extraction was however designed for plant tissue (DNeasy® Plant Handbook, QIAGEN, 2006) and the QIAGEN PCR protocol was designed to allow for sub-optimal design of microsatellite primers and high specificity and sensitivity without prior optimization (DNeasy® Plant Handbook, QIAGEN, 2006). Thus, it is likely that most of the unsuitable loci did not work either because they were developed for other

Vaccinium species or populations and have one allele fixed (i.e. are monomorphic) in Norwegian bilberry populations (Freeland, 2005), or because all individuals and populations have versions of the locus with mutation(s) in the primer sequences (Goldstein and Schlötterer, 1999). This is possible for instance if populations in Norway are genetically differentiated from the domestic blueberries from which most of the microsatellites utilized within the present study have been isolated (Bassil et al., 2010).

4.2.2. Genetic variation within and between populations

After estimating all pairwise F_{ST} values, the locations with the lowest level of genetic differentiation turned out to be Langvatnet and Hedmark, with a negative estimate of F_{ST} (Table 3.1). The two locations that had the highest pairwise level of genetic differentiation were Tromsø and Lierne, with an F_{ST} of 0.055. However, even though there are tendencies for genetic differentiation between the different locations, none of the F_{ST} values were significantly different from zero. Still, it is noticeable that Lierne is the location with the overall highest level of pairwise genetic differentiation with the other locations. Furthermore, it is surprising that Lierne and Langvatnet, which are close to each other geographically (i.e. only separated by approx. 50 km), also show a level of genetic differentiation which is higher than pairs of locations separated by longer geographic distances, such as Tromsø and Hedmark (see Figure 3.6). These results suggest that bilberries in Lierne might have a different origin than bilberries in the three other locations. Indeed, Lierne is the location closest to Norway's neighboring country, Sweden (Figure 2.1), and it may be possible that a reason for this pattern is that after the most recent glacial period the bilberry population in Lierne was founded by immigrants from Sweden (Wille, 1915). Then, the other populations could have been founded through a southern more coastal route such that the locations have a different genetic history (Wille, 1915). Accordingly, Areschoug 1866 (in Wille, 1915) claimed that species had two different migrational routes into Scandinavia and Sweden after the most recent glacial period; one from the south and one from the east. A similar pattern has also been suggested in North America by both Boches et al. (2006) and Camp (1945). Boches et al. (2006) found a grouping in their neighbor joining dendrogram of individuals that had similar origin, e.g.

proximity to the coast, and argued that one reason for morphological variation in highbush blueberry could be the costal versus inland origin of the populations. At the other end of the scale, Langvatnet and Hedmark, and Hedmark and Tromsø were found to have very low levels of pairwise genetic differentiation. Such a pattern was also shown by Boches et al. (2006), which found genetic similarities between their samples from northern and southern populations, and confirmed these clusters through both a dendrogram and PCA. With the above results in mind, it was not surprising that the Mantel's test performed in the present study revealed no significant relationship between the level of genetic differentiation and geographic distances between the locations (see Figure 3.6). To current knowledge this has not previously been examined in bilberry, but it has been explored in other plant species such as e.g. *Arabidopsis thaliana* (Bergelson et al., 1998; Sharbel et al., 2000; Todokoro et al., 1995). Most of the studies on *A. thaliana* have not found a relationship between the geographical distance between populations and the level of genetic differentiation, measured by nucleotide polymorphism in Bergelson et al. (1998) and by microsatellite polymorphism in Todokoro et al. (1995). However, Sharbel et al. (2000) did find significant isolation by distance for populations of *A. thaliana* from Eurasia and southern Europe.

Table A6 in Appendix III shows the number of alleles per locus (N_A), allelic richness (A_R), observed (H_O) and expected heterozygosities (H_E) for the different loci in the four locations. These population-level measures of genetic variability indicated that all locations were similar with respect to the intrapopulation levels of genetic variation, as no significant differences with regards to allelic richness or expected hetrozygosity in the different locations were found (see paragraph 3.2.2). The low genetic variation both between and within the bilberry populations investigated within present study have also been shown in other studies of bilberry (e.g. Albert et al., 2004). These patterns and the predominant lack of isolation by distance is often taken to indicate high gene flow even between populations that are, by our measures, very distant (Flower-Ellis, 1971) or is explained by a recent and rapid spread worldwide from a common evolutionary origin (Aguirre-Planter et al., 2000; Stenøien et al., 2005). Albert et al. (2004) argued that the patterns of the genetic variation in bilberry probably were due to the mating system, with low genetic variation between populations, due to seed dispersal by for instance birds and low genetic variation within

populations due to bilberries vegetative reproduction by horizontal rhizomes (Jacquemart et al., 1994). Based on results from present study, this explanation or that they have recently been colonized from a more or less common source population (but see discussion on Lierne above) can both be reasonable explanations.

4.2.3. Metabolite concentrations versus genetic variation

Variation in the concentrations of metabolites within or between populations may be controlled by genetic, environmental and the interaction of genetic and environmental factors (Harrigan et al., 2007; Skogerson et al., 2010). As environmental factors (e.g. sunlight, temperature or mineral nutrition) were not controlled within the present study, their effects can be confounding and cannot be addressed. This should also increase the amount of variation in the data which cannot be explained and could also make it harder to detect patterns between locations. Both Martinussen et al. (2009) and Åkerström et al. (2010) found samples to display their characteristic metabolite patterns despite being grown under similar environmental conditions. In particular, Åkerström et al. (2010) found that when different genotypes were grown at the same latitude, they nevertheless displayed the same anthocyanidin content as was associated with their parental environments. Thus, these studies indicate a strong contribution from inherent genetic factors in the quantitative composition of compounds. It is expected that differences in metabolite profiles may reflect genetic differences if populations have been sufficiently separated for some period of time and are locally adapted (Åkerström et al., 2010).

In the present study, Lierne and Langvatnet are metabolically very similar (see Figure 3.4 and 3.5), but genetically the non-significant trend was that they were quite differentiated (see Table 3.1). Furthermore, Hedmark and Langvatnet have different metabolite profiles but had by far the lowest F_{ST} value. Thus, in present study there seems to be little accordance between metabolic profiles and F_{ST} estimates. Accordingly, it seems that the variation in chemical compounds discovered among the populations could be environmentally induced differences on a more or less similar genetic background.

4.2.4. Genotyping difficulties

The number of microsatellite loci that was possible to use in the present study was low (i.e. only four). Bootstrap confidence intervals for variance components are generally not recommended for data sets with less than five loci (Raymond and Rousset, 1995a). Thus, the low number of microsatellites are not optimal to quantify the genetic relationships between populations, and may give estimates of both inter- and intrapopulation levels of genetic variation that can have high uncertainty (see the large confidence intervals in Table 3.1) and be in some cases biased (Sjögren and Wyöni, 1994). Consequently, the results based on the genetic analyses in the present study should only be interpreted as suggestive of actual patterns.

All loci were independent from each other when testing for linkage disequilibrium. Furthermore all loci, except CA421F in Lierne, were in Hardy–Weinberg equilibrium. There was a deficiency of heterozygotes in CA421F in Lierne, but if a Bonferroni correction of the level of significance was carried out (i.e. giving critical level of significance equal to 0.05/16 = 0.003) the deviance from Hardy-Weinberg equilibrium was not highly significant. Thus, although this slight deviance from Hardy-Weinberg equilibrium was for 25% of the loci in the most genetically differentiated population it was unlikely that this had major effects on the results (Table A6 in Appendix III). However, in the future, a more robust investigation could be performed by use of multiple loci to further unravel the details of the genetic patterns within and between populations. In addition, microsatellites are neutral genetic markers that do not necessarily reflect adaptive genetic variation (that is variation in genes that do affect the metabolite pattern) and it is still possible that the genetic differences causing compound concentrations to differ could be so minor as to remain undetected. Therefore, utilizing more microsatellite markers or naturally-variable genes that may influence plant metabolism may be an idea for future investigation. For instance, genetic studies have shown that accumulation of malic acid and citric acid is controlled by genes of related enzymes and that this varies between species and cultivars (Berüter, 2004; Etienne et al., 2002). Thus, in this way it would be possible to use genes that are known to be associated with metabolites in order to assess genetic diversity in terms of metabolic diversity. This could be done by use of e.g. quantitative trait locus (QTL) as described by Chan et al. (2010).

Also, new state-of-the-art molecular techniques, such as single nucleotide polymorphisms (SNPs) may also be suitable for such studies.

5. Cultivation in Norway

Based on results from present study there is little need to be very selective on which genotypes to cultivate, as there was low variation in genetics between the four locations, despite that they were fairly geographically separated. Thus, the breeding place or environment in which berries are grown may be more important than the selection of distinct clones. However, due to the limitations in the genetic analyses of the present study, the potential genetic differences between locations should be more thoroughly investigated as there was found significant differences between locations in several metabolic compounds. Knowledge of genetic diversity will promote the use of genetic variation in crop improvement and makes it possible to select genotypes that produce traits which are favorable. However, such production also needs to account for other traits, such as number and size of berries, which is another aspect of cultivation. This was however outside the scope of the present study to investigate, but is still an important trait to consider during cultivation.

6. Conclusion

In the present study, the metabolic profile and level of genetic variation within and across populations is presented from four locations of bilberries grown in Norway along a latitudinal gradient (North to South). Significant differences were found in concentrations of particular metabolites. Furthermore, multivariate statistics revealed a metabolic clustering of the locations in Mid-Norway, and a clustering of the most northern with the southern location. However, genetic analyses using four microsatellites suggested that all locations were very similar both with respect to genetic differentiation and intrapopulation genetic variation. This could be explained by them having similar recent origin and thus that the variation in biochemical compounds discovered among the locations could be environmentally induced differences on a similar genetic background. However, due to the low number of genetic markers utilized, the F_{ST} estimates should be interpreted with care. Also, it is still possible that the genetic differences causing compound concentrations to differ could be so minor as to remain undetected in present study, thus further studies are warranted.

7. References

Aguirre-Planter E., Furnier G.R., Eguiarte L.E. (2000) Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. American Journal of Botany 87:362-371.

Albert T., Raspe O., Jacquemart A.L. (2004) Clonal diversity and genetic structure in *Vaccinium myrtillus* populations from different habitats. Belgian Journal of Botany 137:155-162.

Atlegrim O., Sjøberg K. (1996) Response of bilberry (*Vaccinium myrtillus*) to clear-cutting and single-tree selection harvests in uneven-aged boreal *Picea abies* forests. Forest Ecology and Management 86:39-50.

Ayaz F.A., Kadioglu A., Acar C., Turna I. (2001) Effect of fruit maturation on sugar and organic acid composition in two blueberries (*Vaccinium arctostaphylos* and *V. myrtillus*) native to Turkey. Journal of Crop and Horticultural Science 29:137-141.

Bartish I.V., Jeppsson N., Nybom H. (1999) Population genetic structure in the dioecious pioneer plant species *Hippophae rhamnoides* investigated by random amplified polymorphic DNA (RAPD) markers. Molecular Ecology 8:791-802.

Bassil N., Bunch A., Nyberg A., Hummer K., Zee F.T. (2010) Microsatellite markers distinguish Hawaiian ohelo from other *Vaccinium* L. section *myrtillus* species. Acta Horticulturae 859:81-88.

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

Benzie I.F., Strain J.J. (1996) The ferric reducing ability of plasma (FRAP) as a measure of ''antioxidant power'': The FRAP assay. Analytical Biochemistry 239:70-76.

Bergelson J., Stahl E., Dudek S., Kreitman M. (1998) Genetic variation within and among populations of *Arabidopsis thaliana*. Genetics 148:1311-1323.

Berüter J. (2004) Carbohydrate metabolism in two apple genotypes that differ in malate accumulation. Journal of Plant Physiology 161:1011-1029.

Blumthaler M., Ambach W., Ellinger R. (1997) Increase in solar UV radiation with altitude. Journal of Photochemistry and Photobiology B-Biology 39:130-134.

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

Boches P.S., Bassil N.V., Rowland L.J. (2005) Microsatellite markers for *Vaccinium* from EST and genomic libraries. Molecular Ecology Notes 5:657-660.

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

Bowcock A.M., Ruizlinares A., Tomfohrde J., Minch E., Kidd J.R., Cavallisforza L.L. (1994) High-resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455-457.

Bravo L. (1998) Polyphenols: Chemistry, dietary sources, metabolism, and nutritional significance. Nutrition Reviews 56:317-333.

Buchanan B., Jones R., Gruissem W. (2000) Biochemistry & molecular biology of plants. American Society of Plant Physiologists, Rockville, Maryland. pp. 630-675.

Burdulis D., Ivanauskas L., Dirse V., Kazlauskas S., Razukas A. (2007) Study of diversity of anthocyanin composition in bilberry (*Vaccinium myrtillus* L.) fruits. Medicina-Lithuania 43:971-977.

Burdulis D., Sarkinas A., Jasutiene I., Stackeviciene E., Nikolajevas L., Janulis V. (2009) Comparative study of anthocyanin composition, antimicrobial and antioxidant activity in bilberry (*Vaccinium myrtillus* L.) and blueberry (*Vaccinium corymbosum* L.) fruits. Acta Poloniae Pharmaceutica 66:399-408.

Cacace J.E., Mazza G. (2003) Optimization of extraction of anthocyanins from black currants with aqueous ethanol. Journal of Food Science 68:240-248.

Camire M.E. (2002) Berry beneficial. Dairy Industries International 67:15-16.

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

Chan E.F., Rowe H.C., Hansen B.G., Kliebenstein D.J. (2010) The complex genetic architecture of the metabolome. Plos Genetics 6:1-15.

Connor A.M., Luby J.J., Tong C.B.S., Finn C.E., Hancock J.F. (2002a) Genotypic and environmental variation in antioxidant activity, total phenolic content, and anthocyanin content among blueberry cultivars. Journal of the American Society for Horticultural Science 127:89-97.

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

Debnath S.C. (2009) Development of ISSR markers for genetic diversity studies in *Vaccinium angustifolium*. Nordic Journal of Botany 27:141-148.

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

Etienne C., Moing A., Dirlewanger E., Raymond P., Monet R., Rothan C. (2002) Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity. Physiologia Plantarum 114:259-270.

Flower-Ellis J.K. (1971) Age structure and dynamics in stands of bilberry (*Vaccinium myrtillus* L.). Department of Forest Ecology and Forest Soils, Royal College of Forestry, PhD Thesis.

Formica J., Regelson W. (1995) Review of the biology of quercetin and related bioflavonoids. Food and Chemical Toxicology 33:1061-1080.

Freeland J. (2005) Molecular ecology, Wiley, Chichester. pp. 63-107.

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

Giusti M.M., Wrolstad R.E. (2001) Characterization and measurement of anthocyanins by UV-visible spectroscopy, in: R. E. Wrolstad (Ed.), Current Protocols in Food Analytical Chemistry, John Wiley and Sons, New York. pp. 1.2.1-1.2.13.

Goldstein D.B., Schlötterer C. (1999) Microsatellites evolution and applications. Oxford University Press, Oxford. pp. 2-9.

Goudet J. (1995) FSTAT version 2.9.3.2. A computer software to calculate F-statistics. Journal of Heredity 86:485-486.

Goudet J. (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184-186.

Gross K.C., Sams C.E. (1984) Changes in cell wall neutral sugar composition during fruit ripening - a species survey. Phytochemistry 23:2457-2461.

Halvorsen B.L., Holte K., Myhrstad M.C.W., Barikmo I., Hvattum E., Remberg S.F., Wold A.B., Haffner K., Baugerod H., Andersen L.F., Moskaug J.O., Jacobs D.R., Blomhoff R. (2002) A systematic screening of total antioxidants in dietary plants. Journal of Nutrition 132:461-471.

Harrigan G.G., Stork L.G., Riordan S.G., Reynolds T.L., Ridley W.P., Masucci J.D., MacIsaac S., Halls S.C., Orth R., Smith R.G., Wen L., Brown W.E., Welsch M., Riley R., Mcfarland D., Pandravada A., Glenn K.C. (2007) Impact of genetics and environment on nutritional and metabolite components of maize grain. Journal of Agricultural and Food Chemistry 55:6177-6185.

Havaux M., Kloppstech K. (2001) The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in *Arabidopsis* npq and tt mutants. Planta 213:953-966.

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

Jaakola L., Määttä-Riihinen K., Kärenlampi S., Hohtola A. (2004) Activation of flavonoid biosynthesis by solar radiation in bilberry (*Vaccinium myrtillus* L.) leaves. Planta 218:721-728.

Jaakola L., Määttä K., Pirttilä A.M., Törrönen R., Kärenlampi S., Hohtola A. (2002) Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. Plant Physiology 130:729-739.

Jacquemart A., Mahy G., Raspe O., Desloover J. (1994) An isozyme study in bilberry (*Vaccinium myrtillus*). II: Mating system and genetic structure. Belgian Journal of Botany 127:105-114.

Jurd L., Asen S. (1966) Formation of metal and co-pigment complexes of cyanidin 3-glucoside. Phytochemistry 5:1263-1271.

Kader F., Rovel B., Metche M. (1993) Role of invertase in sugar content in highbush blueberries (*Vaccinium corymbosum* L.). Food Science and Technology-Lebensmittel-Wissenschaft and Technologie 26:593-595.

Kader F., Rovel B., Girardin M., Metche M. (1994) Composition of the fruit of highbush blueberry (*Vaccinium corymbosum* L.) grown in the East of France (Vosges). Sciences Des Aliments 14:281-290.

Li J.Y., Oulee T.M., Raba R., Amundson R.G., Last R.L. (1993) *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell 5:171-179.

Lois R. (1994) Accumulation of UV-absorbing flavonoids induced by UV-B radiation in *Arabidopsis thaliana* L. Planta 194:498-503.

Lätti A.K., Riihinen K.R., Kainulainen P.S. (2008) Analysis of anthocyanin variation in wild populations of bilberry (*Vaccinium myrtillus* L.) in Finland. Journal of Agriculture and Food Chemistry 56:190-196.

Lätti A.K., Jaakola L., Riihinen K.R., Kainulainen P.S. (2010) Anthocyanin and flavonol variation in bog bilberries (*Vaccinium uliginosum* L.) in Finland. Journal of Agricultural and Food Chemistry 58:427-433.

Mantel N. (1967) Detection of disease clustering and a generalized regression approach. Cancer Research 27:209-220.

Martinelli E.M., Baj A., Bombardelli E. (1986) Computer-aided evaluation of liquid-chromatographic profiles for anthocyanins in *Vaccinium myrtillus* fruits. Analytica Chimica Acta 191:275-281.

Martinussen I., Rohloff J., Uleberg E., Junttila O., Hohtola A., Jaakola I., Häggman H. (2009) Climatic effects on the production and quality of bilberries (*Vaccinium myrtillus*). Latvian Journal of Agronomy 12: 71-74.

Martz F., Jaakola L., Julkunen-Tiitto R., Stark S. (2010) Phenolic composition and antioxidant capacity of bilberry (*Vaccinium myrtillus*) leaves in northern Europe following foliar development and along environmental gradients. Journal of Chemical Ecology 36:1017-1028.

Mochida K., Furuta T., Ebana K., Shinozaki K., Kikuchi J. (2009) Correlation exploration of metabolic and genomic diversity in rice. Bmc Genomics 10:1-10.

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 50:519-525.

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

Nohynek L.J., Alakomi H.L., Kähkönen M.P., Heinonen M., Helander K.M., Oksman-Caldentey K.M., Puupponen-Pimiä R.H. (2006) Berry phenolics: Antimicrobial properties and mechanisms of action against severe human pathogens. Nutrition and Cancer 54:18-32.

Ochmian I., Oszmianski J., Skupien K. (2009) Chemical composition, phenolics, and firmness of small black fruits. Journal of Applied Botany and Food Quality 83:64-69.

Paassilta M., Moisio S., Jaakola L., Häggman H. (2009) Voice of the Nordic wild berry industry - a survey among the companies. Oulu University Press, Oulu.

Pemberton J.M., Slate J., Bancroft D.R., Barrett J.A. (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology 4:249-252.

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

Pop M., Lupea A.X., Turcus V. (2008) Studies on the pholyphenolics compounds extraction from *Vaccinium* fruits. Revista De Chimie 59:491-494.

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

Puupponen-Pimiä R., Nohynek L., Alakomi H.L., Oksman-Caldentey K.M. (2005) Bioactive berry compounds - novel tools against human pathogens. Applied Microbiology and Biotechnology 67:8-18.

Puupponen-Pimiä R., Nohynek L., Meier C., Kähkonen M., Heinonen M., Hopia A., Oksman-Caldentey K.M. (2001) Antimicrobial properties of phenolic compounds from berries. Journal of Applied Microbiology 90:494-507.

R Development Core Team. (2010) R: Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Raymond M., Rousset F. (1995a) An exact test for population differentiation. Evolution 49:1280-1283.

Raymond M., Rousset F. (1995b) Genepop (version 1.2) - Population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.

Renner S., Zhang L.B. (2004) Biogeography of the *Pistia* clade (Araceae): Based on chloroplast and mitochondrial DNA sequences and the Bayesian divergence time inference. Systematic Biology 53:422-432.

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

Rieger G., Müller M., Guttenberger H., Bucar F. (2008) Influence of altitudinal variation on the content of phenolic compounds in wild populations of *Calluna vulgaris*, *Sambucus nigra*, and *Vaccinium myrtillus*. Journal of Agricultural and Food Chemistry 56:9080-9086.

Ritchie J.C. (1956) *Vaccinium myrtillus* L. Journal of Ecology 44:291-299.

Rousset F. (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219-1228.

Saeed A.I., Sharov V., White J., Li J., Liang W., Bhagabati N., Braisted J., Klapa M., Currier T., Thiagarajan M., Sturn A., Snuffin M., Rezantsev A., Popov D., Ryltsov A., Kostukovich E., Borisovsky I., Liu Z., Vinsavich A., Trush V., Quackenbush J. (2003) TM4: A free, open-source system for microarray data management and analysis. Biotechniques 34:374-378.

Serafini M., Maiani G., Ferro-Luzzi A. (1998) Alcohol-free red wine enhances plasma antioxidant capacity in humans. Journal of Nutrition 128:1003-1007.

Sharbel T.F., Haubold B., Mitchell-Olds T. (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Molecular Ecology 9:2109-2118.

Singleton V.L. (1965) Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. American Journal of Enology and Vitriculture 16:144-158.

Sjögren P., Wyöni P.I. (1994) Conservation genetics and detection of rare alleles in finite populations. Conservation Biology 8:267-270.

Skogerson K., Harrigan G.G., Reynolds T.L., Halls S.C., Ruebelt M., Iandolino A., Pandravada A., Glenn K.C., Fiehn O. (2010) Impact of genetics and environment on the metabolite composition of maize grain. Journal of Agricultural and Food Chemistry 58:3600-3610.

Stenøien H.K., Fenster C.B., Tonteri A., Savolainen O. (2005) Genetic variability in natural populations of *Arabidopsis thaliana* in Northern Europe. Molecular Ecology 14:137-148.

Sullivan J.T., Phillips T.G., Routley D.G. (1960) Water-soluble hemicelluloses of grass holocellulose. Journal of Agricultural and Food Chemistry 8:152-153.

Sumner L.W., Mendes P., Dixon R.A. (2003) Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry 62:817-836.

Taiz L., Zeiger E. (2006) Plant physiology. 4th ed. Sinauer Associates, Sunderland, Massachusetts. pp. 316-344 and 672-705.

Taulavuori E., Bäckman M., Taulavuori K., Gwynn-Jones D., Johanson U., Laine K., Callaghan T., Sonesson M., Björn L.O. (1998) Long-term exposure to enhanced ultraviolet-B radiation in the sub-arctic does not cause oxidative stress in *Vaccinium myrtillus*. New Phytologist 140:691-697.

Todokoro S., Terauchi R., Kawano S. (1995) Microsatellite polymorphisms in natural populations of *Arabidopsis thaliana* in Japan. The Japanese Journal of Genetics 70:543-554.

Viljakainen S., Visti A., Laakso S. (2002) Concentrations of organic acids and soluble sugars in juices from nordic berries. Acta Agriculturae Scandinavica Section B-Soil and Plant Science 52:101-109.

Wille N. (1915) The flora of Norway and its immigration. Annals of the Missouri Botanical Garden 2:59-108.

Wrolstad R.E., Durst R.W., Lee J. (2005) Tracking color and pigment changes in anthocyanin products. Trends in Food Science and Technology 16:423-428.

Zhang J.J., Wang X., Yu O., Tang J.J., Gu X.G., Wan X.C., Fang C.B. (2011) Metabolic profiling of strawberry (*Fragaria x ananassa* Duch.) during fruit development and maturation. Journal of Experimental Botany 62:1103-1118.

Zheng J., Kallio H., Yang B. (2009a) Effects of latitude and weather conditions on sugars, fruit acids and ascorbic acid in currant (*Ribes* sp.) cultivars. Journal of the Science of Food and Agriculture 89:2011-2023.

Zheng J., Yang B., Tuomasjukka S., Ou S., Kallio H. (2009b) Effects of latitude and weather conditions on contents of sugars, fruit acids, and ascorbic acid in black currant (*Ribes nigrum* L.) juice. Journal of Agricultural and Food Chemistry 57:2977-2987.

Åkerström A., Jaakola L., Bång U., Jäderlund A. (2010) Effects of latitude-related factors and geographical origin on anthocyanidin concentrations in fruits of *Vaccinium myrtillus* L. (bilberries). Journal of Agricultural and Food Chemistry 58:11939-11945.

Appendix I: Standard curves for TPH and FRAP calculations

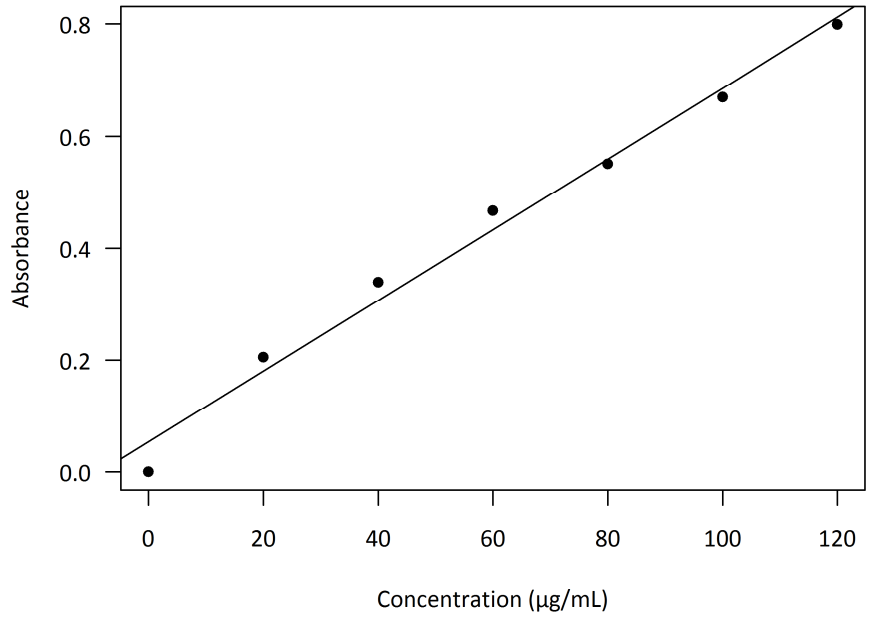

Figure A1: Standard curve of gallic acid (µg/mL) used for calculation of total phenolics (TPH) concentration. The linear regression line is given by, y = 0.0063x + 0.0539 with R^2 = 0.98. Measurements were performed at 595 nm.

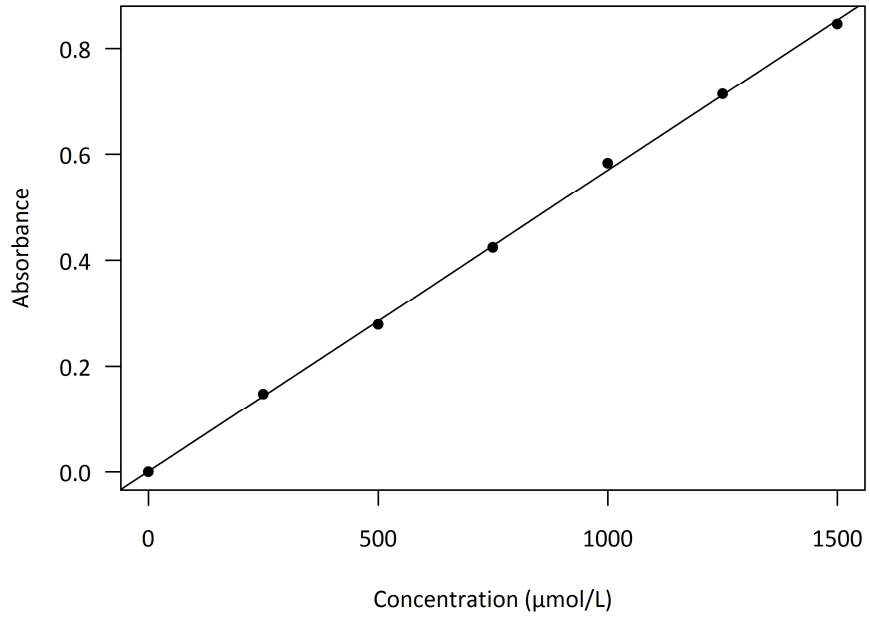

Figure A2: Standard curve of iron (II) sulfate heptahydrate (µmol/L) used for calculations of antioxidant activity (FRAP) concentrations. The linear regression line is given by, y = 0.0006x + 0.0016 with R^2 = 0.99. Measurements were performed at 750 nm.

Appendix II: Concentration of TPH, ACY, FRAP, sugars, acids and simple phenolics

Table A1: Total phenolics (TPH), total anthocyanins (ACY), the ratio between ACY and TPH, and antioxidant activity (FRAP) of the four different locations Tromsø (T), Lierne (LI), Langvatnet (LI) and Hedmark (H). Mean and standard error (SE) shown for individual samples of TPH and FRAP is based on three parallels. Grand mean concentrations (based on fresh weight) for each location are shown together with SE between the different concentrations of the individual samples.

	TPH (mg/100 g ± SE)	**ACY (mg/100 g)**	**ACY/TPH**	**FRAP (mmol/100 g ± SE)**
T1	593.706 ± 0.017	230.958	0.389	5.814 ± 0.014
T2	506.057 ± 0.003	143.797	0.284	4.646 ± 0.024
T3	577.847 ± 0.009	148.594	0.257	5.250 ± 0.021
T4	544.225 ± 0.016	111.406	0.205	5.256 ± 0.011
T5	492.982 ± 0.009	104.514	0.212	4.690 ± 0.006
T6	722.897 ± 0.015	244.800	0.339	6.691 ± 0.005
T7	431.217 ± 0.005	157.493	0.365	4.549 ± 0.010
MEAN	**552.704 ± 35.137**	**163.080 ± 20.701**	**0.293 ± 0.028**	**5.271 ± 0.291**
LI1	538.859 ± 0.006	123.285	0.229	4.982 ± 0.005
LI2	452.618 ± 0.025	219.393	0.485	3.661 ± 0.004
LI3	370.324 ± 0.013	174.425	0.471	3.126 ± 0.001
LI4	417.191 ± 0.008	162.832	0.390	4.085 ± 0.002
LI5	599.843 ± 0.010	157.160	0.262	5.759 ± 0.007
LI6	603.209 ± 0.017	107.282	0.178	5.557 ± 0.012
LI7	398.320 ± 0.019	134.435	0.337	3.616 ± 0.014
LI8	453.059 ± 0.015	96.586	0.213	3.617 ± 0.011
LI9	517.249 ± 0.066	113.574	0.220	3.459 ± 0.006
LI10	488.489 ± 0.012	81.416	0.167	3.946 ± 0.002
MEAN	**483.916 ± 25.480**	**137.039 ± 13.171**	**0.295 ± 0.037**	**4.181 ± 0.291**
LV1	684.259 ± 0.009	96.101	0.140	6.410 ± 0.003
LV2	613.855 ± 0.008	209.829	0.342	6.187 ± 0.005
LV3	458.556 ± 0.014	166.178	0.362	4.045 ± 0.003
LV4	515.964 ± 0.013	230.330	0.446	4.803 ± 0.021
LV5	424.314 ± 0.015	148.797	0.351	3.964 ± 0.011
LV6	536.346 ± 0.010	145.407	0.271	5.138 ± 0.013
LV7	335.002 ± 0.003	110.421	0.330	3.555 ± 0.008
LV8	415.766 ± 0.009	107.389	0.258	3.839 ± 0.005
LV9	728.564 ± 0.030	78.527	0.108	7.238 ± 0.030
LV10	641.051 ± 0.033	113.085	0.176	5.364 ± 0.004
MEAN	**535.368 ± 40.809**	**140.606 ± 15.703**	**0.278 ± 0.034**	**5.054 ± 0.394**
H1	733.085 ± 0.007	159.252	0.217	7.686 ± 0.011
H2	477.394 ± 0.008	203.965	0.427	4.589 ± 0.006
H3	594.553 ± 0.004	67.496	0.114	5.824 ± 0.009
H4	402.064 ± 0.011	131.748	0.328	3.455 ± 0.006
H5	355.964 ± 0.008	103.563	0.291	3.279 ± 0.003
H6	580.392 ± 0.018	112.258	0.193	5.752 ± 0.007
H7	543.057 ± 0.013	156.889	0.289	5.546 ± 0.029
H8	413.984 ± 0.007	158.730	0.383	4.352 ± 0.013
H9	620.663 ± 0.010	108.163	0.174	7.392 ± 0.011
H10	540.987 ± 0.011	133.562	0.247	4.858 ± 0.007
MEAN	**526.214 ± 36.393**	**133.563 ± 12.093**	**0.266 ± 0.031**	**5.273 ± 0.468**

Table A2: Concentrations of metabolites from GC-MS analysis (µg/100 g fresh weight) with mean and standard error (SE) for location Tromsø (T). In NA monosaccharide and NA trisaccharide, NA represent not annotated.

Compound	T1	T2	T3	T4	T5	T6	T7	*MEAN* ± SE
Fructose	27150.11	29550.21	30029.37	27650.59	19021.54	21494.50	17380.28	**24610.94 ± 1968.07**
Glucose	15042.94	15148.71	15289.05	14963.74	10748.65	12279.75	10561.47	**13433.47 ± 818.03**
Sucrose	497.86	878.28	407.16	654.63	710.37	960.85	793.02	**700.31 ± 75.16**
Arabinose	302.59	255.13	267.84	239.09	282.98	174.79	157.41	**239.97 ± 20.61**
Ribose	22.52	18.49	19.09	18.90	56.32	19.30	30.93	**26.51 ± 5.24**
Xylose	3.09	3.50	3.61	3.40	37.92	8.00	1.70	**8.75 ± 4.92**
Citric acid	3562.95	2207.73	1805.88	2205.55	4181.66	3718.40	3587.01	**3038.46 ± 353.50**
Malic acid	764.42	961.42	943.98	1128.42	693.29	904.95	843.32	**891.40 ± 53.83**
Fumaric acid	3.02	3.09	3.05	5.72	4.10	3.38	11.91	**4.90 ± 1.23**
Quinic acid	4361.76	3494.91	2908.43	3059.43	5147.49	6578.62	4744.97	**4327.94 ± 493.54**
Chlorogenic acid	274.96	158.43	271.92	260.82	121.29	340.38	272.25	**242.86 ± 28.65**
Shikimic acid	74.41	93.72	54.42	67.32	58.30	110.40	44.96	**71.93 ± 8.73**
Myo-inositol	448.30	316.98	324.44	306.61	477.96	742.53	538.00	**450.69 ± 59.39**
Galactinol	270.40	477.34	404.35	443.82	138.92	287.92	254.75	**325.36 ± 45.65**
Chiro-inositol	28.18	25.33	27.81	30.78	20.43	27.75	24.54	**26.40 ± 1.26**
NA trisaccharide	888.71	431.38	432.87	444.05	568.10	184.60	863.42	**544.73 ± 95.82**
NA monosaccharide	346.76	285.59	300.43	320.16	297.25	267.12	281.85	**299.88 ± 10.02**

Table A3: Concentrations of metabolites from GC-MS analysis (µg/100 g fresh weight) with mean and standard error (SE) for location Lierne (LI). In NA monosaccharide and trisaccharide, NA represent not annotated.

Compound	LI1	LI2	LI3	LI4	LI5	LI6	LI7	LI8	LI9	LI10	*MEAN* ± SE
Fructose	38209.17	27760.87	34508.63	32059.57	19136.75	24511.28	29168.43	31346.61	17294.11	32495.24	**28649.07 ± 2100.34**
Glucose	14776.09	16412.28	18019.11	16356.37	11448.45	14290.82	15550.57	15629.66	9648.23	17925.69	**15005.73 ± 843.38**
Sucrose	2819.95	2444.29	4396.64	2685.80	1738.49	1710.49	2110.51	2594.84	1839.67	3062.41	**2540.31 ± 254.67**
Arabinose	204.42	209.73	163.28	193.38	136.62	268.85	200.85	241.38	235.98	204.45	**205.89 ± 11.99**
Ribose	24.35	105.88	27.39	26.10	25.03	31.27	22.93	23.81	80.57	24.99	**39.23 ± 9.22**
Xylose	10.92	1.83	7.95	3.09	2.06	3.51	2.49	3.45	53.06	8.42	**9.68 ± 4.92**
Citric acid	3621.28	11.38	2577.68	2918.20	2402.98	2626.61	2707.78	1812.89	2619.95	2415.17	**2371.39 ± 298.67**
Malic acid	1406.98	42.07	1341.38	1108.72	1533.49	861.11	984.77	1310.68	1353.17	1140.20	**1108.26 ± 134.97**
Fumaric acid	4.63	7.48	1.14	2.30	2.04	2.37	2.89	5.15	4.66	3.81	**3.65 ± 0.59**
Quinic acid	5966.09	2670.56	7014.01	6192.39	4933.10	5365.66	4193.08	4894.61	2606.53	5182.31	**4901.83 ± 450.97**
Chlorogenic acid	176.48	21.35	158.53	146.75	406.66	122.07	45.68	280.89	63.29	221.43	**164.31 ± 36.90**
Shikimic acid	533.04	58.15	135.60	72.89	68.04	48.39	45.60	107.21	102.86	102.52	**127.43 ± 46.00**
Myo-inositol	390.25	473.88	618.01	407.44	463.95	505.37	265.15	408.81	142.67	491.16	**416.67 ± 41.94**
Galactinol	150.85	340.76	318.88	315.41	188.76	317.92	396.13	244.54	22.18	268.19	**256.36 ± 34.75**
Chiro-inositol	13.49	29.79	23.37	24.24	22.12	30.99	30.23	36.16	4.38	45.07	**25.98 ± 3.61**
NA trisaccharide	98.57	288.81	815.63	668.88	717.15	1037.20	641.92	682.33	7.41	434.25	**539.22 ± 102.99**
NA monosaccharide	3.36	8.80	91.75	219.82	303.43	307.45	271.03	259.19	2.16	175.14	**164.21 ± 40.15**

Table A4: Concentrations of metabolites from GC-MS analysis (µg/100 g fresh weight) with mean and standard error (SE) for location Langvatnet (LV). In NA monosaccharide and trisaccharide, NA represent not annotated. LV9 was excluded from analyses as the GC-MS results were unsatisfying.

Compound	**LV1**	**LV2**	**LV3**	**LV4**	**LV5**	**LV6**	**LV7**	**LV8**	**LV10**	***MEAN* ± SE**
Fructose	39610.51	28069.86	36788.34	35560.67	31218.12	38344.90	34943.46	42234.15	35290.39	**35784.49 ± 1347.07**
Glucose	18728.85	16823.58	19334.09	19129.37	15743.51	20867.79	18152.05	20457.38	19249.95	**18720.73 ± 515.21**
Sucrose	3449.83	2615.75	3304.50	3071.21	2123.16	3104.87	3290.60	4036.94	3109.14	**3122.89 ± 168.20**
Arabinose	186.84	188.15	226.53	206.62	174.84	185.36	161.02	262.88	224.37	**201.85 ± 9.97**
Ribose	31.46	26.05	23.56	21.59	16.44	17.07	21.01	124.02	23.01	**33.80 ± 10.79**
Xylose	7.61	2.47	3.47	6.74	5.47	4.73	4.39	24.52	7.28	**7.41 ± 2.10**
Citric acid	3488.04	3032.31	2595.28	2576.40	1919.82	2556.91	2068.42	3075.52	2825.80	**2682.05 ± 155.71**
Malic acid	732.03	848.38	736.35	693.22	751.78	912.13	523.29	644.62	581.44	**713.70 ± 38.53**
Fumaric acid	3.00	1.45	1.69	3.07	4.49	2.31	4.40	2.74	1.84	**2.78 ± 0.35**
Quinic acid	5294.19	3848.11	4354.18	4305.53	4524.26	5355.33	3592.87	4874.53	5648.27	**4644.14 ± 221.97**
Chlorogenic acid	81.24	117.74	168.01	134.36	149.33	198.29	137.66	84.54	182.97	**139.35 ± 12.83**
Shikimic acid	54.79	18.09	39.38	39.84	37.07	25.30	69.98	136.52	101.42	**58.04 ± 12.24**
Myo-inositol	524.43	435.94	438.26	328.92	373.07	479.29	377.10	445.55	440.35	**426.99 ± 18.70**
Galactinol	407.60	376.19	470.03	366.50	362.84	412.12	446.47	162.80	462.67	**385.25 ± 29.28**
Chiro-inositol	31.10	30.32	57.46	53.75	40.02	39.84	13.35	32.33	28.70	**36.32 ± 4.25**
NA trisaccharide	772.37	322.39	634.11	560.85	364.66	416.69	870.08	309.55	1140.09	**598.98 ± 89.96**
NA monosaccharide	31.10	311.43	155.34	118.12	205.36	12.42	267.69	4.00	9.12	**123.84 ± 37.43**

Table A5: Concentrations of metabolites from GC-MS analysis (µg/100 g fresh weight) with mean and standard error (SE) for location Hedmark (H). In NA monosaccharide and trisaccharide, NA represent not annotated.

Compound	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	*MEAN* ± SE
Fructose	17613.72	18932.66	22903.94	28612.13	27509.22	27247.60	24582.58	35731.18	20962.31	34973.68	**25906.90 ± 1952.33**
Glucose	11703.66	12738.11	11942.26	14249.67	14985.14	15074.22	15451.95	18951.33	13623.43	16765.40	**14548.52 ± 704.67**
Sucrose	204.85	197.05	372.93	232.42	539.62	229.69	304.37	1488.26	266.33	212.47	**404.80 ± 124.85**
Arabinose	192.23	220.99	212.79	191.13	305.73	265.22	295.16	218.61	206.26	210.71	**231.88 ± 13.15**
Ribose	20.10	20.90	27.95	16.94	21.73	19.18	20.08	138.67	21.05	20.87	**32.75 ± 11.80**
Xylose	7.44	3.13	9.72	3.34	3.24	2.66	2.45	4.79	8.01	3.47	**4.82 ± 0.82**
Citric acid	3052.92	2827.70	2634.52	1674.53	3438.03	1680.13	2319.09	49.81	3130.39	2674.51	**2348.16 ± 314.43**
Malic acid	849.92	1000.01	940.66	964.36	905.00	1089.67	914.58	24.57	918.69	794.10	**840.16 ± 94.08**
Fumaric acid	4.10	3.57	2.61	3.21	3.71	2.78	2.56	8.84	1.83	7.50	**4.07 ± 0.72**
Quinic acid	5054.26	3411.10	4173.20	4116.47	4376.54	3448.36	3960.23	1840.83	4309.97	3996.80	**3868.78 ± 269.47**
Chlorogenic acid	35.09	15.30	30.95	1.73	47.98	58.52	68.25	31.14	63.27	83.16	**43.54 ± 7.99**
Shikimic acid	60.50	31.51	59.26	68.90	155.19	52.26	52.26	43.94	41.08	47.31	**61.22 ± 10.97**
Myo-inositol	552.31	369.87	631.57	466.54	436.16	310.63	507.12	831.85	654.81	531.49	**529.23 ± 47.68**
Galactinol	29.99	196.63	109.03	259.73	255.18	287.99	223.15	159.57	255.74	72.01	**184.90 ± 28.07**
Chiro-inositol	50.60	2.76	16.16	30.44	38.87	26.50	36.33	39.33	29.07	30.09	**30.02 ± 4.19**
NA trisaccharide	180.52	39.77	497.98	392.06	940.44	536.90	654.63	242.37	482.30	783.30	**475.03 ± 87.27**
NA monosaccharide	165.71	264.28	285.13	256.00	314.22	272.18	376.77	6.48	329.58	240.79	**251.11 ± 32.48**

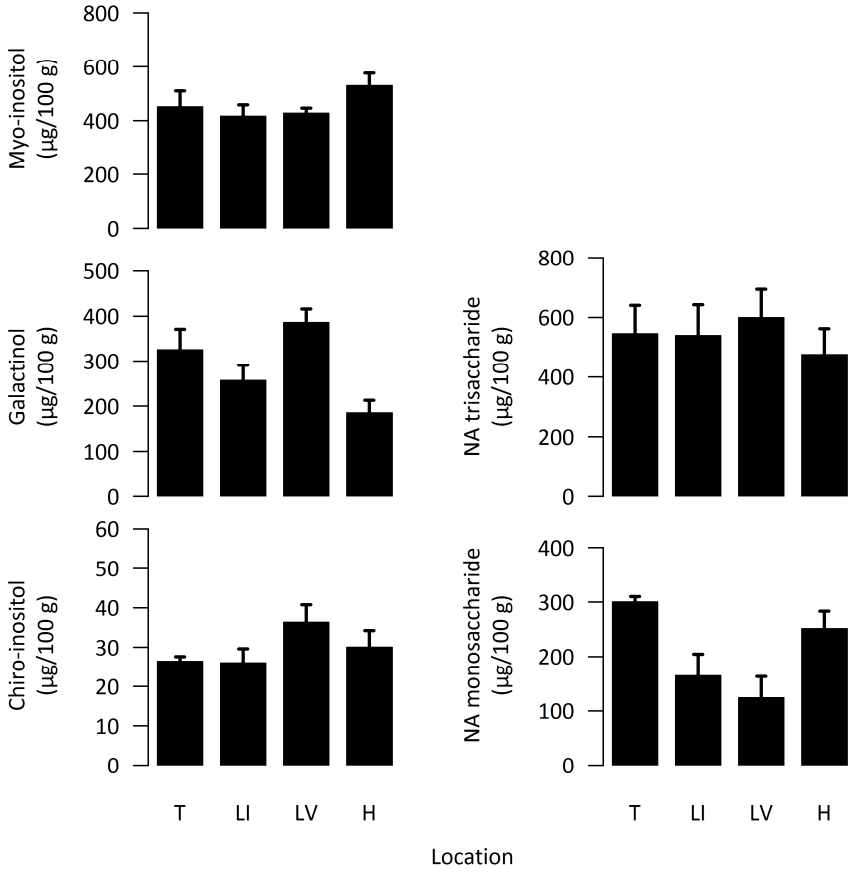

Figure A3: Barplots of the concentrations (µg/100 g fresh weight) of selected compounds (myo-inositol, galactinol, chiro-inositol, not annotated (NA) trisaccharide and not annotated (NA) monosaccharide) for the four different locations Tromsø (T), Lierne (LI), Langvatnet (LV) and Hedmark (H). Means for the four locations ± standard errors (SE) are given.

Appendix III: Descriptive statistics for genetic variation analyses

Table A6: Descriptive statistics for four microsatellites utilized within present study. Degrees of freedom = 8.

				Heterozygosity		Hardy-Weinberg equilibrium		
				Observed	Expected			
Location	**Locus**	**N_A**	**N**	**(H_O)**	**(H_E)**	**χ^2**	***P* value**	**A_R**
Tromsø	NA741	2	7	0.285	0.264	0.000	1.000	1.989
	NA961	2	7	0.571	0.440	4.039	1.000	2.000
	CA421F	8	7	0.714	0.793	21.319	0.505	8.000
	CA483F	3	7	0.571	0.560	4.253	1.000	2.989
Lierne	NA741	2	10	0.500	0.395	0.000	1.000	1.996
	NA961	2	10	0.300	0.453	4.039	0.494	2.000
	CA421F	5	10	0.300	0.688	21.319	0.003*	4.328
	CA483F	3	10	0.700	0.616	4.253	0.636	2.950
Langvatnet	NA741	2	10	0.200	0.189	0.000	1.000	1.853
	NA961	2	10	0.400	0.526	4.039	0.562	2.000
	CA421F	7	10	0.600	0.712	21.319	0.201	5.941
	CA483F	5	10	0.600	0.700	4.253	0.290	4.302
Hedmark	NA741	2	10	0.200	0.189	0.000	1.000	1.853
	NA961	2	10	0.300	0.479	4.039	0.478	2.000
	CA421F	5	10	0.500	0.721	21.319	0.094	4.305
	CA483F	4	10	0.700	0.689	4.253	0.645	3.705

Number of alleles (N_A), number of alleles per locus (N), number of individuals typed (N), observed and expected heterozygosities (H_O and H_E), deviance from Hardy-Weinberg equilibrium significance indicated with a star * (Bonferroni corrected 5% level of significance: $P < 0.003$). Allelic Richness (A_R) per locus and population based on minimum sample size of 6 diploid individuals.