Influential Factors in Nectar Composition and Yield in *Leptospermum scoparium*

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Plant Science

Institute of Agriculture and the Environment
College of Sciences, Massey University
Palmerston North, New Zealand

Elizabeth Mary Nickless

2015

ABSTRACT

Leptospermum scoparium (Mānuka) is the plant nectar source for medically bioactive honey, commercially marketed in New Zealand as Unique Mānuka Factor honey (UMF-honey). Methylglyoxal (MGO) is the unique bioactive component of UMF honey with Mānuka nectar containing significant amounts of the carbohydrate dihydroxyacetone (DHA), the chemical precursor for MGO. Anecdotal evidence and recently published data from nectar samples collected from various cultivars in natural sites or botanical gardens has indicated that the DHA and overall composition of *L. scoparium* nectar varies according to cultivar. The source of this variation is not clearly understood and although there is considerable literature on climatic and genetic influences on nectar composition and yield within various other plant species, there is little published work available on the influence of genetic and environmental factors on the composition and yield of nectar in *L. scoparium.*

Of value to the commercial UMF honey industry in New Zealand is the ability to assess cultivars from breeding programs for the best potential to increase overall UMF honey yield. Predictive modelling of yields is invaluable to the developing honey industry to allow assessment of environmental influences that may affect overall yield along with seasonal influences on nectar production in Mānuka. The research in this thesis establishes the effect of various parameters on overall DHA yield from Mānuka and the beginnings of modelling influencing environmental factors.

To determine influences on dihydroxyacetone (DHA) concentration and yield in the nectar of *L. scoparium* a number of studies were carried out. Methodologies for the collection and analysis of nectar were established. Ten different cultivars of *L. scoparium*

with a range of genetic parentage were studied in controlled glasshouse conditions to assess phenotypic variability in terms of nectar composition and yield as well as plant growth and flowering amongst these cultivars. Significant differences in plant growth and flowering habits were observed amongst the ten cultivars, significant differences in nectar yield and nectar composition with regard to DHA yield were also observed. DHA yields ranged from 2714-7459 mg of DHA/kilogram normalised to 80 ^{o}BRIX, with total nectar sugar yields ranging between 0.7 and 4.8 mg amongst the ten cultivars studied. Preliminary research into the effect of temperature, radiation and humidity on nectar composition and yield were also undertaken.

Effects of soil composition on these same parameters were researched with a subset of three of the ten cultivars grown on ten different soil types. Plant relative growth rates, dry weights and total plant height were measured throughout a 15 month glasshouse trial. Plant growth, flowering phenology, floral density, nectar yield and DHA composition data was gathered. Soils were analysed for various macronutrient and micronutrient levels and these parameters were modelled against plant data to determine which soil components were influencing plant parameters of interest. Soil type was shown to have no significant effect on DHA concentrations in nectar but results did show that soil type had a significant effect on flowering density amongst the three *L. scoparium* cultivars studied in the trial. Results from regression analysis of soil chemistry against measured plant parameters indicate that a fertiliser regime has the potential to increase nectar yields due to increased flower numbers. Multivariate analysis using partial least squares regression of soil composition data against plant parameters of value showed that soil components; phosphorus, sulphate, ferric and chloride were commonly shown to influence plant parameters measured.

Analytical spectroscopy was investigated as a method to chemotype *L. scoparium* cultivars and also as a method for quantifying nectar components sucrose, glucose, fructose and DHA.

Nectar composition was analysed using high pressure liquid chromatography (HPLC) and compared with fourier transform Raman spectroscopy (FT-Raman) and attenuated total reflectance infrared spectroscopy (ATR-FTIR) analytical spectroscopy methods.

FT-Raman spectroscopy was shown to be useful in chemotyping cultivars and in addition proved to be a useful analytical method to predict DHA yield using leaf material from *L. scoparium* plants from the ten cultivars. FT-Raman and ATR-FTIR proved to be relatively accurate techniques to quantify *L. scoparium* nectar components DHA, fructose, glucose and sucrose, compared with HPLC methods which use extensive preparation techniques. R-squared values were very good for all nectar components measured excepting the sucrose model at $R^2 = 0.77$. The R^2 for the FT-Raman predictions of DHA against HPLC data are very good at 0.85. FTIR prediction data against HPLC data was also good at 0.86 R^2. Overall an accurate model is possible for quantifying DHA concentrations in nectar using both FTIR-ATR and FT-Raman spectroscopy.

Overall results show that various factors need to be considered when assessing plants for commercial use in the (UMF) Mānuka honey industry within New Zealand. Due to their large impact on overall nectar yield; floral density and plant growth rate parameters are the two key factors of value for commercial assessment of Mānuka cultivars. This research also highlights the importance of assessing not just DHA concentration in deducing cultivar value, but overall nectar yield. These key features

must be explored when assessing *L. scoparium* plants within breeding programs, prior to selection for large-scale field production of high UMF Mānuka honey.

Acknowledgements

This work was supported by Primary Growth Partnership (PGP) funding (Ministry of Primary Industries, New Zealand) awarded to Mānuka Research Partnership (NZ) Limited (MRPL), and a scholarship from Callaghan Innovation (NZ) to L.M.N. [Contract: MARP1001].

Supervision and staff support from the Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand.

The Primary Growth Partnership (PGP) scheme brings the resources of many people together into one research program. And I feel privileged to have worked on my thesis under this collaborative scheme. The PGP scheme brings an extra focus to the work and the connection and interaction with the commercial partners (MRPL) in the scheme has been an invaluable and enjoyable experience.

There are many people, Dr Steve Holroyd, Professor Keith Gordon, Dr Chris Anderson, Dr Jason Wargent, Dr Jonathan Stephens, Professor Michael McManus, Dr James Milner and Associate Professor Alastair Robertson, who my sincere thanks go to for their scientific input and support for the research in this thesis. My special thanks and appreciation go to Georgie Hamilton for the technical support she has given me, along with our many enjoyable discussions on various other topics of interest while we worked on the Mānuka program. Additional thanks and much appreciation go to my supervisor Dr Steve Holroyd whose unfailing support encouraged me to attempt this thesis.

There are many others who I owe a debt of gratitude for the help and support with the experimental work in this thesis in particular the staff at the plant growth unit at Massey

University, Lindsey, Leslie and Steve and in addition Chris Rawlingson for his help with the analytical work on the HPLC and Geoff Smith at Otago University for his help with using the Raman equipment. And my employer: Fonterra Research and Development Centre (FRDC) Palmerston North who have allowed me the time from work to work on my thesis, and especially the support from FRDC statisticians Barbara Kuhn-Sherlock and Maree Luckman.

Last thanks go to my family and the support of my partner Daryl and son Henry who have suffered my lack of attention when my evenings were spent focussed on writing and my thoughts often elsewhere focussed on data analysis. Special thanks to my son Henry who helped me measure plants, collect leaves and chop up plants when I needed his help and entertained me with his wonderful wit, humour and patience along the way.

For myself, I have greatly enjoyed working on this thesis, it has been a very enjoyable experience and I have learnt an immense amount along the way and have really enjoyed getting to know all those involved and have greatly appreciated the input and knowledge and experience shared with me.

Thank you!!

Table of Contents

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

AAS	atomic absorption spectrophotometry
ACN	Acetonitrile
ANOVA	Analysis of Variance
ATR	Attenuated Total Reflectance
	(Attenuated Total Reflectance Fourier Transform
ATR-FTIR	(Infrared Spectroscopy
B	Boron
BS	Basal Stem
Ca	Calcium
CCD	Charged Couple Device
CCE	Calcium Carbonate Equivalent
CEC	Cation Exchange Capacity
Cl	Chloride
cm^{-1}	Spectral Wavenumbers
Co	Cobalt
Cu	Copper
CV	Cultivar
DHA	Dihydroxyacetone
EDTA	Ethylene-diamine-tetra-acetic acid
Fe	Ferric/Iron
FP	Flowering Period
F	Fructose
FTIR	Fourier Transform Infrared Spectroscopy
FT-Raman	Fourier Transform Raman Spectroscopy
g	Grams
GC	Gas Chromatography
GC-FID	Gas Chromatography with Flame Ionization Detector
GLM	General Linear Model

G	Glucose
HA	Hydroxyacetone
HPLC	High Pressure Liquid Chromatography
HT	Plant Height
ICP-OES	Inductively Coupled Plasma Optical Emission Spectrometry
IR	Infrared
K	Potassium
kg	Kilograms
l	Litre
L. scoparium	*Leptospermum scoparium*
LCMS	Liquid Chromatography Mass Spectrometry
me	milli-equivalents of exchangeable base cation
Mg	Magnesium
mg	Milligrams
MGO	Methylglyoxal
Mj/m^2	Micro-Joules of light energy per meter squared
ml	Milli-litre
mM	Milli-Molar concentration
Mn	Manganese
MRPL	Mānuka Research Partnership Limited
MS	Mass Spectrometry
MSC	Multiplicative Scatter Correction
MSI	Mānuka Soil Index
mW	Milli-Watts
N	Nitrogen
Na	Sodium
NA	Not Applicable
NaOH	Sodium Hydroxide
NIPALS	Non-Linear Iterative Partial Least Squares

NIR	Near Infrared
NMR	Nuclear Magnetic Resonance
NP	Nectar Potential
NPA	Non-Peroxide Antibacterial Activity
P	Phosphate
PC	Principal Component
PCA	Principal Component Analysis
PFBHA	O-(2, 3, 4, 5, 6-pentafluorobenzyl) hydroxylamine
PGP	Primary Growth Partnership
pH	Acidity/Alkalinity level
PLS	Partial Least Squares
PLSR	Partial Least Squares Regression
R	Correlation Value
RGR	Relative Growth Rate
RI	Refractive Index
RMSE	Root Mean Square Error
RMSEC	Root Mean Square Error of Calibration
RMSECP	Root Mean Square Error of the Prediction model.
RMSECV	Root Mean Square Error of the Cross Validated model.
S	Sulphur/Sucrose
SC	Seed Capsule numbers
SD	Start Days = Start of Flowering Period in Days
SNV	Standard Normal Variate
SO_4	Sulphate
µg	Micro-grams
µL	Micro-litres
UMF	Unique Mānuka Factor
UV/VIS	Ultra Violet/Visual Absorbance Spectra)
UV	Ultra Violet

Y, MG, G, B, BS, P, O, LG, PU and R	Cultivar Codes
Yd	Yield
Zn	Zinc

Chapter 1: Introduction

1.1 Project Overview

Leptospermum scoparium is the plant nectar source for medically bioactive honey. The honey produced by bees foraging on Mānuka nectar is a potent antibiotic and a range of medical and nutritional products have been developed as a consequence. In New Zealand this bioactive component is now branded and recognised as the UMF factor (Unique Mānuka Factor). The work presented in this thesis is part of a comprehensive research programme called the Primary Growth Partnership (PGP) research and development scheme that is attempting to research the technical and economic feasibility of combining improved genetics with Mānuka plantation husbandry techniques and the development of predictive techniques and tools to help manage Mānuka honey supply.

Mānuka honey is associated with a broad range of medical applications which command premium prices nationally and internationally. The unique non-peroxide antimicrobial action of Mānuka honey derives from its methylglyoxal (MGO) content and was identified from UMF honey by Mavric et al. in 2008 [1]. The carbohydrate dihydroxyacetone (DHA) found in floral nectar of *L. scoparium* was identified as the chemical precursor for the MGO bioactive component in honey by Adams et al. in 2009 [2]. The concentration of DHA in the floral nectar collected by bees correlates with the concentration of MGO in the resulting honey, with higher MGO content in honey commanding higher prices, UMF honey is graded according to its MGO content and branded accordingly.

Anecdotal evidence and recently published data from various cultivars in field conditions suggest that the DHA and overall nectar composition of *L. scoparium* nectar varies according to cultivar [3, 4], the source of this variation is not clearly understood and although there is considerable literature on climatic and genetic influences on nectar secretion rates and composition within various other plant species, there is no literature available on how genetic and environmental factors impact the composition and yield of nectar in *L. scoparium.*

New Zealand has a varying range of soil types where Mānuka is being considered for plantation planting, with soil nutrition known to affect plant growth and flowering in many species, soil type has been suggested as a possible influence in a recent publication on DHA content in nectar [4]. There is little published work in this field with relevance to *Leptospermum* species and I feel that this body of research adds value to the understanding of the influence of various factors especially soil chemistry on this species which has potentially positive outcomes for managing the future of the developing Mānuka honey industry.

This research project aims to provide identification of key aspects influencing the variability of nectar composition in *L. scoparium* and uses these key aspects to model potential UMF yield using mathematical modelling statistics. In addition a key part of this thesis has used the opportunity to apply facile analytical spectroscopy tools for quantifying metabolites in *L. scoparium* with relevance to nectar yield and quality. Utilising analytical spectroscopy techniques FT-Raman and ATR-FTIR coupled with advanced chemometric modelling tools, a novel application to screen *L. scoparium* cultivars for high UMF potential has been developed. The ability to screen plants for higher DHA content with facile analytical and screening methodologies is an important

capability that the commercial honey sector in New Zealand could take advantage of. Modelling techniques enabling the selection of conditions which could promote an increase in DHA concentrations and yield of *L. scoparium* nectar would be critically advantageous to the commercial honey sector.

1.2 Ministry of Primary Industries Primary Growth Partnership

The Primary Growth Partnership is a government commercial industry initiative that invests in significant programmes of research and innovation to boost the economic growth and sustainability of New Zealand's primary, forestry and food sectors.

The scheme focuses on boosting productivity through on-going investment in innovation, delivering long-term economic growth and sustainability across the primary sectors from producer to consumer.

Investments cover education and skills development, research and development, product development, commercialisation, commercial development and technology transfer.

The current aspiration of the PGP contract relevant to this doctoral research is that by 2035 through productivity gains enabled by this PGP programme, the Mānuka honey industry in New Zealand will be worth $1.2 billion, a 16-fold increase on its value of $75 million as at June 2010 (http://www.mpi.govt.nz). According to the latest Statistics New Zealand data, the current value of Mānuka honey exports jumped 45% in the last 12months to $27.4 million in November 2015. This contributed approximately 10% to the New Zealand annual value revenue from honey exports of $281 million in 2015. This PGP programme will research the technical and economic feasibility of combining

improved genetics with Mānuka husbandry techniques to achieve the following productivity gains:

1. double the hives per hectare carrying capacity on Mānuka throughout New Zealand;
2. double the average yield of a hive;
3. double the proportion of Mānuka honey capable of sale as a medicinal product;
4. double the land area in Mānuka economically accessible to beekeepers

Immediate outputs from the programme include the development of:

1. techniques in Mānuka plantation husbandry; and
2. predictive techniques/tools to help manage Mānuka honey supply.

With reference to the aims of this PGP programme, the output of this thesis aims to provide key information to the UMF honey industry regarding assessment of cultivars from and for breeding programmes, with the plan to select and breed plants with the highest potential for nectar yield and quality with relevance to the Mānuka honey industry, and provide insights into what is influencing the variability in UMF yield within *Leptospermum scoparium sp*. In addition the investigation on effects of soil type on plant parameters of interest, aims to provide information to the industry in regard to plantation siting, giving consideration to soil preferences and possible soil emendation effects worth further investigation. Climatic influences on nectar quality have also been studied to model environmental effects on nectar quality and yield. Alongside this, facile methods for screening plants and assessing nectar quality are investigated with relevance to industry practice with the aim to reduce breeding programme times and analytical assay times.

1.3 Research Objectives

1) Assess phenotypic variability in ten different cultivars of *Leptospermum scoparium* with regard to plant growth, nectar composition and nectar yield.
2) Assess a range of environmental influences on growth and nectar production and composition in *Leptospermum scoparium* with particular reference to the influence of soil type.
3) Assess the capability of analytical spectroscopy techniques to chemotype leaf metabolites in *Leptospermum scoparium* allowing facile identification of relevant cultivars.

1.4 Thesis Outline

This thesis reports the current knowledge of plant morphological characteristics and commercial relevance to New Zealand of *Leptospermum scoparium* along with a review of the methods and application of analytical spectroscopy techniques fourier transform Raman spectroscopy (FT-Raman) and attenuated total reflectance fourier transform infrared spectroscopy (ATR-FTIR). The thesis comprises 9 chapters in total: the first two chapters providing the background and literature review, chapters 3 and 4 consisting of methodology chapters. Chapters 5-8 are research chapters which have either been published or accepted or submitted to journals for review.

Chapter 1 (the current chapter): Introduction: a general introduction to the thesis topic and provision of background information on the format.

Chapter 2: Literature review of current knowledge with regard to the species *Leptospermum scoparium* and its commercial value to New Zealand. This chapter

also gives a general review of plant metabolites linking to metabolites within the nectar of *L. scoparium* and also the link to chemotyping cultivars using leaf metabolites, with a brief introduction to analytical spectroscopy methods.

Chapter 3: Glasshouse experimental methods: An introduction to the experimental design methods used along with breeding information of the proprietary cultivars used in the experiments that is not included within the paper chapters.

Chapter 4: Chemometrics: An outline of the development and theory of chemometric methods applied to analyse data obtained from the analytical spectroscopy methods of FT-Raman and FTIR-ATR applied in this thesis. These methods are important when processing and validating derived data.

Chapter 5: A glasshouse trial on ten different cultivars of *Leptospermum scoparium* presenting data on the variability of plant growth, nectar composition and yield within a range of *L. scoparium* (Mānuka) cultivars.

Chapter 6: A glasshouse trial on the influence of soil composition on plant growth, floral density, nectar composition and yield from three cultivars of Mānuka (*Leptospermum scoparium*) grown on nine different soils collected from New Zealand regional sites where Mānuka is being considered for plantation planting. A control soil made from commercial potting mix is also included in this experiment.

Chapter 7: Analytical FT-Raman spectroscopy to chemotype *Leptospermum scoparium* and generate predictive models for screening dihydroxyacetone levels in floral nectar: This is an analytical experiment researching the usefulness of FT-Raman spectroscopy as a technique for analysing leaf metabolites in a range of *Leptospermum scoparium* cultivars. In addition the possibility of using the data to

chemotype a range of cultivars and develop screening models for dihydroxyacetone levels was investigated.

Chapter 8: Analytical method development comparing FTIR-ATR and FT-Raman spectroscopy to assay the carbohydrates; fructose, sucrose, glucose and dihydroxyacetone, in *Leptospermum scoparium* nectar. Predictive models are developed and used to quantify nectar carbohydrate components.

Chapter 9: A general discussion including an overview of all results and recommendations for future work.

1.5 References

[1] Mavric E., Wittmann S., Barth G., Henle T. Identification and quantification of methylglyoxal as the dominant antibacterial constituent of Manuka (*Leptospermum scoparium*) honeys from New Zealand. Molecular Nutrition & Food Research **(2008)**, 52(4): p. 483-489.

[2] Adams C.J., Boult C. H., Deadman B.J., Farr J.M., Grainger M.N.C., Manley-Harris M. Snow M.J. Isolation by HPLC and characterization of the bioactive fraction of New Zealand manuka (*Leptospermum scoparium*) honey Carbohydrate Research **(2008),** 344(18): p.2609-2609

[3] Adams C.J., Manley-Harris M., Molan P.C. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2009)**, 344(8): p. 1050-1053.

[4] Williams S., King J., Revell M., Manley-Harris M., Balks M., Janusch F., Kiefer M., Clearwater M., Brooks P. Dawson M. Regional, Annual, and Individual Variations in the Dihydroxy4acetone content of the Nectar of Mānuka (*Leptospermum scoparium*) in New Zealand. Journal of Agricultural and Food Chemistry **(2014)**, 344(18): p. 2609-2609.

Chapter 2: Literature Review

2.1 Overview of Leptospermum scoparium

Leptospermum scoparium (Mānuka) is native to New Zealand and Australia and in both countries the species has a variable range of habitats [1], in New Zealand in particular, the species occurs over a wide geographical range and has highly variable morphology and physiology [2]. In habit *L. scoparium* is a considered a shrub and typically grows 2-5 metres (m) high but some varieties can grow up to 10 m tall, it has white to pink flowers, occasionally red, on average 8-15 mm in diameter with stamens shorter than the petals (Figure 2.1). The seed capsules are hard and woody. *Leptospermum scoparium* is an evergreen shrub with dense branching and small leaves, the leaves come to a point and are generally prickly in feel compared to Kanuka which is a very similar looking plant in habit with which Mānuka is often confused but Kanuka is of the *Kunzea* genus [3]. *Leptospermum* is also known commonly in New Zealand and Australia as the tea tree as it was used by Captain Cook and early settlers of both countries as a substitute for tea.

Mānuka colonises open ground and until its recent prominence as an agriculturally important crop was treated as a weed and was subject to vigorous land clearing.

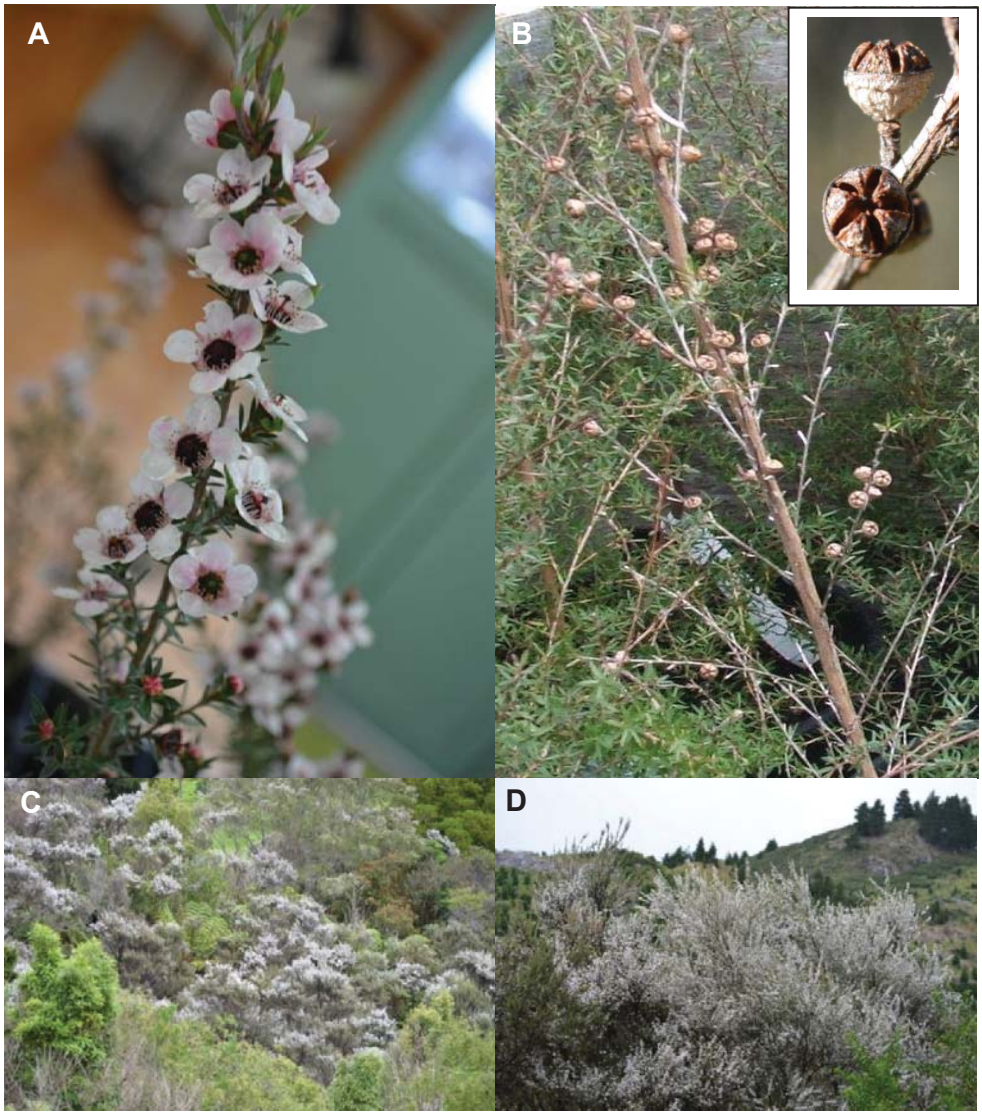

Figure 2.1. *Leptospermum scoparium* (site=Mt. Holdsworth Road, RD 1, Carterton, Wairarapa, New Zealand) A = floral stem, B = seed capsules on mature plant of *L. scoparium* C=ecological habitat growing alongside the Atiwhakitu stream, D = Mānuka plant in peak flowering.

2.2 Overview of Conditions Affecting Vegetative Growth

In New Zealand *Leptospermum scoparium* populations grow in a broad range of ecological environments in terms of altitude, temperature and moisture. *Leptospermum scoparium* grows from Stewart Island to the top of the North Island,

including the volcanic regions of the central plateau. It does not in general grow above 900-1000 m in altitude and it has been observed that it can grow for several years in standing water. In very dry areas it is absent at low to mid altitudes and is ecologically replaced by *Kunzea ericoides* (Kanuka) which is a more drought tolerant species [4]. *Leptospermum scoparium* colonises open ground and does not grow optimally in the shade and is therefore outgrown by forest species. *Leptospermum scoparium* is frost tolerant at temperatures from -4 °C down to -8 °C depending on the cultivar/subspecies. However seedlings of all *L. scoparium* species do not tolerate long periods of low temperature range, overall they are best adapted to the temperate climates in New Zealand [5, 6].

It has been noted that ideal soil conditions for *L. scoparium* are moist, slightly acidic soils with low concentrations of nutrients [7]. Permanent dominance occurs on sites that are unfavourable for the development of climax forest, for example too wet, dry, cold, exposed, infertile or unstable [8]. It is not clear whether *L. scoparium* preferentially grows on poorer soils or is outcompeted on more favourable soils. There are a large number of mycorrhizal associations listed in the literature with reference to *L. scoparium,* the principal benefit of these associations being an increase in phosphorus uptake allowing more rapid growth to exploit expansion into available clear land areas [8]. These associations indicate a possible benefit to the plant of higher nutrition.

Leptospermum scoparium is a relatively salt tolerant shrub so there is potential to irrigate in drier coastal areas with substandard water supplies for commercial cultivation. Cassaniti et al. 2009 [9] showed that *L. scoparium* could tolerate saline levels up to 70 mM irrigation concentrations with a reduction in relative growth rate

(RGR) of 24% at 70 mM over a period of 120 days. Under saline conditions most plants show a reduction in growth exhibited by smaller and fewer leaves and less growth in height as well as a reduction in root mass [9]. At this reduction in RGR *L. scoparium* is considered a salt tolerant plant.

Various sub species of *Leptospermum sp*. flower over different periods from spring through summer indicating that these subspecies require longer day lengths and warmer temperatures [8]. The temperature and photoperiod conditions that best promote flowering in *L. scoparium* were investigated by Zeislin et al. [10] in 1986 and demonstrated that flowering in *L. scoparium* occurs in photoperiods shorter than 12 hours at both high and low overnight temperatures with day temperatures between 20 ^{o}C and 26 ^{o}C. Flowering was completely absent in long day photoperiods of 16hrs plus. Between 12 and 16 hour photoperiods flowering only occurred at lower temperatures. The research showed that *L. scoparium* is a short day plant with a critical day length which is modified by temperature conditions [10].

Pollinators of *L. scoparium* are non-specific and come from a range of genera including Coleoptera (beetles) and Diptera (flies) and the indigenous Hymenoptera (ants, wasps and bees). The introduced honey-bee (*Apis mellifera*) also collects both pollen and nectar. Nocturnal moths have also been observed visiting the flowers of *L. scoparium* [8].

Pathogens of *L. scoparium* should be considered: *L. scoparium* is susceptible to disease infestation principally from scale insects ; *Coelostomidia wairoensis* and introduced *Eriococcus orariensis* and *Eriococcus leptospermi* of the Homoptera order, which leave the plant susceptible to the associated sooty mould infection of the fungus *Capnodium elegans*, the resultant “Mānuka Blight” which can results in the death of the plant is a

combination of insect infestation and the following development of sooty mould. Mānuka beetles (*Pyronota* spp.) can defoliate individual plants [11]. These beetles tend to inhabit Mānuka plants on the margins of bush regions with the larvae living in and feeding on nearby grassland pastures. The beetles are non-specific feeders and so are not isolated to feeding on Mānuka and mainly feed on Mānuka on margins of plant populations and are therefore not considered a serious threat to Mānuka in plantations [8, 11, 12].

2.3 Overview of Commercial Uses of Mānuka

Early settlers to New Zealand including the Maori used *Leptospermum scoparium* for food, medicine and timber [8]. The Maori used a decoction of the leaves as a drink as did the early Europeans who used the leaves as a substitute for tea. *Leptospermum scoparium* sometimes exudes a sugary gum onto its branches and this was used as a sweet and a soother for sore throats, in addition the smoke of burning bark and leaves was used as an inhalant for treating bronchial complaints.

Mānuka is a hardwood and therefore is useful for making handles for tools, the Maori used the hardwood from *L. scoparium* in a range of tools and specifically a named weapon called a “kahikatoa” [8].

Mānuka is currently cultivated in New Zealand for the honey and pharmaceutical industries as well as being propagated for the garden industry due to its popularity as a flowering shrub. In addition the wood is considered a good fuel wood for burning in wood fires and when smoked imparts flavour to fish and meats and is very popular in New Zealand for smoking trout.

Honey derived from Mānuka nectar exhibits relatively high antibacterial activity in comparison with other honeys [13]. Peroxide activity is responsible for the relatively lower antibiotic activity of other commercial honeys, the antibiotic properties of Mānuka honey is non-peroxide and originates directly from the MGO content of Mānuka honey [14] and in New Zealand is marketed as the "Unique Mānuka Factor" UMF.

In addition to its unique contributions to UMF honey, the Mānuka plant has potent antibacterial and antifungal properties due in part to high concentrations of flavonoid and phenolic compounds in the vegetative parts of the plant, the essential oil isolated from *Leptospermum* leaves commonly known as tea tree oil also has medicinal applications as an antifungal agent and powdered leaf material is in use in the food industry. Secondary metabolites of Mānuka also have antiseptic properties and the plant is the basis of several commercial industries promoting health products [1]. The essential oils of *L. scoparium* contain a large range of terpenes of which the majority are sesquiterpenes and flavonoids, these compounds have been shown to have antibacterial, antifungal as well as antioxidant activity. The oil composition is very variable depending on the plant population and seasonal conditions from which the oil is harvested. Standardisation of the oils is required by mixing various population harvests. Cultivation of cultivars that have good quality oils is of commercial importance therefore analysis of essential oils have been used to chemotype various sub-cultivars of *L. scoparium* and also to identify the source cultivar identity of commercial oils [15, 16]. Plant leaf metabolites can be used to identify source cultivars in many plant varieties therefore there is potential to use leaf chemotypes from *L. scoparium* as a fingerprint to identify those cultivars with high UMF potential.

2.4 Bioactive Constituents of Mānuka Honey

Mānuka is known to have many bioactive components of interest to science and medicine; however of particular interest to this study is Dihydroxyacetone (DHA). Dihydroxyacetone is the precursor for MGO the bioactive chemical responsiblefor the antimicrobial action of Mānuka honey, DHA is present in significant amounts in the floral nectar of *Leptospermum scoparium*. In UMF honey produced by the introduced honey bee (*Apis mellifera*) DHA is converted into MGO during storage of the honey [13].

2.4.1 Methylglyoxal and Dihydroxyacetone

The carbohydrate Methylglyoxal (MGO) (Figure 2.2) was determined to be responsible for the potent antimicrobial activity in Mānuka honey by Mavric et al. in 2008. Mavric investigated the metabolite MGO and its precursor DHA using HPLC and LC-MS methods alongside bioassays testing for antibacterial activity against *Staphylococcus aureus* and *Escherichia coli* [14]. Adams et al demonstrated that the formation of MGO occurred during storage of honey and that DHA present in the honey was the precursor [13]. The main component responsible for the antimicrobial action in most other honeys researched is hydrogen peroxide and the antibiotic activity of hydrogen peroxide honeys is not as high as the MGO activity in Mānuka honey [8]. Studies using ornamental tobacco plants indicate that the hydrogen peroxide is produced by the action of nectar proteins [17]. The hydrogen peroxide production is thought to be a natural plant defence system, evolved to protect the plant from the transfer of infectious organisms by visiting pollinators to the gynoecium and therefore also protect the developing seeds in the ovary which are of crucial importance to the

survival of the species [18]. Interestingly DHA in Mānuka floral nectar does not have antimicrobial properties. It is the conversion of DHA to MGO in aged Mānuka honey that adds antimicrobial activity to the honey.

In freshly produced honey, DHA can occur in concentrations measured from 1192-5099 mg/kg of honey. Estimations from Adams work showed a value for concentrations of DHA in the nectar up to the equivalent of 13,600 mg/kg in honey. The New Zealand beekeeping industry recognised that storage of Mānuka honey increases the UMF (Unique Mānuka Factor) rating and consequently the market value. Higher the UMF rating values correlate with higher bioactive activity [19]. Honeys studied by Stephens et al. 2010 showed an approximate 300% increase in DHA concentration over 5 years (Figure 2.3). However, the relationship between storage and developing MGO concentration in Mānuka honey has yet to be established [14]. Figure 2.3 demonstrates the increase in MGO over time as honey matures. The exact pathway by which DHA is converted to MGO in Mānuka honey has yet to be identified [15].

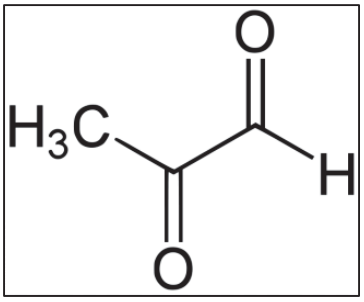

Figure 2.2. Molecular structure of methylglyoxal.

There are a range of chemical pathways by which DHA could be converted to MGO (Figure 2.4). Carbohydrates can degrade during storage to reactive 1, 2-dicarbonyl compounds including MGO [13]. The non-enzymatic reactions involving 1, 2-dicarbonyl compounds are referred to as either caramelisation or if amino containing compounds are present (as they are in honey) then Maillard reactions can also occur [4]. Caramelisation and Maillard processes in food products confer much of the flavour and colour to food. MGO may also be formed directly from DHA by the action of the enzyme methylglyoxal synthetase, in addition it has been shown that DHA dehydrates irreversibly to MGO under acidic conditions [13]. Recent research suggests that the conversion of DHA to MGO results in a protein fractionation effect but the pathway is still as yet unknown [20].

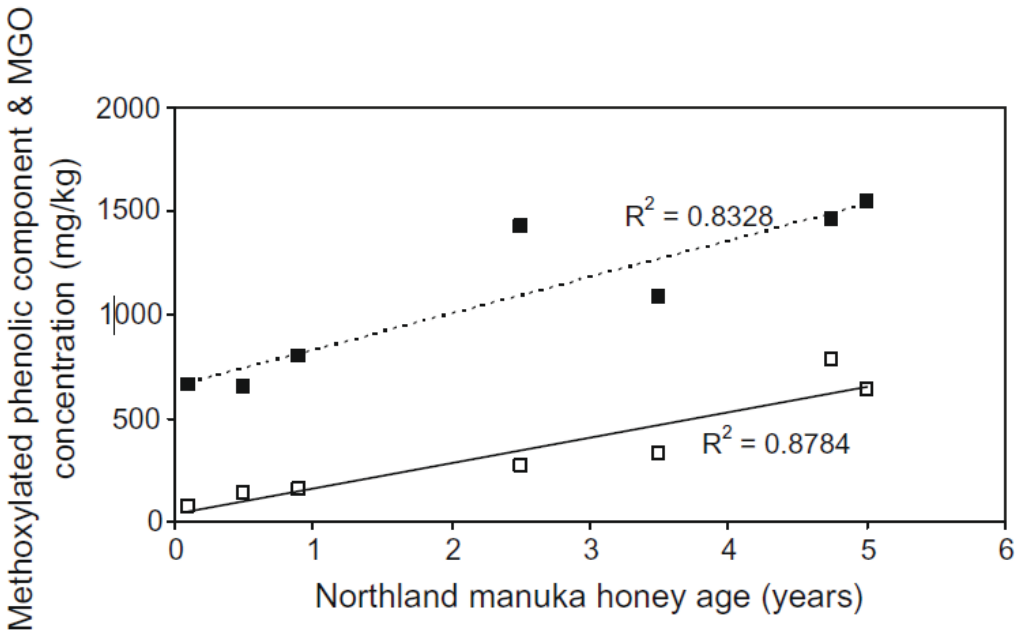

Figure 2.3. Concentrations of MGO (■) and sum of methoxylated phenolic components (□) in Mānuka honeys, fresh collected from the beehive to 5 year old stored honey and harvested in similar areas of Northland. Diagram taken from Jonathan M. Stephens et al. (2010) [24].

2.5 Nectar Chemistry

None of the *L. scoparium* nectars analysed by Adams et al. showed detectable amounts of MGO. In 2008, Stephens et al. found measurable amounts of MGO in nectar from flowers from a range of *L. scoparium* var. *incanum* derived cultivars. The mechanism for the conversion of DHA to MGO in the nectar of *L. scoparium* is as yet unknown and whether MGO is present in nectar across the *Leptospermum* genus needs further investigation. Conversion of DHA to MGO in the plant will be under the control of plant metabolic systems and MGO and DHA levels in nectar would be maintained at as yet unknown functional levels. Possibly when the nectar and therefore DHA and MGO is transferred to the beehive via bee foraging those controls are no longer present or active and so levels of MGO increase as the honey ages as a result of the conversion of DHA into MGO as proposed by Adams et al 2009.

DHA is a monosaccharide (ketotriose) and is metabolised in plants during glycolysis and gluconeogenesis in carbohydrate metabolism via the reversible isomerisation of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate, catalysed by triose phosphate isomerase. It has been proposed that the same systems exist in floral nectaries see Figure 2.4 [21].

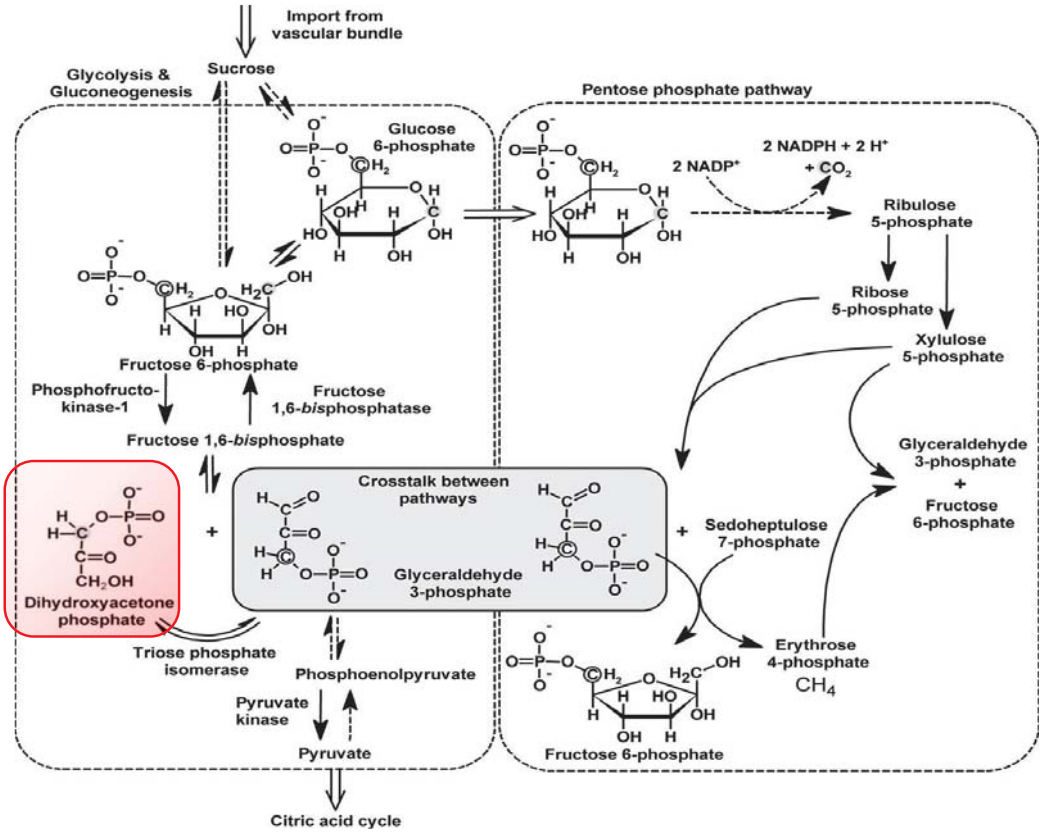

Figure 2.4. Proposed metabolic pathways for sugar metabolism in floral nectaries, note dihydroxyacetone-phosphate (red region). Diagram taken from Wenzler *et al* (2008) [26].

2.5.1 Nectar Composition

The nectar composition of angiosperms is extremely variable across the groups of families within genera and within species [18]. The composition of floral nectar is mainly sugars with small amounts of amino acids and lipids. The main sugar component of most angiosperm nectar is sucrose with varying concentrations of glucose and fructose depending on the plant species with significant intra species variation observed usually dependant on local climate conditions, seasonal variance and pollinator associations [22]. In general nectar and phloem sap contain only the sugars fructose, glucose and sucrose. Pollinators have been recorded depositing micro-flora into nectar which includes yeasts which can hydrolyse the sucrose into additional

concentrations of fructose and glucose. The numerous other carbohydrates are produced by microbial activity and enzymatic reactions [23].

2.6 Review of Secondary Metabolism and Metabolites in Plants

Primary plant functions are defined as growth, photosynthesis and reproduction, roles that are required for the plant to survive. Many thousands of plant metabolites have been identified across the plant kingdoms and are defined as having secondary plant functions not required for the primary function of the plant [24].

Most of the roles secondary metabolites have are in plant defence, for example foraging deterrents and so the chemicals can be toxic to various antagonist species or are precursors to a structural defence system of the plant as in insectivorous plants [25]. They can also have roles as attractants to pollinators such as scented flowers, colour and nectar. They also have regulatory and signalling roles in primary metabolic pathways so the split between primary and secondary metabolism is not always clear cut [17, 36-30]. Although DHA is a carbohydrate and is therefore considered a primary metabolite (Table 2.1) it could be considered a secondary metabolite because its role is not primary in the plant. Plant metabolites are classified on the basis of their chemical structure as shown in Table 2.1.

Table 2.1. This table shows the classification of plant metabolites into primary and secondary groups (table developed from Schulz, H. and M. Baranska, *Identification and quantification of valuable plant substances by IR and Raman spectroscopy,* Vibrational Spectroscopy, 2007) [24].

Primary Plant Metabolite Groupings	
Proteins and Amino Acids	
Lipids and Fatty Acids	
Carbohydrates	
Secondary Plant Metabolite Groupings	
Phenolics	*flavonoids*
	phenylpropanes
	other polyphenols
Terpenoids	*monoterpenes*
	sesquiterpenes
	tetraterpenes
Alkaloids	
Polyacetylenes	
Others	*nitrides*
	iridoids
	chlorophylls

Each family, genus and species of plant produces a characteristic mix of these secondary metabolites and they can be used as taxonomic characters in classifying plants [31]. Various analytical methods are available for investigating these metabolites including spectroscopy and chromatography techniques.

2.6.1 Metabolomics

Metabolomics is the post-genomic study of the molecules (metabolites) and processes that make up metabolism [32]. Metabolism is a cellular process that is the result of multiple chemical pathways and also multiple genes, proteins, and metabolites. Metabolites are the phenotypic end products of cellular regulatory processes, and the composition of metabolites in any system is regarded as a response of the biological entity to genetic or environmental changes [33]. The occurrence of one gene influencing multiple phenotypic traits is described as pleiotropy, inferring its influence

in multiple metabolic pathways contributing to multiple phenotypes [38]. Metabolomic techniques usually apply the various techniques of high pressure liquid chromatography (HPLC) and/or gas chromatography (GC) associated with mass spectrometry (MS) and also nuclear magnetic resonance ^{1}HNMR to elucidate the structure of the metabolite molecules [34, 35]. Chemotype profiles from spectroscopic methods can also identify metabolites and have been used to identify species metabolite profiles from a range of plant tissues [24, 31].

Plants have a complex chemical profile making it difficult to understand plant environment interactions [34], a metabolomic approach has great potential to increase our knowledge of the internal state of the biochemistry of plant cells in response to external stimuli [36]. There have been few metabolomic studies of Myrtaceae. The available literature reports two in the Eucalyptus genus researching the effects of water and/or salt stress and also animal foraging on the metabolic profile in a few species investigating changes in metabolite profiles [34, 37]. Porter et al. [28, 29] and Perry et al. [38] have reported metabolite profiling identifying essential oil components in general in *Leptospermum* and *Kunzea* species. There is no literature reporting the use of FT-Raman spectroscopy to produce metabolomics (chemotype) profiles in *L. scoparium.*

2.7 Overview of Raman and FTIR Spectroscopic Applications

Spectroscopy techniques have recently been used to research plant species to identify various secondary metabolites within plant structures along with fingerprinting cultivars and chemical components and also identification of sources of nectar in honey. Fourier transform Raman (FT-Raman) and fourier transform infrared (FTIR) techniques have proven to be successful techniques for these applications [31, 39-48].

Molecular bonds in biological compounds are not static, but can vibrate in predictive modes when energy is put into the system. FT-Raman and FTIR spectroscopic analyses exploit the various vibrational energy modes of the sample to gain information about the system under investigation. This can be both molecular structure or sample composition. The output data from spectroscopy is in the form of spectral lines and the resonant frequencies are related to the strength of the bond, and the sum of the masses at either end of the chemical bond. Therefore, by analysing the spectral output, the frequency of the vibrations can be associated with a particular bond type, and unique mixes of bond type frequencies are associated with certain chemical compounds [49].

Spectral data is the result of an interaction between a system under investigation (usually atoms, but sometimes molecules or atomic nuclei) and a single photon. When a photon has the right amount of energy to allow a change in the energy state the photon is absorbed. Raman and FTIR spectroscopy techniques use photons from wavelengths in the infrared and visible energy range (usually in the form of a laser) to excite the system under investigation and then detect the molecular vibration output

from the samples molecular structure in the form of an absorbance or emission spectrum.

Simple diatomic molecules have only one bond and only one vibrational band. More complex molecules have many bonds, and their vibrational spectra are correspondingly more complex. Spectra can be analysed visually and also using various statistical and or chemometric techniques to characterise and discriminate various chemicals and components in the sample and if imaging and mapping spectroscopy is applied together chemical maps showing the distribution of any components under investigation can be created [35, 30, 42, 43, 46, 47, 50-52].

2.7.1 FTIR Spectroscopy

FTIR is an analytical technique routinely used to rapidly quantify chemical components and there are many published applications of its use in analytical chemistry. Molecules are said to be IR active if a change in the vibrational state is brought about by absorption or emission of radiation via a change in the dipole of a molecule caused by atoms that are displaced relative to one another [49]. Figure 2.5 is an example of FTIR spectra from nectar, sugar (mock nectar) and DHA in solution. Peaks at the various wavenumber frequencies relate to vibrational states of specfic component molecular bonds.

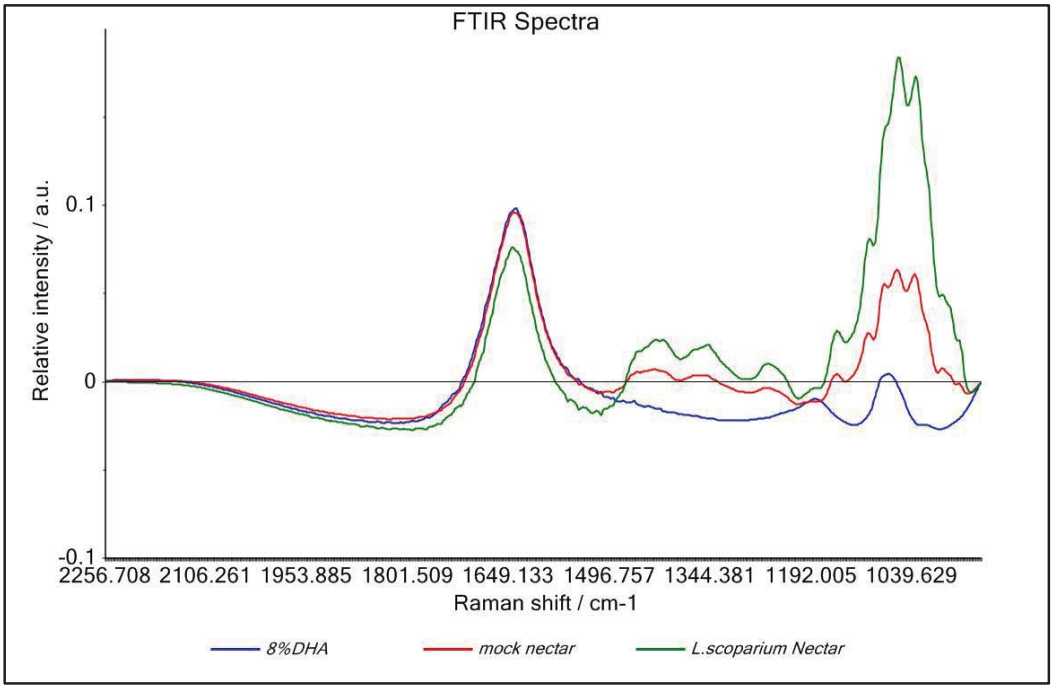

Figure 2.5. Example of FTIR spectra of three carbohydrate components of the nectar of *L. scoparium;* Dark Blue = 8% DHA reference in a mock nectar solution of Fructose and Glucose (ratio 2:1), Green = nectar from wild type *L. scoparium*; Red = a mock nectar solution of Fructose and Glucose (ratio 2:1), (Figure generated by the author).

In this thesis attenuated total reflectance fourier transform infrared spectroscopy (ATR-FTIR) instrumentation is used. An attenuated total reflection accessory measures the changes that occur in an internally reflected infrared beam where the beam comes into contact with a sample. An infrared beam is directed onto an optically dense crystal with a high refractive index (in this case a zinc selenide crystal, Zinc Selenide is a relatively low cost ATR crystal material and is ideal for analysing liquids). The beam protrudes only a few microns (0.5 µm - 5 µm) beyond the crystal surface and into the sample. There must be good contact between the sample and the crystal surface to allow good penetration of the beam into the sample.

In regions of the infrared spectrum where the sample absorbs energy, the evanescent wave will be attenuated or altered and is passed back to the IR beam, which then exits

the opposite end of the crystal and then to the detector in the IR spectrometer. The detector records the attenuated IR beam as an interferogram signal, which can then be used to generate an IR spectrum [53]. The refractive index of the crystal must be significantly greater than that of the sample or else internal reflectance will not occur, the light will be transmitted rather than internally reflected in the crystal. ATR crystals have refractive index values between 2.38 and 4.01 at 2000 cm^{-1}. An infrared background is collected, in this case, from the clean ATR crystal. It is safe to assume that the majority of solids and liquids have much lower refractive indices [54]. Most importantly, the improved spectral acquisition and reproducibility associated with this technique leads to more precise material verification and identification.

2.7.2 Raman Spectroscopy

Raman spectroscopy, similar to infrared spectroscopy, also studies vibrational modes in a system. A molecule is Raman active if the polarisability of the molecule changes as the molecule vibrates [49]. Raman relies on inelastic scattering in which some of the energy of the incident particle is lost or gained. Laser excitation light, usually in the visible range but also the near infrared (NIR) and the near ultra violet (UV) range are also used, interacts with molecular vibrations, resulting in the energy of the laser photons being shifted up or down. The shift in energy gives information about the vibrational modes in the system. Visible and UV excitation wavelengths can induce fluorescent emissions which interferes with the Raman spectrum collected depending on the sample being analysed, in which case NIR wavelengths such as 1024 nm lasers can be used in preference, 1024nm excitation is especially useful when analysing plant leaf material as chlorophyll pigments within leaves are induced to fluoresce by excitation wavelengths in the visible region of the spectrum.

If the final vibrational state of the molecule is higher in energy than the initial state, then the emitted photon will be shifted to a lower frequency in order for the total energy of the system to remain balanced. This shift in frequency is designated as a Stokes shift. If the final vibrational state is less energetic than the initial state, then the emitted photon will be shifted to a higher frequency, and this is designated as an Anti-Stokes shift. Raman scattering is an example of inelastic scattering because of the energy transfer between the photons and the molecules during their interaction [53]. Raman spectroscopy is commonly used in analytical chemistry, since vibrational information is specific to the chemical bonds and symmetry of the molecules in the compound under investigation. Figure 2.6 illustrates the FT-Raman spectra of dihydroxyacetone in solution, *L. scoparium* nectar and a sugar solution (mock nectar). The fingerprint (unique identification) region of organic molecules is in the wavenumber range 500–2000 cm^{-1}.

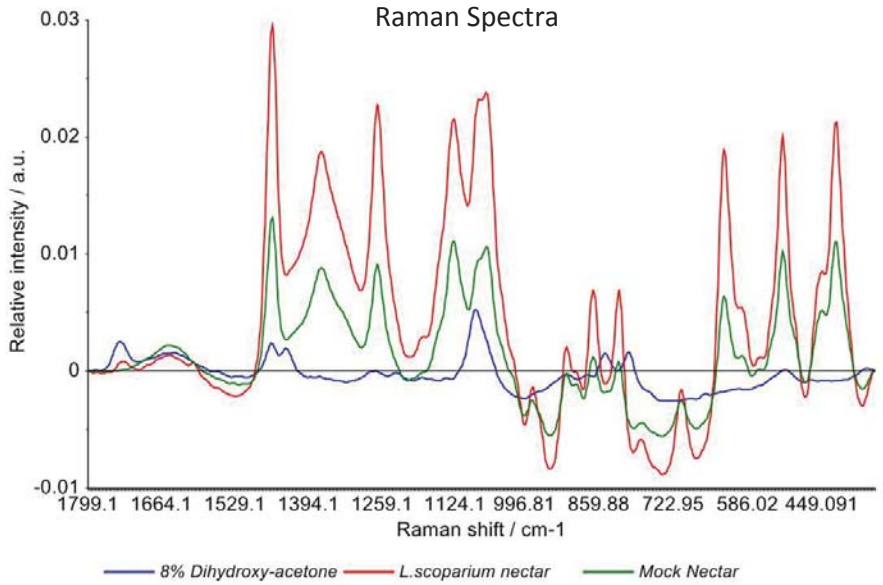

Figure 2.6. FT-Raman spectra of nectar and various carbohydrates in solution. Note that FT-Raman generally detects more bond structures than FTIR illustrated by an increased number of peaks in the spectra. Dark Blue = 8% DHA reference, Green = a mock nectar solution of Fructose and Glucose (ratio 2:1); Red = nectar from wild type *L. scoparium*, (Figure generated by the author).

2.7.3 Application of FTIR and Raman in Analysing Carbohydrates in Plants

FTIR and Raman spectroscopy are complementary analytical techniques applied to analyse chemical components and structure of a range of biological and physical samples. Major components of plant secondary metabolites can be analysed using both of these complementary techniques. There is a large amount of literature covering research into spectroscopic techniques used to analyse various essential oil compounds from various plant species. Spectroscopic analysis is based on the identification of characteristic key bands of individual substances. Spectroscopic techniques allow discrimination between different essential oil profiles allowing plant chemotypes to be distinguished and in most cases identification of individual oil components [31].

Generally both FTIR and FT-Raman spectroscopy techniques allow spectra to be obtained that present some characteristic bands of individual components. These bands provide information about the chemical composition, including both primary and secondary metabolites of the investigated samples [29]. There has been a small amount of work done applying spectroscopy techniques to analyse carbohydrates in plants, these include analysis of mono-saccharides and polysaccharides applying either FTIR or FT-Raman and various chemometric analysis techniques to differentiate individual carbohydrates or fingerprint sample types [52-56].

2.8 Specific Research Objectives

At present gaps exist in many areas with regards to the flowering phenology, nectar composition and yield across all cultivars of *Leptospermum scoparium.* To date there have been no studies to investigate plant growth, nectar production and composition in *L. scoparium* cultivars in a controlled environment. This PhD project uses controlled glasshouse experiments to study a range of cultivars for growth habit, nectar and flowering phenology. These plant parameters are important aspects to consider when assessing plants for best application in the commercial honey industry and there is currently no scientifically based literature available that addresses these keys aspects.

Effects of the local climate and soil composition on plant growth, flowering and nectar yield and composition in *L. scoparium* are also currently unknown and poorly understood. A gap in knowledge also exists in identifying best methods for assessment of the UMF potential of *L. scoparium* and modelling the performance of *L. scoparium* in various environmental conditions.

Facile spectroscopic methods are available to chemotype plant species along with quantification of various metabolites. These applications have not been applied to chemotype cultivars of *L. scoparium*, these methods also offer the opportunity to be of use in analysing plant nectar components.

The specific objectives of this thesis therefore are:

1) Identify the variance of specific nectar and growth parameters within several cultivars of *L. scoparium* to establish the potential of breeding programmes for increasing UMF honey yield from *L. scoparium*.

2) Identify the effect of soil composition if any on nectar, plant growth and flowering parameters to allow decisions to be made regarding best soil sites to establish *L. scoparium* plantations.
3) Model the effects of a range of climatic conditions on nectar composition and yield in several cultivars of *L. scoparium* to allow identification of localities where climatic factors may be beneficial or detrimental to overall UMF honey yield.
4) Establish standard methodologies for nectar collection.
5) Test the capability of FT-Raman spectroscopy to chemotype cultivars of *L. scoparium.*
6) Test the capability of FT-Raman and ATR-FTIR spectroscopy and chemometrics to quantify nectar components of *L. scoparium*
7) Investigate the potential for linking metabolites from chemotype profiles produced by FT-Raman data to UMF potential in *L. scoparium.*

The overall aim of this research was to identify any overall influential factors with the potential of increasing UMF honey yield from *L. scoparium.* And develop modelling techniques for predicting yield based on environmental conditions and cultivars in prospective plantation sites. With additional value to be gained from identifying facile methods for screening plants from breeding programmes or from wild type plants for UMF potential.

2.9 References

[1] Klank C., Pluess A.R., Ghazoul J. Effects of population size on plant reproduction and pollinator abundance in a specialized pollination system. Journal of Ecology **(2010)**, 98(6): p. 1389-1397.

[2] Bond W.J., Dickinson K.J.M., Mark A.F. What limits the spread of fire-dependent vegetation? Evidence from geographic variation of serotiny in a New Zealand shrub. Global Ecology and Biogeography **(2004)**, 13(2): p. 115-127.

[3] Thompson J. A revision of the genus *Leptospermum* (Myrtaceae). Telopea **(1989)**, 3(3): p. 301-449.

[4] Wilson J.B., Yin R.H., Mark A.F., Agnew A.D. Q.A test of the low marginal variance (LMV) theory, in *Leptospermum scoparium* (Myrtaceae). Evolution **(1991)**, 45(3): p. 780-784.

[5] Greer D.H., Muir L.A., Harris W. Seasonal frost hardiness in *Leptospermum scoparium* seedlings form diverse sites throughout New Zealand. New Zealand Journal of Botany **(1991)**, 29(2): p. 207-212.

[6] Greer D.H., Robinson L.A. Temperature control of the development of frost hardiness in 2 populations of *Leptospermum scoparium*. Tree Physiology **(1995)**, 15(6): p. 399-404.

[7] Hall, I.R. Effect of applied nutrients and endomychorrhizas on Metrosiderous umbellata and *Leptospermum scoparium*. New Zealand Journal of Botany **(1977),** 15(2): p. 481-484.

[8] Stephens J.M.C., Molan P.C., Clarkson B.D. A review of *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany **(2005)**, 43(2): p. 431-449.

[9] Cassaniti C., Leonardi C., Flowers T.J. The effects of sodium chloride on ornamental shrubs. Scientia Horticulturae **(2009)**, 122(4): p. 586-593.

[10] Zieslin N. Environmental Factors involved in growth, flowering and post harvest flowers of *Leptospermum scoparium*. Israel Journal of Botany **(1986)**, 35: p. 101-108.

[11] Thomson N.A., Miln A.J., Kain W.M. Biology of manuka beetle in Taranaki. Proceedings of the Thirty-second Weed and Pest Control Conference (1979), 32: p. 80-85.

[12] Hodgson C.J., Henderson R.C. Coccidae (Insecta: Hemiptera: Coccoidea). Fauna of New Zealand (2000), 41: pp. 264.

[13] Adams C.J., Manley-Harris M., Molan P.C. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2009)**, 344(8): p. 1050-1053.

[14] Mavric E., Wittmann S., Barth G., Henle T. Identification and quantification of methylglyoxal as the dominant antibacterial constituent of Manuka (*Leptospermum scoparium*) honeys from New Zealand. Molecular Nutrition & Food Research **(2008)**, 52(4): p. 483-489.

[15] Lis-Balchin M., Hart S.L. An investigation of the actions of the essential oils of Manuka (*Leptospermum scoparium*) and Kanuka (*Kunzea ericoides*), Myrtaceae

on guinea-pig smooth muscle. Journal of Pharmacy and Pharmacology **(1998)**, 50(7): p. 809-811.

[16] Lis-Balchin M., Hart S.L., Deans S.G. Pharmacological and antimicrobial studies on different tea-tree oils (*Melaleuca alternifolia*, *Leptospermum scoparium* or Manuka and Kunzea ericoides or Kanuka), originating in Australia and New Zealand. Phytotherapy Research **(2000)**, 14(8): p. 623-629.

[17] Park S., Thornburg R.W. Biochemistry of Nectar Proteins. Journal of Plant Biology **(2009)**, 52(1): p. 27-34.

[18] De la Barrera E., Nobel P.S. Nectar: properties, floral aspects, and speculations on origin. Trends in Plant Science **(2004)**, 9(2): p. 65-69.

[19] Stephens J.M., Morris B.D., Yang D., Fearnley L., Greenwood D.R., Loomes K.M. Phenolic compounds and methylglyoxal in some New Zealand manuka and kanuka honeys. Food Chemistry **(2010)**, 120: p. 78–86.

[20] Rogers K.M., Grainger M., Manley-Harris M. The Unique Manuka Effect: Why New Zealand Manuka Honey Fails the AOAC 998.12 C-4 Sugar Method. Journal of Agricultural and Food Chemistry **(2014)**, 62(12): p. 2615-2622.

[21] Wenzler M., Holscher D., Oerther T., Schneider B. Nectar formation and floral nectary anatomy of *Anigozanthos flavidus*: a combined magnetic resonance imaging and spectroscopy study. Journal of Experimental Botany **(2008)**, 59(12): p. 3425-3434.

[22] Canto, A., Herrera C.M., Garcia I.M., Perez R., Vaz M. Intraplant variation in nectar traits in *Helleborus foetidus* (Ranunculaceae) as related to floral phase,

environmental conditions and pollinator exposure. Flora **(2011)**, 206(7): p. 668-675.

[23] Canto, A., Perez R., Medrano M., Castellanos M.C., Herrera C.M. Intra-plant variation in nectar sugar composition in two Aquilegia species (Ranunculaceae): Contrasting patterns under field and glasshouse conditions. Annals of Botany **(2007)**, 99(4): p. 653-660.

[24] Schulz H., Baranska M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vibrational Spectroscopy **(2007)**, 43(1): p. 13-25.

[25] Kennedy D.O., Wightman E.L. Herbal extracts and phytochemicals: plant secondary metabolites and the enhancement of human brain function. Advances in Nutrition **(2011)**, 2(1): p. 32-50.

[26] Haberlein H., Tschiersch K.P. Triterpenoids and flavanoids from *Leptospermum scoparium*. Phytochemistry **(1994)**, 35(3): p. 765-768.

[27] Lucas-Barbosa D., van Loon J.J.A., Dicke M. The effects of herbivore-induced plant volatiles on interactions between plants and flower-visiting insects. Phytochemistry **(2011)**, 72(13): p. 1647-1654.

[28] Porter, N.G., Smale P.E., Nelson M.A., Hay A.J., Van Klink J.W., Dean C.M. Variability in essential oil chemistry and plant morphology within a *Leptospermum scoparium* population. New Zealand Journal of Botany **(1998)**, 36(1): p. 125-133.

[29] Porter N.G., Wilkins A.L. Chemical, physical and antimicrobial properties of essential oils of *Leptospermum scoparium* and *Kunzea ericoides*. Phytochemistry **(1999)**, 50(3): p. 407-415.

[30] Baranski R., Baranska M., Schulz H. Changes in carotenoid content and distribution in living plant tissue can be observed and mapped in situ using NIR-FT-Raman spectroscopy. Planta **(2005)**, 222(3): p. 448-457.

[31] Baranska M., Schulz H., Kruger H., Quilitzsch R. Chemotaxonomy of aromatic plants of the genus Origanum via vibrational spectroscopy. Analytical and Bioanalytical Chemistry **(2005)**, 381(6): p. 1241-1247.

[32] Levin N., Multivariate statistics and the enactment of metabolic complexity. Social Studies of Science **(2014)**, 44(4): p. 555-578.

[33] Fiehn O. Metabolomics - the link between genotypes and phenotypes. Plant Molecular Biology **(2002)**, 48(1-2): p. 155-171.

[34] Tucker D.J., Wallis I. R., Bolton J.M., Marsh K.J., Rosser A.A., Brereton I.M., Nicolle D., Foley W.J. A Metabolomic Approach to Identifying Chemical Mediators of Mammal-Plant Interactions. Journal of Chemical Ecology **(2010)**, 36(7): p. 727-735.

[35] Hall R., Beale M., Fiehn O., Hardy N., Sumner L., Bino R. Plant metabolomics: The missing link in functional genomics strategies. Plant Cell **(2002)**, 14(7): p. 1437-1440.

[36] Taylor J., King R.D., Altmann T., Fiehn O. Application of metabolomics to plant genotype discrimination using statistics and machine learning. Bioinformatics **(2002)**, 18: p. S241-S248.

[37] Warren C.R., Aranda I., Javier Cano F. Metabolomics demonstrates divergent responses of two Eucalyptus species to water stress. Metabolomics **(2012)**, 8(2): p. 186-200.

[38] Perry, N.B., VanKlink J.W., Brennan N.J., Harris W., Anderson R.E., Douglas M.H., Smallfield B. M.Essential oils from New Zealand manuka and kanuka: Chemotaxonomy of Kunzea. Phytochemistry **(1997)**, 45(8): p. 1605-1612.

[39] Agarwal U.P. Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (Picea mariana). Planta **(2006)**, 224(5): p. 1141-1153.

[40] Baranska M., Proniewicz L.M. Raman mapping of caffeine alkaloid. Vibrational Spectroscopy **(2008)**, 48(1): p. 153-157.

[41] Baranska M., Schulz H., Joubert E., Manley M. In situ flavonoid analysis by FT-Raman spectroscopy: Identification, distribution, and quantification of aspalathin in green rooibos (*Aspalathus linearis*). Analytical Chemistry **(2006)**, 78(22): p. 7716-7721.

[42] Gamsjaeger, S., Baranska M., Schulz H., Heiselmayer P., Musso M. Discrimination of carotenoid and flavonoid content in petals of pansy cultivars (*Viola x wittrockiana*) by FT-Raman spectroscopy. Journal of Raman Spectroscopy **(2011)**, 42(6): p. 1240-1247.

[43] Gierlinger N., Schwanninger M. The potential of Raman microscopy and Raman imaging in plant research. Spectroscopy-an International Journal **(2007)**, 21(2): p. 69-89.

[44] Kruger H., Schulz H. Analytical techniques for medicinal and aromatic plants. Stewart Postharvest Review **(2007)**, 3(4): p. 4-8.

[45] Pierna J.A.F., Abbas O., Dardenne P., Baeten V. Discrimination of Corsican honey by FT-Raman spectroscopy and chemometrics. Biotechnologie Agronomie Societe Et Environnement **(2011)**, 15(1): p. 75-84.

[46] Rozo, J.I.J., Zarow A, Zhou B., Pinal R., Iqbal Z., Romanach R. J. Complementary Near-Infrared and Raman Chemical Imaging of Pharmaceutical Thin Films. Journal of Pharmaceutical Sciences **(2011)**, 100(11): p. 4888-4895.

[47] Coimbra M. l A., Barros A., Rutledge D.N., Delgadillo I. FTIR spectroscopy as a tool for the analysis of olive pulp cell-wall polysacchride extracts. Carbohydrate Research **(1999)**, 317(1999): p. 145-154.

[48] Dokken K.M., Davis L.C. Infrared imaging of sunflower and maize root anatomy. Journal of Agricultural and Food Chemistry **(2007)**, 55(26): p. 10517-10530.

[49] Camargo S.C., Garcia R.C., Feiden A., Vasconcelos E.S., et al. Implementation of a geographic information system (GIS) for the planning of beekeeping in the west region of Parana. Anais Da Academia Brasileira De Ciencias **(2014)**, 86(2): p. 955-971.

[50] Willis L.B., Gil G.A., Lee H.L.T., Choi D., Schoenheit J., Ram R.J., Rha C., Sinskey A.J. Application of spectroscopic methods for the automation of oil palm culture. Journal of Oil Palm Research **(2008)**, p. 1-13.

[51] Baranska M., Schulz H., Reitzenstein S., Uhlemann U., Strehle M. A., Kruger H., Quilitzsch R., Foley W., Popp J.Vibrational spectroscopic studies to acquire a quality control method of Eucalyptus essential oils. Biopolymers **(2005)**, 78(5): p. 237-248.

[52] Armenta S., Garrigues S., de la Guardia M., Rondeau P.Attenuated Total Reflection-Fourier transform infrared analysis of the fermentation process of pineapple. Analytica Chimica Acta **(2005)**, 545(1): p. 99-106.

[53] Technical Note: Thermo Fischer Scientific.

[54] Technical Note: Perkin Elmer FTIR Spectroscopy (007024B_01) *Perkin Elmer Life and Analytical Sciences* (2005).

[55] Gavina M. K. A., Rabajante J. F., Cervancia C.R. Mathematical Programming Models for Determining the Optimal Location of Beehives. Bulletin of Mathematical Biology **(2014)**, **76**(5): p. 997-1016.

[56] Killeen D.P., Sansom C. E., Ross E. L., Eason J. R., Gordon K. C., Perry N. B. Quantitative Raman Spectroscopy for the Analysis of Carrot Bioactives. Journal of Agricultural and Food Chemistry **(2013)**, 61(11): p. 2701-2708.

Chapter 3: Glasshouse experimental methods

3.1 Overview

The ten cultivars of *Leptospermum scoparium* used in these experiments are proprietary cultivars bred by Comvita New Zealand Limited. Their breeding history is important to note and the parental information of each cultivar used in the experiments is displayed along with photographs of the habit of each, which illustrates flower structure and plant growth habit for each cultivar.

Experimental designs allowing for any microclimatic variation within glasshouse conditions were established for the controlled glasshouse experiments. The decision was taken to use ten replicates (clone plants/cuttings) from each cultivar for each experiment to allow for a good confidence interval for the statistical tests on data results. Plant basal stem and plant heights were recorded at regular intervals to assess plant growth rates. Flowering periods, nectar production and nectar composition were assessed and quantified in each experiment.

3.2 Introduction

This chapter outlines the basics of experimental design, soil science information and additional plant breeding parentage information regarding the ten *Leptospermum scoparium* cultivars used in this thesis.

3.2.1 Genotypes

Ten different cultivars of *Leptospermum scoparium* were provided from the breeding programmes of Comvita New Zealand Limited. These cultivars are proprietary cultivars

with each cultivar having a different genotype bred from a range of *Leptospermum* parental stock. The ten cultivars are compared in the experiments in this thesis to determine what range of variation there is in terms of nectar composition, plant growth and flowering between different genotypes of *L. scoparium*.

3.2.2 Soil Composition

This research attempts to research the response of *L. scoparium* to ten different New Zealand soil types and establish whether soil type affects phenotypic expression in these ten different genotypes. New Zealand has a large range of soil types and climate for the size of the country due to the geographical nature of the country; having mountainous regions, volcanic regions and lowland plains along with its variation in climatic conditions from west to east and north to south. New Zealand is considered as having a temperate climate, although there is a large range of regional climates; regions where annual droughts are common, the far north regions of New Zealand are considered sub-tropical and other regions are considered as sub alpine in climate.

Formation and fertility of soil structure is a complex procedure involving properties of the bedrock beneath the soil, the climate and the historic geographical nature of the soil. For example; deposits from volcanic eruptions, flood erosion or deposits from floods such as silt and organic matter from the historical and current ecosystem will change the soil structure in an affected region. For these reasons, soil conditions can greatly vary throughout the country and are of important consideration for Mānuka husbandry. The level of success in Mānuka plantation planting is potentially influenced by soil conditions as it is with many crop plants.

Plant nutrition largely depends on the efficient cycling of nutrients within the soil to available mineral forms, such as nitrate, that the plant can easily absorb. High nitrogen

nutrition for example causes greater protein synthesis in the plant and results in an increase in vegetative growth, an important factor in the successful establishment of crops. Soil pH is also an important factor, in very acidic soils nitrification may be inhibited by deficiencies of nutrients such as calcium and magnesium and high levels of aluminium can cause toxicity in plants at a low soil pH. Plant-soil chemistry interactions and plant nutrition is a complex process and can be plant specific. Macro-nutrient and micronutrient composition of the soil, along with soil pH and the availability of these nutrients to the plant i.e. the cation exchange capability (CEC) of the soil is important information required to determine the potential success of crops grown on the range of soil types where Mānuka is being considered for plantation planting. Therefore analysis of soil chemistry along with researching plant responses to these soil variables is important when considering plantation sites for any plant species proposed for commercial use [5].

3.2.3 Experimental Design

Experimental design is the term used to describe the structure of an experiment. Nearly every scientific experiment is subject to random variation, some of which is predictable and some non-predictable, i.e. variation can be identified in some cases but can be unidentifiable in others. There can be spatial or temporal patterns within an experimental field site, glasshouse or growth chamber. Identifying the scale and sources of variation relative to the specific hypothesis of interest in the experiment is a critical component of good design, the design must also allow for any possible unknown variation that may occur during the length of the experiment enabling the generation of meaningful data that leads to meaningful inferences from that data [1, 2].

Glasshouse environments commonly have microclimate variations due to the aspect of the glasshouse position, structural parts of the glasshouse such as the building structure itself and the position of temperature control systems such as cooling fans and/or heating systems, all depending on the glasshouse design and the materials it is made from. Systematic blocks designs for the layout of plants within an experimental system are generally more effective in accounting for microclimate variation than rearrangement approaches of plant positioning and provide the optimal solution in experimental design for glasshouses that have microclimatic variability within them. Blocking designs can be grouped into two categories: one-way blocking (randomized complete block, lattice, and balanced incomplete block designs) or two-way blocking (Latin squares, Youden squares, Trojan squares, lattice squares, and row-column designs). In the experiments in this thesis latin blocking designs and latin square designs were used. Blocking was used for the three cultivars in the soils experiment with a latin square arrangement for the soils. A 10 x 10 latin square design was used for the experimental setup for assessing phenotypic differences between the ten different cultivars [2, 3, 4].

3.3 Plant Breeding History

Ten cultivars of *Leptospermum scoparium* were evaluated in controlled glasshouse conditions. The ten proprietary cultivars (Figure 3.1) were bred by Comvita, (Te Puke, New Zealand), who hold the parental stock. Each cultivar is represented by arbitrary code letters: Y, MG, G, B, BS, P, O, LG, PU and R. The parental genotypes of the cultivars analysed in these sets of experiments were generally from crosses of

Leptospermum scoparium var *scoparium* crossed with a range of other parental lines consisting of *L. scoparium* var *linifolium* and *L. scoparium* var *incanum*, (Table 3.1).

Figure 3.1. Floral Density habits of 10 cultivars BS, B, LG, MG, O, P, PU, R, Y and G.

Table 3.1. *Leptospermum scoparium* cultivar parentage information.

Massey Code	Plant code	Mother	Pollen donor	SUGGESTED DESCRIPTION
LG	27043-940	N6B305	C7VGSU	***L. scoparium* var. *scoparium***
MG	27033-908	N6A107	C8VSAN	***L. scoparium* var. *incanum***
O	26011-820	C6VELR	N6B304	***L. scoparium* var. *scoparium***
P	27043-931	N6B305	C7VGSU	***L. scoparium* var. *scoparium***
Y	27059-944	H1	Selfed	***L. scoparium* var. *scoparium***
B	26049-811	N6B304	N6A105	***L. scoparium* var. *scoparium***
BS	26012-824	C6VELR	N6C101	***L. scoparium* var. *linifolium***
PU	28040-012	C6VELR	CV8MER	***L. scoparium* var. *scoparium***
G	28076-029	C6VKEA	Open	***L. scoparium* var. *incanum***
R	28060-006	C8VSAN	C6VKEA	***L. scoparium* var. *incanum***

To follow are sets of photographs of the habit and flowers each of the ten cultivars of *Leptospermum scoparium* (Figures 3.2 to 3.11).

Figure 3.2. Cultivar 26012-824 Cultivar X *L. scoparium* var *linifolium* selection.

Figure 3.3. Cultivar 26049-811 *L. scoparium* var *scoparium* selection.

Figure 3.4. Cultivar 28076-029 *L. scoparium* var *incanum* cultivar (open cross).

Figure 3.5. Cultivar 27033-908 *L. scoparium* var *incanum* selection.

Figure 3.6. Cultivar 27043-940 *L. scoparium* var *scoparium*.

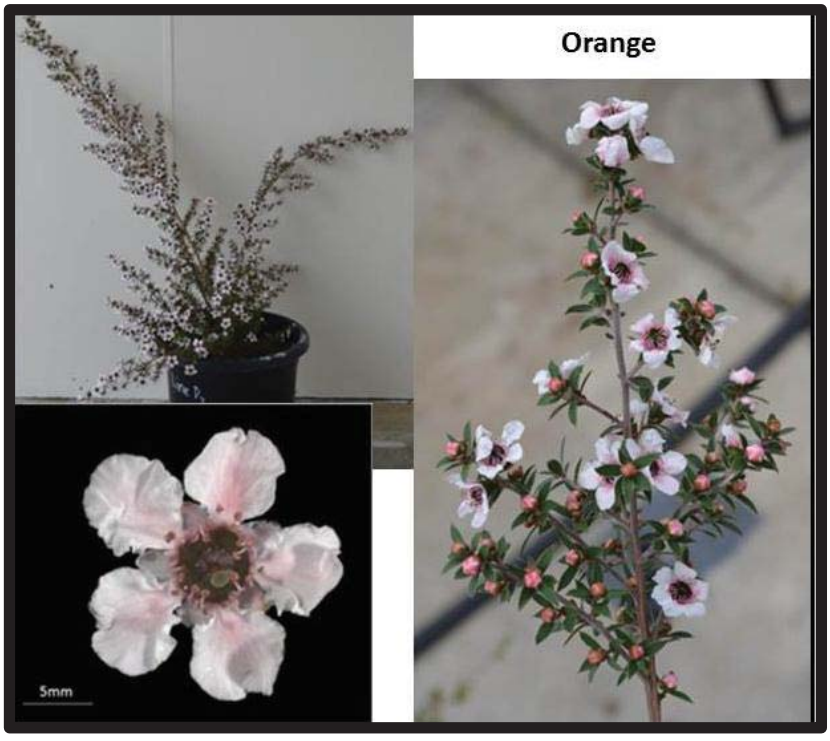

Figure 3.7. Cultivar 26011-820 *L. scoparium* var *scoparium* selection.

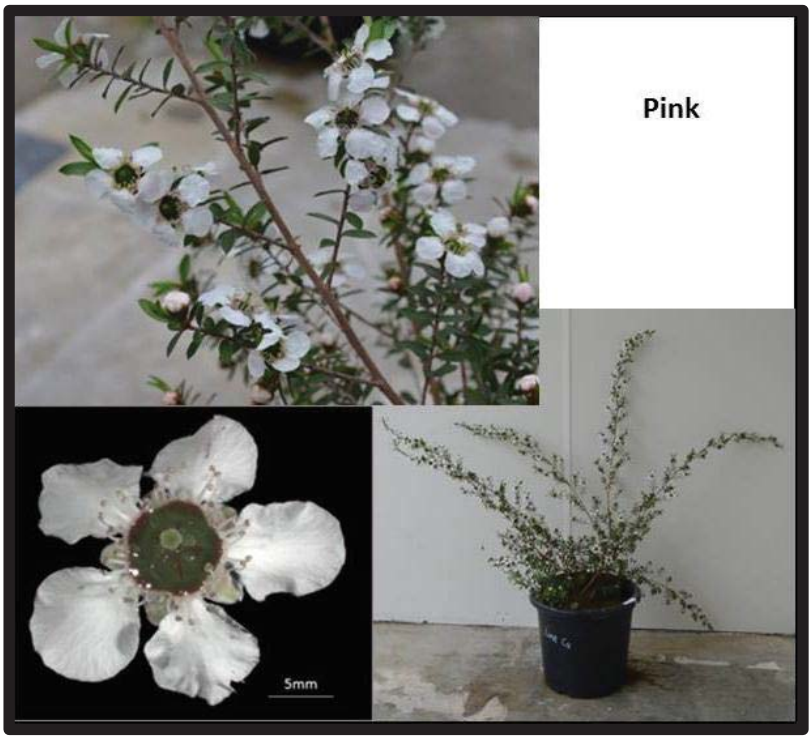

Figure 3.8. Cultivar 27043-931 *L. scoparium* var *scoparium.*

Figure 3.9. Cultivar 28040-012 *L.scoparium var. scoparium.*

Figure 3.10. Cultivar 28060-006 *L. scoparium* var *incanum* selection.

Figure 3.11. Cultivar 27059-944 *L. scoparium* var *scoparium* field selection.

3.4 Glasshouse Plant Conditions Experiment 1: Genotype Comparisons

10 different Cultivars of *Leptospermum* were grown on standard tree and shrub mix with a peat and pumice base with added fertiliser at rates below per 100 litres of soil mix: Per 100 L of peat and pumice base was added: 300 g Long-Term Fertiliser, 60 g Short-Term Fertiliser, and 150 g Dolomite. Tables 3.2, 3.3 and 3.4 show the chemical composition of the added fertilisers.

Table 3.2. Chemical composition of Long-Term Fertiliser.

Total Nitrogen (N)	**18%**
	2.0% Ammoniacal Nitrogen 5.0% Water Insoluble Nitrogen 2.2% Urea Nitrogen 8.8% Other Water Soluble Nitrogen
Available Phosphate (P_2O_5)	**5%**
Soluble Potash (K_2O)	**10%**
Magnesium (Mg)	**0.5%**
	0.02% Water Soluble Magnesium (Mg)
Sulphur (S)	**3.2%**
	3.2% Combined Sulfur (S)
Iron (Fe)	**1.6%**
	0.2% Water Soluble Iron (Fe)
Manganese (Mn)	**0.3%**
	0.04% Water Soluble Manganese
Chlorine (Cl) not more than	**2%**

Table 3.3. Chemical composition of Short-Term Fertiliser.

Total Nitrogen (N) 14%
2.8% Ammoniacal Nitrogen
5.8% Water Insoluble Nitrogen
2.7% Urea Nitrogen
2.7% Other Water Soluble Nitrogen
Available Phosphate (P_2O_5) 14%
Soluble Potash (K_2O) 14%
Magnesium (Mg) 1.0%
1.0% Water Soluble Magnesium (Mg)
Sulphur (S) 4.0%
4.0% Combined Sulfur (S)
Iron (Fe) 1.0%
1.0% Water Soluble Iron (Fe)
Manganese (Mn) 0.5%
0.5% Water Soluble Manganese (Mn)
Chlorine (Cl) not more than 2.0%

Table 3.4. Ingredients of Dolomite.

Calcium (Ca)		**21%**
Magnesium(Mg)		**10%**
Calcium Carbonate Equivalent (CCE)		**89%**
Inert Binder		**8%**
Total Carbonates Derived from Dolomitic Limestone		**96%**
Calcium Carbonate (CaCO3)……	57%	
Magnesium Carbonate (MgCO3)…	39 %	

Plants were grown in 30 cm diameter pots, the soil mix was weighed into each pot so that each pot contained an equal volume of soil. 10 replicate plants of each cultivar were planted and plants were irrigated via inline feeds with a line into each individual pot to ensure equal watering to each pot (Figure 3.12).

Figure 3.12. Picture showing in-line irrigation system to each individual pot.

Figure 3.13. Glasshouse plant layout for experiment 1.

The experiment ran for a time length of 15 months. At the 9 month stage each pot was in addition irrigated by an individual plant feed system supplemented with liquid feed using Peters Liquid feed formulae (Table 3.5) via the same irrigation lines at 0.5 g/L. Plants were therefore grown with optimal nutrition in glasshouse conditions.

Table 3.5. Peters Liquid feed formulations.

20 : 9 : 17 NPK + Trace Elements
20% Nitrogen (N): 4.5% as Nitrate nitrogen 2.4% as Ammonium nitrogen 13.1% as Urea nitrogen
8.7% Phosphorus (p) as ammonium phosphate
16.5% Potassium (K) as potassium nitrate
Trace elements are completely water soluble
0.12% Iron (Fe) chelated by DTPA
0.06% Manganese (Mn) chelated by EDTA
0.02% Boron (B)
0.015% Copper (Cu) chelated by EDTA
0.015% Zinc (Zn) chelated by EDTA
0.01% Molybdenum (Mo)

3.5 Glasshouse Experimental Design

Plant pots were arranged inside the glasshouse in a latin square design (Figures 3.13 and 3.14). The blocking in the latin square design occurs along two axes. It therefore enabled control for environmental variability along two gradients rather than one. In controlled environments, mainly in greenhouses, it is often the case that one environmental factor varies in this condition and the latin square design is considered the best design to account for variation in glasshouse conditions. It assumes no

interaction between the variables i.e. the experimental units. The design requires the same number of rows and columns. We used a 10 x 10 grid with each treatment factor found in each row and each column [3, 4].

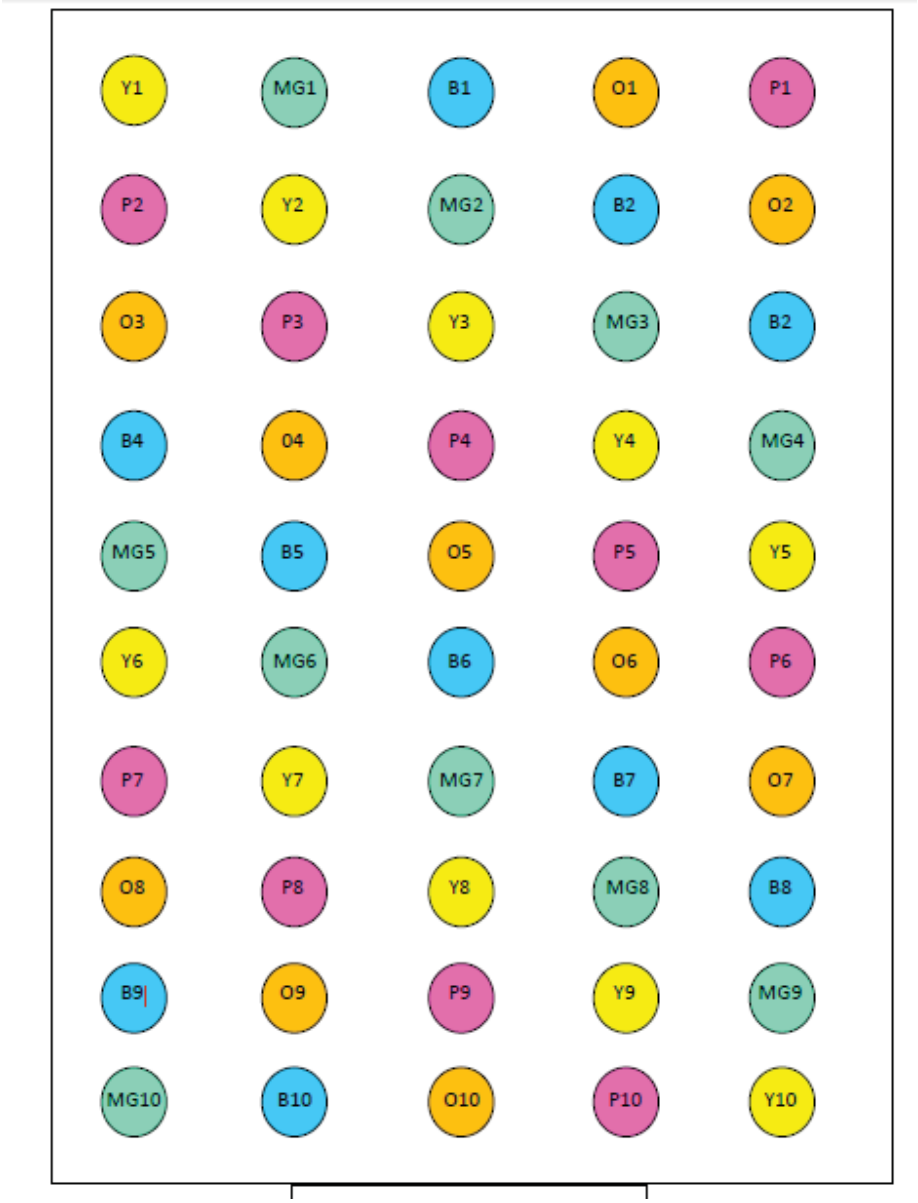

Figure 3.14. Example of a Latin Square Design used in experiment 1, in this case with two blocks of a 5 x 5 grid. In the actual experiment a 10 x 10 grid was used.

The glasshouse was fan ventilated and insect protected from bees and wasps and pots sprayed with insect repellent to prevent crawling insect visitation.

Plant growth parameters of basal stem measurement and plant height data were gathered at 3 month intervals throughout the experiments, and fresh weight and dry weight biomass data was collected at experiment end with flowering data and nectar were collected throughout the experiment when plants were flowering.

Temperature and humidity data was collected using HortPlus Microloggers installed inside the glasshouse. Total solar radiation and hours of sunshine data were downloaded from the National Institute of Water and Atmospheric Research weather station located 500 m from the glasshouse experiment at site 2193 grid reference T24315885; Latitude -40.38195; Longitude 175.60915.

3.6 Glasshouse Experiment 2: Soils

This research was conducted under controlled glasshouse conditions using three of the available *Leptospermum scoparium* cultivars representing a range of DHA potential based on Comvita preliminary data and grown on a range of New Zealand soils (10) over a 15 month period.

Soil was collected from nine different locations within New Zealand that were representative of regions that might be suitable for Mānuka plantations, these locations are Lincoln, Ahu Ahu Valley, Pahiatua, Ruatiti, Takaka, Rangitatau, Te Awaiti, Wharekohe and Waikato peat-land.

At each location a composite bulk sample (200 kg) was collected from 10 random points within a one hectare location. One square meter was dug for each collection point and soil was sampled down to 30 cm. The top 5 cm of soil was discarded from collection to remove any grass or herbage. A commercial potting media was used as the control soil in the study as an example of optimum nutrition. Soil collected from

each location was sieved and roto-mixed to ensure even homogenous distribution before transfer into 300 mm diameter plant pots. The total pot volume for each soil type was matched at 8 litres per pot to standardise soil volume across all soils. Three proprietary cultivars of Mānuka, labelled R, G and Y were planted into each soil type and grown for 15 months. All treatments (cultivar and soil combinations) were replicated ten times (300 pots in total). Each cultivar set were genetically identical clones of the parent material. Cuttings were supplied by Comvita New Zealand and were transferred at age 3 months to the potted soil. The soil experiment was conducted at the Massey University Plant Growth Unit; Palmerston North, New Zealand. Plants were watered with in-line irrigation to ensure each pot received the same amount of water (Figure 3.15). All experimental units were planted into one greenhouse in a randomised latin block design (Figure 3.18). The glasshouse was ventilated by a fan system to provide a climate of mild temperate conditions. Plant response parameters measured were: basal stem diameter, plant height, relative growth rate (RGR), nectar DHA concentration normalised to 80oBRIX, nectar volume, start of flowering, flowering period and seed capsules numbers. The glasshouse trial ran for a period of 15 months (Figures 3.16 and 3.17), growth parameters were measured 3 monthly over a 15 month period from the 1st of October 2012 until January 2014. Nectar was collected when plants flowered and seed capsules numbers were counted near the end of the experiment in December 2014.

Figure 3.15. In-line watering system for soil experiment and later photos showing grown plants and flowering in progress.

Figure 3.16. Experiment 2 mid experiment showing plants growing on the ten different soils.

Figure 3.17. Experiment 2 mid experiment showing plants in flower during the soils experiment.

	Soil Type									
Cultivar R	1	2	3	4	5	6	7	8	9	10
Cultivar G	10	1	2	3	4	5	6	7	8	9
Cultivar Y	9	10	1	2	3	4	5	6	7	8
Cultivar R	8	9	10	1	2	3	4	5	6	7
Cultivar G	7	8	9	10	1	2	3	4	5	6
Cultivar Y	6	7	8	9	10	1	2	3	4	5
Cultivar R	5	6	7	8	9	10	1	2	3	4
Cultivar G	4	5	6	7	8	9	10	1	2	3
Cultivar Y	3	4	5	6	7	8	9	10	1	2
Cultivar R	2	3	4	5	6	7	8	9	10	1
Cultivar G	1	2	3	4	5	6	7	8	9	10
Cultivar Y	10	1	2	3	4	5	6	7	8	9
Cultivar R	9	10	1	2	3	4	5	6	7	8
Cultivar G	8	9	10	1	2	3	4	5	6	7
Cultivar Y	7	8	9	10	1	2	3	4	5	6
Cultivar R	6	7	8	9	10	1	2	3	4	5
Cultivar G	5	6	7	8	9	10	1	2	3	4
Cultivar Y	4	5	6	7	8	9	10	1	2	3
Cultivar R	3	4	5	6	7	8	9	10	1	2
Cultivar G	2	3	4	5	6	7	8	9	10	1
Cultivar Y	1	2	3	4	5	6	7	8	9	10
Cultivar R	10	1	2	3	4	5	6	7	8	9
Cultivar G	9	10	1	2	3	4	5	6	7	8
Cultivar Y	8	9	10	1	2	3	4	5	6	7
Cultivar R	7	8	9	10	1	2	3	4	5	6
Cultivar G	6	7	8	9	10	1	2	3	4	5
Cultivar Y	5	6	7	8	9	10	1	2	3	4
Cultivar R	4	5	6	7	8	9	10	1	2	3
Cultivar G	3	4	5	6	7	8	9	10	1	2
Cultivar Y	2	3	4	5	6	7	8	9	10	1

Figure 3.18. Glasshouse layout for investigating the influence of soil type on growth and flowering in *Leptospermum scoparium.* Ten soil types were investigated and three different cultivars. Ten replicate plants of each cultivar arranged in a blocked latin grid. Each treatment factor has to be found in each row and each column. The treatment appears once in each block in a latin square position. We used 10 blocks, each containing the 3 cultivars in ten replicates.

3.6.1 Soil Analysis

A soil corer with an internal volume of 80 cm^3 was used to collect a soil sample from each pot in the glasshouse at the end of the experiment (January 2014). Replicates 1-3; 4-6 and 7-10 for each treatment unit were combined, mixed then air dried and sieved to less than 2 mm to generate 3 replicate soil samples for each soil/cultivar combination.

Soil samples were analysed for soil volume (g/ml), pH, cation exchange capacity (CEC), macronutrients; phosphorus (P), sulphate ($SO_{4)}$, potassium (K), calcium (Ca), magnesium (Mg), sodium (Na) by the Fertilizer and Lime Research Centre, Palmerston North according to the methods of Blakemore et al. 1987 [6]. pH was measured in water at a soil: water ratio of 10.25 g/ml. Plant available phosphorus was assessed using the Olsen P method with bicarbonate extraction followed by the Murphy and Riley (1962) [7] colorimetric test. S was assessed according to the method of Johnson and Nishita (1952) [8] calcium-phosphate extractable S but with a modern auto-analyser. CEC and cations; Ca^{2+}, Mg^{2+}, Na^{2+} and K^{+} were assessed by semi-micro leaching with 1M ammonium acetate H^{+} approximated by pH drop elemental analysis by atomic absorption spectrophotometry (AAS).

The plant available micronutrients boron (B), chlorine (Cl), iron (Fe), manganese (Mn), zinc (Zn), copper (Cu) and cobalt (Co) were quantified by Hills Laboratories; Waikato Innovation Park, Ruakura Lane, Hamilton. Soils for micronutrient analysis were air dried at 35-40^{o}C overnight with residual moisture typically 4%. Boron was assessed by extraction with boiling dilute calcium chloride followed by ICP-OES (Inductively coupled plasma atomic emission spectroscopy), Chloride was assayed by saturated calcium sulphate extraction followed by Potentiometric Titration. The plant-available

concentration of iron, manganese, zinc, copper and cobalt was determined by 0.05M EDTA extraction followed by ICP-OES analysis of the extract.

3.6.2 Soil Chemistry

Soil collection sites for the soils experiment are listed in table 3.6, the table also gives general information on the soil type from each collection site and a brief rationale for choosing the site. Table 3.7 displays the soil chemistry data from micro and macro nutrient analysis of each soil type. The fertility parameter Olsen P quantifies the concentration of plant available phosphorus in soil. The parameter CEC defines cation exchange capacity, and quantifies the ability of soil to retain a range of important nutrients. CEC is a function of the structure of the soil and is highly dependent on organic matter content. Fe, Mn, Zn, Cu, Co are plant-available EDTA extractable measures. B, Cl, Mo, Se are total soil measures. Units of me/100 g represent milli-equivalents of exchangeable base cation per 100 g of soil, and reflect the degree of saturation of total sediment CEC that is associated with these essential macro-nutrients. Units of me/100 g are standard for soil fertility testing in NZ, and represent plant-available concentrations of nutrient rather than total soil concentration.

Table 3.6. Soil collection sites with information on soil types.

Code	Soil Location NZTM Co-ordinates	Soil Classification	Soil profile texture group	Parent Material
S1	MAXWELL: NZTM 1760836E 5596919N	Orthic Brown	Brown Soil	Sandstone and limestone: Iron Oxide
S2	LINCOLN: NZTM 1554648E 5168342N	Typic Immature Pallic Soils	Silty	Alluvium
S3	WANGANUI: NZTM 1774789E 5610367N	Orthic Brown	Brown Soil	Sandstone/siltstone
S4	PAHIATUA: NZTM 1834148E 5519656N	Pallic Orthic Brown	Silty Loam	Loess on rock
S5	CONTROL	Commercial growth media (potting mix)	Organic	Peat and Pumice
S6	RUATITI: NZTM 1790106E 5647255N	Allophanic Brown	Brown Soil	Sandstone and Volcanic ash
S7	TAKAKA: NZTM 598371E, 5450593N	Sandy Recent	Sandy Loam	Alluvium
S8	TE AWAITI: NZTM821261E 5409918N	Orthic Brown	Brown Soil	Sandstone and limestone: Iron Oxide
S9	WHAREKOHE: NZTM 1699578E 6042684N	Densipan Ultic	Clay	Quartz rich Rock
S10	WAIKATO PEAT: NZTM 808553E, 5800648N	Fibric Organic	Organic	Peat

Table 3.7. Soil composition data from site collected soils.

	Soil	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	NZ medium range
Fertility Parameter	unit											
pH		6.10	5.50	5.30	5.60	5.20	5.40	6.00	5.70	5.30	5.10	5.8-6.2
Olsen P	mg/L	12.70	30.60	9.20	5.70	99.80	23.90	43.60	11.30	33.10	41.90	15-25
SO_4	mg/kg	6.80	53.70	14.20	8.20	1200.00	26.50	14.70	25.90	38.80	62.80	
K	me/100g	0.36	0.45	1.17	0.61	5.99	0.50	0.51	0.81	0.15	0.30	0.3-0.5
Ca	me/100g	9.70	70	3.70	2.40	30.60	3.00	4.40	3.10	6.80	69.90	3.0-9.0
Mg	me/100g	2.12	1.10	1.13	0.76	7.21	0.42	1.40	2.05	0.96	2.35	0.8-1.5
Na	me/100g	0.20	0.20	0.13	0.06	1.02	0.07	0.14	0.29	0.06	0.09	0.2-0.4
CEC	me/100g	17.00	16.00	20.00	17.60	77.00	18.00	9.00	13.00	15.00	123.00	
Volume	g/mL	1.00	1.10	0.91	0.77	0.51	0.85	1.15	1.02	0.96	0.44	0.6-1.00
B	mg/kg	0.64	0.44	0.62	0.93	4.07	0.63	0.28	0.52	0.39	1.48	CL >0.4
Cl	mg/kg	19.88	17.89	30.33	54.22	441.00	47.89	6.67	16.67	21.88	no data	-
Fe	mg/kg	764.75	758.33	385.00	329.89	1392.78	314.78	761.67	411.78	597.25	713.71	-
Mn	mg/kg	187.63	161.78	121.33	22.11	138.22	59.44	62.33	57.00	5.88	17.29	CL>50
Zn	mg/kg	3.94	3.23	3.22	1.46	27.50	3.57	2.51	1.99	2.76	28.10	CL>1.1
Cu	mg/kg	3.04	1.59	3.90	1.24	46.87	4.09	7.60	0.94	1.31	7.69	CL>0.8
Co	mg/kg	1.34	1.27	0.91	0.33	0.80	0.40	2.14	0.50	0.00	0.50	CL>1.0

3.7 References

[1] Casler M.D., Vermerris W., Dixon R.A. Replication Concepts for Bioenergy Research Experiments. Bioenergy Research **(2015)**, 8(1): p. 1-16.

[2] Casler M.D. Fundamentals of Experimental Design: Guidelines for Designing Successful Experiments. Agronomy Journal **(2015)**, 107(2): p. 692-705.

[3] Gurevitch J., Scheiner S.M. Design and Analysis of Ecological Experiments.Oxford University Press. Inc. New York. **(2001)**.

[4] Krebs C. J. Ecological Methodology. 2nd Edition. Benjamin/Cummings. **(1999)**.

[5] McLaren R.G., Cameron K.C. Soil Science. 2nd Edition. Oxford University Press. Inc. **(1996)**.

[6] Blakemore L.C., Searle P.L., Daly B.K. Methods for Chemical Analysis of Soils. New Zealand Soil Bureau Scientific Report. **(1987)**, 80(103).

[7] Murphy J., Riley J.P. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta. **(1962)**, 27: p. 31-36.

[8] Johnson C.M., Nishita H. Micro-estimation of sulphur in plant materials, soils and irrigation waters. Analytical Chemistry. **(1952)**, 24: p. 736-42.

Chapter 4: Chemometrics: An outline of the development of the chemometric workflow applied in the analysis of Mānuka plant material

4.1 Overview

The workflow and methods used to collect and investigate data from analytical spectroscopic techniques is an important aspect in chemometric analysis. Knowledge is derived from experiments and measurements made on experimental systems and the analysis of the resulting data. This chapter describes the multivariate methods developed and applied to analyse the plant metabolite and nectar data generated from the analytical techniques used in the research experiments, particularly spectroscopy methods but also multivariate datasets from other experimental work.

4.2 Introduction

Fourier Transform Infrared spectroscopy (FTIR) and Fourier Transform Raman spectroscopy (FT-Raman) are complementary spectroscopy techniques commonly used in food analysis and chemotyping of plant species, they both utilise the capability to detect chemical bond vibrations within a chemical component and present a spectral signature from the compound of interest. Chemical bond types vibrate at unique wavelengths causing a spectrum of vibrational peaks that are unique to individual compounds. Both FTIR and FT-Raman techniques are rapid and non-destructive requiring minimal sample preparation before analysis [1-6].

Chemometrics is the name given to the science of extracting data from chemical systems such as FTIR and FT-Raman spectroscopy using a series of multivariate statistical data analysis applications [7]. Chemometric techniques are heavily used in analytical spectroscopy and multivariate statistical analysis has become a powerful technique for analysing data with many and/or complex variables simultaneously, enabling identification of patterns and relationships between those variables and their relevance to the system under investigation [7]. With the development of analytical spectroscopy instrumentation and chemometric methods, multivariate predictive modelling techniques such as principal component analysis (PCA) and partial least squares (PLS) can be applied to reduce the dimensionality and reveal the structure of data allowing identification of the minimum number of variables that describe a sample set [8-12]. PLS methods in addition allow quantification of chemical constituents in samples, reducing the demand for traditional analytical methods such as high pressure liquid chromatography (HPLC) that are more time consuming and expensive [10, 12].

Knowledge is generated through the management, analysis, and interpretation of these large volumes of data [13]. Multivariate analysis is a data reduction technique that can be used to simplify the analysis of complex datasets, mainly classification and grouping of observations, allowing ease of investigating the relationships between data and modelling those relationships [14].

In spectroscopic analysis, typical spectral data are collected across a range of wavenumbers that are dependent on the analytical system used. Thus the X variables represent wavelengths (wavenumber cm^{-1}) and in this case are from FT-Raman spectra or FTIR-attenuated total reflectance (ATR) absorbance spectra and the Y variables are

either the relative intensity in arbitrary units (a.u.) of the Raman scattering or absorbance from infrared methods of the analyte of interest. In partial least squares (PLS) analysis in this thesis there is X data of the spectral data matrix i.e. the wavenumbers and the relative intensity and the Y data matrix is the variables of various quantified values of plant components being analysed and compared, i.e. nectar sugars and leaf metabolites [8].

The workflow and methods used to collect and investigate analytical data from spectroscopic techniques is an important aspect when applying chemometric analysis to spectral data.

4.3 Workflow and Methods

The important aspects of workflow when using multivariate data analysis techniques are designing the experiment, collecting data, making a model from correlative information, and applying the model to new samples and testing its predictive capability [7, 9, 15]. A number of steps encompassing mathematical processing of the spectral data are typically carried out in sequence in order to develop robust predictive models.

Complex and powerful computer analysis software are used to analyse correlative information and related statistics. In this investigation, the software package “The Unscrambler® X version 10.2” (CAMO Software AS, Oslo, Norway) was used to analyse ATR-FTIR and FT-Raman spectral data.

4.3.1 Preprocessing Spectral Data

Spectral data pre-processing is an important first step in the workflow of FTIR and FT-Raman spectral analysis, it involves specific processing procedures performed on the

raw spectral data after optimal data collection on relevant instrumentation. Pre-processing of spectral data is the most important step before chemometric modelling techniques such as principal component analysis (PCA) and partial least squares (PLS) are applied [15]. Chemometric models developed from pre-processed data perform better than models that simply use raw spectral data, pre-processing has been shown to be of crucial importance for subsequent data mining tasks [15, 16].

Pre-processing spectral data removes artefacts that are interfering with the structure of the spectral data, figures 4.1 and 4.2 display two examples of artefact noise that interfere with spectral data structure. Pre-processing improves the robustness and subsequent analysis of data especially with respect to quantification modelling and classification [8, 10, 12, 17, 18]. Pre-processing can detect and remove outlying spectral data and allows the removal of redundant spectral data. Redundant spectral data is wavenumber regions of the spectrum that are not contributing to the analysis, such as regions where there are no spectral peaks or regions where the noise levels are too high and are therefore interfering with the data analysis i.e. we do not want to be modelling noise ratios [15].

Spectral data consists of two elements, the spectral structure which is the relevant peak wavenumber information plus any noise that overlays that information. The aim of pre-processing is to remove as much of the noise interference as possible to allow the structural integrity of the spectral data to be the dominant feature analysed [8, 12, 15].

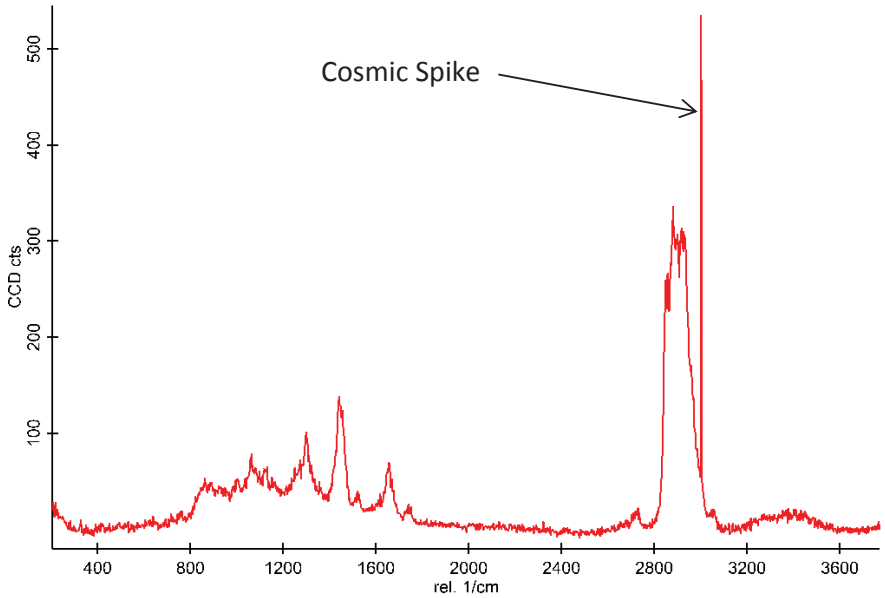

Figure 4.1. Cosmic spikes are visible as very narrow, high relative intensity peaks in your spectral data. Charge-coupled device detectors (CCD) utilised by Raman instruments detect occasional spikes caused by cosmic rays. Cosmic spikes can disturb the chemical information expressed by normal Raman spectra.

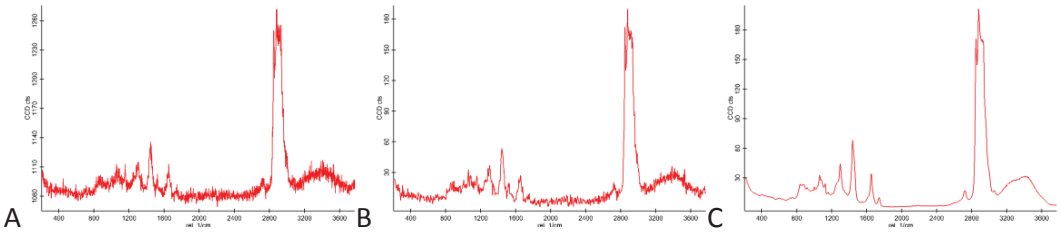

Figure 4.2. Illustration of high frequency noise in spectral data; (A) raw spectrum, (B) spectrum after a smoothing algorithm is applied and (C) after averaging several individual spectra of the same sample.

The first step in pre-processing spectral data is to visually assess the data for spectral artefacts using line plots of the full spectrum of data or spectral regions of interest. The spectral signature of artefacts can be obviously different from your main data set, for example, broad high intensity peaks from interference of fluorescent signal (Figure 4.3) or they can be flat-lined spectra showing no spectral features (Figure 4.3a).

Artefacts can be caused by heat from the laser power source on the instrument burning samples during scanning, optimisation of data collection by reducing laser power settings and adjusting scan integration times can overcome artefacts caused by burning, a compromise between the signal and noise ratio (Figure 4.2) can be a result of optimisation for some sample types. Other artefacts can be from scattering effects caused by particle size [7, 16], or in the case of ATR-FTIR lack of full contact between the sample and the ATR crystal increases scattering effects. Spectral data with obvious artefacts can be assessed and removed from the dataset prior to pre-processing if it is assessed that the artefacts are instrumental and not due to component variation in the sample.

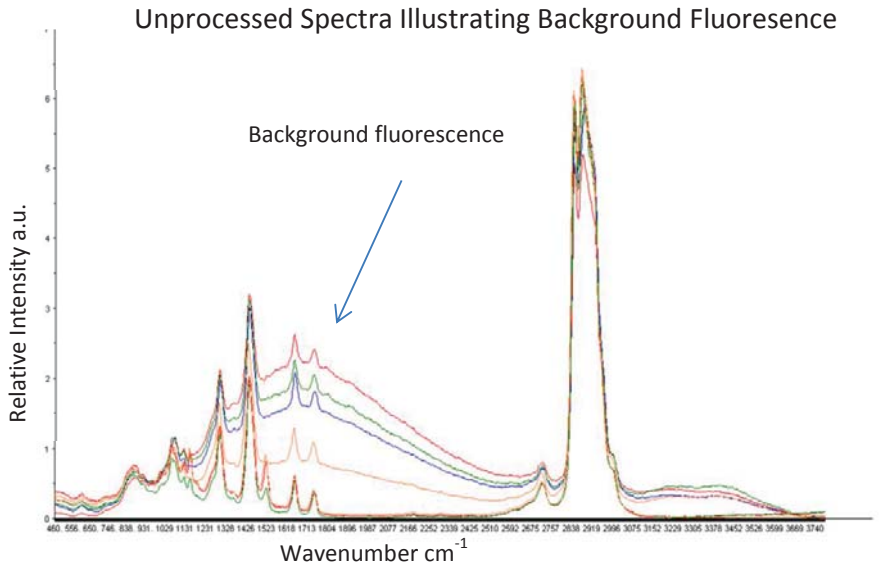

Figure 4.3. Example of spectral data of an organic sample set showing some spectra with fluorescent signal overlaying in the background resulting in a broad band of increased height through the middle region of the spectrum with the FT-Raman peaks still showing above the background band.

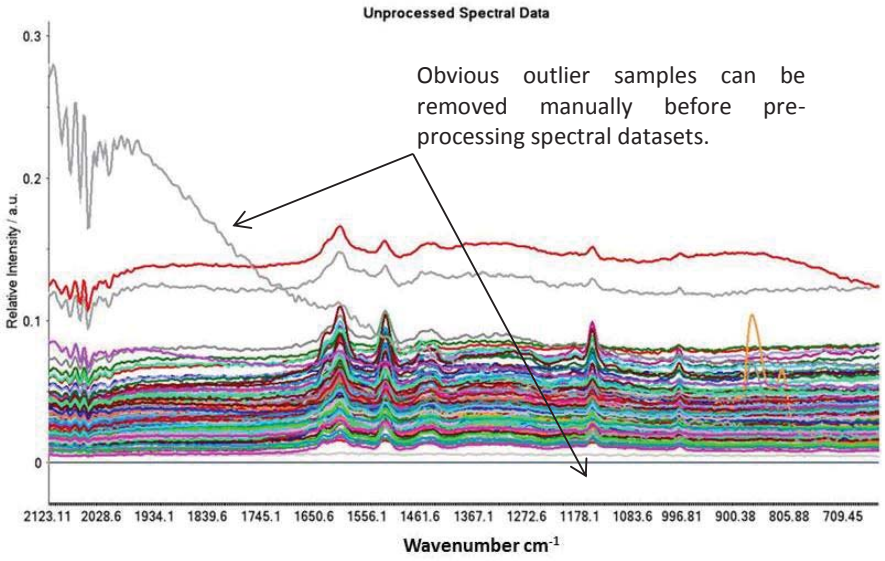

Figure 4.3a. Full FT-Raman spectral data set plot from a large range of samples. Note two obvious artefact outliers in the sample set.

Pre-processing uses chemometric techniques that are applied to the full body of spectral data or subsets of the spectral data i.e. sub-regions of the spectrum. There are many options for processing data in multivariate analytical software. In this section I will discuss several options applied within this thesis.

4.3.2 Baseline Adjustments

Baseline processing of spectral data can be applied to remove the vertical shift (additive scatter) in the spectra that is caused by thermal effects from the instrumentation or particle size effects. Baseline corrections are used to adjust the spectral offset by either adjusting the data to the minimum point in the data or by making a linear correction based on user defined variables, linear baselines can transform a sloped baseline to a horizontal baseline [15, 16, 18]. Compare figures 4.4a and 4.4b before and after baseline processing, post removal of artefact spectral data.

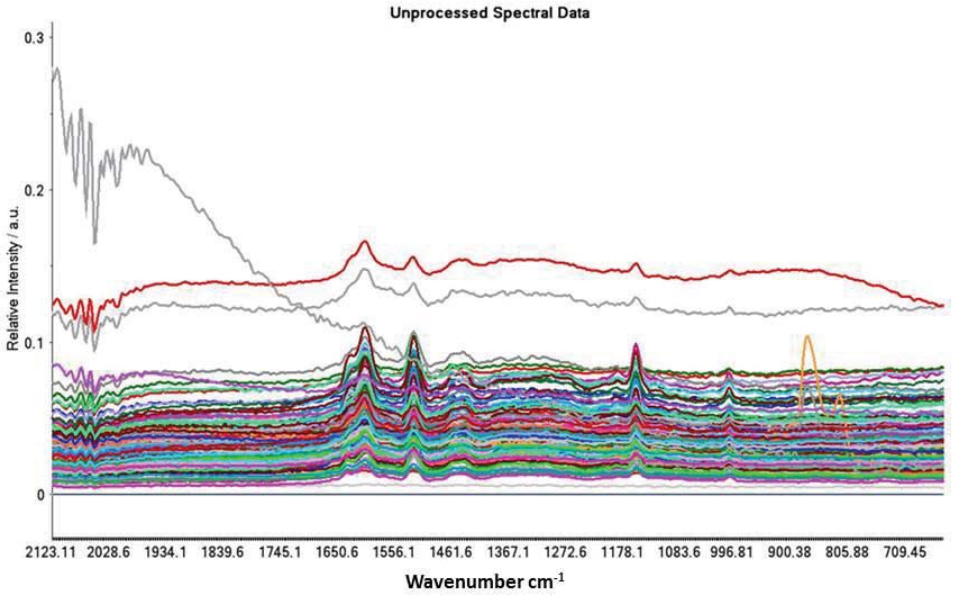

Figure 4.4a. Un-processed FT-Raman spectral data showing vertical shift in the spectral dataset.

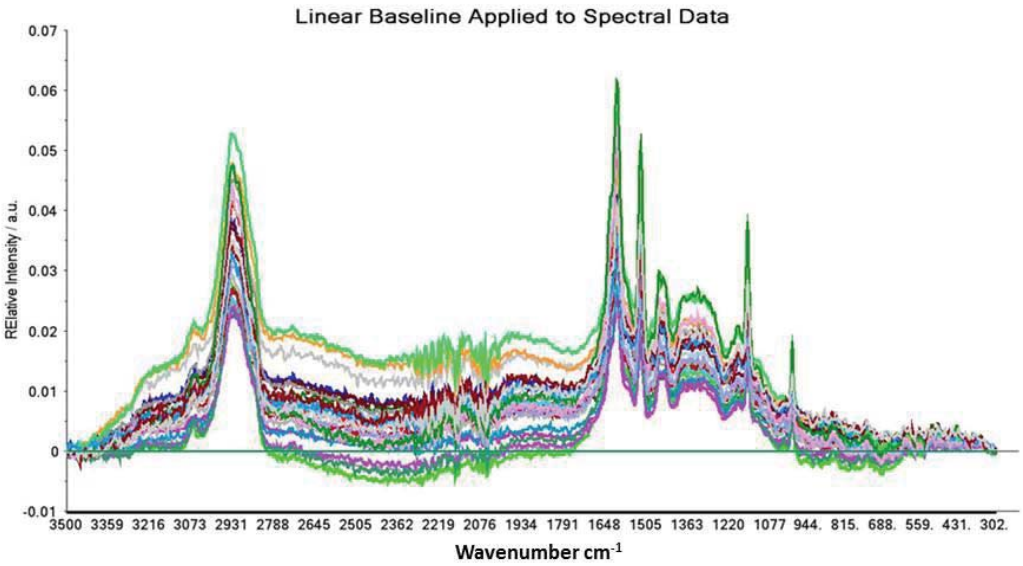

Figure 4.4b. FT-Raman spectral data set after linear baseline processing. Note alignment of most spectral data to the zero baseline.

4.3.3 Scatter Correction

Light scatter effects in spectroscopy are caused by the physical parameters of the sample, such as particle size, rather than its chemical properties. Scatter interferes with the relationship between chemical properties and the shape of the spectrum. There are two types of scatter in spectral data: additive (offset), which are generally vertical baseline shifts in spectral data and multiplicative (amplification) which show up as amplifying linear shifts in the baseline within the spectral region, they can dominate over the chemical signatures [7]. In these experiments, generally standard normal variate (SNV) processing rather than multiplicative scatter correction (MSC) was applied as the scatter did not fit the classic fan-shaped spread of spectral data that is indicative of multiplicative scatter, figures 4.5 and 4.6 illustrate the multiplicative scatter effects application (CAMO, 2012).

4.3.4 Standard Normal Variate (SNV)

SNV is a processing transformation that serves to remove the interference of scatter and particle size effects across the spectral range, separating the informative absorbance/reflectance data of the analyte and the scattering signal in the spectral data. SNV can correct for both additive and multiplicative spectra, whether MSC or SNV is applied requires investigation of the scatter effects in the samples and the effects of the pre-processing transformations on the data as to which best removes scatter effects.

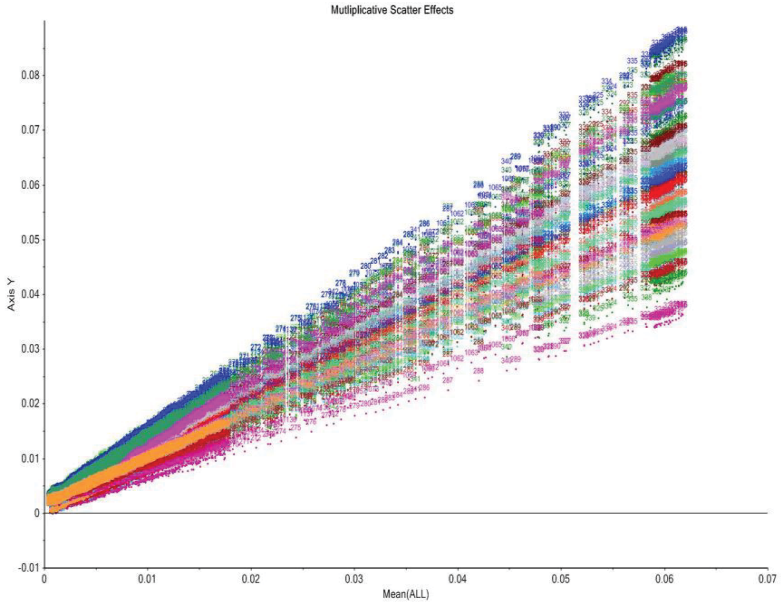

Figure 4.5. CAMO software data analysis showing MSC scatter effects in spectral data, note the classic fan shape in the descriptive data. This would indicate that application of the MSC transformation may be more useful than SNV.

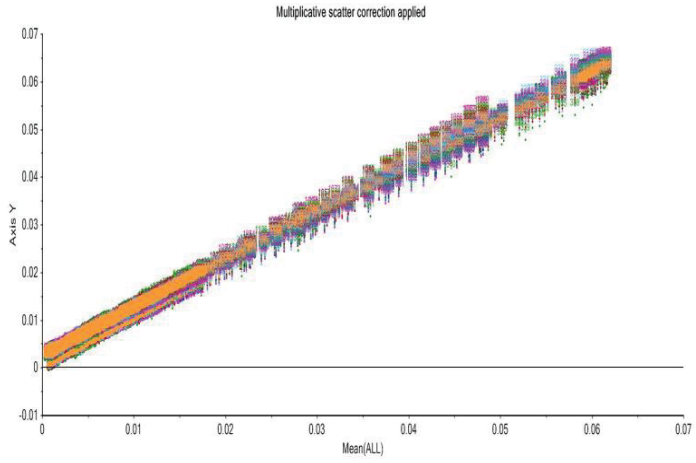

Figure 4.6. CAMO software data analysis showing data scatter after MSC transformation has been applied. Note that the data is now aligned.

SNV processing is applied to every spectrum individually. The average and the standard deviation of all data points for each spectrum are calculated. Every data point of the spectrum is then subtracted by the mean and divided by the standard deviation.

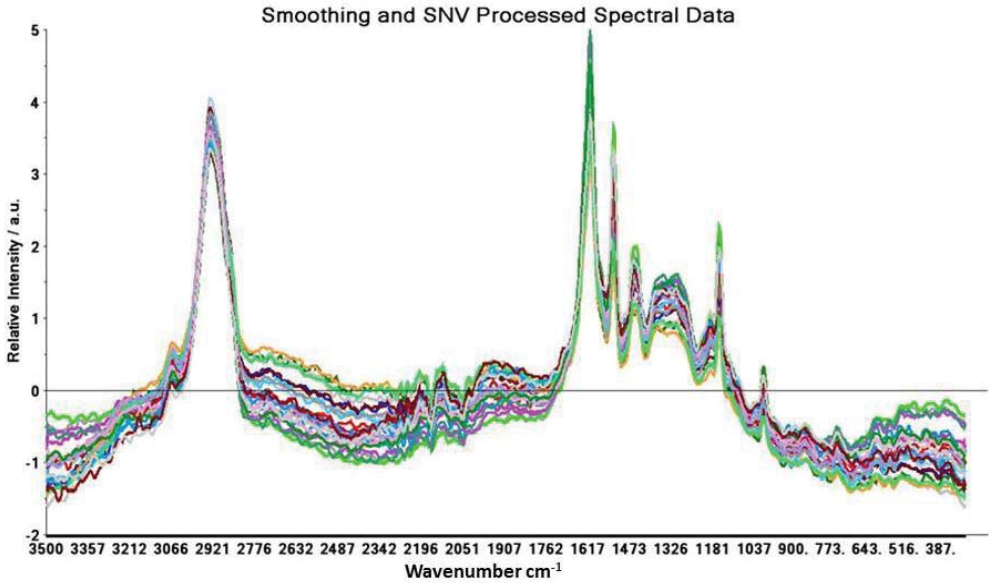

Figure 4.7. Line plot of spectral data after smoothing and alignment using SNV processing.

4.3.5 Detection of Outliers in Spectral Data

Various methods are available in statistical analytical software to detect sample outliers. In this case, hotelling with a T^2 statistical analysis was used. Hotelling has a linear relationship with the leverage of the samples and a *P* value critical limit can be applied. In this case, a *P* value of 5% was applied equalling a 95% confidence level, indicated by a red line, marking samples outside the 95th percentile (Figure 4.8). Each principal component (PC) was examined separately, all samples above the 95% confidence level are marked and the PCA was then performed without the marked samples and results compared with the previous PCA model. The PCA can also be performed by down-weighting the influence of the outliers instead of removing them all together.

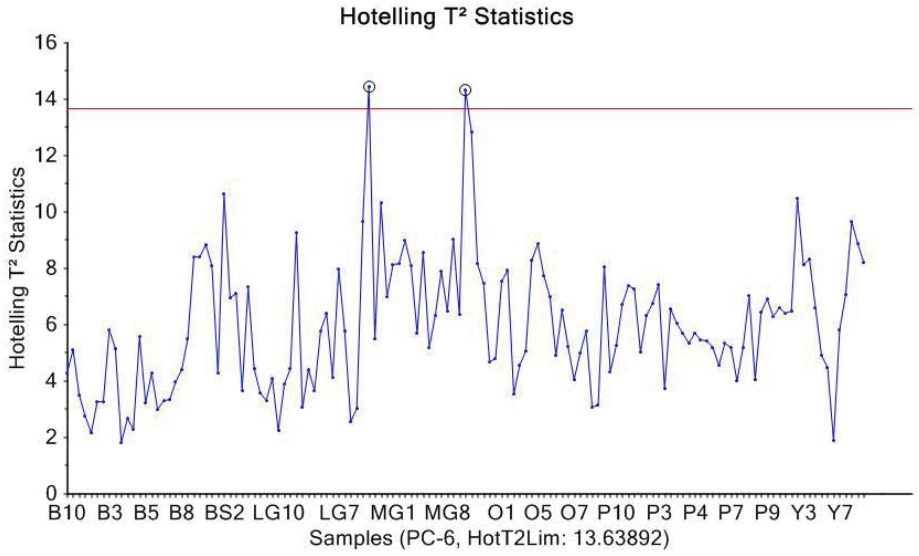

Figure 4.8. Hotelling with a T^2 statistical analysis, outliers are displayed above the percentile line (Red line in graph) and can be marked as above for removal from the dataset before recalculating the PCA. The choice can be made whether to download or remove outlying data along with defining what percentile to use to interrogate for outliers.

4.3.6 Use of Derivative Spectra

Computing a derivative of spectra is also called differentiation, it can help resolve overlapping bands in your spectral data. Derivatisation of spectra can be useful to improve peak resolution especially from broad peaks of either or any of infrared, Raman, fluorescent or UV/VIS (ultra violet/visual absorbance spectra) spectral data gathered from the instrumentation depending on the sample type. In this research derivatisation was applied to ATR-FTIR spectral data only. Figures 4.9A, 4.9B and 4.9C show the effects of 1st and 2nd derivation processing on ATR-FTIR spectral data. The definition of broad peaks is enhanced, and the derivatisation baselines sloped data, compare Figure 4.9A which shows the un-derivatised data with figures 4.9B and 4.9C which show 1st and 2nd derivatisation of the spectral data.

A B C

Figure 4.9. Examples of the effect of 1st and 2nd derivatives on FTIR-ATR spectral data. A: un-derivatised data, B: 1st derivation, C: 2nd derivation. Ringed area in 4.9 C shows that derivatisation of spectral data will amplify frequencies beyond useful parameters in some regions of the spectrum where there is greater noise.

First and second derivatives can increase noise in high definition regions, and drop peak troughs below zero, 2nd derivatives further enhance overlapping and hidden peaks compared with using first derivatives [19].

4.4 Multivariate Data Analysis as a Tool to Explore Data

4.4.1 Principal Component Analysis (PCA)

This is a useful multivariate statistical technique to investigate and model the dimensional differences in a multivariable data set by simplifying the data to fewer dimensions and removal of noise for ease of analysis and is a useful application to test for similar groupings of samples in experimental data [8, 12].

Experimental data consists of a number of variables (***p***) for each sample (***n***). The data are gathered in a matrix *X* (Table 4.1), with *n* rows and *p* columns [12, 14].

Table 4.1. Example of a data matrix ***X***. Individual objects or samples denoted by ***n*** rows and the variables denoted by ***X*** columns.

Sample	Variable						
	X_1	X_2	X_3	.	.	.	X_p
1	$X1_1$	$X1_2$	$X1_3$				$X1_p$
2	$X2_1$	$X2_2$	$X2_3$	.	.	.	$X2_p$
3	$X3_1$	$X3_2$	$X3_3$	.	.	.	$X3_p$
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	Xn_1	Xn_2	Xn_3	.	.	.	Xn_p

In general for spectral data, principal components analysis algorithms such as the NIPALS (non-linear iterative partial least squares) algorithm are used to group some given spectra that are defined by a set of numerical properties in such a way that the spectra within a group are more similar than the spectra in different groups.

Each spectrum is represented as a point in the variable space, so that the data set presents as a swarm of sample points in the variable space (CAMO, 2012) (Figure 4.11), the density of the swarm depends on the number of data points.

Principal component analysis finds the eigenvalues from the covariant matrix of the data variables, the variations in and between data are defined as variance and covariance, where variance is the measure of the spread of variable values. In the case of spectral data the variables are the wavenumbers (cm^{-1}) and the variance is the

relative intensity of the Raman scattering at each wavenumber. The association between two variables, X and Y or X and X, is a measure of their covariance [7]. The covariance of two variables X_1 and X_2 for example is:

$$\boldsymbol{cov}\left(\boldsymbol{x_y}, \boldsymbol{x_z}\right) = \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \left(\boldsymbol{x_{iy}} - \overline{\boldsymbol{x}}_{\boldsymbol{y}}\right)\left(\boldsymbol{x_{iz}} - \overline{\boldsymbol{x}}_{\boldsymbol{z}}\right)}{\left(\boldsymbol{n-1}\right)}$$

The covariant matrix of a data set with n samples and p variables is denoted by:

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C_{11}} & \cdots & \boldsymbol{C_{1p}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{C_{p1}} & \cdots & \boldsymbol{C_{pp}} \end{bmatrix}$$

Calculated by the sample co-variance between variables.

Eigenvectors (z) and eigenvalues (λ) calculated from the covariance matrix and are described by

$$\boldsymbol{Cz} = \boldsymbol{\lambda z}$$

Where $\boldsymbol{Cz}$ is the eigenvector of the covariance matrix $\boldsymbol{C}$ and $\boldsymbol{\lambda z}$ is the eigenvalue of the corresponding eigenvector $\boldsymbol{z}$. PCA converts a set of observations of possibly correlated variables into a set of variables called principal components [7, 12]. The number of principal components possible will equal the number of variables (*p*), PCA does this by picking the direction of the largest elongation of the swarm of data points in the multivariate space and plots a line through the centre point along this direction, this line is the direction of the first principal component (PC1) and this line points in the direction of maximum variation ("The Unscrambler® X version 10.2", CAMO Software

AS, Oslo, Norway). The second principal component will be orthogonal to the first and in the direction of the second most variation in the data set and so on with each following PC orthogonal to the previous PC. If the first PC does not explain 100% of the variation in the data, then further PC's are transformed depending on the complexity of variation in the dataset. The maximum number of principal components will be less than or equal to the number of original samples in the dataset, with each succeeding PC accounting for the next largest variance.

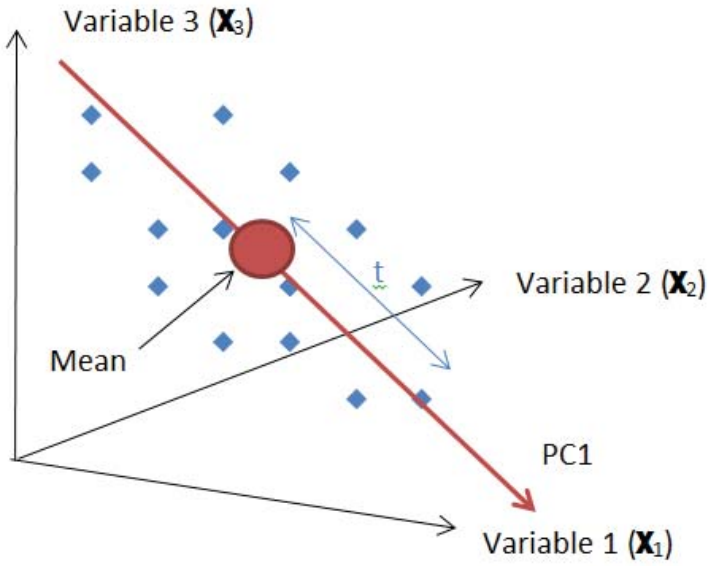

Figure 4.10. Representation of data points in the variable space. The red line passes through the centre and direction of the largest variance and is PC1 (an eigenvector). *T* = the t-score value of each object point.

Each PC is an eigenvector (Figure 4.10) with an eigenvalue representing the spread of the variance of the data points along the vector (i.e. comparative length of the vector). The order of value of the eigenvalue determines the order of influence for each PC. For each sample (points) the PC score (**t**) is the distance of any point from the mean,

parallel to the PC. The loadings are coefficients of the vector (Figure 4.11) which represent the values of the parallel distance of the PC vector to the variable axis' and plots the vector in the variable space and reflects the relative weight given to that variable in describing the location of the vector.

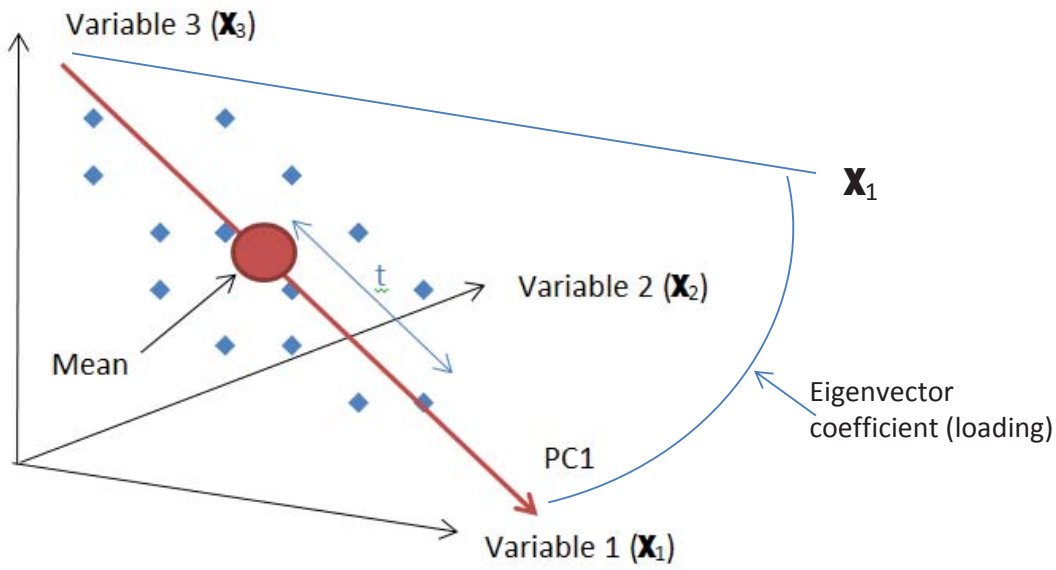

Figure 4.11. Representation of data points in the variable space. The red line passes through the centre and direction of the largest variance and is PC1. *T* = the t-score value of each object point. The eigenvalue is the value of the variance along the eigenvector and the associated coefficients are the angle value between the vector and any variable axis.

PCA deconstructs the initial data set matrix ($\boldsymbol{X}$) into scores ($\boldsymbol{T}$), loadings ($\boldsymbol{P^T}$) and residuals ($\boldsymbol{E}$ noise) (Figure 4.12). Scores and loadings (coefficient) values are defined by calculating eigenvector and eigenvalue data from the variable space (figures 4.10 and 4.11) [8, 12]. The residuals in this case are the PC's that account for very little variance in the data matrix whereas the product of $\boldsymbol{T}\ \boldsymbol{x}\ \boldsymbol{P^T}$ contains the main structure of your data.

$$\boldsymbol{X} = \boldsymbol{T} \times \boldsymbol{P}^{\boldsymbol{T}} + \boldsymbol{E}$$

Where $\boldsymbol{X}$ is the matrix of the data set, $\boldsymbol{T}$ is the scores matrix, $\boldsymbol{P}^{\boldsymbol{T}}$ is the loadings matrix and $\boldsymbol{E}$ represents the residuals matrix of the data set [11].

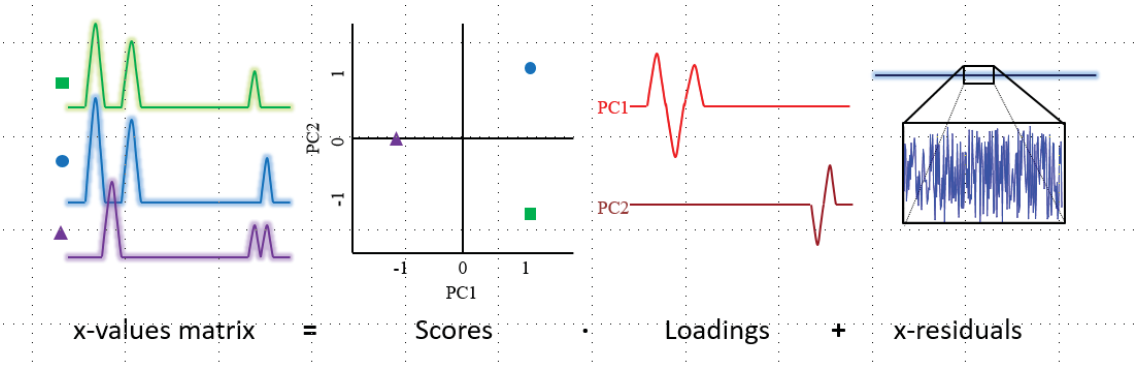

Figure 4.12. Example of PCA on spectral data. Figure taken from G.P.S. Smith et. al. 2015 "Raman imaging of drug delivery systems".

In PCA non-linear iterative partial least squares (NIPALS) is the most common algorithm used for computing the eigenvector components in a principal component or partial least squares analysis [9, 10, 18].

Herman Wold invented the NIPALS algorithm in 1966. NIPALS iterative PC calculation calculates one PC component at a time. The main mathematical features of this algorithm are as follows:

1. Centre and scale the covariance matrix appropriately, as necessary = **C.**

(Index initialization, *f*: *f*=1; $\mathbf{C}_f$ = C)

2. For $\mathbf{t}_f$ choose any column in **C** (initial proxy **t** vector)

3. $\mathbf{p}_f = \mathbf{C}^{\mathrm{T}}\mathbf{t}_f / |\mathbf{C}^{\mathrm{T}}\mathbf{t}_f|$

4. $\mathbf{t}_f = \mathbf{C}\mathbf{p}_f$

5. Check convergence: if $|\mathbf{t}_f.\mathbf{new} - \mathbf{t}_f.\mathbf{old}| <$ criterion, stop; else go to step 2 (iteration steps)

6. $\mathbf{C}_{f+1} = \mathbf{C}_f - \mathbf{t}_f \mathbf{p}_f^T$

7. $f = f+1$

Repeat steps until $f = \mathbf{A}$ the optimum number of principal components i.e. minimum PC's required to identify the main structure of your data with removal of residual data.

Where

1. The start of the process with the centred covariance matrix

2. Requirement to start the algorithm with a proxy $\mathbf{t}$ vector, any column vector of $\mathbf{C}$ will do but it is best to use the column with the largest variance.

3. Calculation of the loading vector, $\mathbf{P}_f$ for iteration number f.

4. Calculation of the score vector $\mathbf{t}_f$ for iteration number f. This step represents a projection of the vector down onto the fth PC component in variable space. These projections also correspond to the regression formulae for calculation regression coefficients used in PLS analysis (section 6).

5. This is the convergence test, NIPALS usually converges to a stable vector solution in less than 40 iterations, Convergence means that the proxy PC in variable space has stabilized to the maximum variance sought and the 1st eigenvector has been found.

6. Updates $\mathbf{C}_{f+1} = \mathbf{C}_f - \mathbf{t}_f \mathbf{p}_f^T$, i.e. subtraction of component number f (in the first case =PC1). The principal component model: $\mathbf{T}*\mathbf{P}^T$ is calculated for one component at a time. After convergence the rank one model $\mathbf{t}_f \mathbf{p}_f^T$ is subtracted from $\mathbf{C}$ [8].

4.4.2 Analysing Loading Results

Analysis of loading results from PCA and PLS methods are critically important. Loadings are estimated modelling methods in which information carried by the variables in the data is concentrated on to a few components. Each variable has a loading value for

each model component, the loadings are analysed to understand how much each X-variable contributes to the data and to interpret relationships between variables [7, 18]. Higher loading values (as indicated in Figures 4.13 and 4.14) indicate a higher contribution to the principal components calculated. In spectral data, this shows the wavenumbers that are responsible for influencing the variation between PC's. IR and Raman spectra are obtained by detecting vibrations from molecular bonds in the chemical components in the samples and these bond vibrations give peaks at specific wavenumbers, therefore the greater the loading value of specific wavenumbers in a loading plot indicates the chemical components that may be responsible.

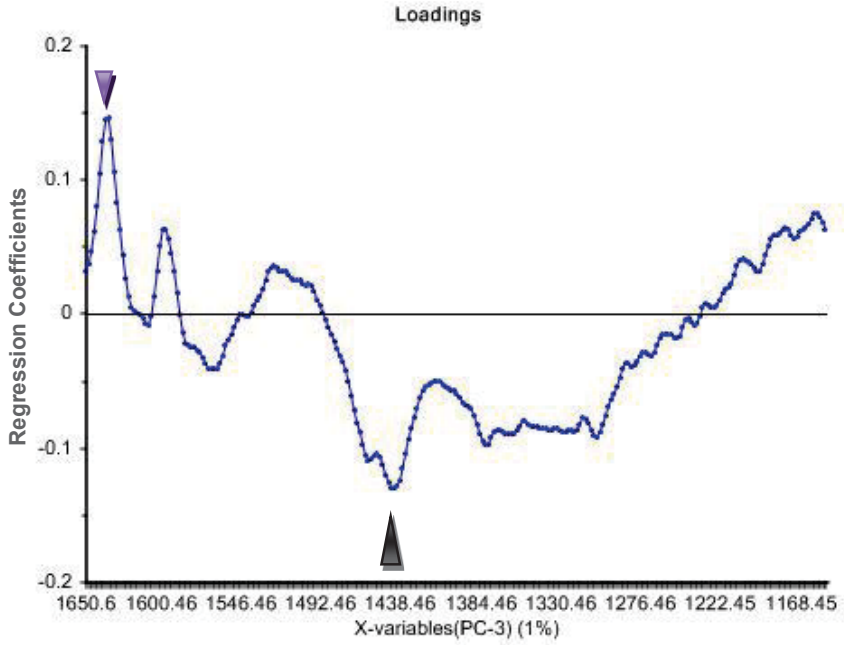

Figure 4.13. Loadings plot of PC3 for the sub-region spectrum (1650−1150 cm^{-1}). The arrows show the wavenumbers with the most influence = the largest coefficient values (-ve or +ve) on PC3.

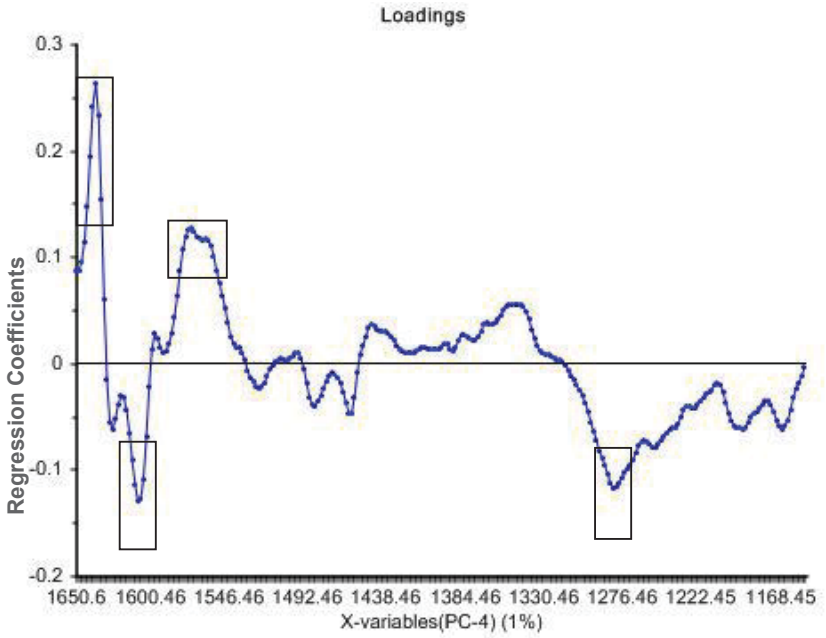

Figure 4.14. Loadings plot of PC4 for the sub-region spectrum (1650−1150 cm^{-1}). Regions with the large magnitude coefficients (-ve or +ve) are shown in boxes.

A correlation loadings plot (Figure 4.15) shows the distribution of all wavenumbers that were considered in plotting one PC against another. This can be used to determine which wavenumbers have positive and negative correlations in the dataset and which have the most influence on the PC under investigation. This type of plot gives a more visual interpretation of the influence of the various spectral wavenumbers. The outer ellipse in figure 4.16 indicates the 100% correlation mark, i.e. the data point has 100% of the explained variance, which would equal an R^2 of 1, and the inner ellipse indicates the 50% correlation mark.

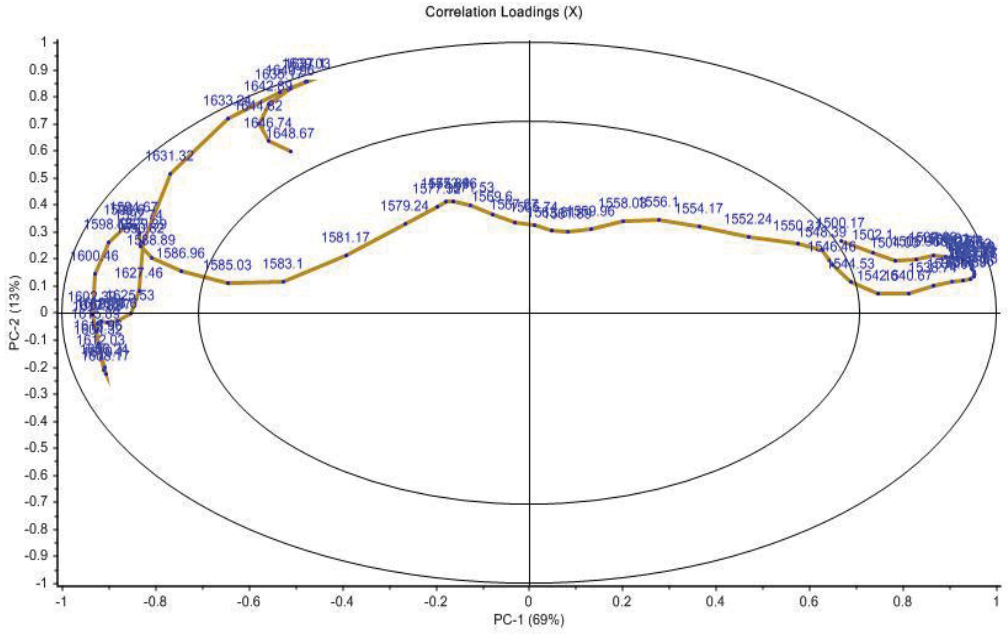

Figure 4.15. Correlation loadings plot. Wavenumbers in the outer ellipse have a higher influence on the corresponding model than wavenumbers in the inner ellipse. Outermost wavenumbers have the highest influence and the relative position in the X and Y axis indicate whether the correlation is positive or negative.

4.4.3 Explained Variance

Explained variance is a critical statistic for modelling the spectral data. An explained variance plot shows how much of the variation in the model is explained by each principal component (PC). Figure 4.16 shows that PC1 explained 95% of the variation in the data and that the remaining PCs were responsible for the last 5% of variance in the sample. However, this does not indicate that the PCs with less influence are not of value; plotting various PC's against each other show which PCs differentiate your data. Typically in spectral data the first PC contains the broad variance in many sample types and the following PC's contain the differentiating data [8, 12].

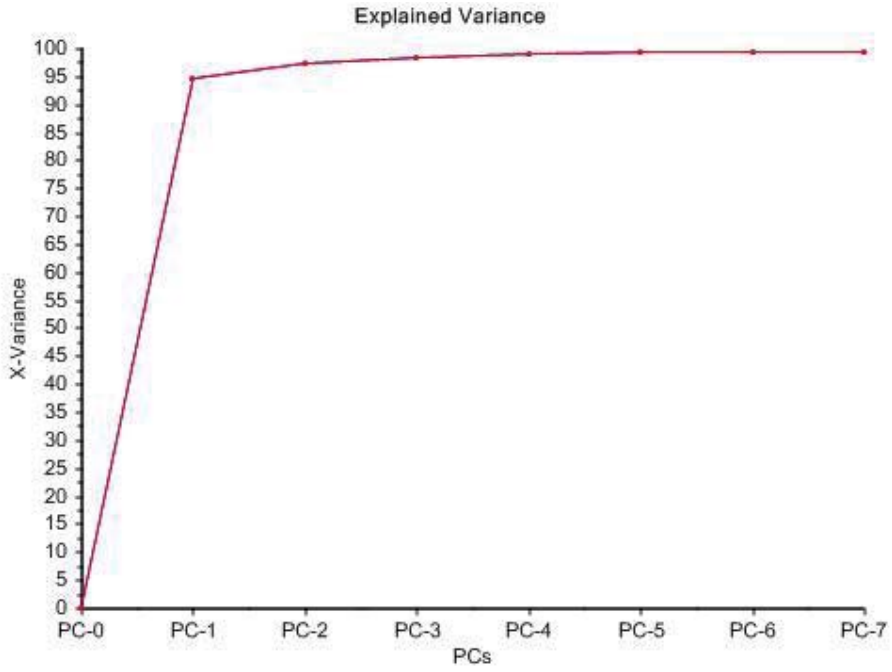

Figure 4.16. Explained variance PC plot showing how much each PC contributes to the variation in the data set.

After examination of regression coefficients from loadings plots, correlation loadings plots and explained variance plots, information on possible compounds that may be responsible for the differences in the variables can be determined.

4.4.4 Prediction Modelling Using Partial Least Squares (PLS)

Partial Least Squares (PLS) regression techniques were applied, using the data from spectroscopic analysis and chemical composition data collected from HPLC analysis. PLS is a multivariate method that is used to measure covariance in the system under investigation. The variations in and between data are defined as variance and covariance, where variance is the measure of the spread of variable values. The association between two variables, X and Y or X and X, is a measure of their covariance. If large values of variable X1 occur together with large values of X2, the

covariance will be positive; conversely, if large values of variable X1 occur together with small values of variable X2, and vice versa, the covariance will be negative.

The correlation between two variables (x, y) is calculated by dividing the covariance value by the product of their respective standard deviations (S); this gives a unit-less scaled covariance measure:

$$\boldsymbol{R = \frac{cov(x, y)}{SxSy}}$$

Where R = the correlation value. The correlation value always lies between –1 and +1; a correlation of 0 means that there is no relationship, whereas a correlation of +1 or −1 means that there is an exactly linear positive or negative relationship. R^2 is the most common form of expressing correlation [8, 12]. In general, as a high R^2 value denotes a high correlation between two variables, one variable can accurately estimate the value of the other variable. In many cases, several variables in the dataset contribute to the property of interest Y. Correlation between variables does not ensure causality, i.e. if any variables are correlated with the property of interest, this does not necessarily mean that they are the cause of the value of the property of interest. Further investigation of the system under investigation and knowledge of contributing causes is required to establish cause and effect.

In this thesis PLS regression techniques were applied, using the data from the spectroscopic analysis and DHA results collected from HPLC analysis, to determine whether there was a correlation between the Raman spectral data from *L. scoparium* leaves and the DHA levels in *L. scoparium* nectar. This method was investigated because DHA, anthocyanins and carotenoids are plant secondary metabolites and it was hypothesised that the metabolic processes responsible for carbohydrate

synthesis, anthocyanins etc. may be correlated in *L. scoparium.* The normalised DHA concentration values obtained from the nectar of the cultivars using HPLC are shown in Figure 4.17.

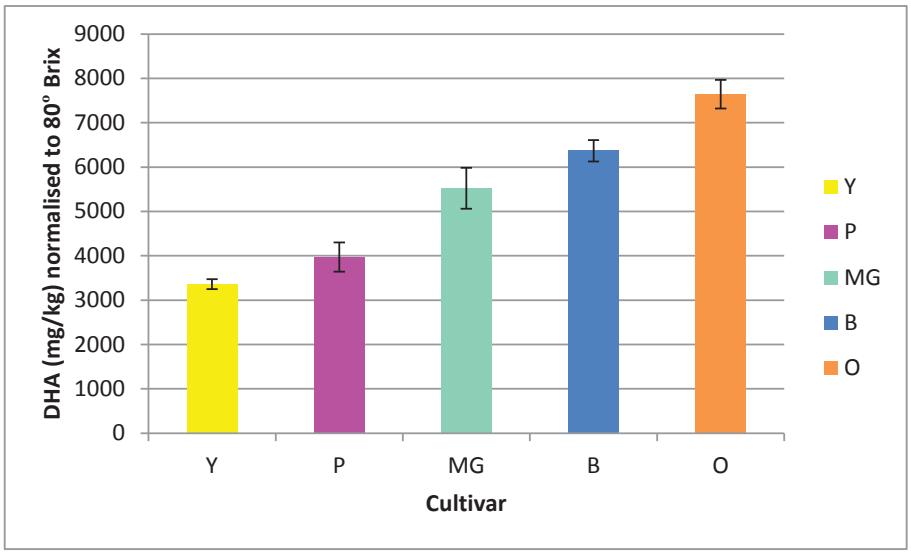

Figure 4.17. HPLC data showing DHA values for the cultivars normalised to 80° BRIX. It should be noted that only five out of the seven initial cultivars were used in this model because only five cultivars from which nectar was collected flowered during this experiment.

Figures 4.18-4.20 show the various overview plots from the resulting PLS model. Figure 4.18 is the variance plot, which shows how much of the model is explained by the variance in the data for each factor. The factors plotted represent percentages of the variance in the data for each variable, i.e. spectral features and DHA levels. PLS methods use the non-iterative partial least squares (NIPALS) algorithm to calculate the components in PLS and PCA models. PLS factors are similar to PCA components. Total values below 50% indicate that a model is not well explained. The explained variance plot indicates that this model was well explained by the variance in the data and that three or four factors gave the best model including 78% of the explained variance well

validated shown by the closeness of the red validation line to the modelled line. If the validated explained variance line plot starts to fall away from the modelled line it indicates that those factors are not well validated for inclusion in the model.

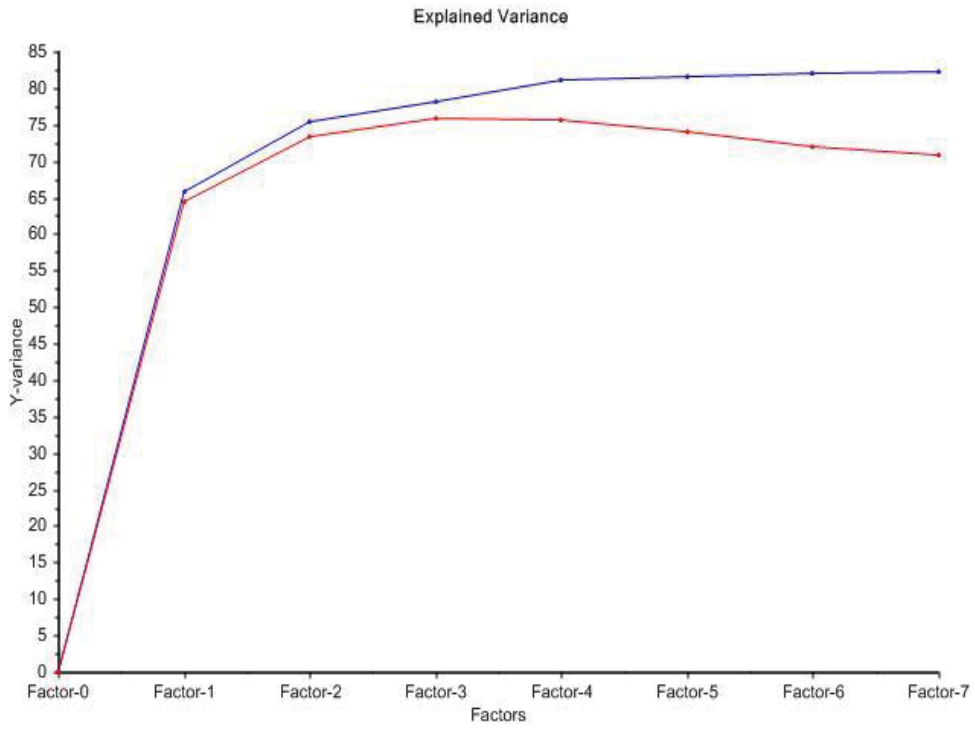

Figure 4.18. Explained variance plot for the PLS model.

Figure 4.19 illustrates the PLS score plot of factor 1 against factor 3 showing PLS component grouping best separating the five different cultivars. These groupings according to cultivar validate the cumulative distribution plot that the spectral data did not have a normal distribution and suggest that this distribution is due to the differences in leaf components between the cultivars.

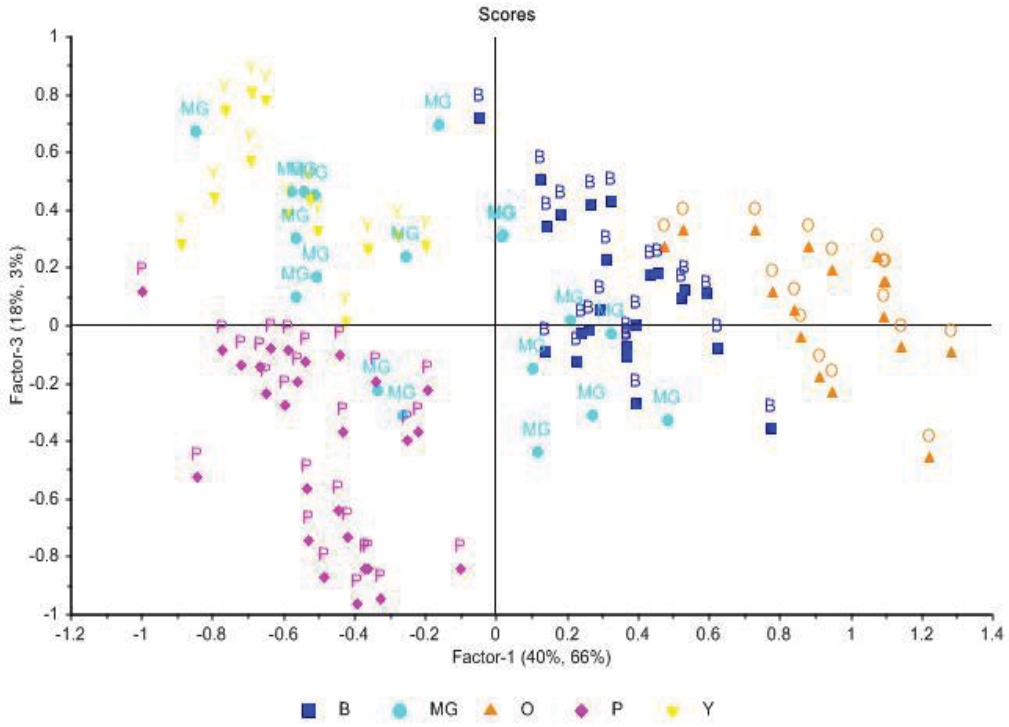

Figure 4.19. PLS component plot of factor 3 versus factor 1 separates the P and Y cultivars better than the plot of PC1 vs PC2.

The graph in Figure 4.20 shows the PLS model itself with predicted line versus the reference line of the model and an R^2 value of 0.78, indicating a validated model with a predictive error component of 764.77 mg/kg across a range for DHA of 2800-8000 mg/kg.

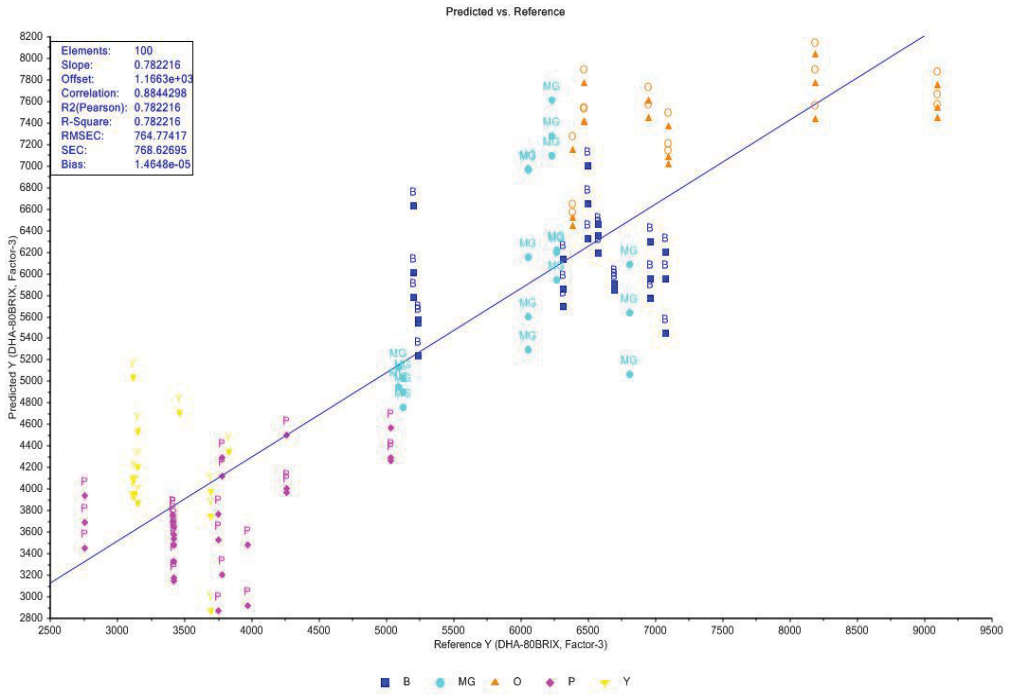

Figure 4.20. Regression graph of PLS model, note the R^2 value of 0.78 and the inclusion of the first three factors for this model.

In the case of PLS models it is worth noting the weighted regression co-efficients which relate to the loading information and show the relative contribution and influence from wavenumbers from each factor to the model, and also illustrates information on which component spectra are having the most influence on the PLS model. In the case above the first factor shown in the explained variance plot (Figure 4.20) accounts for approximately 65% of the model and therefore has the largest influence of all the factors and so the related coefficients from that factor have the most importance. As stated the first four factors gave the best model and so all four factor co-efficients need to be analysed to investigate which wavelengths are having the most influence in the overall model. Higher values in the weighted regression co-efficient plots indicate higher relative importance in the model.

4.5 Summary

This chapter described the chemometrical approach used in this thesis to model spectroscopy data from the Mānuka plant material. In addition the multivariate methods of PCA and PLS are also used in the analysis of non-spectroscopic variables such as the climate data variables and the soil chemistry variables in chapters 5 and 6.

4.6 References

[1] Cozzolino D., Corbella E., Smyth H. E. Quality Control of Honey Using Infrared Spectroscopy: A Review. Applied Spectroscopy Reviews **(2011)**, 46(7): p. 523-538.

[2] Anjos O., Campos M.G., Ruiz P.C., Antunes P. Application of FTIR-ATR spectroscopy to the quantification of sugar in honey. Food Chemistry **(2015)**, 169: p. 218-223.

[3] Butchweitz M., Gudi G., Carle R., Kemmerere D., Schulz H. Systematic investigations of anthocyanin-metal interactions by Raman spectroscopy. Journal of Raman Spectroscopy **(2012)**, 43(12): p. 2001-2007.

[4] Pierna J.A.F., Abbas O., Dardenne P., Baeten V. Discrimination of Corsican honey by FT-Raman spectroscopy and chemometrics. Biotechnologie Agronomie Societe Et Environnement **(2011)**, 15(1): p. 75-84.

[5] Sultanbawa Y., Cozzolino D., Fuller S., Cusack A., Currie M., Smyth H. Infrared spectroscopy as a rapid tool to detect methylglyoxal and antibacterial activity in Australian honeys. Food Chemistry **(2015)**, 172: p. 207-212.

[6] Uysal R.S., Soykut E.A., Boyaci I.H., Topcu A. Monitoring multiple components in vinegar fermentation using Raman spectroscopy. Food Chemistry **(2013)**, 141(4): p. 4333-4343.

[7] Esbensen K.H. Multivariate Data Analysis − In Practice. An Introduction to Multivariate Data Analysis and Experimental Design 5th Edition. CAMO Process AS. **(2004)**.

[8] Esbensen K.H., Guyot D., Westad F., Houmoller L.P. Multivariate Data Analysis in Practise 5th Edition. CAMO Process AS. **(2002)**.

[9] Eriksson L., Trygg J., Wold S. A chemometrics toolbox based on projections and latent variables. Journal of Chemometrics **(2014)**, 28(5): p. 332-346.

[10] Esbensen K.H., Geladi P. Principles of Proper Validation: use and abuse of re-sampling for validation. Journal of Chemometrics **(2010)**, 24(3-4): p.168-187.

[11] Smith G.P.S., McGoverin C.M., Fraser S.J., Gordon K.C. Raman imaging of drug delivery systems. Advanced Drug Delivery Reviews **(2015)**, 89: p. 21-41.

[12] Manly B.F.J. Multivariate Statistical Methods: A Primer, 3rd Edition. Chapman and Hall/CRC **(1998)**.

[13] Levin N. Multivariate statistics and the enactment of metabolic complexity. Social Studies of Science **(2014)**, 44(4): p. 555-578.

[14] Cordella C.B.Y. Principal Component Analysis: The Basic Building Block of Chemometrics. Intech **(2012)**.

[15] Lasch P. Spectral Pre-processing for Biomedical Vibrational Spectroscopy and Microspectroscopic Imaging. Chemometrics and Intelligent Laboratory Systems **(2012)**, 117: p. 100-114.

[16] Rinnan A., van den Berg F., Engelsen S. B. Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry **(2009)**, 28(10): p. 1201-1222.

[17] Balss K.M., Long F.H., Veselov V., Orana A., Akerman-Revis E., Papandreou G., Maryanoff C. A. Multivariate analysis applied to the study of spatial distributions found in drug-eluting stent, coatings by confocal Raman microscopy. Analytical Chemistry **(2008)**, 80(13): p. 4853-4859.

[18] McGoverin C. PhD Thesis (Otago University): Raman spectroscopy of complex mixtures **(2008)**.

[19] de Aragão B. J. G. and Messaddeq Y. Peak Separation by Derivative Spectroscopy Applied to FTIR Analysis of Hydrolized Silica. Journal of Brazilian Chemical Society **(2008)**, 19(8): p. 1582-1662.

Chapter 5: Genetic and environmental influences on nectar composition and yield in *Leptospermum scoparium* (Mānuka)

This chapter is currently under final preparation for submission to a scientific journal.

Elizabeth M. Nickless[1], Georgie Hamilton[1], Christopher W. N. Anderson[1], Jonathan M. Stephens[2], Jason J. Wargent[1] *

[1] Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand

[2] Comvita Innovations, Institute for Innovation in Biotechnology, University of Auckland, Auckland, New Zealand

* Corresponding Author: J. J. Wargent

E.M.N and G.H. carried out the experimental work. E.M.N. analysed the data and authored the paper. All authors designed the experiments, and J.M.S and J.J.W. were project supervisors to E.M.N. All authors approved the final manuscript.

5.1 Abstract

Floral nectar *of Leptospermum scoparium* (Mānuka) is the source for the unique non-peroxide antibacterial activity (NPA) component in Mānuka honey. We investigated the influence of genetic and environmental factors on plant growth and nectar yield in several different cultivars of *L. scoparium*. Plant growth, nectar composition and nectar yield of ten different cultivars, were quantitatively measured and analysed from plants grown in controlled conditions. Nectar components; sucrose, fructose, glucose and DHA were analysed using HPLC analytical methods. Multivariate statistical approaches were applied to analyse for environmental influences on nectar production and composition*,* with increased sunlight hours and solar radiation positively correlated with nectar production, but on an intraspecific basis. Equally, flowering phenology (when plants flowered, floral density and flowering period) and therefore overall nectar yield, were significantly different between genotypes ($P \leq 0.05$). DHA, sucrose, glucose, fructose concentrations and also total sugar yield and nectar volume were significantly different between genotypes ($P \leq 0.05$). All cultivar nectars were hexose dominant with fructose being the dominant hexose. Plant growth rate were also significantly different between the 10 genotypes. Overall, nectar composition in high value *L. scoparium* cultivars appears to have a strong genetic determinacy, yet within some cultivars nectar yield is also significantly influenced by environmental factors.

Keywords: dihydroxyacetone; NPA honey; *Leptospermum scoparium,* cultivar, biomass, nectar composition.

5.2 Introduction

The nectar of *Leptospermum scoparium* is the source for the carbohydrate dihydroxyacetone (DHA) which is the precursor chemical for methylglyoxal (MGO) the unique non-peroxide antibacterial activity (NPA) component of Mānuka honey [1-3]. Greater concentrations of DHA collected by foraging honey bees, metabolises to higher concentrations of MGO in Mānuka honey. The retail price of Mānuka honey is directly and positively dependent upon increasing MGO content and honey products are graded accordingly [2, 4], therefore there is high commercial value in assessing *L. scoparium* nectar components with the prospect of breeding plants with higher DHA potential.

Considerable variation in DHA levels has been noted between different *L. scoparium* plants at various field sites [2, 5], although the source of this variation is not yet well understood. Williams et al. (2014) [5] showed variation according to cultivar type and regional location in the DHA content of field-collected nectars of *L. scoparium*. In addition, possible climatic influences were indicated by a difference in DHA concentration noted from nectar collected from the same plants over different years. Analysis of nectar composition from various other plant species has shown similar results, with significant intraspecific variation in nectar composition [6]. It is apparent that the nectar composition of angiosperms is extremely variable across groups of families, within genera, and within species [7]. It has been reported in the literature that environmental factors such as climatic conditions and insect visitation dynamics also influence nectar composition and production in angiosperms, however, the extent of the influence of environmental factors versus genetic variation, relating to nectar composition and nectar production in angiosperms is not fully understood. Nectar

composition has been shown to vary with: soil type, climatic conditions, the presence of microbes such as yeasts and fungi, insect visitation inducing subsequent changes in nectar composition mainly due to deposition of microbes, plant age and plant genetics [5-12].

Successful bee-keeping and apiary management is dependent on assessing plant resources to maximise honey production. Bee preferences and nectar yield are factors to consider when breeding plants suitable for the honey production industry. Various authors have demonstrated that honeybees make choices based on the sweetness of nectar and volumes available, therefore, nectar volume, total sugar content and floral density are important variables in terms of palatability to foraging bees [10, 13-15], When assessing *L. scoparium* plant varieties for inclusion in the high value Mānuka honey industry, DHA concentrations, sugar yield, sugar composition, floral density and plant growth are all important aspects for consideration [16, 17].

In New Zealand, *L. scoparium* has been noted to flower predominantly during spring and early summer [18]. *L. scoparium* is a short-day flowering plant with a critical day length modified by temperature [19], although plants have been observed to flower in autumn and winter in milder climates.

Here we have grown plants of the same age from cloned cultivars in one control soil in the same environment without insect visitation to try and further elucidate the influence of genetics on nectar composition in *L. scoparium*. To analyse nectar composition we adapted existing High Performance Liquid Chromatography (HPLC) methods currently used for honey analysis [3] to quantify DHA and the key nectar components sucrose, fructose, glucose. Given that net DHA yield is dependent upon a range of aspects related to plant size and growth, plant physiological parameters and

flowering period were also measured to assess cultivars with the best overall performance when considering large-scale *L. scoparium* cultivation. Finally, we investigated possible influences of environment on nectar composition by recording daily environmental conditions at nectar collection, to see if climate variables such as temperature, radiation and humidity influence DHA concentrations and nectar production.

5.3 Materials and Methods

5.3.1 Plant Material and Growth Conditions

Growth and nectar parameters of ten different cultivars of *Leptospermum scoparium* were evaluated in controlled glasshouse conditions. The ten proprietary cultivars were bred by Comvita, (Te Puke, New Zealand), who hold the parental stock. Each cultivar is represented by arbitrary code letters: Y, MG, G, B, BS, P, O, LG, PU and R. 10 replicate plants of each cultivar were grown from cuttings of a single parent plant and so were genetically identical, i.e. clones of the parent material, with all plants at the same age. Plants were supplied by Comvita, (Te Puke, New Zealand) and were transferred into 30 cm pots containing standard tree and shrub potting mix at 3 months from cutting initiation. Plants were housed in a standard glasshouse facility (Massey University, Palmerston North, New Zealand); the glasshouse was evenly ventilated by a fan system to approximate ambient external temperature ranges and effectively provided a climate of mild temperate conditions. To prevent insect visitation by bees, moths and flies, the ventilation systems were covered with mesh and insect repellent was used to prevent ant visitation. Plant pots were watered and supplemented with feeding solution in-line to each individual pot to maintain plants at optimum growing

conditions over the 15 month experimental period. The feeding solution was diluted to a concentration of 0.5 g/L and consisted of 20% nitrogen, 8.7% phosphorus, 16.5% potassium and trace elements iron, manganese, boron, copper, zinc and molybdenum.

Plant height and basal stem diameter were measured at 3 monthly intervals for 15 months. Relative growth rates (RGR) were measured as the average change in the basal diameter per month. Plant biomass, a measure of plant growth, was determined by a destructive harvest at the end of the experiment. Dry weights of the roots and aerial biomass of the plants were measured separately with the aerial parts separated at the first root junction on the basal stem.

5.3.2 Nectar Collection

We observed that *L. scoparium* flowers, although fully open, do not start secreting nectar until anthers are mature and fully extended from the hypanthium edge. We denoted this stage as stage IV in line with Davis and colleagues' description of flower stage development in *Eucalyptus*, another member of the Myrtaceae family [20]. Nectar was collected at randomly selected times between 10 am and 2 pm on any day, from flowers at development stage IV, thereby standardising nectar collection at the optimum health of the flower.

Initial analysis of variance (ANOVA) of the nectar composition of the cultivars in the experiment showed that at least 15 flowers needed to be sampled to allow for the intra-plant variability of nectar composition within each plant (data not shown). For our study we collected 20 flowers from each plant at each flower collection time, with flowers removed from plants using forceps and placed on a tray on ice. Nectar was removed from individual flowers by rinsing the hypanthium with 5 µl of distilled H_2O using a pipette then removing all nectar. Nectar removal and collection was performed

using a stereomicroscope and pipette to ensure accurate rinsing and maximum nectar retrieval. Nectar from the 20 flowers from each plant replicate was pooled into eppendorf tubes for weighing to obtain total sugar yield and nectar volume data. The tubes and water were weighed prior to collection to allow for accurate nectar volume measurements. Nectar from each plant was collected immediately following flower collection and stored at -80^{o}C. DHA levels in the nectar samples were measured using aqueous extraction, derivatisation and analysis by high pressure liquid chromatography (HPLC) adapted from the method used by Windsor et al [3] for analysing DHA in honey samples with minor changes in the method. The standards range and reaction chemical dilutions were changed to allow for the lower concentrations of these components in floral nectar and the smaller amount (100-200 ul) of each sample available to be analysed. DHA levels in nectar samples from four different cultivars were validated using commercial LCMS methods available at R.J. Hills Laboratories (Hamilton, New Zealand).

DHA and sugar levels were normalized against total sugar concentrations to allow comparisons of nectar DHA concentrations between cultivars and a Ternary diagram was generated from the main sugar data i.e. sucrose, glucose, and fructose to allow comparisons of relative sugar composition between the ten cultivars.

5.3.3 High Pressure Liquid Chromoatography (HPLC) Conditions for DHA Analysis

Analyses were performed on a Perkin Elmer Series 200 Pump and Auto sampler with a Flexar photo diode array detector (λ = 263 nm). HPLC separations were performed on a Synergi Fusion column (75 × 4.6 mm, 4 µm particle size). The column was held at 30 ^{o}C to maintain stable run conditions. Mobile phase A was water: acetonitrile (ACN), 70/30, v/v and mobile phase B was 100% ACN. The following 23 minute gradient

elution was employed: A: B = 90:10 (isocratic 2.5 min), graded to 50:50 (8.0 min), graded to 0:100 (1.5 min), 0:100 (isocratic 7.0 min), graded to 90:10 (1.0 min), 90:10 (isocratic 3.5 min), detection at 263 nm.

5.3.3.1 Preparation of Reaction Solutions

Hydroxyacetone (HA) (3.01 mg/ml) formed the HA internal standard solution. The *O*-(2, 3, 4, 5, 6-pentafluorobenzyl) hydroxylamine (PFBHA) derivatising reagent was 19.8 mg/ml in citrate buffer (0.1 M) adjusted to pH 4 with sodium hydroxide (NaOH) (4 M). DHA (3.88 mg/ml) formed the DHA standard solution.

5.3.3.2 Sample Preparation

For the preparation of standards, DHA standard stock solution (100, 80, 60, 40, 20, 10 and 0 µl) was added to tubes 1 to 7 respectively and made up to 100 µl with nanopure water. For sample analysis, 20 µl of nectar or standard was pipetted into a mix tube and 25 µl of the HA was added. Derivatisation steps were performed at 25 ^{o}C in a controlled temperature room. Each of the HPLC samples and standards was thoroughly mixed and placed in a rack on a rotating table for 1 hour to allow complete dissolution. PFBHA derivatising solution (100 µl) was added to each test tube, which was mixed and placed in a rack on a rotating table for 1 hour to allow for complete derivatisation. ACN (1.5 ml) was added to each test tube and mixed. Nanopure water (0.5 ml) was added to each test tube and mixed. Samples were then syringe filtered with a 0.22 µm filter into HPLC vials. Vials were placed into the auto sampler and run overnight and repeat analysis of standards were analyzed through each run to check stability of the analysis. DHA calibration curves were generated from standards by linear regression using the HPLC peak area ratios of DHA: HA plotted against the mass of the DHA. Mass DHA content of the nectar samples were determined against these calibration curves.

5.3.3.3 HPLC Conditions for Sucrose, Glucose and Fructose Analysis

Individual sugar analysis was performed on an Alliance Waters 2690 Separations Module HPLC and auto sampler using a Waters 2410 Refractive Index (RI) detector. Sucrose, glucose and fructose were separated using an Aminex HPX-87C Carbohydrate Column (300 x 7.0 mm). The column was held at 65 °C to maintain stable conditions. A standard run method was used with one phase solution of milliQ water at a flow rate of 0.6 ml/minute. The sample injection size was 20 µl and each sample run was 18 minutes.

1% stock solutions of each of the sugars; sucrose, glucose and fructose were made in nano-pure water. Five standards were used to create a calibration curve for each sugar, the standards contained equal proportions of sucrose, glucose and fructose, and had concentrations of 0%, 0.01%, 0.05%, 0.17%, and 0.33% sugar.

30 µl of each nectar sample was added to 1470 µl of nano-pure H_2O and mixed. The diluted nectar samples as well as the standards were filtered using 0.22 µm syringe filters into HPLC vials. Vials were loaded into the auto-sampler along with the standards and analysed overnight. Standards were repeat assayed throughout the run to reference for any variability during the analysis.

5.3.4 Environmental Data

Daily environmental data was recorded for nectar collection dates. Minimum, maximum, daily average temperatures, radiation, sunshine hours and humidity parameters were logged (Table 5.7) to analyse for possible environmental influences on nectar yield and composition (Table 5.1). Temperature and humidity data was collected using HortPlus Microloggers installed inside the glasshouse. Total solar radiation and hours of sunshine data were downloaded from the National Institute of

Water and Atmospheric Research weather station located 500 m from the glasshouse experiment at site 2193 grid reference T24315885; Latitude -40.38195, Longitude 175.60915.

Floral density was assessed as average flower numbers per 20 cm length of stem from five replicate plants of each cultivar. Growth parameters, plant height and basal stem diameter were measured every three months for 15 months. Biomass was measured by harvesting five plants from each cultivar at the start and end of the experiment and assessing root mass and aerial mass via fresh and dry weight measures.

The influence of environmental factors at nectar collection was assessed using Partial Least Squares (PLS) regression analysis of environmental data against DHA concentrations, total sugars and nectar volumes using the software application The Unscrambler® X (CAMO, NORWAY). The environmental data, for example temperature and humidity and sunshine hours, are collinear and PLS regression is the recommended method for analysing co-linearity in multivariate data [21].

5.4 Results and Discussion

The mean concentration of the nectar sugars; sucrose (S), fructose (F), glucose (G) and DHA for each of the ten cultivars normalised to 80 °BRIX are shown in Table 5.1, along with relevant sugar ratios. S/ (F+G) ratios are standardly used to classify nectar and nectars are considered hexose dominant at a ratio of 0.1 and lower [11]. In the ten *L. scoparium* cultivars investigated in this experiment S/ (F+G) ratios range from 0.03 to 0.002 between the cultivars and demonstrate that the nectar in all ten cultivars is hexose dominant. The F/G ratios showed that fructose is the dominant hexose sugar in the nectar of *L. scoparium* and averages over all ten cultivars at a ratio of 1.5:1 fructose

to glucose with a range of 1.20 to 1.70 across all cultivars. Figure 5.1 shows that nectars are grouped in one main group, indicating that nectar composition in the cultivars, in terms of the main sugars, are similar. However the statistical analysis of fructose, glucose and sucrose data (Table 5.1) detected significant differences between some cultivars and PCA analysis (Figure. 5.1B) displays that, although there is some spread along PC1 (which accounts for 98% of the variation in the nectar composition), the nectars are grouping largely according to cultivar type. The loadings data from the PCA analysis indicates that PC1 mainly represents the variation in fructose and glucose concentrations whereas PC2 represents mainly the variance in sucrose concentration, illustrated by cultivar G, which has a significantly higher percentage of sucrose in its nectar, grouping separately from the other 9 cultivars, along PC2.

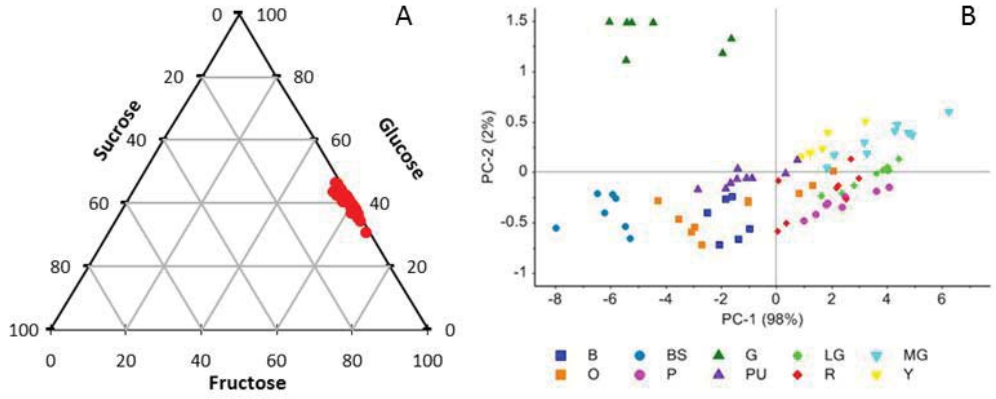

Figure 5.1. A: Ternary diagram of the dominant sugars in *L. scoparium* nectar of ten different cultivars. Fructose, glucose and sucrose concentrations from each plant in the experiment (100 data points) are normalised to 100% total sugars for the ternary diagram. B: PCA score plot of main sugar composition data from the same ten cultivars, separate colours represent different cultivars. The PCA score plot shows that although there is some overlap, the groupings indicate that the main nectar sugar composition is conserved according to cultivar, implying a predominantly genetic influence on nectar composition.

The nectar of all ten cultivars contained sucrose at levels with concentrations less than 2%. In contrast Williams et al (2014) observed no sucrose in the nectar of the *Leptospermum* cultivars they investigated using GC-FID analysis. Insect visitation influence on nectar composition has been proven to contaminate nectar with microbes such as yeast which can metabolise sugars by invertase activity [6, 12, 14, 22, 23]. This could change the composition of *L. scoparium* nectar collected from plants without insect protection, with further metabolism in the nectar of sucrose into fructose and glucose, and could explain the lack of sucrose in field collected samples compared to nectar collected from glasshouse insect protected plants. Sucrose levels in nectar samples were confirmed independently from our samples using commercial LCMS methods available at R.J. Hills Laboratories (Hamilton, New Zealand) (Data not shown).

Table 5.1. Means of the main sugars; sucrose, glucose and fructose and DHA in *L. scoparium* nectar normalised to 80°BRIX ± one standard error. Average total sugar yield as mg per flower is also shown. Cultivars (CV) sharing the same letter are not significantly different (Tukey post-hoc comparisons, P≤0.05). S/ (F+G) ratios show that all nectars are hexose dominant and the F/G ratio shows that fructose is the dominant hexose in the nectar.

CV	Fructose % 80°BRIX	Glucose % 80°BRIX	Sucrose % 80°BRIX	Ratio S/(F+G)	Ratio F/G	Total Sugars %	Sugar Yield per Flower	Nectar Volume ml	DHA 80°BRIX mg/kg
BS	43.22^{f} ± 0.22	35.94^{a} ± 0.19	0.85^{b} ± 0.06	0.01	1.2	$76.61^{a,b,c}$ ± 7.52	$1.5mg^{b,c}$ ± 0.05	$0.0025^{b,c,d}$ ± 0.0008	$4252.20^{b,c,d,e}$ ± 804.5
G	$44.90^{e,f}$ ± 0.73	33.24^{b} ± 0.59	1.87^{a} ± 0.20	0.02	1.36	$65.93^{b,c}$ ± 8.62	$3.0mg^{a,b}$ ± 0.05	$0.0044^{a,b,c,d}$ ± 0.0006	$6153.12^{a,b}$ ± 583.25
O	$46.73^{d,e}$ ± 0.51	$32.82^{b,c}$ ± 0.49	$0.45^{c,d}$ ± 0.04	0.006	1.43	$79.70^{a,b,c}$ ± 6.84	$4.0mg^{a}$ ± 0.05	$0.0064^{a,b,c}$ ± 0.0022	$5177.02^{b,c}$ ± 510.2
PU	$46.87^{c,d,e}$ ± 0.27	$32.47^{b,c,d}$ ± 0.25	$0.66^{b,c}$ ± 0.02	0.008	1.44	60.53^{c} ± 5.01	$4.5mg^{a}$ ± 0.03	$0.0078^{a,b}$ ± 0.0009	$4843.71^{b,c,d}$ ± 404.31
B	$46.97^{b,c,d,e}$ ± 0.45	$32.68^{b,c,d}$ ± 0.41	$0.35^{c,d}$ ± 0.07	0.004	1.44	$93.42^{a,b}$ ± 3.36	$1.4mg^{b,c}$ ± 0.02	$0.0016^{c,d}$ ± 0.0003	2714.35^{e} ± 358.57
R	$48.93^{a,b,c,d}$ ± 0.40	$30.78^{c,d,e}$ ± 0.37	0.29^{d} ± 0.05	0.004	1.59	97.51^{a} ± 1.52	$0.7mg^{c}$ ± 0.02	0.0007^{d} ± 0.0002	$5981.54^{a,b}$ ± 346.09
P	$49.25^{a,b,c}$ ± 0.46	$30.55^{d,e}$ ± 0.4	0.20^{d} ± 0.07	0.003	1.62	61.59^{c} ± 4.97	$4.8mg^{a}$ ± 0.04	0.0082^{a} ± 0.001	$3266.70^{c,d,e}$ ± 135.01
Y	$49.41^{a,b}$ ± 0.73	29.90^{e} ± 0.73	$0.69^{b,c}$ ± 0.02	0.009	1.67	$69.77^{b,c}$ ± 5.9	$4.0mg^{a}$ ± 0.07	$0.0061^{a,b,c}$ ± 0.0007	$3058.88^{d,e}$ ± 110.36
MG	49.70^{a} ± 0.86	29.79^{e} ± 0.84	$0.51^{b,c,d}$ ± 0.02	0.006	1.68	61.76^{c} ± 6.86	$4.7mg^{a}$ ± 0.05	0.0094^{a} ± 0.0019	7459.07^{a} ± 500.9
LG	50.25^{a} ± 0.20	29.51^{e} ± 0.20	0.24^{d} ± 0.01	0.003	1.7	$68.80^{b,c}$ ± 5.74	$4.3mg^{a}$ ± 0.06	$0.0068^{a,b,c}$ ± 0.001	$4116.01^{b,c,d,e}$ ± 386.7

Regression analysis of fructose versus glucose as nectar constituents gave a negative correlation (R^2 = 0.97) (Figure 5.2). The variance from a 1:1 ratio of F and G is not a result of insect visitation effects in *L. scoparium* nectar since insects were excluded in this study. These findings agree with indications from other authors that nectar composition is genetically controlled and Wenzler et al (2008) suggested that variations from a 1:1 ratio might be the result of invertase action on sucrose in the nectaries activating the glycolysis pathway, while Lin et al (2014) speculated that hexose reuptake genes in the cell walls of the nectary could be responsible [22, 24, 25].

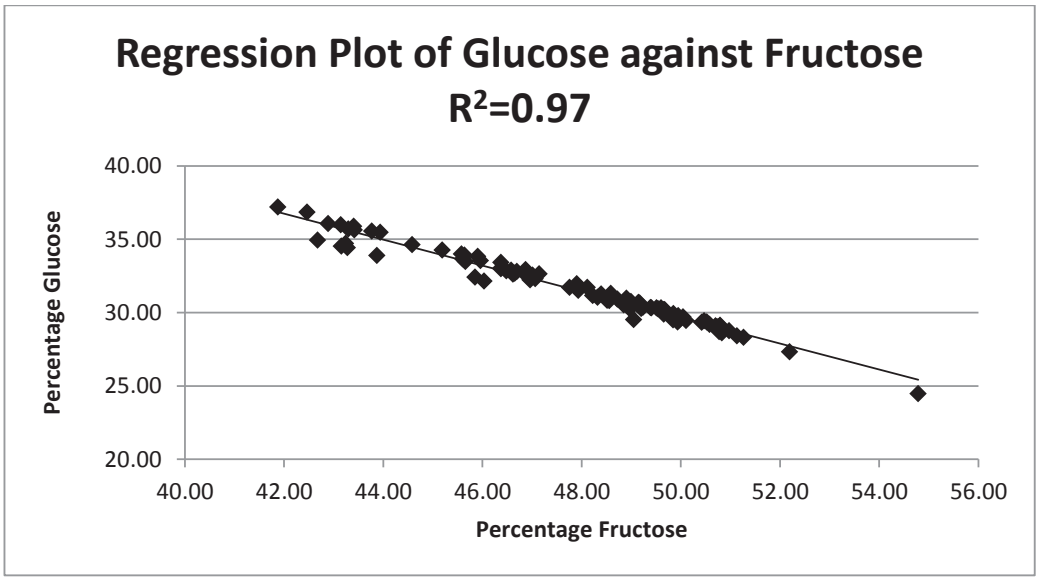

Figure 5.2. Regression analysis of nectar constituents: fructose versus glucose illustrating the negative correlation between fructose and glucose in *L. scoparium* nectar.

Analysis of DHA showed significant differences between the cultivars with the lowest DHA concentrations at 2714.35 mg/kg and the highest at 7459 mg/kg. The comparison of means indicated that there were four significantly different genetic groups in regard to DHA concentration. Partial Least Squares regression analysis of the sugar data

indicated that there was no significant correlation between concentrations of DHA and the sugars; fructose, glucose and sucrose: R^2 0.16 RMSECV 986.1 mg/kg. We hypothesise that *L. scoparium* has unusually high concentrations of DHA due to further glycolytic activity within the nectaries before nectar secretion to the hypanthium surface. DHA derives from an as yet unknown pathway, from Dihydroxyacetone phosphate (DHA-P), which is a product of glycolysis, therefore higher concentrations of DHA-P would need to be available to allow for higher concentrations of DHA.

Total sugar percentages (Table 5.1), in the nectar were very high in some of the cultivars due to very low production of nectar in terms of volume, because of this they were often almost dry, probably due to some evaporation in addition to low production, resulting in a high sugar concentration. Total sugar yield and floral density are important factors to consider, in addition to nectar composition, when calculating total nectar yield from any cultivar. Nectar from some cultivars e.g. R and BS were high in DHA concentration but because of the low overall yield (Table 5.1) they have low potential in the NPA honey industry.

5.4.1 Floral Density as a Component of DHA Yield

Floral density was observed to vary significantly between the different cultivars which, when combined with the DHA content per flower gave very divergent DHA yields per stem (Table 5.2). We predict that this measure will be a good indicator of the potential honey yield during establishment years in a young plantation of Mānuka.

Table 5.2. Potential DHA yield for each cultivar calculated as the product of µg of DHA per flower and the flower density (flower number per 20 cm stem length at peak flowering). Cultivars sharing the same letter are not significantly different (Tukey post-hoc comparisons, P≤0.05).

Cultivar	µg of DHA per flower	Floral density	Potential yield per cultivar: mg of DHA/20cm flowering stem length
MG	75.55± 8.04[a]	30	2.27±0.22
LG	29.66± 7.62[b]	30	0.89±0.21
G	31.27± 8.04[b]	23	0.72±0.22
O	32.36± 7.62[b]	20	0.65±0.21
PU	39.31± 8.04[b]	16	0.63±0.22
P	28.33± 8.04[b]	20	0.57±0.22
BS	14.67± 8.04[b]	28	0.41±0.22
Y	18.87± 8.04[b]	14	0.26±0.22
B	3.79± 8.04[b]	48	0.18±0.22
R	4.36± 8.04[b]	40	0.17±0.22

Floral density will not be as easy to measure in a mature plantation [16, 26] as many variables need to be considered, such as peak flowering times, length of anthesis, number of annual flowering events and an assessment of the number of flowers per plant and plants per hectare. Flowering period and floral density together will affect overall nectar yield. Flowering period is an important aspect of floral density, and in this investigation was calculated by denoting flowering start as the point where there were at least 20 flowers at stage IV development (i.e. enough flowers to collect nectar for analysis), with flowering end denoted as less than 20 flowers at stage IV. Cultivars were significantly different on the basis of flowering start time, and duration of

flowering (Table 5.3 and Figure 5.2.) at P<0.005 and F values at 9.2 and 28.03 for Start of flowering and Flowering period respectively.

Table 5.3. Flowering period in potted plants where Start Days = the start of flowering in days numbered from the 1st of January in the main experimental year, and Flowering Period = the average total flowering period for each cultivar. Cultivars sharing the same letter are not significantly different (Tukey post-hoc comparisons, P≤0.05).

Cultivar	Start Days	S.E.	Flowering Period	S.E.
Y	260.33[a]	±1.54	18.33[d]	±1.13
P	259.00[a]	±1.50	27.44[d]	±1.73
B	251.44[a]	±1.60	25.89[d]	±1.87
LG	250.90[a]	±2.25	40.40[d]	±3.21
PU	249.44a	±1.80	26.78[d]	±2.03
BS	233.33[a,b]	±9.17	43.11[d]	±9.20
O	229.90[a,b]	±9.20	45.22[c,d]	±6.66
R	207.22[b,c]	±5.07	76.67[b,c]	±5.29
G	197.00[c]	±11.63	87.50[b]	±13.74
MG	139.78[d]	±10.41	140.11[a]	±10.66

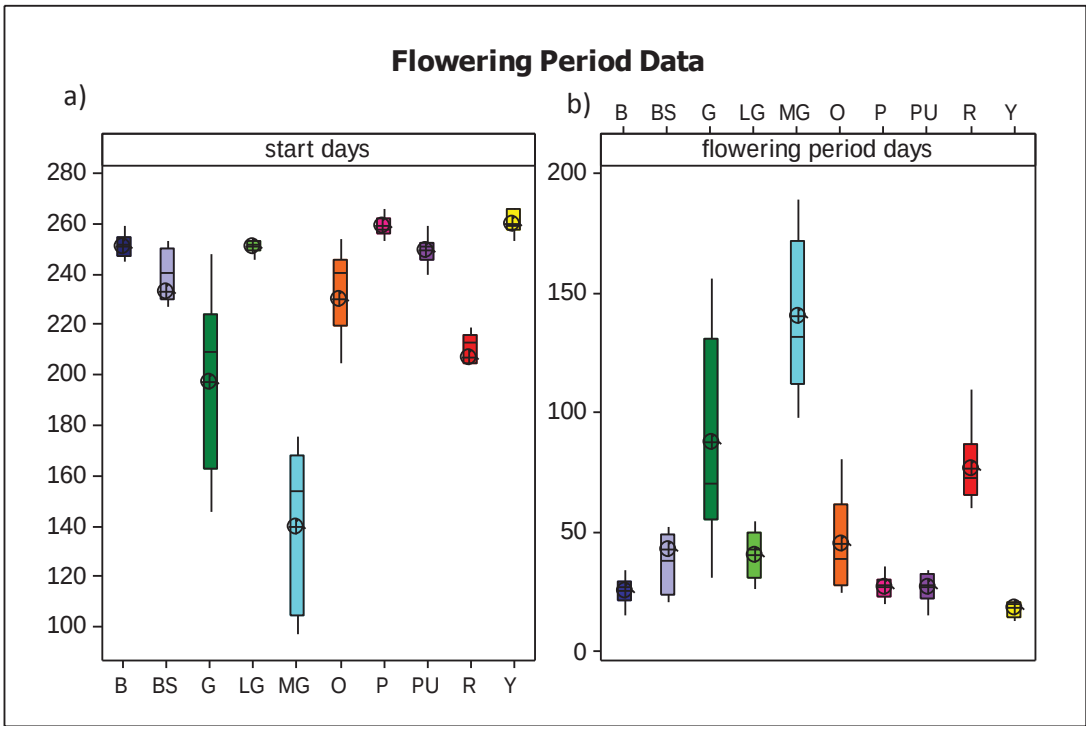

Figure 5.3. Flowering period in potted plants a) the start of flowering in days numbered from the 1st of January in the main experimental year, b) the average total flowering period for each cultivar.

Ideal cultivars would be those that flower while honey bees are active, and those which flower for long periods during suitable weather conditions for honey bee foraging activities. Two of our cultivars, MG and G, flowered through winter in the mild growing conditions experienced, but higher floral density flushes were observed in warmer autumn and spring months. The remaining eight cultivars flowered in the spring months and four of these cultivars flowered for only 7-10 days.

5.4.2 Variability in Plant Growth

Significant differences in growth and habit were noted in the various cultivars in this study, although with a significant amount of overlap amongst groups of cultivars (Tables 5.4 & 5.5).

Table 5.4. Growth data parameters measured over 15 months for each cultivar in the glasshouse experiment. Plant height (mm), basal stem diameter (mm), and the average relative growth rate (RGR) mm/month. Cultivars sharing the same letter are not significantly different (Tukey post-hoc comparisons, P≤0.05).

Cultivars	height (mm)	S.E.	Basal stem diameter	S.E.	RGR	S.E
LG	1181.30[a]	±86.89	12.74[a]	±0.62	1.14[a]	±0.07
BS	1157.80[a]	±42.08	8.45[c,d]	±0.34	0.57[c,d]	±0.03
B	1126.67[a,b]	±19.73	8.06[c,d]	±0.64	0.53[d]	±0.03
O	1108.56[a,b]	±42.53	12.60[a]	±1.18	0.88[a,b,c]	±0.05
MG	1061.89 [a,b]	±65.93	8.61[b,c,d]	±0.57	0.87[a,b,c]	±0.06
Y	967.00 [a,b,c]	±40.84	12.21[a,b]	±0.83	1.14[a]	±0.07
P	960.44[a,b,c,d]	±99.14	11.07[a,b,c]	±0.60	0.88[a,b,c]	±0.05
G	832.50[b,c,d]	±85.37	11.07[a,b,c]	±0.60	0.67[b,c,d]	±0.12
PU	706.70[c,d]	±113.03	13.20[a]	±1.55	0.97[a,b]	±0.08
R	628.89[d]	±60.03	6.68[d]	±0.22	0.56[c,d]	±0.08

We calculated an index of overall cultivar potential as a useful value to assess cultivars for DHA value, estimated as a factor of nectar yield versus plant growth (Table 5.4), this data gives a rank value, with consideration to biomass and nectar yield, to compare the DHA potential of the various cultivars. The proposed model of potential nectar yield and growth in Table 5.6 is calculated by: basal stem growth rate (BS RGR)*Potential yield per 20 cm stem (mg DHA)* Biomass (dry weight). An overall relative rank of the data from the plants was assigned based on value ranges in the rating column. On this basis, cultivars MG, LG, P, and O are assessed as holding the most potential for increased DHA yield, and would be worthy of future testing in field conditions. To model performance of mature plantation plants, a ratio of the

percentage of the biomass that flowers in a mature plant would need to be established.

Table 5.5. Comparison of the final biomass data between the 10 cultivars in the study, illustrating root biomass and aerial biomass and total biomass values in grams. Cultivars sharing the same letter are not significantly different (Tukey post-hoc comparisons, P≤0.05).

Cultivar	Dry weight roots (g)	S.E	Dry weight shoots (g)	S.E.	Total (g)	S.E.
Y	48.51	±10.12	287.50	±13.98	336.01[a]	±23.37
P	34.09	±3.85	298.67	±20.18	332.76[a]	±23.61
LG	27.56	±4.32	270.73	±16.42	298.29[a]	±20.03
O	30.80	±5.26	226.35	±15.07	257.14[a,b]	±19.08
G	14.29	±2.17	191.60	±19.05	205.89[b,c]	±20.65
B	19.94	±2.89	185.94	±11.70	205.88[b,c]	±11.56
MG	5.13	±0.77	158.71	±10.13	163.84[c,d]	±10.56
PU	17.58	±4.12	120.48	±22.06	138.06[c,d]	±25.49
BS	21.44	±2.57	87.03	±4.90	108.46[d,e]	±5.62
R	0.58	±0.24	53.47	±4.15	54.06[e]	±4.16

Table 5.6. Proposed estimate of potential nectar yield and growth i.e. overall cultivar rating, calculated by: basal stem relative growth rate (BS RGR)*Potential yield (mg DHA)* Biomass (dry weight). An overall relative rank from the data from the plants is assigned based on value ranges in the rating column.

Cultivar	RGR	Potential cultivar yield mg DHA(Y)	Biomass= average Total DWT (B)	Overall cultivar rating RGR*Y*B	Rank
MG	0.82	2.27	163.84	305	1
LG	1.14	0.89	298.29	303	1
P	0.88	0.57	332.76	167	2
O	0.87	0.65	257.14	145	2
Y	1.14	0.26	336.01	100	3
G	0.67	0.72	205.89	99	3
PU	0.97	0.63	138.06	84	4
BS	0.57	0.41	108.46	25	5
B	0.53	0.18	205.88	20	5
R	0.56	0.17	54.06	5	6

5.4.3 Environmental Influences on Nectar Production

Although results from nectar composition and yield data showed significant differences between cultivars, it must be kept in mind that some cultivars were observed to flower at significantly different times of the year (Table 5.3 and Figure 5.2). The possible effects of environmental influences due to the varying climatic influences at different times of the year when flowering occurred should be interrogated. We attempted to model the effect of these parameters using partial least squares regression (PLSR) modelling of the nectar data with respect to the environmental data collected and presented in Table 5.7.

Table 5.7. Data ranges of environmental parameters logged at the time of nectar collection. Differences between cultivar are due to differences in flowering dates and the conditions that prevailed at the time of sampling. Minimum and maximum temperatures, average temperature, radiation in micro-joules per meter squared (MJ/m^2), daily sunshine hours and percentage humidity are tabulated.

DATA	Min Temp ^oC		Max Temp ^oC		Average Temp ^oC		Radiation MJ/m^2		Daily Sunshine Hours		Humidity %	
RANGES	***min***	***max***	***min***	***max***	***min***	***max***	***min***	***max***	***min***	***max***	***min***	***max***
Y	13	14	20.5	24	15.87	18.3	3.59	16.63	0	8.5	62.48	86.88
P	13	14	20.5	24.5	15.87	18.3	3.59	16.63	0	8.5	62.48	86.88
MG	4.5	15	21	25.5	12.24	19.08	2.29	10.92	0	9.2	57.92	89.28
LG	13	14	16.5	24.5	14.37	18.26	3.59	15.04	0.2	5.9	66.85	85.58
R	12.5	14.5	20.5	23.5	14.89	17.63	4.68	10.3	0.1	9.4	66.4	80.7
G	3	14	14	23.5	8.9	17.1	2.24	13.18	0.1	5.9	67.51	89.28
PU	13	14.5	16.5	23.5	14.37	17.69	3.59	13.99	0.2	6.3	65.68	85.29
O	12.5	14	17.5	24.5	14.44	18.02	4.42	15.26	0.2	10	62.93	81.01
BS	12.5	14.5	20.5	24.5	16	17.78	3.59	12.79	0.3	5.6	65.6	85.58
B	13	14.5	16.5	24	14.37	17.75	4.42	16.63	0.2	10	65.33	79.36

Nectar data for each and all cultivars were regressed against environmental data using PLSR methods. Nectar data was also regressed against the individual genotypes to distinguish any added effect from environmental data on individual cultivars. Data analysis using all cultivars within the models showed that environmental influences explained further variability in the data for DHA concentrations and total sugars only. R^2 values increased from 0.47 to 0.61 for DHA explaining a further 15% of the variation in the data explained than from genotype alone and an R^2 of 0.44 increased to 0.68 for total sugar percentages a further 23% of the variance explained. Nectar volume had low correlations against cultivars and environmental data with a low R-squared value of 0.31 increasing only to 0.35 with environmental parameters included in the model, a further 3.5% explained variance. We therefore deduced that environmental effects measured were not significantly affecting nectar volumes we therefore suggest that the significant differences observed in nectar volumes between cultivars are mainly due to the influence of genotype. Table 5.8 shows the calibration results of the PLSR models using the individual cultivars in comparison with all cultivars kept in the model and show that there are some significantly high R-squared values for some individual cultivars in these models, indicating a clear effect of climate on the nectar data with respect to individual cultivars that is worth further investigation. Interestingly the concentration of DHA seemed to correlate with environmental parameters in some cultivars namely B, G and P indicating that the sugar composition itself was responding to environmental signals. Calibrated models showed that specifically, increased temperatures seem to have positive correlation with DHA concentration. Total sugars, increased on days of higher temperature and lower humidity and has been noted by other research investigating uptake and composition of nectar from bees as humidity and temperature increase [27]. There was visibly more nectar production in some cultivars

during the experiment i.e. on sunny days there was substantially more nectar on the flowers of some cultivars, however, as the data shows, the nectar volumes are generally robust against environmental changes with regard to the cultivars, indicating that evaporation rates are largely counteracting any increased volume. We would expect that volume would be affected the same as total sugars if nectar production is up-regulated on warmer days. Nectar metabolism and secretion is not yet fully understood, research into nectar secretion show that jasmonates can regulate nectar secretion [28-31], and that jasmonate levels are regulated by light levels in lima bean [30] with extra floral nectar secretion increasing with increased light levels. Research by Radhika et al. (2010) indicated that light composition also influenced nectar secretion with Red Far-Red ratios indicated [31].

Not all parameters measured were affected by environmental conditions for all cultivars, and we suggest that future analysis of nectar collected at a greater range of environmental parameters would further elucidate environmental influences within this species.

Table 5.8. Partial Least Squares Regression (PLS) analysis of climatic influences on nectar production and composition. Modelling data for each parameter interrogated for climatic influences is displayed. R-squared values are tabled for DHA (normalised to 80°BRIX), nectar volume and total sugars. Root Mean Square Error of Calibration (RMSEC) represents the error values in the calibrated data. RMSECV represents root mean square error of the cross validated model. (The model is not well validated in terms of using the models for prediction purposes where RMSEC and RMSECV are not similar values).

	DHA 80°BRIX		Nectar Volume per Flower		Total Sugars	
Cultivar	R^2	RMSEC:RMSECV mg/kg	R^2	RMSEC:RMSECV µl	R^2	RMSEC:RMSECV %
All	0.61	1220:1320	0.35	0.004:0.004	0.68	4.70:5.33
Y	NA		0.36	0.004/0.005	0.72	2.80/3.61
P	0.89	128.46:179.45	NA		0.65	3.58:4.51
MG	NA		NA		0.57	2.04:2.53
LG	NA		0.49	0.002:0.003	NA	
R	NA		NA		NA	
G	0.76	813.9:1196	NA		NA	
PU	NA		0.86	0.0008:0.002	NA	
B	0.99	54.74:160.96	NA		NA	
O	0.66	893.3:1184	NA		0.75	2.73:4.22
BS	NA		NA		0.67	3.24:4.24

This study evaluates important floral aspects to consider when measuring floral density and comparing the potential nectar yield of different cultivars for the honey industry. Results demonstrate that nectar composition in *L. scoparium* appears to have a strong genetic determinacy, with variability in nectar yield in some cultivars further influenced by environmental factors such as solar radiation. Additional factors need to be considered when assessing plants for commercial use in the Mānuka honey industry, particularly floral

density and plant growth, yet we also show here the importance of assessing not just DHA concentration in deducing cultivar value, but overall nectar yield. Our findings suggest that these key features must be explored when assessing *L. scoparium* plants within breeding programs, prior to germplasm selection for large-scale field production of high NPA Mānuka honey.

5.5 Acknowledgements

This work was supported by a Primary Growth Partnership (Ministry of Primary Industries, New Zealand) awarded to Mānuka Research Partnership (NZ) Limited, and a scholarship from Callaghan Innovation (NZ) to L.M.N. [Contract: MARP1001].

We would like to acknowledge the staff at the Plant Growth Unit at Massey University; Palmerston North, New Zealand for support with the glasshouse experimental setup. Chris Rawlingson, Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand for assistance with HPLC instrumentation and supporting staff from Massey University; Palmerston North, New Zealand, who assisted with nectar collection and plant growth data measurements.

5.6 References

[1] Adams C.J., Boult C.H., Deadman B.J., Farr J.M., Grainger M.N.C., Manley-Harris M. et al. Isolation by HPLC and characterization of the bioactive fraction of New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2008),** 344(18): p. 2609-2609.

[2] Adams C.J., Manley-Harris M., Molan P.C. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research. **(2009),** 344(8): p. 1050-1053.

[3] Windsor S., Pappalardo M., Brooks P., Williams S. and Manley-Harris M. A convenient new analysis of dihydroxyacetone and methylglyoxal applied to Australian Leptospermum honeys. Journal of Pharmacognosy and Phytotherapy **(2012)**, 4(1): p. 6-11.

[4] Mavric E., Wittmann S., Barth G., Henle, T. Identification and quantification of methylglyoxal as the dominant antibacterial constituent of Manuka (*Leptospermum scoparium*) honeys from New Zealand. Molecular Nutrition. Food Research **(2008)**, 52(4): p. 483-489.

[5] Williams S., King J., Revell M., Manley-Harris M., Balks M., Janusch F., Kiefer M., Clearwater M., Brooks P., and Dawson M. Regional, Annual, and Individual Variations in the Dihydroxyacetone Content of the Nectar of Mānuka (*Leptospermum scoparium*) in New Zealand. Journal of Agricultural and Food Chemistry **(2014)**, 62(42): p. 10332-10340.

[6] Canto A., Herrera C. M., Garcia I. M., Perez R., Vaz M. Intraplant variation in nectar traits in Helleborus foetidus (Ranunculaceae) as related to floral phase, environmental conditions and pollinator exposure Flora **(2011)**, 206(7): p. 668-675.

[7] De la Barrera E., Nobel P.S. Nectar: properties, floral aspects, and speculations on origin. Trends in Plant Science. **(2004)**, 9(2): p. 65-69.

[8] Nardone E., Dey T., Kevan P.G. The effect of sugar solution type, sugar concentration and viscosity on the imbibition and energy intake rate of bumblebees. Journal of Insect Physiology **(2013)**, 59(9): p. 919-933.

[9] Morrant D.S., Petit S., Schumann R. Floral nectar sugar composition and flowering phenology of the food plants used by the western pygmy possum, Cercartetus concinnus, at Innes National Park, South Australia. Ecological Research **(2010)**, 25(3): p. 579-589.

[10] Afik O., Hallel T., Dag A., Shafir S. The components that determine honeybee (*Apis mellifera*) preference between Israeli unifloral honeys and the implications for nectar attractiveness. Israel Journal of Plant Sciences **(2009)**, 57(3): p. 253-261.

[11] Chalcoff V.R., Aizen M.A., Galetto L. Nectar concentration and composition of 26 species from the temperate forest of South America. Annals of Botany **(2006)**, 97(3): p. 413-421.

[12] de Vega C., Herrera C.M. Micro-organisms transported by ants induce chnages in floral nectar composition of an ant-pollinated plant. American Journal of Botany **(2013)**, 100(4): p. 792-800.

[13] Gonzalez-Teuber M., Heil M. Nectar chemistry is tailored for both attraction of mutualists and protection from exploiters. Plant Signaling and Behavior **(2009)**, 4(9): p. 809-813.

[14] Herrera C.M., Perez R., Alonso C. Extreme intraplant variation in nectar sugar composition in an insect-pollinated perennial herb. American Journal of Botany **(2006)**, 93(4): p. 575-581.

[15] Gonzalez A., Rowe C.L., Weeks P.J., Whittle D., Gilbert F.S., Barnard C.J. Flower choice by honey-bees (*Apis mellifera*) - sex-phase of flowers and preferences among nectar and pollen foragers. Oecologia **(1995)**, 101(2): p. 258-264.

[16] Camargo S.C., Garcia R.C., Feiden A., de Vasconcelos E.S., Pire, B.G., Hartleben et al. Implementation of a geographic information system (GIS) for the planning of beekeeping in the west region of Parana. Anais Da Academia Brasileira De Ciencias. **(2014)**, 86(2): p. 955-971.

[17] Gavina M.K.A., Rabajante J.F., Cervancia C.R. Mathematical Programming Models for Determining the Optimal Location of Beehives. Bulletin of Mathematical Biology **(2014)**, 76(5): p. 997-1016.

[18] Stephens J.M.C., Molan P.C., Clarkson B.D. A review of *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany **(2005)**, 43(2): p. 431-449.

[19] Zieslin N. Environmental-factors involved in growth, flowering and postharvest behaviour of flowers of *Leptospermum scopariuim* Forst, J.R. and Forst, G. Israel Journal of Botany **(1986)**, 35(2): p. 101-108.

[20] Davis A.R. Influence of floral visitation on nectar-sugar composition and nectary surface changes in Eucalyptus. Apidology **(1997)**, 28: p. 27-42.

[21] Esbensen. K.H. Multivariate Data Analysis − In Practice. An Introduction to Multivariate Data Analysis and Experimental Design **(2004).**

[22] Lin I.W., Sosso D., Chen L.Q., Gase K., Kim S.G., Kessler D., Klinkenberg et al. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. Nature **(2014)**, 508(7497): p. 546-562.

[23] Ruhlmann J.M., Kram B.W., Carter C.J. CELL WALL INVERTASE 4 is required for nectar production in Arabidopsis. Journal of Experimental. Botany **(2010)**, 61(2): p. 395-404.

[24] Wenzler M., Holscher D., Oerther T., Schneider B. Nectar formation and floral nectary anatomy of Anigozanthos flavidus: a combined magnetic resonance imaging and spectroscopy study. Journal of Experimental. Botany **(2008)**, 59(12): p. 3425-3434.

[26] Slot M.; Janse-ten Klooster S.H., Sterck F.J., Sass-Klaassen U., Zweifel R. A lifetime perspective of biomass allocation in Quercus pubescens trees in a dry, alpine valley. Trees-Structure and Function **(2012)**, 26(5): p. 1661-1668.

[27] Fidalgo A. de O., Kleinert A. de M.P. Floral preferences and climate influence in nectar and pollen foraging by *Melipona rufiventris Lepeletier* (Hymenoptera: Meliponini) in Ubatuba, Sao Paulo state, Brazil. Neotropical Entomology **(2010)**, 39(6): p. 879-884.

[28] Heil M. Nectar: generation, regulation, and ecological functions. Trends in Plant Science **(2011),** 16(4): p. 191-200.

[29] Radhika V., Kost C., Boland W., Heil M. Towards elucidating the differential regulation of floral and extrafloral nectar secretion. Plant Signaling and Behavior **(2010),** 5(7): p. 924-926.

[30] Radhika V., Kost C., Mithoefer A., Boland W. Regulation of extrafloral nectar secretion by jasmonates in lima bean is light dependent. Proceedings of the National Academy of Sciences of the United States of America **(2010)**, 107(40): p. 17228-17233.

[31] Radhika V., Kost C., Boland W., Heil M. The Role of Jasmonates in Floral Nectar Secretion. Plos One (**2010)**, 5(2) e9265.

Chapter 6: Soil type influences floral density and growth in Mānuka (*Leptospermum scoparium*)

This chapter has been accepted for publication by the New Zealand Journal of Botany as "Soil influences on plant growth, floral density and nectar yield in three cultivars of Mānuka (*Leptospermum scoparium*)"

Elizabeth M. Nickless[a*], Georgie Hamilton[b], Jonathan M. Stephens[c], Jason Wargent[b] Christopher W. N. Anderson[b].

[a] Fonterra Research and Development Centre; Palmerston North, New Zealand; [b] Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand; [c] Comvita Innovations, Institute for Innovation in Biotechnology, University of Auckland, Auckland, New Zealand.

* Correspondence to: E. M. Nickless, E-mail: elizabeth.nickless@fonterra.com

All authors designed the experiments. E.M.N, C.W.N.A and G.H. designed and established the controlled glasshouse experiment. E.M.N. and G.H. carried out the experimental work. E.M.N. analysed the data and authored the paper. J.M.S., C.W.N.A and J.J.W. were project supervisors to E.M.N. All authors approved the final manuscript.

Keywords: Dihydroxyacetone; *Leptospermum scoparium*; soil composition; nectar; Partial least squares regression (PLS); phenology

6.1 Abstract

Honey derived from the nectar of *Leptospermum scoparium* J.R. et G. Forst. Myrtaceae (mānuka) is a high-value product and there is considerable potential for economic growth in honey-growing regions of New Zealand through increased nectar yield from mānuka plantations. *Leptospermum scoparium* exhibits a significant amount of phenotypic plasticity throughout regions in New Zealand where it has established although the influences on this plasticity are unknown. When assessing *L. scoparium* as a nectar source for honey in marginal land areas, the possible effect of soil on nectar chemistry and yield should be considered. We investigated whether phenological patterns of flowering, plant growth, nectar composition and nectar yield were influenced by soil composition.

Three different cultivars of *L. scoparium* were grown on 10 different soils in glasshouse conditions. The soils chosen were representative of the range of New Zealand soils where mānuka is being considered as a commercial crop for the honey industry. ANOVA and General Linear Model (GLM) models revealed no significant effect of soils on nectar composition or production, however significant but complex interactions between cultivars and soils influenced plant growth and flowering ($P \leq 0.05$). Accordingly, the overall nectar yield was influenced depending on the cultivar and soil interaction. Measured attributes of the soil such as CEC, sulphate, iron, manganese, calcium and chloride were shown to influence the plant parameters assessed.

Results allowed modelling of nectar potential (NP) against each soil type and established a mānuka soil index (MSI) to determine the most appropriate soil for each cultivar. The results indicated that potential nectar yield increases will be dependent on cultivars being deployed according to the nature of the soil present. Furthermore, the mānuka cultivars displayed

significantly greater growth in response to increased nutrients and some cultivars increased floral density, suggesting potential to improve nectar yield by greater plant growth using targeted fertilisation.

6.2 Introduction

Commonly known as mānuka, *Leptospermum scoparium* J.R. et G. Forst. Myrtaceae is an economically important plant in New Zealand. The main commercial products derived from mānuka are honey and essential oils. In addition the plant is commonly used as an ornamental shrub. The unique non-peroxide antibacterial (NPA) bioactive component in mānuka honey is due to methylglyoxal (MGO), which is derived from the precursor dihydroxyacetone (DHA); DHA is only present in the nectar and MGO forms in maturing honey solutions [1]. Mānuka honey is a valuable product with considerable potential for economic growth from plantations that produce elevated nectar yields. The value of mānuka honey is currently driven by the NPA which is intrinsically linked to the concentration of DHA in the nectar collected by foraging bees. Hence, investigating the influence of genotype and plant nutrition on plant growth, nectar composition in particular DHA concentrations, and total nectar yield is of interest to the mānuka honey industry in New Zealand.

The various taxonomic characteristics of *L. scoparium* have been described [2-10] and some publications have investigated honey and nectar components responsible for the NPA [1, 11-13]. Williams et al. 2014 [13] studied the composition of soil samples collected from underneath trees sampled for nectar and suggested no effect of soil chemistry on nectar quality (specifically DHA concentrations) from plants in field locations.

Leptospermum scoparium is a highly polymorphic species [2] and exhibits significant phenotypic plasticity and genotypic variation. Polymorphism is common in plants and is related to biodiversity, genetic variation and adaptive responses to the environment [10, 14, 15]. The reported ideal soil conditions for *L. scoparium* colonisation are moist, slightly acidic soils with low nutrient content [16]. However, the extent to which soil environment or plant competition exerts control on *L. scoparium* establishment and growth is not clearly understood. Permanent dominance has been observed in positions that are unfavourable for the development of climax forest, for example situations that are too wet, dry, cold, exposed, infertile or unstable [9] and this indicates probable competition effect from climax species. Additionally, *L. scoparium* forms a large number of mycorrhizal associations, the principal benefit of these is thought to be increased phosphorus uptake allowing more rapid growth to exploit expansion into available cleared land areas [9].

The establishment of *L. scoparium* plantations in New Zealand marginal agricultural land is receiving more attention in land-use planning particularly in some regions of the country. Marginal land areas in New Zealand are often steep hill country that is prone to erosion with low-fertility soils that exhibit low grassland productivity. *Leptospermum scoparium* is seen as a commercially viable plant species for use in protecting hill country from erosion. Afforestation in erosion susceptible areas mitigates soil loss and flood damage and reduces negative effects on water quality [17, 18]. Mānuka for honey production is seen as a greater value land-use than other options in these areas.

New Zealand marginal land areas carry a variety of soils (http://data.gns.cri.nz/geology/) and soil chemistry and structure have been shown to affect growth and flowering in many plant species [19-23]. Hence the effect of New Zealand soils on plant growth and flowering in *L. scoparium* is an important field of research. The various sub-species of *Leptospermum*

have been observed to flower over different periods from spring through summer in New Zealand [9] and Yin et al (1984) [15] reported significant differences in flowering phenology in a range of *L. scoparium* seedling populations grown under uniform environmental conditions, this indicates a significant genotypic effect on flowering. However the effects of soil type on flowering phenology have not been reported to date. Research that investigates optimum conditions to promote growth and flowering, and integrates this with predictive tools that assess the likely performance of *L. scoparium* plantations in different regions, is seen as beneficial to the economic development of the mānuka honey industry.

The current study was conducted in the context of this higher level objective and specifically investigated the influence of soil chemistry on *L. scoparium* through measurement of plant growth (measured by; plant height, basal stem diameter, biomass and relative growth rate), flowering phenology (onset of flowering, flowering period and flower numbers), nectar yield and nectar quality (DHA concentrations), as a function of the soil chemistry from soils collected from a range of geographic locations within New Zealand.

6.3 Materials and Methods

This research was conducted under controlled glasshouse conditions using three *L. scoparium* cultivars (R: cultivar variety *L. scoparium* var *scoparium* cultivar X *L. scoparium* var *incanum*; G: cultivar variety *L. scoparium* var *incanum*; and Y: cultivar variety *L. scoparium var scoparium*) on a range of New Zealand soils (10) over a 15 month period.

6.3.1 Soil collection and experimental design

Nine locations within New Zealand (Table 6.1) that were representative of regions that might be suitable for mānuka plantation were selected to investigate possible effects of soil

type on the nectar quality, growth and flowering of this species. At each location a composite bulk sample (200 kg) was collected from 10 random points within a one hectare location. One square metre was dug for each collection and soil was sampled down to a depth of 30 cm. The top 5 cm of soil was discarded from collection to remove any grass or herbage. A commercial potting media developed specifically for shrubs and trees was used as the control soil in the study. Soil collected from each location was sieved and roto-mixed to ensure homogenous distribution before transfer into 300 mm diameter plant pots. The total pot volume for each soil was matched at 8 litres per pot. Three proprietary cultivars of mānuka, labelled R, G and Y were planted into each soil and grown for 15 months. All treatments (cultivar and soil combinations) were replicated ten times (300 pots in total). Each cultivar set were genetically identical clones of the parent material; they were the same age and grown from cuttings of a single parent plant. Cuttings were supplied by Comvita New Zealand and were transferred at age 3 months to the potted soil. The soil experiment was conducted at the Massey University Plant Growth Unit; Palmerston North, New Zealand. Plants were watered with in-line irrigation to ensure each pot received the same volume of water. All experimental units were planted into one greenhouse in a randomised latin block design. The glasshouse was ventilated by a fan system to provide a climate of mild temperate conditions. Plant response parameters measured were: basal stem diameter, plant height, relative growth rate (RGR), nectar DHA concentration normalised to 80 °BRIX, nectar volume, start of flowering, flowering period and seed capsules numbers. The glasshouse trial ran for a period of 15 months, growth parameters were measured 3 monthly over a 15 month period from the 1st of October 2012 until January 2014. Nectar was collected when plants flowered and seed capsules numbers were counted near the end of the experiment in December 2014.

Table 6.1. Description of soil used for the experimental analysis. Soils are from a range of locations within the North and South Island of New Zealand plus commercial potting mix as a control soil.

Code	Soil Location and NZTM Co-ordinates	Soil Classification	Soil profile texture group	Parent Material
S1	MAXWELL: NZTM 1760836E 5596919N	Orthic Brown	Brown Soil	Sandstone and limestone: Iron Oxide
S2	LINCOLN: NZTM 1554648E 5168342N	Typic Immature Pallic Soils	Silty	Alluvium
S3	WANGANUI: NZTM 1774789E 5610367N	Orthic Brown	Brown Soil	Sandstone/siltstone
S4	PAHIATUA: NZTM 1834148E 5519656N	Pallic Orthic Brown	Silty Loam	Loess on rock
S5	CONTROL	Commercial growth media (potting mix)	Organic	Peat and Pumice
S6	RUATITI: NZTM 1790106E 5647255N	Allophanic Brown	Brown Soil	Sandstone and Volcanic ash
S7	TAKAKA: NZTM 598371E, 5450593N	Sandy Recent	Sandy Loam	Alluvium
S8	TE AWAITI: NZTM 1821261E 5409918N	Orthic Brown	Brown Soil	Sandstone and limestone: Iron Oxide
S9	WHAREKOHE: NZTM 1699578E 6042684N	Densipan Ultic	Clay	Quartz rich Rock
S10	WAIKATO PEAT: NZTM 808553E, 5800648N	Fibric Organic	Organic	Peat

6.3.2 Nectar Collection

Leptospermum scoparium flowers, although fully open, do not start secreting nectar until the anthers are mature and fully extended from the hypanthium edge. This can be denoted as stage IV in flower stage development in Eucalyptus, another member of the Myrtaceae family [24]. Stage IV of flower development in *L. scoparium* lasts on average 4-5 days. In the later stage V, anthers start to shrivel and detach from the hypanthium. Nectar production continues from stage IV to Stage V, however, these flower stages were not included in the collection protocol to allow standardisation of the nectar collection at optimum health of the flower. Flowers were sampled at randomly selected times between 10 am and 2 pm on sampling days. Flowers were removed from plants using forceps and placed on ice. Nectar from each plant was collected immediately following flower collection and stored at -80oC. Nectar collection was performed using a stereo microscope and pipette to ensure accurate rinsing and maximum nectar retrieval. Nectar was removed from individual flowers by rinsing the hypantheum with 5 µl of distilled H2O using a pipette then removing all liquid. Nectar from each plant was pooled into eppendorf tubes for weighing for total yield measurements. The tubes and water were weighed prior to collection to allow for accurate nectar volume measurements. Preliminary experimentation revealed that at least 15 flowers needed to be sampled to allow for the intra-variability of nectar composition within each plant (data not shown), thus for the current study 20 flowers were taken from each plant. Cultivar Y did not flower in significant numbers for nectar collection on any of the natural New Zealand soils, and therefore no comparison on nectar yield and quality could be made for cultivar Y in this experiment.

6.3.3 Analysis of Dihydroxyacetone

DHA levels in the nectar samples were measured using aqueous extraction, derivatisation and analysis by an adapted HPLC method [25] for analysing DHA in honey samples. The standard range and reaction chemical and sample dilutions were modified to allow for the analysis of the lower concentration and the smaller volume (100-200 µL) of these components in floral nectar for each sample in the current study.

6.3.4 HPLC Conditions

Analyses were performed on a Perkin Elmer Series 200 Pump and Auto sampler with a Flexar photo diode array detector (λ = 263 nm). HPLC separations were performed on a Synergi Fusion column (75 × 4.6 mm, 4 µm particle size) held at 30^{o}C. Mobile phase A was water: acetonitrile (ACN), 70/30, v/v and mobile phase B was 100% ACN. The following 23 minute gradient elution was employed: A: B = 90:10 (isocratic 2.5 min), graded to 50:50 (8.0 min), graded to 0:100 (1.5 min), 0:100 (isocratic 7.0 min), graded to 90:10 (1.0 min), 90:10 (isocratic 3.5 min), detection wavelength 263 nm.

6.3.4.1 Preparation of Reaction Solutions

Hydroxyacetone (HA) (3.01 mg/ml) formed the HA internal standard solution. The *O*-(2, 3, 4, 5, 6-pentafluorobenzyl) hydroxylamine (PFBHA) derivatising reagent was 19.8 mg/ml in citrate buffer (0.1M) adjusted to pH 4 with sodium hydroxide (NaOH) (4M). DHA (3.88 mg/ml) formed the DHA standard solution.

For the preparation of standards, DHA standard stock solutions (100, 80, 60, 40, 20, 10 and 0 µl) were added to tubes 1 to 7, respectively and made up to 100 µl with nanopure water. For sample analysis, 20 µl of nectar or standard was pipetted into a mix tube and 25 µl of the HA was added. Derivatisation steps were performed at 25^{o}C in a controlled temperature room. Each of the HPLC samples and standards was

thoroughly mixed and placed in a rack on a rotating table for 1 hour to allow complete dissolution. PFBHA derivatising solution (100 µl) was added to each test tube, which was mixed and placed in a rack on a rotating table for 1 hour to allow for complete derivatisation. ACN (1.5 ml) was added to each test tube and mixed. Nanopure water (0.5 ml) was added to each test tube and mixed. Samples were then syringe filtered with a 0.22 µm filter into HPLC vials. Vials were placed into the auto sampler and run overnight with repeat analysis of standards analyzed through each run to check stability of the analysis. DHA calibration curves were generated from standards by linear regression using the HPLC peak area ratios of DHA: HA plotted against the mass of the DHA. Mass DHA content of the nectar samples were determined against these calibration curves.

6.3.5 Soil Analysis

A soil corer with internal volume of 80 cm^3 was used to collect a soil sample from each pot in the glasshouse at the end of the experiment (January 2014). Replicates 1-3; 4-6 and 7-10 for each treatment unit were combined, mixed then air dried and sieved to less than 2 mm to generate 3 replicate soil samples for each soil/cultivar combination. Soil samples were analysed for soil volume (g/ml), pH, cation exchange capacity (CEC), macronutrients; phosphorus (P), sulphate ($SO_{4)}$, potassium (K), calcium (Ca), magnesium (Mg), sodium (Na) by the Fertilizer and Lime Research Centre, Palmerston North according to the methods of Blakemore et al. (1987) [26]. pH was measured in water at a soil : water ratio of 10.25 g/ml. Plant available phosphorus was assessed using the Olsen P method with bicarbonate extraction followed by the Murphy and Riley (1962) colorimetric test [27]. S was assessed according to the method of Johnson and Nishita (1952) [28] but with a Technicon modern auto-analyser. CEC and cations;

Ca^{2+}, Mg^{2+}, Na^{2+} and K^{+} were assessed by semi-micro leaching with 1M ammonium acetate H^{+} approximated by pH drop elemental analysis.

The plant available micronutrients boron (B), chlorine (Cl), iron (Fe), manganese (Mn), zinc (Zn), copper (Cu) and cobalt (Co) were quantified by Hills Laboratories; Waikato Innovation Park, Ruakura Lane, Hamilton. Soils for micronutrient analysis were air dried at 35-40 ^{o}C overnight with residual moisture typically 4%. Boron was assessed by extraction with boiling dilute calcium chloride followed by ICP-OES, Chloride was assayed by saturated calcium sulphate extraction followed by Potentiometric Titration. The plant-available concentration of iron, manganese, zinc, copper and cobalt was determined by 0.05 M EDTA extraction followed by ICP-OES analysis of the extract.

6.3.6 Plant Growth Measurement Methods

Plant height and basal stem were measured every 3 months, using callipers to measure the basal stem diameter at the root stem node junction, denoted as the point where the first lateral root appears. Plant height was measured using the standing height at highest point with a meter rule. Relative growth rate (RGR) in this work is defined as the average increase in basal stem diameter per month.

Statistics

6.3.7 Soil Composition

Multivariate statistical methods using principal component analysis (PCA) was used to analyse the overall differences in the composition of the 10 soils in order to establish the main differences between the soils. PCA results give an overview of how similar or different the samples are with regard to each other and also the variations within the samples themselves is shown by how well they group together in any number of PC plots. Loadings plots show the influence of the various parameters measured on the

separation and groupings in the relevant in PC plot. Components close to the X-Y intercept in the loadings plots have a low influence on the separation and groupings in the PCA plots.

6.3.8 Plant data

Minitab statistical software Version 16.2.4 was used to determine whether the means of the response variables measured differed between the cultivars and/or the soils. Statistical methods, ANOVA, General Linear Model (GLM) and comparison of means using a Tukey pair-wise comparison at the 95% confidence level ($P \leq 0.05$) were used to analyse the response data and assess for any significant differences between each cultivar, each soil type and any interactions between the cultivars and the soils. Any parameters showing significant differences were analysed against soil composition using partial least squares (PLS) regression techniques and analysis of relevant associated loadings. Soil composition data is highly co-linear due to the inter-relation of the parameters assayed. Co-linearity in data results in errors when using multiple regression analysis, in comparison PLS regression analysis allows for data of this type and was therefore chosen as the best method [29]. Outliers were hotelled and removed at the 95^{th} percentile. R-squared values, model validation, weighted loadings and PLS score plots were analysed for influencing soil composition factors in the PLS results. PLS loadings were analysed to understand how much each variable contributes to the data and to interpret relationships between variables [30, 31].

Statistical results for the analysis of plant data indicated that the 3 cultivars varied in their responses to the soils and therefore each line was also analysed for any correlation against soil composition separately. Interactions between the cultivars and soils were also analysed. Cultivar Y could not be sampled for nectar analysis as

minimal number (12%) of plants flowered and those plants did not flower in substantial numbers to allow nectar collection.

6.4 Results and discussion

6.4.1 Soil Chemistry

The chemical properties of the ten soils used in this experiment are summarised in Table 2. Multivariate analysis of the soil properties in Table 6.2 using PCA illustrated differences between the 10 soils (Figure 6.1).

The distribution of the soils in the PCA plot shows groupings of soils with similar properties (Figure 6.1). The loadings plots (Figure 6.1B and D) show which variables are influencing the positioning of the soils in the PCA plot. The soils are generally tightly grouped showing similar properties, except soil 5 (Figure 6.1A). The spread in the PCA for soil 5 is variable indicating that the properties of soil 5 are heterogeneous. Soil 5 is a commercial potting mix, with nutrients in the form of slow-release fertiliser grains. Data heterogeneity for this soil may represent random inclusion of fertiliser grains in the soil cores sampled. Soil 5 is grouped separately to the right of the PC plot away from the other soils indicating that the properties of soil 5 is substantially different from the natural New Zealand soils.

Soils 1, 2, 7 and 10 are similar in PC1 (Figure 6.1A). Soil 1 and 2 have the same position in PC1 and PC2 and are essentially the same with respect to soil chemistry. Soils 3, 6, 4 and 8 form another grouping along PC1, while Soil 9 spreads across both of these two major groupings along PC1. Figure 6.1A shows the loadings plot for PC1 versus PC2 and indicates that the groupings along PC1 are due mainly to differences in plant-available Fe concentration (94% of the variation). The distribution of soils across PC2 (Y axis) are

mainly due to the plant-available manganese and sulphate concentration in the soils, with other soil chemistry parameters measured not spreading in the X or Y direction in the loadings plot but located around the central X-Y axis. This indicates that they have very little influence on soil composition differences within the 10 soil types (Table 6.2). The plot of PC2 against PC3 (Figure 6.1C) shows minor differences in the chemistry of the ten soils due to sulphate and plant-available manganese concentration, with some additional influence from phosphorus, plant-available copper, calcium and plant-available iron concentration (Figure 6.1D).

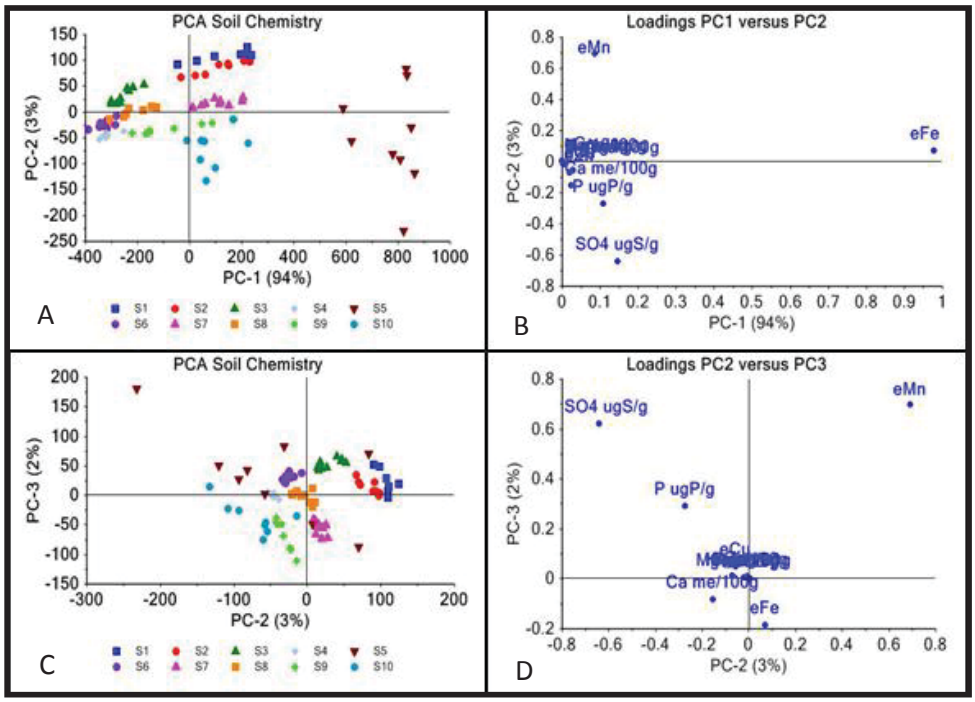

Figure 6.1. PCA plots and related loading plots of the soil chemistry data for all ten soils. Individual soils are labelled and coloured separately to show soil groupings. PCA loadings plots show the influence of soil compositional data related to the separations in the PCA plots **A**, PC1 vs PC2, **B**, Loadings PC1 vs PC2, **C**, PC2 vs PC3, **D**, Loadings PC2 vs PC3.

Table 6.2. Average macronutrient and micronutrient concentration in the ten soils used for the glasshouse trial. Soil codes S1-10 are defined in Table 6.1. New Zealand (NZ) medium range values are taken from the Hill Laboratories database and represent the mean concentration of these macronutrients and micronutrients in productive New Zealand soils (http://www.hill-laboratories.com).

	Soil	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	NZ medium range
Fertility Parameter	**unit**											
pH		6.10	5.50	5.30	5.60	5.20	5.40	6.00	5.70	5.30	5.10	5.8-6.2
Olsen P	mg/L	12.70	30.60	9.20	5.70	99.80	23.90	43.60	11.30	33.10	41.90	15-25
SO_4	mg/kg	6.80	53.70	14.20	8.20	1200.00	26.50	14.70	25.90	38.80	62.80	
K	me/100g	0.36	0.45	1.17	0.61	5.99	0.50	0.51	0.81	0.15	0.30	0.3-0.5
Ca	me/100g	9.70	70	3.70	2.40	30.60	3.00	4.40	3.10	6.80	69.90	3.0-9.0
Mg	me/100g	2.12	1.10	1.13	0.76	7.21	0.42	1.40	2.05	0.96	2.35	0.8-1.5
Na	me/100g	0.20	0.20	0.13	0.06	1.02	0.07	0.14	0.29	0.06	0.09	0.2-0.4
CEC	me/100g	17.00	16.00	20.00	17.60	77.00	18.00	9.00	13.00	15.00	123.00	
Soil Volume	g/mL	1.00	1.10	0.91	0.77	0.51	0.85	1.15	1.02	0.96	0.44	0.6-1.00
B	mg/kg	0.64	0.44	0.62	0.93	4.07	0.63	0.28	0.52	0.39	1.48	CL[1] >0.4
Cl	mg/kg	19.88	17.89	30.33	54.22	441.00	47.89	6.67	16.67	21.88	no data	-
Ext[2]. Fe	mg/kg	764.75	758.33	385.00	329.89	1392.78	314.78	761.67	411.78	597.25	713.71	-
Ext. Mn	mg/kg	187.63	161.78	121.33	22.11	138.22	59.44	62.33	57.00	5.88	17.29	CL>50
Ext. Zn	mg/kg	3.94	3.23	3.22	1.46	27.50	3.57	2.51	1.99	2.76	28.10	CL>1.1
Ext. Cu	mg/kg	3.04	1.59	3.90	1.24	46.87	4.09	7.60	0.94	1.31	7.69	CL>0.8
Ext. Co	mg/kg	1.34	1.27	0.91	0.33	0.80	0.40	2.14	0.50	0.00	0.50	CL>1.0

6.4.2 Plant Response to Soil Chemistry

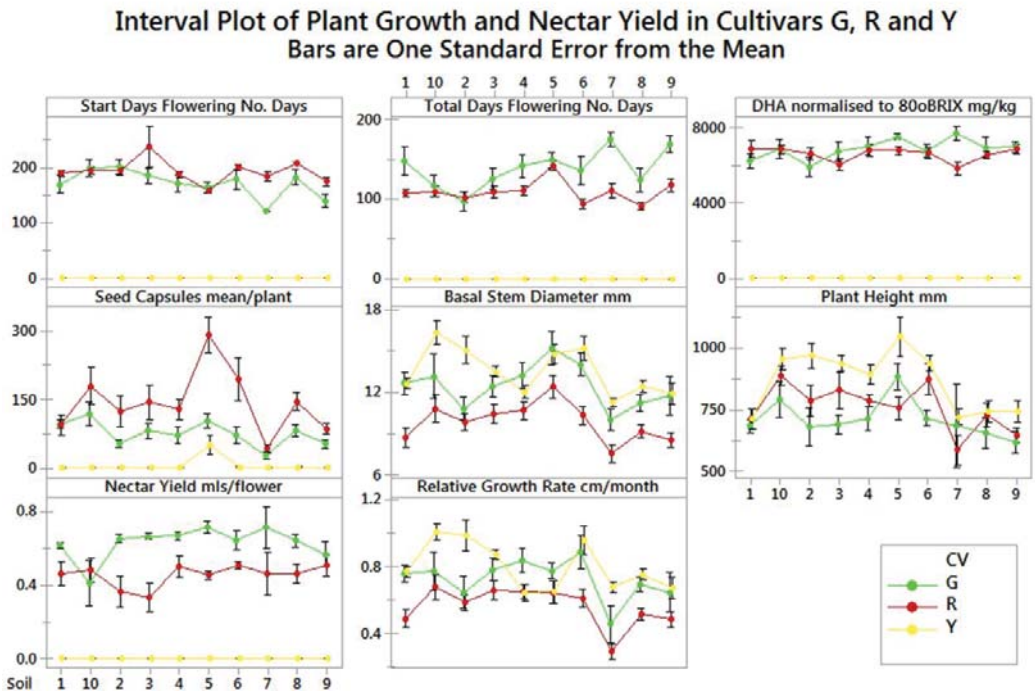

Figure 6.2. Relative start day of flowering, total days flowering, DHA concentration normalised to 80^{o}BRIX, numbers of seed capsules, basal stem diameter, plant height, nectar yield and relative growth rates for each cultivar R, G and Y. Plots are mean value and standard error.

6.4.2.1 Plant Growth

Basal stem diameter and plant height were significantly affected by soil type for each of the three cultivars (Figure 6.2 and Tables 6.3-6.6) indicating that soil type has a significant influence on the growth of *L. scoparium.* For cultivar R, basal stem growth was significantly affected by soil composition (P≤0.001). Plants on soil 5 had a greater basal stem diameter than plants on soils 8, 1, 9 and 7. Plant height was also affected by soil composition (P≤0.002) with plants on soils 10 and 6 having greater height than on soils 9 and 7 (Figure 6.2, Table 6.3). For Cultivar G, the differences in basal stem diameter were not statistically significant however the trend showed a difference in diameter of 5.27 mm with soil 5 having the largest stem diameter and soil 7 the

smallest diameter. Cultivar G growing on soil 5 had the greatest plant height (884.2 mm) at the end of the experiment, however this was only significantly greater than soil 9 (616.6 mm), ($P \leq 0.046$) (Figure 6.2 and Table 6.5). For cultivar Y, plants on soil 10 had a significantly greater stem diameter than plants on soils 1, 8, 4, 9 and 7 ($P \leq 0.001$). Height of Cultivar Y was greater on soils, 5, 2, and 10 relative to the other soils ($P \leq 0.001$) (Figure 6.2 and Table 6.5).

The relative growth rate (RGR) of the cultivars showed a significant interaction between soil and cultivars with the effect of soil composition on the RGR being dependant on the cultivar ($P \leq 0.005$). For example the RGR for Cultivar Y was higher on soil 10 than soils 7, 9 and 5, whereas for cultivar G, RGR was higher on soils 6 and 4 than soil 7 (Figure 6.2 and 6.5).

Overall, soils 5 and 10 had, on average, the most growth with plants on soil 7 consistently exhibiting less growth than on any other soil (soil 7 is a sandy loam with low fertility; Table 6.2). PCA plots highlight that soils 5 and 10 are mainly characterised by higher sulphate concentration relative to the other soils (negative scores along PC2; Figure 6.1; Table 6.2).

PLS regression was performed on growth and floral data against soil chemistry data to investigate if any individual or groups of soil compositional components were significantly influencing the response parameters measured. Plant response parameters for each cultivar were individually correlated against the soil properties in Table 6.1. Table 6.6 lists all parameters correlated against each response parameter measured for each cultivar.

6.4.2.2 Flowering Phenology

Onset of Flowering. The onset of flowering (Start Days) varied as a function of soil for cultivar R but not cultivar G. Cultivar Y flowered poorly on most soils except soil 5 and therefore no quantitative assessment of start of flowering against soil type for this cultivar is made. Cultivar R began flowering significantly earlier on soils 5 and 9 than on soil 3 ($P \leq 0.002$). Plants on soils 5 and 9 flowered up to 5 weeks longer than plants on soils 3, 8 and 2 and 10. PCA analysis of the soil properties (Figure 6.1) does not fully elucidate what is influencing flowering onset although regression analysis (Table 6.6) of all properties show a low but significant correlation ($R^2 = 0.37$) between manganese, chloride and sulphate concentration and onset of flowering (start days).

Flower Numbers. To assess flower numbers we assumed seed capsule numbers were a direct function of flower numbers and that all plants had an equal chance of being pollinated. The interaction of soils with cultivars was significant for flower numbers and depended on the cultivar ($P \leq 0.032$). For Cultivar Y, plants on soil 5 had significantly higher flower numbers than any of the other 9 soils with no significant difference between the other 9 soils. For cultivar R, plants on soil 5 had significantly higher numbers of flowers than those on soils 8, 3, 4, 2, 1, 7 and 9 but were the same as soils 6 and 10. Flower numbers on soils 7, 1 and 9 were significantly lower than the other soils. Flower numbers for cultivar G were not significantly affected by soil chemistry. We therefore suggest that soil properties have a significant effect on *L. scoparium* flowering which is important with respect to overall nectar yield (more flowers will lead to more nectar). Regression analysis in Table 6.8 suggests that sulphate, manganese, iron and calcium concentration and CEC have a significant effect on flower numbers ($R^2 = 0.31$ for cultivar R and $R^2 = 0.94$ for cultivar Y).

Flowering Period. Cultivar G had a significantly longer flowering period than cultivar R. However, the onset of flowering varied as a function of soil for cultivar R but not cultivar G. Cultivar R showed a significant difference ($P \leq 0.001$) where plants on soil 5 flowered for significantly longer periods than those on soils 4, 3, 1, 2, 6 and 8 (Figure 6.2). On average, cultivar R flowered for periods between 30 and 50 days longer on soil 5 than the same cultivar on soils 4, 3, 1, 2, 6 and 8. Cultivar Y flowered poorly and therefore no assessment of the effect of soil on the flowering period of cultivar Y is made. Regression analysis of flowering period for cultivar R against the soil parameters indicates a significant influence of sulphate, manganese and chloride concentration and CEC on flowering period ($R^2 = 0.65$).

Table 6.3. Averaged data is displayed for plant response parameters for Cultivar R: Plant Height (HT) Dihydroxyacetone (DHA), Nectar Yield in g/flower Relative Growth Rate (RGR) is the average change in basal stem diameter per month over the length of the experiment, Basal Stem diameter (BS), seed capsule numbers (Mean Caps), start of flowering (Mean Start Days) and flowering period. Results from a Tukey Comparison of means at the 95.0% confidence level are displayed in the columns next to the parameter value means that do not share a letter are significantly different. Plant growth parameters are those measured at the end of the experiment. SC = soil code.

SC	Mean Start Days	SC	Mean Flowering Period	SC	Mean DHA	SC	Mean Caps	SC	Mean nectar Yield g/flower	SC	Mean BS (mm)	SC	Mean HT (mm)	SC	Mean RGR (mm/month)
3	237.7 A	5	142.00 A	1	6910 A	5	292.11 A	6	0.0051 A	5	12.421 A	10	889.3 A	10	0.68 A
8	207.4 A B	9	117.90 A B	9	6881 A	6	195.30 A B	9	0.0051 A	10	10.823 A B	6	872.7 A	3	0.66 A
6	200.4 A B	4	110.80 B	10	6873 A	10	180.75 A B	4	0.0051 A	4	10.673 A B	3	828.3 A B	4	0.65 A
10	194.8 A B	7	110.40 A B	4	6828 A	8	146.00 B	10	0.0048 A	3	10.391 A B	4	787.0 A B	5	0.65 A
2	193.8 A B	10	109.56 B	5	6796 A	3	144.67 B	8	0.0046 A	6	10.305 A B	2	784.6 A B	6	0.61 A
1	189.4 A B	3	109.22 B	6	6668 A	4	128.90 B	7	0.0046 A	2	9.815 A B	5	755.6 A B	2	0.59 A B
4	188.0 A B	1	107.89 B	2	6630 A	2	125.00 B	1	0.0046 A	8	9.157 B	8	727.4 A B	8	0.52 A B
7	184.4 A B	2	102.25 B	8	6561 A	1	95.33 B	5	0.0046 A	1	8.680 B	1	711.0 A B	1	0.49 A B
9	174.9 B	6	94.30 B	3	6083 A	9	86.00 B	2	0.0037 A	9	8.517 B	9	645.7 B	9	0.49 A B
5	158.9 B	8	91.70 B	7	5854 A	7	41.80 B	3	0.0033 A	7	7.504 B	7	584.4 B	7	0.29 B

Table 6.4. Averaged data is displayed for plant response parameters for Cultivar G: Plant Height (HT), Dihydroxyacetone (DHA), Nectar Yield in g/flower Relative Growth Rate (RGR) is the average change in basal stem diameter per month over the length of the experiment, Basal Stem diameter (BS), seed capsule numbers (Mean Caps), start of flowering (Mean Start Days) and flowering period. Results from a Tukey Comparison of means at the 95.0% confidence level are displayed in the columns next to the parameter value; means that do not share a letter are significantly different. Plant growth parameters are those measured at the end of the experiment. SC = soil code. Parameters values are listed from the highest to lowest inn each column with the relevant soil code listed alongside.

SC	Mean Start Days	SC	Mean Flowering period	SC	Mean DHA	SC	Mean Seed Capsules	SC	Mean Nectar yield g/flower	SC	Mean BS(mm)	SC	Mean HT(mm)	SC	Mean RGR (mm/month)
2	201.2 A	7	175.50 A B	7	7713 A	10	118.50 A	5	0.0078 A	5	15.27 A	5	884.2 A	6	0.89 A
10	198.9 A	9	169.14 A	5	7554 A	5	103.00 A	7	0.0072 A B	6	14.06 A	10	791.4 A B	4	0.84 A
3	185.6 A	5	149.40 A B	4	7045 A	1	96.29 A	4	0.0067 A B	4	13.20 A	6	715.2 A B	3	0.78 A
8	182.5 A	1	148.86 A B	9	6994 A	8	82.25 A	3	0.0067 A	10	13.18 A	4	713.5 A B	5	0.78 A
6	179.0 A	4	141.50 A B	8	6959 A	3	81.38 A	2	0.0065 A B	1	12.67 A	3	688.4 A B	10	0.78 A
4	171.2 A	6	136.40 A B	10	6882 A	6	71.67 A	6	0.0065 A B	3	12.46 A	7	682.0 A B	1	0.76 A
1	169.4 A	3	125.78 A B	6	6786 A	4	71.50 A	8	0.0065 A B	9	11.75 A	1	681.7 A B	8	0.70 A
5	164.0 A	8	124.63 A B	3	6780 A	2	53.75 A	1	0.0062 A B	8	11.19 A	2	678.8 A B	2	0.65 A
9	139.6 A	10	117.00 A B	1	6271 A	9	52.43 A	9	0.0057 A B	2	10.82 A	8	654.0 A B	9	0.65 A
7	121.0 A	2	97.29 B	2	5953 A	7	27.50 A	10	0.0041 B	7	10.00 A	9	616.6 B	7	0.46 A

Table 6.5. Averaged data is displayed for plant response parameters for Cultivar Y: Plant height (HT), Relative Growth Rate (RGR) is the average change in basal stem diameter per month over the length of the experiment, Basal Stem diameter (BS) and seed capsule numbers (Mean Caps). Results from a Tukey Comparison of means at the 95.0% confidence level are displayed in the columns next to the parameter value; means that do not share a letter are significantly different. Plant growth parameters are those measured at the end of the experiment. SC = soil code. Parameters values are listed from the highest to lowest inn each column with the relevant soil code listed alongside.

SC	Mean BS(mm) Y	SC	Mean HT (mm) Y	SC	Mean CAPS	SC	Mean RGR(mm/month)
10	16.38 A	5	1046.60 A	5	51.80 A	10	1.01 A
6	15.25 A B	2	968.67 A	6	2.30 B	2	0.99 A
2	15.10 A B C	10	955.80 A	8	1.20 B	6	0.96 A
5	14.80 A B C D	3	938.90 A B	1	1.00 B	3	0.88 A B
3	13.50 A B C D E	6	938.20 A B	10	0.60 B	1	0.77 A B
1	12.65 B C D E	4	893.44 A B C	3	0.30 B	8	0.75 A B
8	12.43 B C D E	9	743.00 B C	4	0.11 B	7	0.68 B
4	12.04 C D E	8	743.00 B C	2	0.11 B	9	0.67 B
9	11.91 D E	7	719.33 C	9	0.00 B	5	0.65 B
7	11.27 E	1	712.00 C	7	0.00 B	4	0.65 B

6.4.2.3 Nectar DHA concentration and yield

There was no significant effect of soil type or cultivar on DHA concentration (Figure 6.2 and Tables 6.3 & 6.5). However, there was a significant difference in overall nectar yield (g sugar per flower) between cultivars R and G. Cultivar G had the highest sugar yield at 0.0063 g of sugar per flower, compared with R, at 0.0046 g of sugar per flower (Figure 6.2). The nectar yield for Cultivar Y could not be measured due to limited flowering. ANOVA and GLM showed a significant interaction ($P \leq 0.038$) between soils and cultivars for yield, indicating that the difference between the cultivars is affected by soil. Cultivar G had a significantly higher nectar yield on soils 5 and 3 than soil 10; up to 0.0037 g on soil 5 which is a 90% increase in yield. Yield on soil 10 is up to 90% lower than other soils and indicates that soil type is an important factor to consider in terms of potential overall yield to the honey industry. Indications from the PCA plot of PC2 versus PC3 illustrates that the increased yield in G on soil 5 and 3 is influenced by calcium levels although PLS regression analysis shows no significant influence on yield from soil properties. The factor influencing nectar yield in this cultivar is not clear from these results. Possibly soil structure and relative water availability due to soil structure could be a possible confounding factor in overall nectar yield and warrants further investigation.

6.4.2.4 Regression analysis of response parameters against each other

Correlations of response parameters against each other were also assessed. Linear regression analysis of the various growth factors against the various flowering factors showed no correlation indicating that aspects of flowering (start of flowering, flowering period and flower numbers) are not associated with the growth rate or

amounts of growth in *L. scoparium*. Using simple linear regression for response factors against each other there is a significant correlation R^2 = 0.87 between the start of flowering and how long the plants flowered for. There was no correlation between start of flowering or flowering period against flower numbers. A simple regression of soil structure type against response parameters showed some correlation of soil structure with the start of flowering and plant height but no correlation with the other parameters measured.

6.4.2.5 Interaction effects between the cultivars grown on the ten soils

Interaction effects represent the combined effects of factors on the response parameter measured. When an interaction effect is present, the impact of one factor depends on the level of the other factor. Differences in the growth and flowering response parameters on the soils were noted between the cultivars. ANOVA analysis showed a significant interaction for some parameters between the cultivars and the soils, indicating that response of the plants to the soil depended on the cultivar. The three cultivars R, G and Y did not respond to soils in the same pattern; some parameters were the same for a particular soil type and some were not. This has an important impact to the mānuka honey industry indicating that overall yield will depend in some cases on the cultivar chosen and the soil in the region chosen for planting. Table 6.7 summarises the significance of the interactive effect of soils and cultivar for the measured parameters. For the response parameters that showed no interaction between the soil type and cultivar, table 6.8 presents the relative ranking of the cultivars for each parameter. Cultivar G had a significantly longer flowering period than R, however the onset of flowering was earlier for R than G. For basal stem and height, Cultivar Y was greater than R and G, however G had a greater basal stem

than R but R had greater height than G. In terms of DHA composition there was no significant difference between cultivars G and R. For RGR, flower numbers and nectar yield in grams of sugar per flower, a significant interaction between the soil and the cultivar was observed and that the effect of soil depended on the cultivar (Table 6.7). Table 6.8 describes the significant interactions among soils and cultivars and quantifies the similarity between the cultivars for those parameters.

The interactions described in Table 6.8 show that nectar yield for cultivars G and R was the same on all soils except for soils 3 and 5. Relative growth rates (RGR) for all three cultivars were the same on soils 3, 4, 5, 8, 9 and 10 (i.e. Y=G=R). However, on soils 1, 6 and 7, the RGR for Y and G were the same (Y=G) and G and R were the same (G=R), however Y≠R. On soil 2 the RGR of G=R however Y≠G, and Y≠R. The interaction of cultivar and soil for seed capsules was complex. Cultivar G=R for seed capsules for all soils except soil 5. For soil 5 all three cultivars had a different response in terms of seed capsule numbers. For the other nine soils whether G=Y or R=Y was very variable in all the soils (Table 6.8). This data reiterates that the response of the plants to the soil range depends on the cultivar and that overall yield will depend in some cases on the cultivar chosen and the soil in the region chosen for planting.

6.4.2.6 Ranking soils for best performance based on growth and floral data

Data on the growth and floral performance of the cultivars tested in this work were used to model soils which had the highest potential to increase overall yield for the mānuka honey industry. Two key parameters of this model are Plant Nectar Yield (Yd);

Equation 1:

$$Yd = S\ X\ DHA\ X\ SC$$

Where S = g of sugar per flower, SC = seed capsule numbers (a measure of floral density) and DHA = dihydroxyacetone mg/kg and Nectar Potential (NP); Equation 2:

$$NP = FP\ X\ Yd$$

Where NP = nectar potential, FP = flowering period and Yd = Plant DHA Yield (Yd) = (ug sugar/flower*DHA conc.)*FD where FD = a measure of floral density.

Once calculated these parameters are used in this work to define the Mānuka Soils Index (MSI) (Equation 3):

$$\boldsymbol{MSI}_{value} = (NP*RGR)*BF$$

Where NP = Nectar Potential, RGR = Relative Growth Rate and BF = Bee Factor: i.e. * 1 if all bees out *0.5 if 50% bees out *0 if no bees out.

Calculated NP and MSI values are displayed in Table 6.9. The greatest nectar potential was recorded for soil 5 for both cultivars. Calculated nectar potential (NP) provides data on the potential nectar resources available to bees in terms of a food source during a flowering season. For cultivar R, NP for soil 5 was significantly greater than all other soils with no significant difference in NP between the other soils. For cultivar G, NP was only significantly greater for soil 5 relative to soil 2.

The MSI values in Table 6.9 quantify the soils which offer the highest potential for overall NPA honey yield for each cultivar. The MSI value shows that soil 5 has the highest potential for both cultivar R and G, whereas soil 7 has the lowest potential for cultivar R and second lowest for Cultivar G. The MSI for Cultivar R on soil 5 was significantly greater than each of the other soils ($P \leq 0.001$), whereas the MSI for

Cultivar G on soil 5 was only significantly greater than on soil 2 ($P \leq 0.001$). Cultivar Y could not be assessed for an MSI value due to the lack of nectar data, but statistical analysis of data for seed capsules and RGR show that plants on soil 5 had the best overall growth and flowering. Soil analysis PCA results in Table 2 indicate that phosphorus, sulphate, chloride and iron are the most obvious elements that have significantly different concentrations in Soil 5. However PLS analysis results in Table 6.6 show that, iron, manganese and calcium are the common soil nutrients significantly influencing flowering in cultivars R and Y with sulphate and manganese the common soil nutrients influencing overall plant growth in all three cultivars.

Overall annual NPA honey potential would include assessment of bee activity, local and seasonal weather conditions, for example, reduced daily temperatures and rain at the time of flowering will affect bee activity and if plants flower when there is no bee activity then the potential MSI value is reduced to zero.

Table 6.6. PLS regression results from plant growth and nectar parameter data of cultivars R, G and Y regressed against soil chemistry. Sig = whether the PLS model was significant or not. RMSEC = relative mean square error of the calibration model. R^2 is the regression correlation value and Influence lists the main chemicals influencing the regression correlation.

CV	Response	Sig	R^2	RMSEC	Influence
R	start days	Y	0.37	29.88 days	Chloride ions, Manganese and Sulphate
R	length	Y	0.65	19.24 days	Chloride ions, CEC, Sulphate and Manganese
R	Seed Caps	Y	0.31	90.62 flowers	Calcium, CEC, Sulphate, Manganese and Iron,
R	DHA	N			
R	yield	N			
R	BS	Y	0.65	0.82mm	Calcium, CEC, Chloride, Sulphate and Zinc
R	HT	Y	0.73	63.95mm	Calcium, CEC, Sulphate and Manganese
R	RGR	Y	0.28	0.15mm	Calcium, CEC, Sulphate and Manganese
G	start days	N			
G	length	N			
G	Seed Caps	N			
G	DHA	N			
G	Yield	N			
G	BS	N			
G	HT	Y	0.7	63mm	Iron, Manganese and Sulphate
G	RGR	N			
Y	start days	NA			
Y	length	NA			
Y	Seed Caps	Y	0.94	4.5 flowers	Calcium, Iron and Manganese
Y	DHA	NA			
Y	volume	NA			
Y	BS	Y	0.68	1.11mm	Calcium, CEC, Chloride ions, Manganese and Sulphate
Y	HT	Y	0.51	86mm	Calcium, CEC, Chloride, Iron, Manganese and Sulphate
Y	RGR	Y	0.39	0.17mm	Calcium, CEC, Manganese, PO_4 and Sulphate

Note 1. Nectar parameters for Cultivar Y are not applicable (NA) as there was insufficient flower density to afford reliable measurement.

Table 6.7. Summary of analysis of Variance (ANOVA) to assess the differences between the cultivars in terms of response to measured parameters. Interaction category: Y= there was a significant interaction between the cultivar and the soils for that parameter in which case further data analysis is shown in Table 6.8; and N = no significant interaction between the soils and the cultivars and the overall differences between the three cultivars is listed in column CV.

Note 1. *= see interaction data Table 6.8.

Response	Interaction	CV
Flowering Period	N	G>R
Start day	N	R>G
Basal Stem	N	Y>G>R
Height	N	Y>R>G
DHA	N	G=R
RGR	Y	*
Yield	Y	*
Caps	Y	*

Table 6.8. Interaction results for the parameters that showed a significant interaction between the cultivar (CV) and the soil type (Table 6.7). This data shows that the effect of soil type on; seed capsule numbers, relative growth rate (RGR) and nectar yield depended on the cultivar and differed for each cultivar depending on the soil type. The interaction column designates which the cultivars had the same (=) or different (≠) interaction with the soils.

Soil Type	Parameter	Interaction	Soil Type	Parameter	Interaction	Soil Type	Parameter	Interaction
1	Seed Capsule #'s	G=R, G≠Y, R≠Y	1	Yield (g of sugar)	G=R	1	RGR	Y=G, G=R, Y≠R
2	Seed Capsule #'s	G=R, G≠Y, R≠Y	2	Yield (g of sugar)	G≠R	2	RGR	Y≠G, G=R, Y≠R
3	Seed Capsule #'s	G=R, G≠Y, R≠Y	3	Yield (g of sugar)	G≠R	3	RGR	Y=G=R
4	Seed Capsule #'s	G=R, G≠Y, R≠Y	4	Yield (g of sugar)	G=R	4	RGR	Y=G=R
5	Seed Capsule #'s	G≠R≠Y	5	Yield (g of sugar)	G≠R	5	RGR	Y=G=R
6	Seed Capsule #'s	G=R, G≠Y, R≠Y	6	Yield (g of sugar)	G=R	6	RGR	Y=G, G=R, Y≠R
7	Seed Capsule #'s	G=R, R≠Y, G=Y	7	Yield (g of sugar)	G=R	7	RGR	Y=G, G=R, Y≠R
8	Seed Capsule #'s	G=R, G≠Y, R≠Y	8	Yield (g of sugar)	G=R	8	RGR	Y=G=R
9	Seed Capsule #'s	G=R, G≠Y, R≠Y	9	Yield (g of sugar)	G=R	9	RGR	Y=G=R
10	Seed Capsule #'s	G=R, G≠Y, R≠Y	10	Yield (g of sugar)	G=R	10	RGR	Y=G=R

Table 6.9. Calculated nectar potential (NP) and MSI value for cultivars R and G. Ranking is defined by Tukey Comparison of means at the 95.0% confidence level. Means that do not share a letter are significantly different.

General Linear Model: NP versus Soil Type R			General Linear Model: NP versus Soil Type G			General Linear Model: MSI versus Soil Type R			General Linear Model: MSI versus Soil Type G		
Soil Type	Mean	Grouping	**Soil Type**	**Mean**	**Grouping**	**Soil Type**	**Mean**	**Grouping**	**Soil Type**	**Mean**	**Grouping**
S5	**8906.4**	**A**	S5	6926	A	S5	4554	A	S5	6775	A
S6	**4582.3**	**B**	S1	4727	A B	S6	2762	A B	S4	4178	A B
S3	**3498.6**	**B**	S9	4723	A B	S3	2356	B C	S1	3664	A B
S4	**3018.6**	**B**	S6	4578	A B	S4	2119	B C	S6	3657	A B
S10	**2956.5**	**B**	S8	4510	A B	S10	1915	B C	S9	3351	A B
S2	**2698.8**	**B**	S10	3975	A B	S2	1716	B C	S10	3338	A B
S8	**2604.3**	**B**	S3	3427	A B	S8	1315	B C	S8	3180	A B
S9	**2525.3**	**B**	S7	3338	A B	S9	1283	B C	S3	2504	A B
S1	**2193.2**	**B**	S4	2190	A B	S1	1157	B C	S7	1421	A B
S7	**961.2**	**B**	S2	1639	B	S7	271	C	S2	790	B

6.5 Discussion

Our results indicate that soil chemical characteristics in land areas being examined for high-yielding mānuka plantations is an important factor for consideration when selecting cultivars for the honey industry in New Zealand.

Our data indicate that contrary to previously reported studies, *L. scoparium* has improved growth and floral density on soils with increased nutrient concentration. For example, improved performance was recorded on the control soil which contained significantly higher concentrations of the nutrients sulphate, phosphorus, iron and chloride than the field NZ soils in this experiment and greater than the mean range of natural NZ soils (Table 6.2). Common soil components across all cultivars that support further investigation are CEC, sulphate, manganese, calcium, chloride, and iron concentration for plant growth; sulphate, manganese and chloride concentration for flowering period; and iron, manganese and calcium concentration for flower numbers.

The calcium content of soil has been associated with flowering in tropical trees, as well as lilac, hazel and red currant [20, 32]. While a relationship is demonstrated between

calcium and flower numbers in the current work, PLS analysis also indicates that the manganese and iron concentration of the soil may also affect flowering and growth of mānuka. Many woody shrubs are affected by iron and manganese deficiencies and these elements are more readily available to plants in soils with more acidic pH [32, 33]. The improved flowering on soils 5 & 9 may have been in response to greater plant-available manganese and iron resulting from the higher acidity (<pH 6) in those soils. However superior flowering performance on soil 1 with a pH of 6.6 and lower calcium levels is inconsistent with this observation and indicates some other confounding factor, possibly soil structure having an additional influence. Soil 1 is a sandy loam whereas soils 5 and 7 are peat based. An effect of soil structure on the growth performance of *Salix caprea* and *Prunus padus* has been reported previously (Wielgolaski*,* 2001), where plant development in *Prunus* and flowering in *Salix caprea* was correlated with the bulk density and gravel content of soil, this correlation was hypothesised to be related to the water holding potential of the soil or possibly the effect of the soil structure on the availability of nutrients.

Wielgolaski (2001) [32] suggests that soil factors, along with climatic factors, influence the timing of plant development including flowering on a number of plant species. In agreement, our observations showed that flowering phenology, nectar yield and plant growth was significantly different in the examined mānuka cultivars grown on a number of soils under constant and controlled environmental conditions, strongly suggesting that soil factors influence plant development and flowering in mānuka.

Furthermore additional plantation worth may be realised through targeted fertilisation of mānuka stands and plantations. The potential of fertilisation strategies should be

further researched through the correlation of field collected soil, plant data and field trials.

The usefulness of modelling mānuka growth and flowering using the MSI formulae to assess cultivar potential for various sites in New Zealand has the potential to inform the selection of locations for new mānuka plantations and additionally select cultivars relevant to soil characteristics of the various regions of New Zealand. Whilst *L. scoparium* does establish on nutrient-deficient soils and in marginal areas where there is little competition from climax forest, improved growth and flowering may be facilitated with improved commercial plantation development, and it is likely there is a degree of dependence on soil nutrition factors.

6.6 Acknowledgements

This work was supported by a Primary Growth Partnership (Ministry of Primary Industries, New Zealand) awarded to Mānuka Research Partnership (NZ) Limited, and a scholarship from Callaghan Innovation (NZ) to L.M.N. [Contract: MARP1001]. We would like to acknowledge the staff at the Plant Growth Unit at Massey University; Palmerston North, New Zealand for support with the glasshouse experimental setup. Chris Rawlingson, Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand for assistance with HPLC instrumentation and supporting staff from Massey University; Palmerston North, New Zealand, who assisted with nectar collection and plant growth data measurements.

6.7 References

[1] Adams CJ, Manley-Harris M, Molan PC. 2009. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydr. Res. 344:1050-1053.

[2] Dawson M. 1997 A history of *Leptospermum scoparium* in cultivation-Discoveries from the wild. The New Plantsman 4(1):51-59.

[3] Douglas MH, van Klink JW, Smallfield BM, Perry NB, Anderson RE, Johnstone P, Weavers RT. 2004. Essential oils from New Zealand manuka: triketone and other chemotypes of *Leptospermum scoparium*. Phytochemistry. 65:1255-1264.

[4] Forst JRG. 1986. Environmental Factors involved in growth, flowering and post harvest flowers of *Leptospermum scoparium*. Israel Journal of Botany 35:101-108.

[5] Greer DH, Muir LA, Harris W. 1991: Seasonal frost hardiness in *Leptospermum scoparium* seedlings from diverse sites thorughout New Zealand. New Zealand Journal of Botany 29:207-212.

[6] Haberlein H, Tschiersch KP. 1998. On the occurrence of methylated and methoxylated flavonoids in *Leptospermum scoparium*. Biochem. Syst. Ecol. 26:97-103.

[7] Morris JD. 1984. Establishment of trees and shrubs on a saline site using drip irrigation. Australian Forestry 47:210-217.

[8] Perry NB, Van Klink JW, Brennan NJ, Harris W, Anderson RE, Douglas MH, Smallfield BM. 1997. Essential oils from New Zealand mānuka and kanuka: Chemotaxonomy of *Kunzea*. Phytochemistry 45:1605-1612.

[9] Stephens JMC, Molan PC, Clarkson BD. 2005. A review of *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany. 43:431-449.

[10] Wilson JB, Yin RH, Mark AF, Agnew ADQ. 1991. A test of the low marginal variance (LMV) theory in *Leptospermum scoparium* (Myrtaceae). Evolution. 45:780-784.

[11] Atrott J, Haberlau S, Henle T. 2012. Studies on the formation of methylglyoxal from dihydroxyacetone in Manuka (*Leptospermum scoparium*) honey. Carbohydr. Res. 361:7-11.

[12] Donarski JA, Roberts DPT, Charlton AJ. 2010. Quantitative NMR spectroscopy for the rapid measurement of methylglyoxal in manuka honey. Ana. Methods. 2:1479-1483.

[13] Williams S, King J, Revell M, Manley-Harris M, Balks M, Janusch F, Kiefer M, Clearwater M, Brooks P, Dawson M. 2014. Regional, annual, and individual variations in the dihydroxyacetone content of the nectar of mānuka (*Leptospermum scoparium*) in New Zealand. J. Agric. Food Chem. 42:10332-40.

[14] Alonso-Blanco C, Mendez-Vigo B, Koornneef M. 2005. From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. International Journal of Developmental Biology. 49:717-732.

[15] Yin R, Mark AF, Wilson JB 1984. Aspects of the ecology of the indigenous shrub *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany 22: 483–507.

[16] Hall IR. 1977. Effect of applied nutrients and endomycorrhizas on *Metrosideros umbellata* and *Leptospermum scoparium*. New Zealand Journal of Botany 15:481-484.

[17] Bergin DO, Kimberley MO. 2014. Factors influencing natural regeneration of totara (*Podocarpus totara* D.Don) on grazed hill country grassland in Northland, New Zealand. Journal of Forestry Science 44:13-23

[18] Funk JM, Field CB, Kerr S, Daigneault A. 2014. Modeling the impact of carbon farming on land use in a New Zealand landscape. Environmental Science & Policy. 37:1-10.

[19] Brun LA, Le Corff J, Maillet J. 2003. Effects of elevated soil copper on phenology, growth and reproduction of five ruderal plant species. Environmental Pollution 122:361-368.

[20] Cardoso FCG, Marques R, Botosso PC, Marques MCM. 2012. Stem growth and phenology of two tropical trees in contrasting soil conditions. Plant and Soil 354:269-281.

[21] Ryser P, Sauder WR. 2006. Effects of heavy-metal-contaminated soil on growth, phenologyand biomass turnover of Hieracium piloselloides. Environmental Pollution 140:52-61.

[22] Sperens U. 1997. Long-term variation in, and effects of fertiliser addition on, flower, fruit and seed production in the tree *Sorbus aucuparia* (Rosaceae). Ecography. 20:521-534.

[23] Staggemeier VG, Diniz-Filho JAF, Morellato LPC. 2010. The shared influence of phylogeny and ecology on the reproductive patterns of *Myrteae* (Myrtaceae). Journal of Ecology. 98:1409-1421.

[24] Davis AR. 1997. Influence of floral visitation on nectar-sugar composition and nectary surface changes in Eucalyptus. Apidology 28:27-42.

[25] Windsor SAM, Brooks P, Williams S, Manley-Harris M. 2012. A convenient new analysis of dihydroxyacetone and methylglyoxal applied to Australian *Leptospermum* honeys. Journal of Pharmacognosy and Phytotherapy. 4:6-11.

[26] Blakemore LC, Searle BK, Daly BK. 1987. Methods for chemical analysis of soils. Lower Hutt, NZ: NZ Soil Bureau, Dept. of Scientific and Industrial Research, NZ Soil Bureau scientific report.

[27] Murphy J, Riley JP. 1962. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta. 27:31-36.

[28] Johnson CM, Nishita H. 1952. Micro-estimation of sulfur in plant materials, soils and irrigation water. Anal. Chem. 24:736-742.

[29] Mattila M, Koskinen K, Saloheimo K. 2004. Novel multivariate calibration methods for x-ray analyzer. New Technologies for Automation of Metallurgical Industry 221-226

[30] Eriksson L, Trygg J, Wold S. 2014. A chemometrics toolbox based on projections and latent variables. J. Chemometr. 28: 32-346.

[31] Esbensen KH, Guyot D, Westad F, Houmoller LP. 2004. Multivariate Data Analysis − In Practice. An Introduction to Multivariate Data Analysis and Experimental Design. 4th ed. CAMO: Oslo, Norway.

[32] Wielgolaski FE. 2001. Phenological modifications in plants by various edaphic factors. International Journal of Biometeorology 45:196-202.

[33] Riki G, Mobasser HR, Ganjali HR. 2014. Effect of iron and manganese foliar spraying on some quantitative characteristics of canola. International Journal of Biosciences. 5:61-68.

Chapter 7: Analytical FT-Raman spectroscopy to chemotype *Leptospermum scoparium* and generate predictive models for screening for dihydroxyacetone levels in floral nectar

This chapter has been published in the Journal of Raman Spectroscopy as: "Analytical FT-Raman spectroscopy to chemotype *Leptospermum scoparium* and generate predictive models for screening for dihydroxyacetone levels in floral nectar"

Elizabeth M. Nickless[1], Steve Holroyd[2], Jonathan M. Stephens[4], Keith C.Gordon[3§], Jason J. Wargent[1*]

[1] Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand

[2] Fonterra Co-operative Group, Fonterra Research & Development Centre, Palmerston North, New Zealand

[3] Department of Chemistry and Dodd-Walls Centre, University of Otago, Dunedin, New Zealand

[4] Comvita Innovations, Institute for Innovation in Biotechnology, University of Auckland, Auckland, New Zealand

*Corresponding Author: J. J. Wargent

Citation: Nickless et al. "Analytical FT-Raman spectroscopy to chemotype *Leptospermum scoparium* and generate predictive models for screening for dihydroxyacetone levels in floral nectar". Journal of Raman Spectroscopy **(2014).**

E.M.N. and J.J.W. designed the experimental plant work. E.M.N. carried out the experimental work, analysed the data and authored the paper. S.E.H. and K.C.G. provided technical guidance and expertise in FT-Raman and analysis. S.E.H., J.M.S. and J.J.W. were project supervisors to E.M.N. All authors approved the final manuscript.

7.1 Acknowledgements

This work was supported by a Primary Growth Partnership (Ministry of Primary Industries, New Zealand) awarded to Mānuka Research Partnership (NZ) Limited. We are grateful to Georgie Hamilton, Massey University, for assistance in preparing experiment setup, nectar collection and leaf samples, and also Sara Fraser, Department of Chemistry, University of Otago, for help with FT Raman use and support in using Unscrambler software.

7.2 Abstract

Leptospermum scoparium (Mānuka) is the source of nectar for Unique Mānuka Factor (UMF) honey. Spectroscopic techniques, in combination with principal component analysis and partial least squares regression analysis, were investigated as an analytical tool for building a predictive model of specific chemical components in the nectar of *L. scoparium* that are linked to UMF levels in honey. The chemical component of interest to this study was dihydroxyacetone (DHA). DHA is the precursor for the chemical methylglyoxal which is the main chemical responsible for the UMF activity in Mānuka honey, therefore research and development of predictive tools relating to DHA synthesis in *L. scoparium* is a critical factor in growing the Mānuka Honey industry in New Zealand. Leaf samples of seven cultivars of the species *L. scoparium* were collected in an attempt to correlate metabolic factors in the plant with DHA synthesis

in the nectar of *L. scoparium* flowers. The leaf material was analysed using Fourier Transform-Raman spectroscopy (FT-Raman). The DHA levels in nectar samples of the same cultivars were measured using standard high performance liquid chromatography methods. This study showed that the application of multivariate analysis of FT-Raman spectra from leaf material is a useful tool to predict DHA potential in *L. scoparium*

7.3 Introduction

Leptospermum scoparium (Mānuka) is a commercially significant plant in the New Zealand economy, it is the only representative of the genus *Leptospermum* in New Zealand and grows in a varied range of climates and soil conditions [1]. Mānuka has several current commercial uses. The Mānuka plant has potent antibacterial and antifungal properties due in part to high flavonoid and phenolic concentrations in the vegetative parts of the plant [2, 3, 4, 5], the wood when smoked imparts flavour to fish and meats and is very popular in New Zealand for smoking trout. *L. scoparium* leaves are also rich in essential oils, however the oil composition varies significantly across various species and cultivars within the *Leptospermum* populations in New Zealand [4, 5, 6]. The composition of these oils are mainly sesquiterpenes, monoterpenes and triketones [7], these compounds have application in the herbal healthcare industry due to their antimicrobial activity [1, 5, 8, 9].

L. scoparium is mainly cultivated in New Zealand for the honey and pharmaceutical industries although it has been cultivated as an ornamental plant since the early 1800's due to its range of white to pink and red flowers either single or double which flower in abundance, it is a popular ornamental plant and has also been investigated for

potential as a cut flower for the florist industry. Deliberate hybrid breeding in the 20th century has resulted in a large range of cultivars in use in the ornamental plant industry [10]. Recent interest in the unique antimicrobial properties of honey produced from *L. scoparium* plantations has resulted in research into the source and improvement of this potential in the *L. scoparium* population in New Zealand.

Leptospermum scoparium exhibits significant phenotypic plasticity and genotypic variation in wild populations in New Zealand. The capability to rapidly assess plants for essential oil composition and nectar composition is seen as beneficial to the economic development of this plant in New Zealand, of particular interest to this study is the potential for assessing plants for high antimicrobial activity of use in breeding plants that would benefit the commercial honey sector [10]. This study has assessed the potential of FT-Raman spectroscopy as a rapid non-invasive technique to fingerprint a range of species of cultivars of *L. scoparium* using leaf material and assess these cultivars for their potential for the unique Mānuka factor (UMF). The chemical component of interest to this study is the carbohydrate dihydroxyacetone (DHA) which is found in the nectar of *L. scoparium*. DHA is the precursor chemical which is metabolised into the chemical methylglyoxal (MGO) in honey which is the compound responsible for the UMF activity in Mānuka honey [11, 12].

There is a large amount of literature available covering research into spectroscopic techniques used to analyse various essential oil compounds and phenolics from various plant species, in addition many plant secondary metabolites have been analysed using this technique [13, 14, 15, 16, 17]. Spectroscopic techniques have been shown to discriminate different essential oil profiles allowing plant chemotypes to be distinguished and in most cases identification of individual oil components [15].

Investigated in this study was the ability of FT-Raman spectroscopy and multivariate data analysis to analyse the secondary metabolites in the leaves of *Leptospermum scoparium*, with the possibility of establishing a chemotype profile of each cultivar to be established from the leaf material sampled. FT-Raman spectroscopy techniques allow spectra to be obtained that identify some characteristic bands of individual chemical components from biological samples. These bands provide information about the chemical composition, including both primary and secondary metabolites of the investigated samples [18]. Each family, genus and species of plant produces a characteristic mix of these secondary metabolites and they have been used as taxonomic markers for classifying [16]. Multivariate data analysis is a well-known successful technique of use in differentiating samples containing many variables such as spectral data along with distinguishing correlating factors within complex data [19, 15, 20].

Leptospermum scoparium is known to have many bioactive components (e.g. secondary metabolites) of interest to science [7], and such compounds typically give complex Raman spectra. In this study, we applied FT-Raman spectroscopy in combination with multivariate statistical analysis to analyse plant material. We used this approach to provide insight into the chemical composition of various cultivars of *L. scoparium*.

7.4 Methods

7.4.1 Experimental Design

Seven different cultivars of *L. scoparium* (Figure 7.1) with varying levels of DHA concentrations in the nectar were investigated using FT-Raman spectroscopic

techniques, each cultivar set was grown from cuttings of a single parent plant and so were genetically identical. Plants were supplied by Comvita New Zealand and were grown in standard tree and shrub potting mix in 10 cm pots in a courtyard area at the Plant Growth Unit at Massey University; Palmerston North; New Zealand. DHA levels in the nectar of the cultivars were determined using high performance liquid chromatography (HPLC). (Note only 5 out of the 7 cultivars flowered during this experiment.) Leaf samples were collected from 10 replicate plants of each cultivar. The seven cultivars were labelled Y, MG, B, BS, O, P and LG. For FT-Raman spectroscopy, three analytical repeats of each replicate sample were analysed. Only fully developed healthy leaves were harvested. Leaves were collected from the mid-section of each plant between 10 cm of the tip of the plant and 10 cm above the soil level.

Figure 7.1. *L. scoparium* genotypes used in the experiments. Genotypes are identified by unique codes.

For FT-Raman analysis, leaf samples were ground with a pestle and mortar in liquid nitrogen, ground leaf samples were packed into divots and were analysed by Raman

spectroscopy on a Bruker Equinox 55 interferometer equipped with the Bruker FRA 106/S FT-Raman accessory and (Ettlingen, Germany) and a D418 T liquid-nitrogen-cooled germanium detector was used, controlled by the Bruker OPUS v6.0 software and with a Nd:YAG 1064 nm excitation source. A 1 mm diameter laser spot size and 150 mW power setting were used, 120 scans per spectrum with a spectral resolution of 4 cm^{-1} in the wave-number range 3500–0 cm^{-1} were produced. Samples were taken in triplicate. Acidified Methanol extracts of the leaves were prepared by weighing out 0.2 g of the ground leaf material. This powdered material was then ground with 3 mls of acidified methanol (70% absolute Methanol: 29% H_2O:1% Conc. HCL). The ground liquid and particulate material was then transferred to centrifuge tubes, tubes were centrifuged at 6000 rpm for 15 minutes, the supernatant was then removed, transferred to a clean eppendorf tube and stored at -20 ^{o}C. Aliquots of the acidified methanol extracts were pipetted into divots and analysed with FT-Raman using the same settings as the ground leaf material. All spectra were processed and analysed using The Unscrambler® X software (CAMO, NORWAY), Unscrambler is a multivariate analytical software package specifically designed to analyse spectral data.

For HPLC analysis of the DHA levels in the nectar samples, nectar was collected from 30 flowers from each of the 10 plants of each cultivar and was pooled. Nectar was collected at randomly selected times between 10 am and 2 pm on any day from flowers at development stage III-IV [21]. The sugar concentration in the nectar was measured using a hand-held refractometer (Bellingham and Stanley; UK) and was expressed as BRIX, the DHA level is represented as the amount per milligram of sucrose normalised to 80° BRIX. Normalisation to 80 °BRIX was used as this is the method of use in the commercial honey industry in New Zealand (pers. comm. J. Stephens, Comvita New Zealand, Te Puke, New Zealand). DHA levels in nectar samples

were measured using a commercial HPLC method at R.J. Hills Laboratories (Hamilton, New Zealand).

The DHA values obtained from the nectar of the cultivars using LCMS are shown in Figure 7.2. Statistical analyses of the comparison of means of each cultivar applying a Tukey pair-wise comparison of the genotypes show that there are three significantly different groups in regard to DHA concentration in the five cultivars analysed (Figure 7.2). There is no significant difference between cultivars Y and P and no significant difference between cultivars MG and B see Figure 7.2.

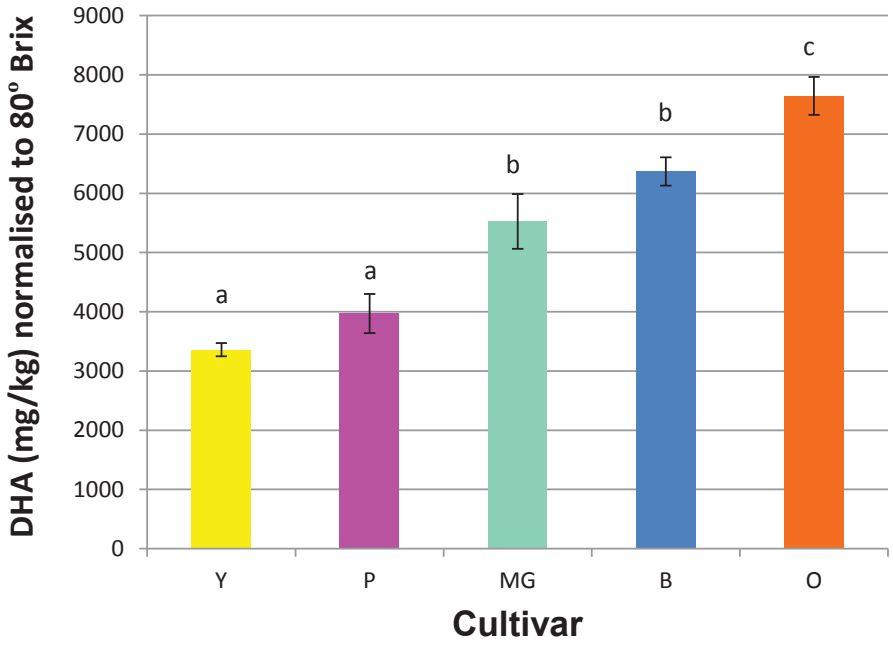

Figure 7.2. LCMS data showing DHA values for the cultivars normalised to 80 °BRIX Error bars show standard error of the means. Letters above the bars indicate the results from the analysis of the comparison of means of each cultivar applying a Tukey pair-wise comparison of the cultivar lines using Minitab statistical software. Same letter indicates no significant difference between the cultivars at the 95% confidence level and P-values ≤ 0.05. Y=Yellow; P=Pink; O=Orange; MG=Mint Green; B=Blue.

7.4.2 Multivariate Analysis of the Spectral Data

The assumption underlying the use of multivariate data analysis is that the measured data collected carries information about the system under investigation. For example, to predict the concentration of a particular chemical substance, in this case the levels of dihydroxyacetone (DHA) in the nectar of *Leptospermum scoparium*, the measurements performed on the biological system involved must in some way reflect the quantification of that substance. The variables measured must have a quantitative relationship to the property of interest.

Mathematically, the property of interest is termed “Y” and the variable is termed “X”. In most systems under investigation, especially biological systems, many X variables are usually investigated and invariably more than one X variable influences any one Y variable, i.e. the data are covariant. In spectroscopic analysis, we are typically dealing with spectral data collected across a range of wavelengths that is dependent on the analytical system used. Thus, the X variables represent wavelengths and the X data in this case are from Fourier transform (FT)-Raman reflectance spectra. The Y variables are the various components being analysed and compared, in this case DHA [22, 23].

7.5 Results and discussion

7.5.1 Multivariate Analysis of Spectral Data from *L. scoparium* Genotypes

An initial principal component analysis (PCA) of the leaf spectral data was performed to test the ability of multivariate methods to differentiate the cultivars. A linear baseline was applied to all spectra to remove the vertical shift in the spectra that is caused by thermal effects from the instrumentation, after the baseline correction, standard normal variate (SNV) processing of the spectra was applied to all spectra to

remove scatter in the spectral data unconnected to the analytes of interest [24]. Spectra from the replicate samples of each cultivar were colour grouped and obvious outliers were removed from the spectral database, less obvious outliers were identified using hotelling statistical methods and those outside the 95% confidence level were removed. The total spectral range was analysed including various specific spectral regions of interest (Figure 7.3).

Using principal components analysis (PCA), adequately separated groups were achieved using the 1650−1150 cm^{-1} region of the Raman spectrum and plots of principal components 2 vs 3 and 3 vs 4, (see Figures 7.4 and 7.5). This region contains information from flavonoids, carotenoids and essential oils, compounds common to many plants [13]. Clearly, even though the plants analysed in this experiment are sub-cultivars of the same species there are still sufficient differences in chemical composition to observe spectral differences. Observation of various PC score plots from the PCA plots illustrated that PC1 did not differentiate the cultivars, PC1 contained most of the variability but did not differentiate the cultivars and so was removed from consideration. PC1 did contain similar compounds i.e. similar peaks to the further PC's, this is explained by the fact that many secondary metabolites are common to all plants, the results indicate that it is a subset of a group of compounds that is responsible for differentiating the cultivars.

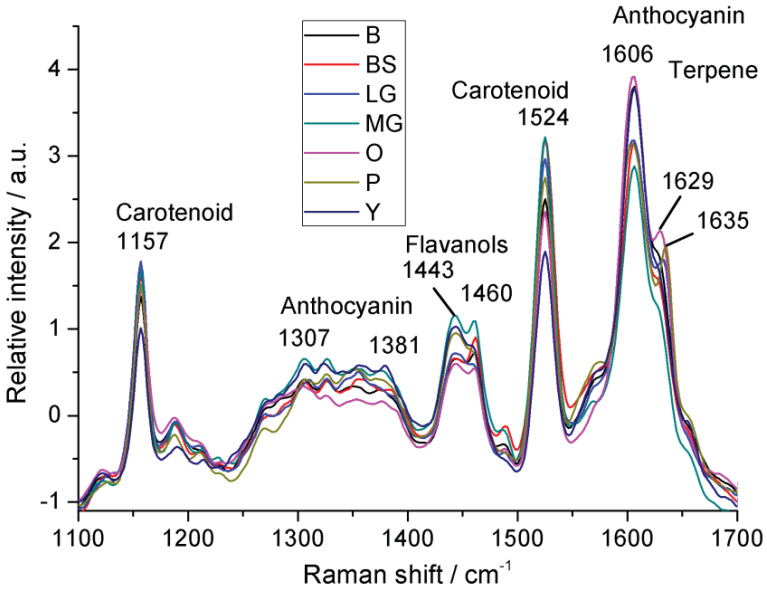

Figure 7.3. The sub-region 1650−1150 cm^{-1} of Raman spectra of leaf samples of *L. scoparium* after linear baseline correction. Y=Yellow; LG=Lime Green; P=Pink; O=Orange; BS=Blue Stripe; MG=Mint Green; B=Blue.

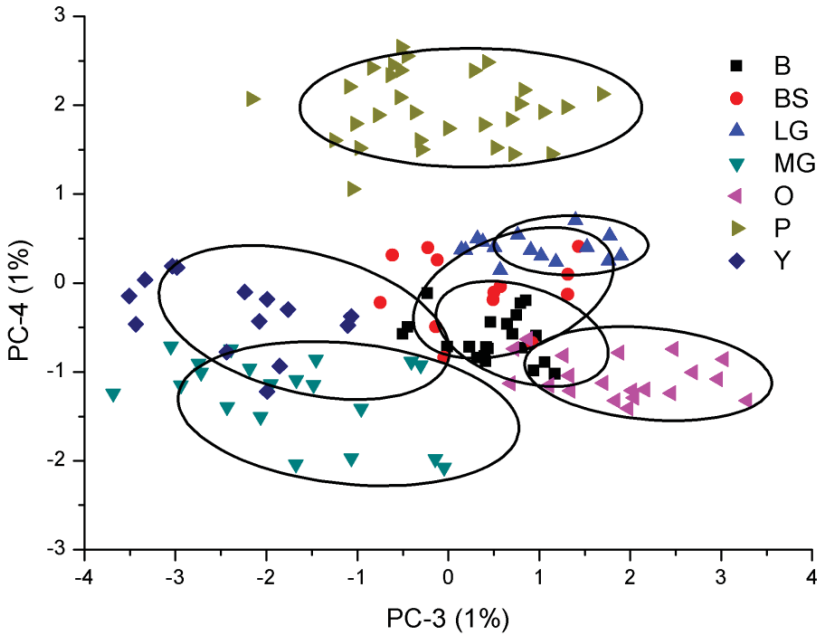

Figure 7.4. Scores plot (PC3 and PC4) obtained using the sub-region spectrum (1650−1150 cm^{-1}), plots shows the initial separation between leaf spectra of the seven cultivars.

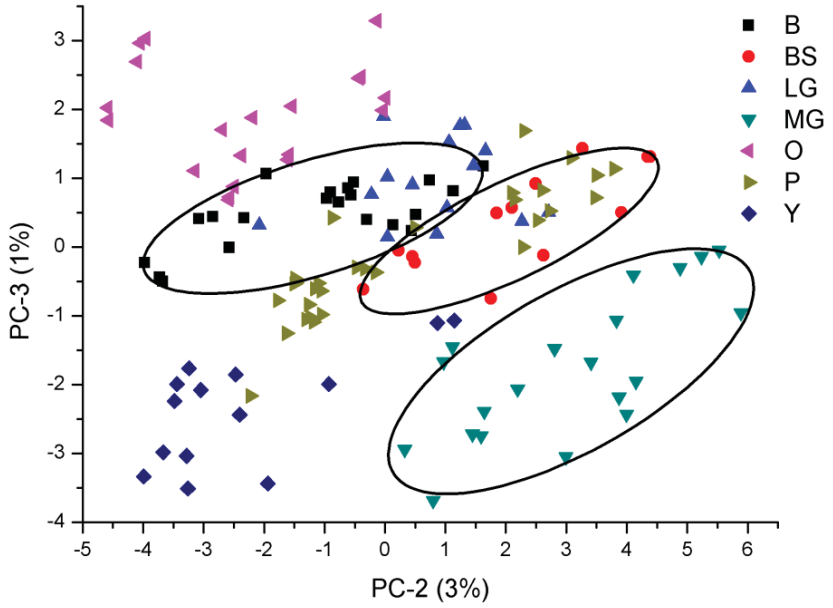

Figure 7.5. Scores plot (PC2 and PC3) obtained using the same spectral region. Note the further separation of B from BS and MG from Y. P appears to have 2 groupings.

The various principal components (PC's) can be plotted against each other in different combinations to further separate groupings that overlap. In Figure 7.5 plotting PC2 against PC3 scores show that this combination exhibits good differentiation between the B group and the BS group. Also the MG group is distinct from the Y group, whereas there is distinct overlap in the PC3 versus PC4 scores plot. Other groups are well differentiated in the PC3 versus PC4 plot see Figure 7.4. As these plots attempt to depict the position of the principal components in three-dimensional space, various combinations are necessary to get a visual impression of how the cultivars are separated. Chemical information on possible compounds that may be responsible for the differences in the PCs and how the cultivars separate in PC space can be determined from analysing the main peaks in relevant loading plots, see Figure 7.6.

Analysis of the plot of PC3 against PC4, shows that the cultivars are possibly splitting according to terpene and flavonoid composition, whereas better differentiation from

PC2 and PC3 for cultivars MG vs Y and B vs BS could possibly be due to carotenoid composition, as PC2 has the largest influence from features at 1525 cm^{-1} and 1156 cm^{-1} as noted in the PC loadings plots (Figure 7.6). Carotenoids have been shown to have a characteristic conjugated C=C bond which has a distinct peak between 1509 cm^{-1} and 1530 cm^{-1} due to the stretching modes of the conjugated C=C bond, they also have a second distinct peak at 1156 cm^{-1} associated with the C-C bonds in the central chain. Here we observe this peak to be present in all PC loadings considered except PC4, showing that the increased separation in PC4 is not influenced by the carotenoid composition [13, 25]. PC4 also contains distinct peaks in the wavenumber region 1600-1650 cm^{-1} attributed to flavonoids [26] and terpenes [27, 28, 29, 30]. *Leptospermum scoparium* contains essential oils that are rich in monoterpenes and sesquiterpenes in particular α-pinene, myrcene, trans-methyl cinnamate [6], and many of the peaks noted could be assigned to these compounds. In addition other peaks noted in the PC loadings are consistent with flavonoid peaks particularly anthocyanins. However the pigment composition of *L. scoparium* has not been studied and therefore it is not possible to fully assign all the relevant peaks observed in the PC loadings. Further work elucidating the flavonoid composition of *L. scoparium* is required. Analysis of pigment extracts from the leaves of two cultivars of *L. scoparium* using acidified methanol extraction, and analysis with FT-Raman demonstrated strong peaks at 1606 cm^{-1}, 1609 cm^{-1} and 1632 cm^{-1} with a medium peak at 1625 cm^{-1} and a shoulder at 1635 cm^{-1}. These peaks are consistent with peaks exhibited in the various loadings plots from the PC's that are differentiating the various cultivars under study in this investigation. In addition there are many smaller peaks in the 1200-1400 cm^{-1} region as noted by Gamsjaeger et al. 2011, these minor peaks are attributed to inter-ring stretching

vibrations of anthocyanin molecules [13]. *L. scoparium* leaf material also contains elevated levels of the unique triketone Leptospermone. While the FT-Raman spectra of triketones are not available in the literature, ketones in general exhibit strong peaks in the wavenumber region 1650-1800 cm^{-1}, this region did not add to the differentiation when included in previous principal component analysis, it is therefore inferred that triketone composition is not a major influence in differentiating the cultivars in this experiment.

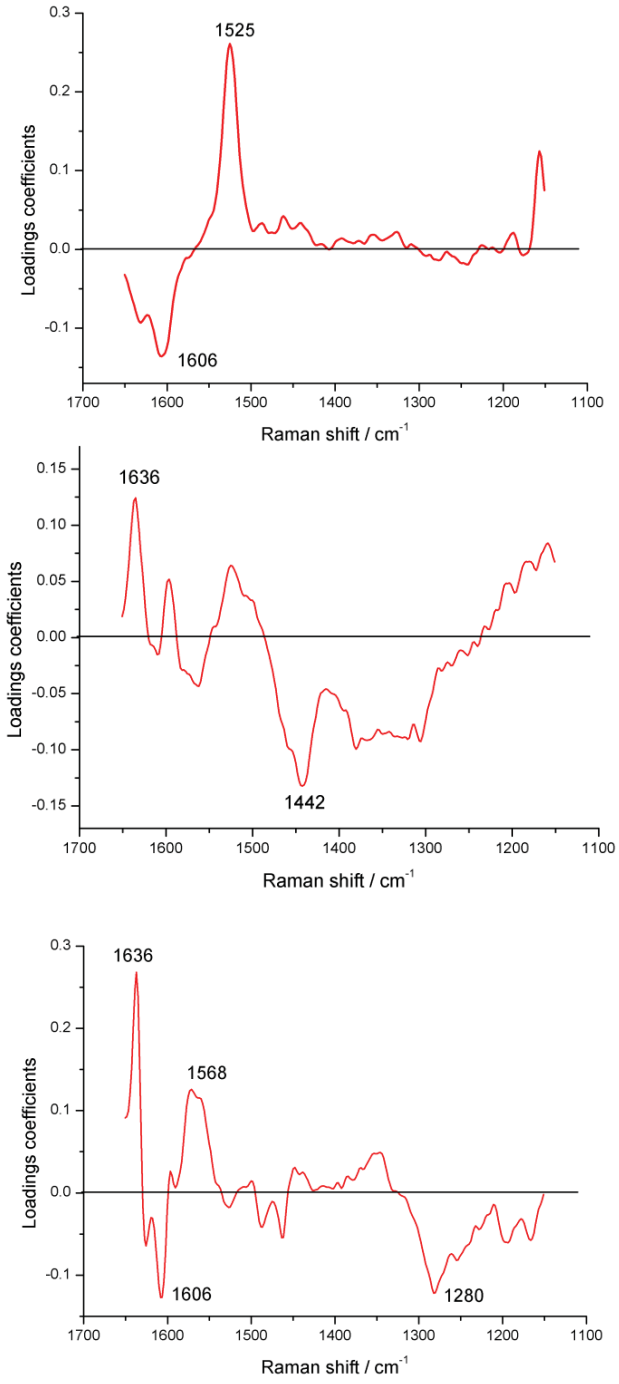

Figure 7.6. Loadings plots for PCs 2, 3 and 4. The main peak wavenumbers from these plots are used to indicate what compounds, based on their molecular structure or reference spectra, could be responsible for differentiating the cultivars.

7.5.2 Multivariate Partial Least Squares (PLS) Regression

In this study, FT-Raman spectroscopy and multivariate analysis were used to differentiate cultivars and to identify those groups of compounds which may underpin this differentiation. Further work aimed at correlating the chemical composition of the leaves against the various DHA levels in the nectar of the same cultivars was performed. Multivariate Partial Least Squares (PLS) regression techniques were applied to the leaf and DHA data using the data from the spectroscopic analysis of the leaves and DHA results collected from HPLC analysis.

The PLS results (Figures 7.7 and 7.8) indicate that the chemical composition of the leaves provide some correlation to the DHA levels in the nectar of the flowers of each cultivar and suggest that FT-Raman along with multivariate statistics could be used to develop a predictive model to assess DHA potential in *Leptospermum scoparium* from leaf material. Partial least squares regression results show that R^2 values are high at 0.75 (with an RMSE of 836.45 mg/kg DHA 80 ^{o}BRIX) at the 95% confidence level and a P value ≤ 0.005. Groupings in the model Figure 7.7 show the range of DHA levels for each cultivar group and show that this model could be used to predict the DHA levels within a range in the cultivars of *Leptospermum scoparium* and matches well with the statistics for the DHA analysis shown in Figure 7.2. The score plot in Figure 7.8 indicates that factor 1 is the main influence in differentiating the cultivars, analysis of the loadings from this factor gives similar information as the principal components analysis and indicates that carotenoid, flavonoid and terpene structures are influencing the correlation between leaf composition and nectar composition in *L. scoparium* [6, 28, 29, 30, 31].

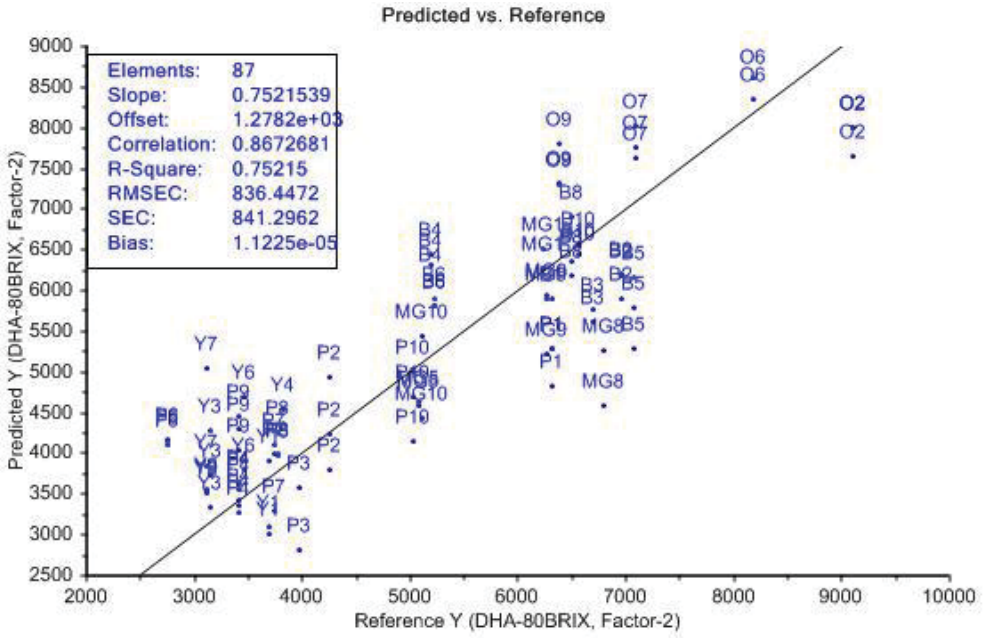

Figure 7.7. Regression graph of the PLS model, note the R^2 value of 0.75 and the inclusion of the first two factors for this model.

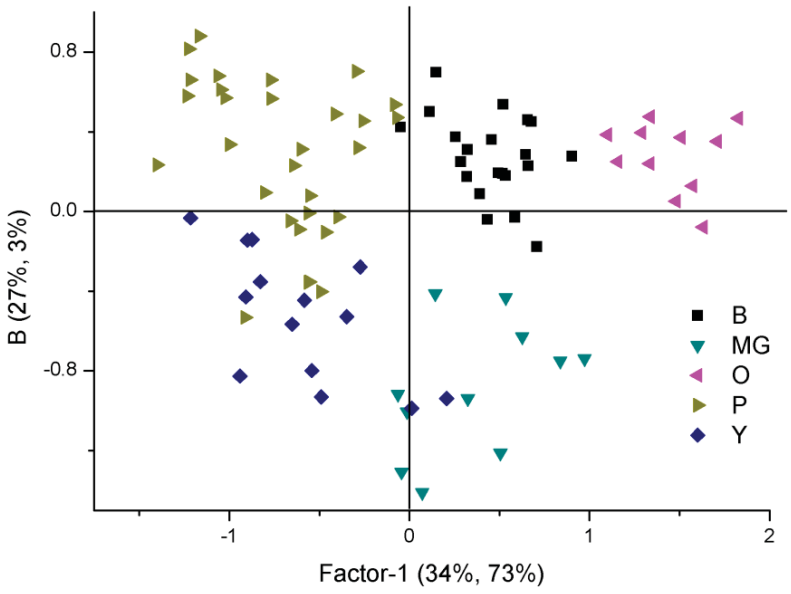

Figure 7.8. Graph of score plot from the PLS regression of FT-RAMAN spectra of leaves against DHA levels in the nectar of the flowers from these five cultivars. Score plots show groupings if they are present, which indicate that a correlation is possible using further PLS analytical methods. This figure is a plot of factor 1 against factor 2 and shows well separated groupings for 4 out of 5 of the cultivars.

7.6 Conclusion

These experiments show that FT-Raman spectroscopy is a useful tool to chemotype cultivars within the species *Leptospermum scoparium*, using FT-Raman spectra from the leaf grinds, in depth analysis of the loadings that are influencing the principal component analysis show that the main chemical components influencing the separation of the cultivars are carotenoids, flavonoids, and terpenes. PLS regression techniques demonstrated that FT-Raman analysis of leaf material could provide a rapid screening method to assess DHA potential of *L. scoparium* plants. Future work could include a full analysis and isolation of leaf pigment composition and FT-Raman analysis on these pigments. This coupled with isolation of individual terpenes and the triketone Leptospermone with associated FT-Raman analysis would elucidate exactly which chemicals are correlating with the nectar composition in *Leptospermum scoparium*.

7.7 References

[1] Haberlein H., Tschiersch K.P. On the occurrence of methylated and methoxylated flavonoids in *Leptospermum scoparium*. Biochemical Systematics and Ecology **(1998),** 26(1): p. 97-103.

[2] Lis-Balchin M., Hart S.L. An investigation of the actions of the essential oils of Manuka (*Leptospermum scoparium*) and Kanuka (*Kunzea ericoides*), Myrtaceae on guinea-pig smooth muscle. Journal of Pharmacy and Pharmacology **(1998),** 50(7): p. 809-811.

[3] Lis-Balchin M., Hart S.L., Deans S.G. Pharmacological and antimicrobial studies on different tea-tree oils (*Melaleuca alternifolia, Leptospermum scoparium* or Manuka and *Kunzea ericoides* or Kanuka), originating in Australia and New Zealand. Phytotherapy Research **(2000),** 14(7): p. 623-629.

[4] Porter N.G., Smale P.E., Nelson M.A., Hay A.J., Van Klink J.W., Dean C.M. Variability in essential oil chemistry and plant morphology within a Leptospermum scoparium population. New Zealand Journal of Botany **(1998),** 36(1): p. 125-133.

[5] Porter N.G., Wilkins A.L. Chemical, physical and antimicrobial properties of essential oils of *Leptospermum scoparium* and *Kunzea* ericoides. Phytochemistry **(1999),** 50(3): p. 407-415.

[6] Perry N.B., VanKlink J.W., Brennan N.J., Harris W., Anderson R.E., Douglas M.H., Smallfield B.M. Essential oils from New Zealand manuka and kanuka: Chemotaxonomy of Kunzea. Phytochemistry **(1997),** 45(8): p. 1605-1612.

[7] Douglas M.H., van Klink J.W., Smallfield B.M., Perry N.B., Anderson R.E., Johnstone P. Weavers R.T. Essential oils from New Zealand manuka: triketone and other chemotypes of *Leptospermum scoparium*. Phytochemistry **(2004),** 65(9): p. 1255-1264.

[8] Jeong E.Y., Kim M.G. & Lee H.S. Acaricidal activity of triketone analogues derived from *Leptospermum scoparium* oil against house-dust and stored-food mites. Pest Management Science **(2009),** 65(3): p. 327-331.

[9] van Klink J.W., Brophy J.J., Perry N.B., Weavers R.T. beta-triketones from Myrtaceae: Isoleptospermone from *Leptospermum scoparium* and papuanone from *Corymbia dallachiana*. Journal of Natural Products **(1999),** 62(3): p. 487-489.

[10] Stephens J.M.C., Molan P.C., Clarkson B.D. A review of *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany **(2005),** 43(2): p. 431-449.

[11] Adams C.J., Manley-Harris M., Molan P.C. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2009),** 344(18): p. 1050-1053.

[12] Mavric E., Wittmann S., Barth G., Henle T. Identification and quantification of methylglyoxal as the dominant antibacterial constituent of Manuka (*Leptospermum scoparium*) honeys from New Zealand. Molecular Nutrition & Food Research **(2008),** 52(4): p. 483-489.

[13] Gamsjaeger S., Baranska M., Schulz H., Heiselmayer P., Musso M. Discrimination of carotenoid and flavonoid content in petals of pansy cultivars

(Viola x *wittrockiana*) by FT-Raman spectroscopy. Journal of Raman Spectroscopy **(2011),** 42(6): p. 1240-1247.

[14] Baranska M. & Proniewicz L.M. Raman mapping of caffeine alkaloid. Vibrational Spectroscopy **(2008),** 48(1): p. 153-157.

[15] Baranska M., Schulz H., Kruger H. & Quilitzsch R. Chemotaxonomy of aromatic plants of the genus Origanum via vibrational spectroscopy. Analytical and Bioanalytical Chemistry **(2005),** 381(6): p. 1241-1247.

[16] Kruger H. & Schulz H. Analytical techniques for medicinal and aromatic plants. Stewart Postharvest Review **(2007),** 3(4): p. 1241-1247

[17] Rozo J.I.J., Zarow A., Zhou B., Pinal R., Iqbal Z., Romanach R.J. Complementary Near-Infrared and Raman Chemical Imaging of Pharmaceutical Thin Films. Journal of Pharmaceutical Sciences **(2011),** 100(11): p. 4888-4895.

[18] Schulz H., Baranska M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vibrational Spectroscopy **(2007),** 43(1): p. 13-25.

[19] Balss K.M., Long F.H., Veselov V., Orana A., Akerman-Revis E., Papandreou G., Maryanoff C.A. Multivariate analysis applied to the study of spatial distributions found in drug-eluting stent, coatings by confocal Raman microscopy. Analytical Chemistry **(2008),** 80(13): p. 4853-4859.

[20] Killeen D.P., Sansom C.E., Lill R.E., Eason J.R., Gordon K.C., Perry N.B. Quantitative Raman Spectroscopy for the Analysis of Carrot Bioactives. Journal of Agricultural and Food Chemistry **(2013),** 61(11): p. 2701-2708.

[21] Davis A. R. Influence of floral visitation on nectar-sugar composition and nectary surface changes in Eucalyptus. Apidology **(1997),** 28: p. 27-42.

[22] Wolthuis R., Tjiang G.C.H., Puppels G.J., Schut T.C.B. Estimating the influence of experimental parameters on the prediction error of PLS calibration models based on Raman spectra. Journal of Raman Spectroscopy **(2006),** 37(1-3): p. 447-466.

[23] Pierna J.A.F., Abbas O., Dardenne P., Baeten V. Discrimination of Corsican honey by FT-Raman spectroscopy and chemometrics. Biotechnologie Agronomie Societe Et Environnement **(2011),** 15(1): p. 75-84.

[24] Esbensen K.H., Guyot D., Westad F., Houmoller L.P. Multivariate Data Analysis in Practise 5th Edition **(2002).**

[25] Collins A.M., Jones H.D.T., Han D., Hu Q., Beechem T.E., Timlin J.A. Carotenoid Distribution in Living Cells of Haematococcus pluvialis (Chlorophyceae). Plos One **(2011),** 6(9): p. e24302

[26] Butchweitz M., Gudi G., Carle R., Kemmerere D.R. and Schulz H. Systematic investigations of anthocyanin-metal interactions by Raman sepctroscopy. Journal of Raman Spectroscopy **(2012),** 43(12): p. 2001-2007.

[27] Kalinowska M., Swislocka R., Lewandowski W. Zn(II), Cd(II) and Hg(I) complexes of cinnamic acid: FT-IR, FT-Raman, H-1 and C-13 NMR studies. Journal of Molecular Structure **(2011),** 993(1-3): p. 404-409.

[28] Kalinowska M., Lewandowski W., Swislocka R., Regulska E. The FT-IR, FT-Raman, H-1 and C-13 NMR study on molecular structure of sodium(I),

calcium(II), lanthanum(III) and thorium(IV) cinnamates. Spectroscopy-an International Journal **(2010),** 24(3-4): p. 277-281.

[29] Merlin J.C., Statoua A., Cornard J.P., Saidiidrissi M., Brouillard R. Resonance raman spectroscopic studies of anthocyanins and anthocyanidins in aqueous solutions. Phytochemistry **(1994)** 35(1): p. 227-232.

[30] Merlin J.C., Statoua A., Brouillard R. Investigation of the invivo organistaion of anthocyanins using resonance raman microspectrometry. Phytochemistry **(1985)** 24(7): p. 1575-1581.

[31] Buchweitz, M., Gudi, G., Carle, R., Kammerer, D.R., Schulz, H. Systematic investigations of anthocyanin-metal interactions by Raman spectroscopy. Journal of Raman Spectroscopy **(2012)** 43(12): p. 2001-2007.

Chapter 8: Analytical method development using FTIR-ATR and FT-Raman spectroscopy to assay the carbohydrates; fructose, sucrose, glucose and dihydroxyacetone, in *Leptospermum scoparium* nectar

This chapter has been published with the "Journal of Vibrational Spectroscopy" as: Analytical method development using FTIR-ATR and FT-Raman spectroscopy to assay the carbohydrates; fructose, sucrose, glucose and dihydroxyacetone, in *Leptospermum scoparium* nectar.

Elizabeth M. Nickless[1], Stephen E. Holroyd[2], Georgie Hamilton[1] Keith C. Gordon[3]; Jason J. Wargent[1] *

[1] Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand

[2] Fonterra Co-operative Group, Fonterra Research & Development Centre, Palmerston North, New Zealand

[3] Department of Chemistry and Dodd-Walls Centre, University of Otago, Dunedin, New Zealand

*Corresponding Author: J. J. Wargent

Citation: E.M. Nickless et al. "Analytical method development using FTIR-ATR and FT-Raman spectroscopy to assay the carbohydrates; fructose, sucrose, glucose and dihydroxyacetone, in *Leptospermum scoparium* nectar." Vibrational Spectroscopy 84 **(2016):** p. 38–43.

E.M.N. designed and carried out the FTIR and Raman spectroscopy and HPLC experiments, analysed the data and authored the paper. G.H. assisted with HPLC experimental development and assisted with sample prep for HPLC sample analysis. S.E.H. and K.C.G. provided technical guidance and expertise in FT-Raman and FTIR experiments. S.E.H and J.J.W. were project supervisors to E.M.N. All authors approved the final manuscript.

8.1 Abstract

The carbohydrate dihydroxyacteone (DHA) occurs in significant levels in *Leptospermum scoparium* (mānuka) nectar and is the precursor of methylglyoxal the unique non-peroxide antibacterial activity (NPA) component in mānuka honey. The nectar of ten different cultivars of *Leptospermum scoparium* was assayed quantitatively for fructose, glucose, sucrose and DHA with high pressure liquid chromatography (HPLC) for comparison with FT-Raman and FTIR-ATR spectroscopic methods. FT-Raman spectroscopy and ATR-FTIR spectroscopy alongside chemometric methods, principal component analysis (PCA) and partial least squares (PLS) prediction were shown to be useful techniques to quantify and compare the nectar composition in a range of cultivars of *L. scoparium*.

Keywords: Dihydroxyacetone, *Leptospermum scoparium*, FT-Raman, FTIR-ATR, nectar composition.

8.2 Introduction

The nectar composition of *Leptosperum scoparium* (mānuka), as the plant nectar source for mānuka honey, is of significant interest to the honey industry in New Zealand. The nectar of *L. scoparium* contains the carbohydrate dihydroxyacetone (DHA) which is the precursor chemical for methylglyoxal (MGO), the unique non-peroxide antibacterial activity (NPA) component of mānuka honey [1–3]. *Leptospermum* honeys are valued for their therapeutic application in wound healing of skin infections [4]. The traditional method of applying the anthrone assay and colorimetric analysis [5] to measure total sugars in *L. scoparium* nectar is a time consuming and laboratory-intensive method. This study presents the potential of attenuated total reflectance (ATR) Fourier transform Infrared (FTIR) and Fourier transform (FT)- Raman spectroscopy combined with the chemometric tool Partial Least Squares Regression analysis (PLSR) to develop a method suitable for the rapid non-destructive discrimination of nectar composition, in particular DHA and saccharide sugars in *L. scoparium*. Most chromatographic techniques are based on solvent extraction followed by high performance liquid chromatography (HPLC) separation and/or colorimetric and enzymatic analysis [5, 6]. Traditional analytical HPLC methods for quantification of analytes of interest require expensive and extensive sample preparation, with additional derivatisation techniques for small molecules such as DHA [7]. FTIR techniques have been shown to be useful in analysing and quantifying sugars in a range of samples such as cereal, baby foods, honey and fruit juices during product processing [6, 8–11]. Sultanbawa et al. recently demonstrated that mid infrared techniques provided a good model for predicting methylglyoxal levels in

Leptospermum polygalifolium honeys [4]. Infrared spectroscopy is useful in identifying compounds based on the vibration frequencies of their molecular structure and is a sensitive technique for analysing the chemical composition of various sample types [6]. The application of spectroscopic techniques for the quantification of carbohydrates in-situ in plants has received some previous attention from other authors. Past work includes analysis of monosaccharides and polysaccharides applying either FTIR or Raman, and various chemometric analysis techniques to differentiate individual carbohydrates or fingerprint sample types [9]. Raman spectroscopy and FTIR spectroscopy have been used to quantify fructose, glucose, sucrose and saccharose in fruit juices and fermentation stages in vinegar and pineapple juice along with quantifying various sugars and acids for grading fruit quality [6, 9–11]. Other findings have been reported using FT-Raman to quantify sugars in carrot tissue along with carotenoids and polyacetylene [12–14]. FT-Raman has also been used to analyse the chemical composition of floral honey [15]. Both FTIR and FT-Raman analytical spectroscopy methods are widely applied to identify substances from characteristic spectral patterns ("fingerprinting") and to ascertain quantitatively or qualitatively the amount of a substance in a sample [16, 17]. Infrared spectroscopy has been used successfully in a preliminary study on nectar using principal component analysis to distinguish various different plant species nectars [18]. To date there has been no research published using these methods to quantify individual plant nectar components.

8.3 Experimental Methods

The floral nectars of ten different proprietary cultivars of *L. scoparium* were investigated using both attenuated total reflectance (ATR), Fourier transform Infrared (FTIR), and Fourier transform (FT)-Raman spectroscopic techniques, along with high pressure liquid chromatography (HPLC) as a reference analysis. Each cultivar set contained ten replicate plants grown from cuttings of a single parent plant and so were genetically identical i.e. clones of the parent material. Plants were supplied by Comvita New Zealand, and were grown in standard tree and shrub potting mix in 30 cm pots in a glasshouse at the Plant Growth Unit at Massey University, Palmerston North, New Zealand. Nectar was collected in a consistent fashion from 20 flowers from each of the 10 plants of each cultivar and was pooled. Nectar was collected at randomly selected times between 10am and 2pm on any day from flowers at development stage IV [19]. DHA levels in nectar samples were measured using aqueous extraction, derivatisation and analysis by HPLC adapting the method used by Windsor et al. to analyse DHA in honey samples [7]. In addition, DHA levels in nectar samples were validated using a commercial analytical service (R.J. Hills Laboratories Hamilton, New Zealand: http://www.hill-laboratories.com/page/pageid/2145845743/Honey_Testing). HPLC analysis of the sugars fructose, glucose, and sucrose were performed at Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand. Three replicates from each nectar sample were used for the development of the spectroscopic methods. ATR-FTIR was used to analyse the same components and quantify using Partial least Squares regression (PLSR), and results and error values

were compared to test the accuracy of using FTIR. FT-Raman spectra from a separate nectar set was collected for comparative DHA quantification capability.

8.3.1. HPLC conditions for DHA analysis

Analyses were performed on a PerkinElmer Series 200 Pump and Auto sampler with a Flexar photo diode array detector (λ = 263 nm). HPLC separations were performed on a Synergi Fusion column (75 x 4.6 mm, 4 mm particle size). The column was heated and held at 30 ^{o}C to maintain stable run conditions. Mobile phase A was water: ACN, 70/30, v/v and mobile phase B was 100% ACN. The following 23 min gradient elution was employed: A: B = 90:10 (isocratic 2.5 min), graded to 50:50 (8.0 min), graded to 0:100 (1.5 min), 0:100 (isocratic 7.0 min), graded to 90:10 (1.0 min), 90:10 (isocratic 3.5 min), detection at 263 nm.

8.3.1.1 Preparation of Reaction Solutions

Hydroxyacetone (HA) (3.01 mg/ml) formed the HA internal standard solution. The O-(2, 3, 4, 5, 6-pentafluorobenzyl) hydroxylamine (PFBHA) derivatising reagent was 19.8 mg/ml in citrate buffer (0.1 M) adjusted to pH 4 with sodium hydroxide (NaOH) (4 M). DHA (3.88 mg/ml) formed the DHA standard solution.

8.3.1.2 Sample Preparation

For the preparation of standards, DHA standard solution (100, 80, 60, 40, 20, 10 and 0 ml) was added to tubes 1–7 respectively, and made up to 100 ml with nanopure water. For sample analysis, 20 ml of nectar or standard was pipetted into a mix tube and 25 ml of the HA was added. Derivatisation steps were performed at 25 ^{o}C in a controlled temperature room. Each of the HPLC samples and standards was thoroughly mixed and placed in a rack on a rotating Table for 1 hour to allow complete dissolution. PFBHA derivatising solution (100 ml) was added to each test tube, which was mixed

and placed in a rack on a rotating Table for 1 hour to allow for complete derivatisation. Acetonitrile (ACN) (1.5 ml) was added to each test tube and mixed. Nanopure water (0.5 ml) was then added to each test tube and mixed. Samples were then syringe filtered with a 0.22 mm filter into HPLC vials. Vials were placed into the auto sampler and run overnight, and repeat analysis of standards were analysed through each run to check stability of the analysis. DHA calibration curves were generated from tubes 1–7 by linear regression using the HPLC peak area ratios of DHA: HA plotted against the mass of the DHA mass content of the nectar samples were determined against these calibration curves.

Table 8.1. HPLC data showing DHA and sugar component values for the cultivars normalised to 80°BRIX. Error statistics show standard error of the means. Letters adjacent to column values in superscript indicate the results from the analysis of the comparison of means of each cultivar applying ANOVA with a Tukey pair-wise comparison of the cultivar lines using Minitab statistical software, same letter indicates no significant difference between the cultivars at the 95% confidence level and P-values ≤ 0.05

Cultivar	Fructose % 80°BRIX	STD error	Glucose % 80°BRIX	STD error	Sucrose % 80°BRIX	STD error	DHA 80°BRIX mg/kg	STD error
BS	43.22 ^{f}	0.22	35.94 ^{a}	0.19	0.85 ^{b}	0.06	5177.02 ^{b,c,d,e}	677.10
G	44.90 ^{e,f}	0.73	33.24 ^{b}	0.59	1.87 ^{a}	0.20	4116.01 ^{a,b}	521.43
O	46.73 ^{d,e}	0.51	32.82 ^{b,c}	0.49	0.45 ^{c,d}	0.04	5981.54 ^{b,c}	442.12
PU	46.87 ^{c,d,e}	0.27	32.47 ^{b,c,d}	0.25	0.66 ^{b,c}	0.02	2714.35 ^{b,c,d}	380.44
B	46.97 ^{b,c,d,e}	0.45	32.68 ^{b,c,d}	0.41	0.35 ^{c,d}	0.07	6153.12 ^{e}	469.22
R	48.93 ^{a,b,c,d}	0.40	30.78 ^{c,d,e}	0.37	0.29 ^{d}	0.05	7459.07 ^{a,b}	373.56
P	49.25 ^{a,b,c}	0.46	30.55 ^{d,e}	0.40	0.20 ^{d}	0.07	3058.88 ^{c,d,e}	352.04
Y	49.41 ^{a,b}	0.73	29.90 ^{e}	0.73	0.69 ^{b,c}	0.02	4843.71 ^{d,e}	111.35
MG	49.70 ^{a}	0.86	29.79 ^{e}	0.84	0.51 ^{b,c,d}	0.02	3266.70 ^{a}	554.31
LG	50.25 ^{a}	0.20	29.51 ^{e}	0.20	0.24 ^{d}	0.01	4252.20 ^{b,c,d,e}	366.02

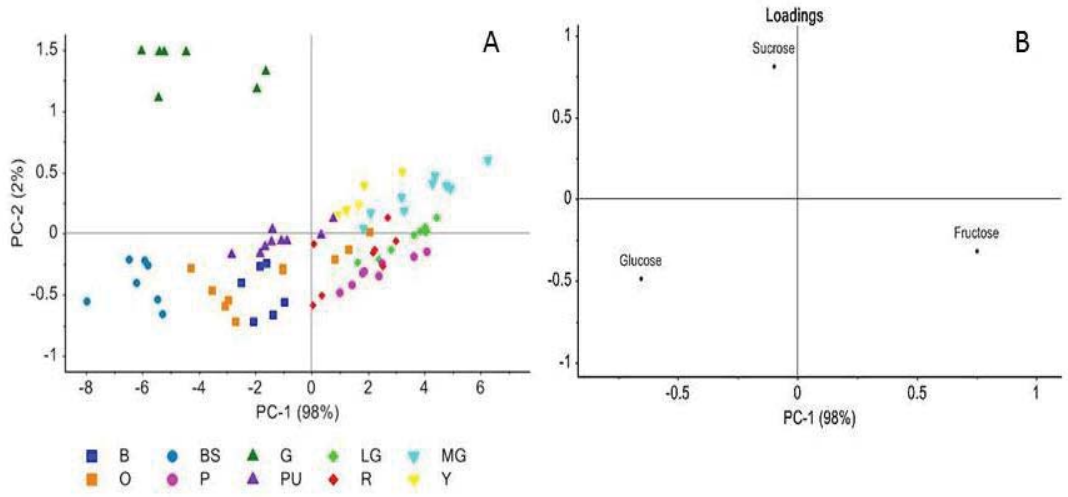

Figure 8.1. PCA of the main sugar composition (fructose, glucose and sucrose) of the nectar of the ten cultivars; B, BS, G, LG, O, P, PU, R and Y. Graph A= PCA score plot B = Loadings plot

8.3.1.3 HPLC conditions for Sucrose, Glucose and Fructose Analysis

Individual sugar analysis was performed on an Alliance Waters 2690 Separations Module HPLC and auto sampler using a Waters 2410 Refractive Index (RI) detector. Sucrose, glucose and fructose were separated using an Aminex HPX-87C Carbohydrate Column (300 x 7.0 mm). The column was heated and held at 65 ^{o}C to maintain stable conditions. A standard run method was used with one phase solution of milliQ water at a flow rate of 0.6 ml/minute. The sample injection size was 20 ul and each sample run was 18 min.

8.3.1.4 Preparation of Standards

1% stock solutions of each of the sugars sucrose, glucose, and fructose were made in nano-pure water. Five standards were used to create a calibration curve for each sugar; the standards contained equal proportions of sucrose, glucose and fructose, and had concentrations of 0%, 0.01%, 0.05%, 0.17%, and 0.33% sugar.

8.3.1.5 Sample Preparation

30 ml of each nectar sample was added to 1470 ml of nano-pure H_2O and mixed. The diluted nectar samples as well as the standards were filtered using 0.22 um syringe filters into HPLC vials. Vials were loaded into the auto-sampler along with the standards and analysed overnight. Standards were repeat assayed throughout the run to reference for any variability during the analysis. HPLC analysis of the sugars fructose, glucose and sucrose was performed at Institute of Agriculture & Environment, Massey University; Palmerston North, New Zealand.

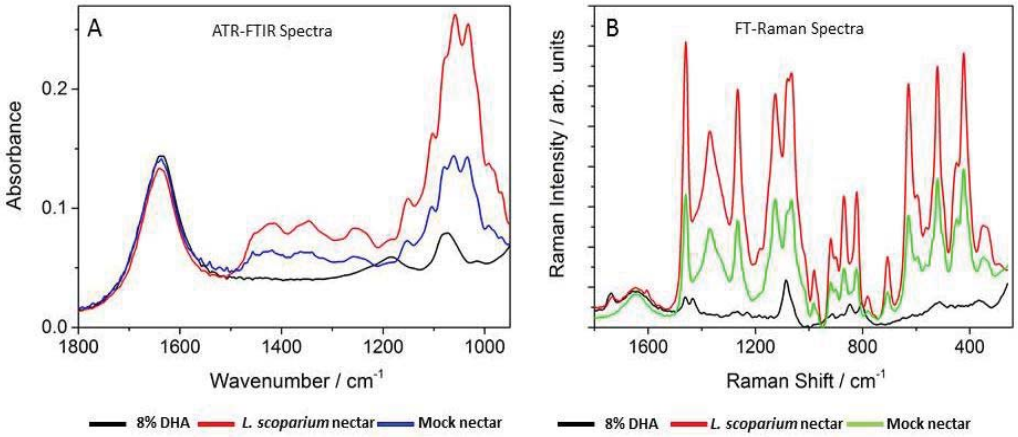

Figure 8.2. Comparison of FTIR (A) and FT-RAMAN (B) spectra of 8% DHA, *L. scoparium* nectar and a composition of mock nectar made from the saccharide solutions fructose and glucose (ratio 2:1).

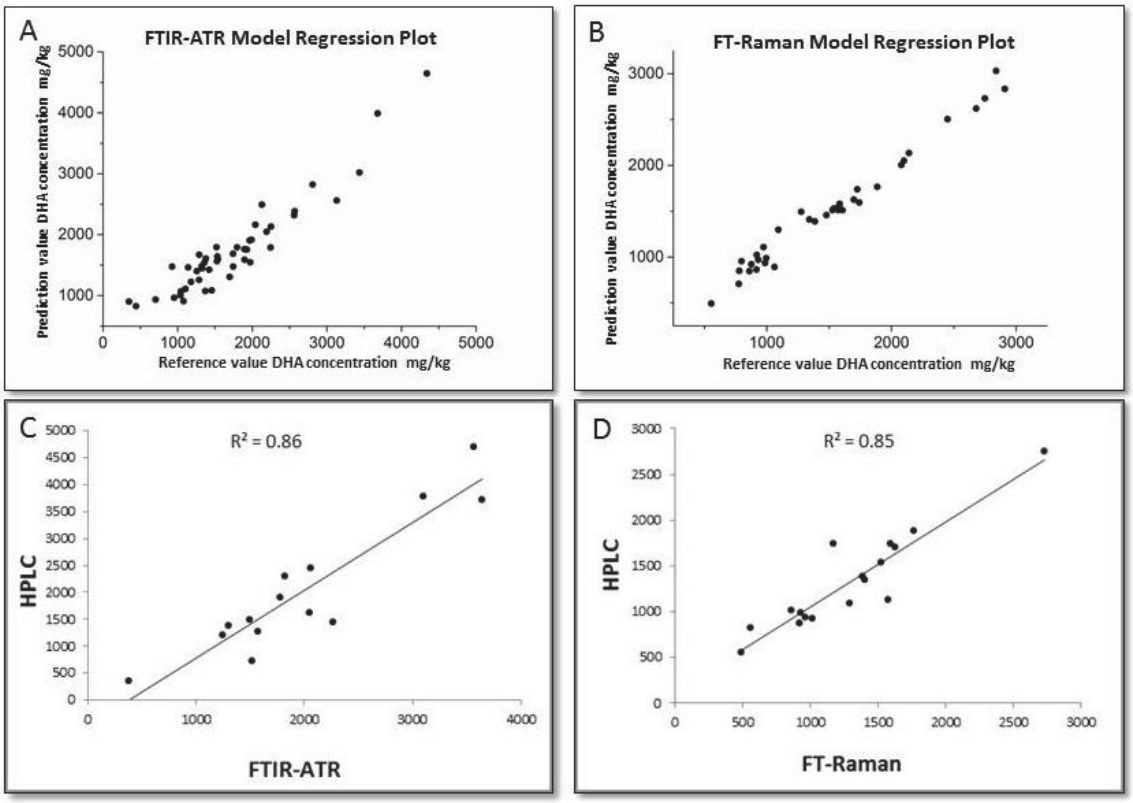

Figure 8.3. DHA regression model plots from FTIR-ATR (A) and FT-Raman (B) spectral data. Also regression plots of predicted data against HPLC data for FTIR-ATR (C) and FT-Raman (D)

8.4. Spectroscopic methods

For FTIR analysis of nectar, three replicates from each standard were analysed using a Bruker Tensor 37 MICRO-ID mid-IR source FTIR (Ettlingen, Germany). Samples were scanned at a resolution of 4 cm^{-1} and a scan average of 64 scans; wavelengths scanned 4000 to 0 cm^{-1}. Five micro-litres of each sample were placed on the zinc selenide ATR crystal for scanning. FTIR Spectra were analysed using PCA and PLS methods using statistics software Unscrambler V10.2 (CAMO, Norway). For FT-Raman analysis of nectar, samples were pipetted into aluminium divots and were analysed by Raman spectroscopy on a Bruker Equinox 55 interferometer equipped with the Bruker FRA 106/S FT-Raman accessory (Ettlingen, Germany) and a D418 T liquid-nitrogen-cooled

germanium detector was used, controlled by the Bruker OPUS v6.0 software and with a Nd:YAG 1064 nm excitation source. A 1 mm diameter laser spot size and 400 mW power setting were used, 120 scans per spectrum with a spectral resolution of 4 cm^{-1} in the wave-number range 3500 to 0 cm^{-1} were produced. Samples were analysed in triplicate. Raman spectra were analysed using PCA and PLS methods using statistics software Unscrambler V10.2 (CAMO, Norway). Reference spectra of 8% DHA-phosphate made up in milliQ H_2O and a mock nectar reference solution of fructose: glucose 2:1 were analysed with both ATR-FTIR and FT-Raman to compare spectral data between the two analytical methods.

Table 8.2. PLS models and Regression analysis results for nectar sugars in *L. scoparium* using ATR-FTIR and FT-Raman.

Sugar vs Method	R-squared	RMSEC|RMSEP	R^2 Model prediction vs HPLC
Fructose FTIR	0.97	0.63%|0.80%	0.97
Glucose FTIR	0.97	0.45%|0.65	0.95
Sucrose FTIR	0.94	0.06%|0.11	0.77
Total Sugars FTIR	0.99	1.16%|1.62%	0.95
DHA FTIR	0.88	264.52mg/kg|357.28	0.86
DHA FT-Raman	0.98	92.87mg/kg|397.62	0.85

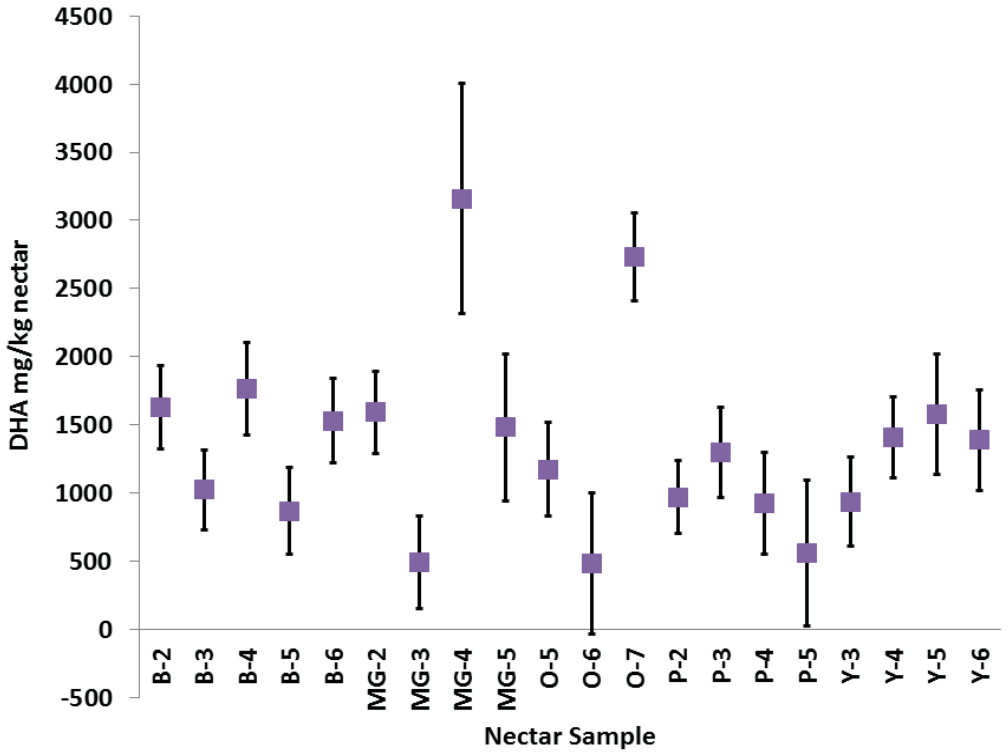

Figure 8.4. Prediction data plot from the FT-Raman regression model. Note samples with a larger standard deviation represented by the box around the mean are considered outliers in the prediction i.e. MG-4, O-6, MG-5 and P-5. Removal of these outliers gives a regression model of FT-Raman predicted values against HPLC data with an $R^2 = 0.85$.

8.5 Results and Discussion

The DHA, fructose, glucose, sucrose, and total sugar values obtained from the nectar of the cultivars using HPLC are shown in Table 1. Results show that the composition of *L. scoparium* nectar consists mainly of the monosaccharides fructose and glucose, in agreement with previous research [20], and also low concentrations of the disaccharide sucrose (see Table 1). *L. scoparium* also contains the triose sugar dihydroxyacetone (DHA) in significant concentrations [21–26]. Statistical analyses of the comparison of means of each cultivar show that there are five significantly different groups in regard to DHA concentration in the ten cultivars analysed, along with significant differences in monosaccharide and disaccharide composition. The saccharide data was normalised against the total sugar content to allow comparisons of the composition between the cultivars. Principal component analysis of the HPLC data of the saccharide nectar sugars glucose, fructose and sucrose show significant groupings according to cultivar (CV) type (Fig. 1A). In this plot, cultivars are grouped and spread mainly along PC1, which is influenced by glucose and fructose concentrations. Cultivars with higher fructose concentration are shown towards the positive end of PC1, and higher glucose concentration towards the negative end, as shown by the loadings for PC1 in Fig. 1B. Cultivar G has a significantly higher sucrose concentration compared to the other nine cultivars, as highlighted in the PC plot with CV G in a separate grouping along PC2 compared to the position of the other CV groups along PC2. PC2 is mainly influenced by sucrose concentrations as indicated by the loadings plot in Fig. 1B. Regression analysis of the two mono-saccharides (fructose against glucose) showed a negative correlation, indicating that as fructose

concentration increases, the glucose concentrations decrease in the nectar of *L. scoparium*. These data suggest an equilibrium not due to random effects within the sugar metabolism of floral nectar in *L. scoparium*. FTIR and FT-Raman spectral signatures for the saccharides fructose, glucose, and the sucrose are well documented [6, 9, 13, 15]. However there is a paucity of literature describing the spectral signature for the monosaccharide DHA. The Raman spectra of DHA has a distinct peak at 1740 cm^{-1} from the carbonyl bond *v*(C=O) [13] (Fig. 3), which is not present in other saccharides in the floral nectar of *L. scoparium* (Fig. 2B). However, at low concentrations of DHA, as are present in nectar assayed, this peak is not as distinct. The 1740 cm^{-1} peak is visible as a smaller peak in the Raman spectra in nectars of *L. scoparium* (Fig. 2B), where the DHA concentrations are higher; for example, Cultivars O and B (Table 1). Other peaks typical of saccharide sugar bonds are observed e.g. C-O stretch vibrations at 1077 cm^{-1} and 1062 cm^{-1}, C-C-H, O-C-H and C-O-H deformations at 1268 cm^{-1}, 1124 cm^{-1} and 915 cm^{-1}, CH_2 asymmetric and symmetric modes at 1457 cm^{-1}, 2946 cm^{-1} and 1365 cm^{-1}, C-H deformation at 777 cm^{-1} and 915 cm^{-1}. With various skeletal modes C-C-O and C-C-C at 524 cm^{-1} and 426 cm^{-1} and ring deformations showing as a weak band at 624 cm^{-1} [15]. The FTIR-ATR spectral signature for DHA is much less distinct (Fig. 2A). In the FTIR spectrum broad peaks are visible underneath the main distinct peaks representing the various (C-O), (C-C) and (OCH-) bonds present in sugars.

8.5.1 Regression Analysis

For FTIR spectra the wavenumber region 1650–700 cm^{-1} was used for regression analysis. PLSR analyses were cross validated with the leave-one-out method [27]. Spectra were either used raw or processed to the 1^{st} or 2^{nd} derivative depending on

the sugar being analysed for quantification. For FT-Raman, the region 1800 to 300 cm^{-1} was used. Standard spectral pre-processing was applied using Savitsky-Golay smoothing, SNV and linear baseline applications prior to partial least squares regression (PLSR). PLS regression models for both FTIR-ATR and FT-Raman were generated using spectral data against HPLC quantification data for each sugar (non-normalised) along with total sugars. The predictive models generated from the FTIR spectral data gave excellent correlations at ranges 0.88–0.99 R-squared values (Table 2). Sucrose and DHA are in much lower concentrations in the floral nectar, and the spectral data required further pre-processing to improve the model. Sucrose required 1^{st} order derivation and DHA data also required 1^{st} order derivation, both with a polynomial order of two of the spectra to optimise the models for these two sugars. The prediction performance was evaluated by mean squares error (RMSE) of calibration (RMSEC) and prediction (RMSEP) [28]. FT-Raman techniques used to analyse the nectar for quantification of DHA demonstrated that the improved contrast, with additional peaks, in Raman spectra could generate a more accurate model for DHA quantification (Table 2). Using FT-Raman, better results were provided for analysing DHA with a lower calibration error (RMSEC) from the model compared to FTIR, with an R^2 of 0.98 and RMSEP of 397.62 mg/kg sugar compared with R^2 of 0.88 and RMSEP of 357.28 mg/kg when using FTIR spectra (Fig. 4). Typically larger differences indicate a less reliable model. These error levels compare favourably with those obtained with HPLC techniques which gave a means standard error range of 111.35– 677.10 mg/kg depending on the cultivar nectar analysed. The error between the calibration error and prediction error for the Raman model is much larger than the differences in the FTIR model, this indicates that although the regression model for FT-

Raman is more accurate; the model has a larger chance of predicting errors, contrasted with the FTIR model which has a less accurate model but less chance of prediction-based errors. However, on analysis of the predicted data set from the models, outlier samples such as MG-4, O-6, MG-5 and P-5 can be detected from the method via larger standard deviation values and can be removed from the dataset (Fig. 4). Outliers could likely be due to interference from possible contaminants in the nectar such as nectar pigmentation, or possibly pollen contamination. Further analysis using predicted values from the PLSR models against the original HPLC data values for some samples was performed. Table 2 shows the regression results of the spectral analysis using PLS and the regression results showing predicted data from the PLS models against the HPLC data. R-squared values are very good for all nectar components measured excepting the sucrose model at R^2 = 0.77. The R^2 for the FT-Raman predictions against HPLC data are very good (0.85 R^2) after removing outliers. FTIR prediction data against HPLC data was also good (0.86 R^2). Overall an accurate model is possible for quantifying DHA concentrations in nectar using both FTIR-ATR and FT-Raman spectroscopy. In evaluating the model data against raw data for DHA, the range of standard error of the means (111.35 to 677.10 mg/kg DHA) is quite large for nectar samples, yet the prediction errors using spectral data with PLSR methods is well within the range of intra-plant variation in terms of nectar composition. Both ATR-FTIR and FT-Raman spectroscopy can be used as a quantitative model for screening *L. scoparium* nectar for DHA. Fructose, glucose, sucrose, and total sugars can be also accurately quantified using ATR-FTIR techniques, alongside multivariate analysis models using PLS. The results suggest that ATR-FTIR is a sensitive technique to analyse the sugars in the nectar of *L. scoparium*, including DHA.

8.6 Conclusion

It is of benefit to the mānuka honey industry to access facile methods such as analytical spectroscopy techniques, in combination with chemometrics, to screen for DHA in the nectar of *L. scoparium*, along with related applications to quantify nectar components of interest rapidly and non-destructively. ATR FTIR and FT-Raman techniques have both proven to have strong potential to be of use in this application.

8.7 Acknowledgements

This work was supported by a Primary Growth Partnership (Ministry of Primary Industries, New Zealand) awarded to Manuka Research Partnership (NZ) Limited.

8.8 References

[1] Adams C.J., Boult C.H., Deadman B.J., Farr J.M., Grainger M.N.C., Manley-Harris M., Snow M.J. Isolation by HPLC and characterization of the bioactive fraction of New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2008),** 344(18): p. 2609.

[2] Adams C.J., Manley-Harris M., Molan P.C. The origin of methylglyoxal in New Zealand manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2009)**, 344(8): p. 1050-1053.

[3] Windsor S., Pappalardo M., Brooks P., Williams S. and Manley-Harris M. A convenient new analysis of dihydroxyacetone and methylglyoxal applied to Australian *Leptospermum* honeys. Journal of Pharmacognosy and Phytotherapy **(2012)**, 4(1): p. 6-11.

[4] Sultanbawa Y., Cozzolino D., Fuller S., Cusack A., Currie M., Smyth H. Infrared spectroscopy as a rapid tool to detect methylglyoxal and antibacterial activity in Australian honeys. Food Chemistry **(2015)**, 172: p. 207-212.

[5] McKenna M. A., and Thomson J. D. A technique for sampling and measuring small amounts of floral nectar. Ecology **(1988)**, 69(4): p. 1306-1307.

[6] de Oliveira G. A., de Castilhos F., Renard C. M-G. C., Bureau S. Comparison of NIR and MIR spectroscopic methods for determination of individual sugars, organic acids and carotenoids in passion fruit. Food Research International **(2014)**, 60: p. 154-162.

[7] Windsor S., Brooks P.A convenient new analysis of dihydroxyacetone and methylglyoxal applied to Australian Leptospermum honeys. Chemistry Central Journal **(2012)**, 4(1): p. 6-11.

[8] Lin C-A., Ayvaz H., Rodriguez-Saona L.E. Application of Portable and Handheld Infrared Spectrometers for Determination of Sucrose Levels in Infant Cereals Food Analytical Methods **(2014)**, 7(7): p. 1407-1414.

[9] Armenta S., Garrigues S., de la Guardia M., Rondeau P. Application of Portable and Handheld Infrared Spectrometers for Determination of Sucrose Levels in Infant Cereals. Analytica Chimica Acta **(2005)**, 545(1): p. 99-106.

[10] Anjos O., Campos M.G., Ruiz P.C., Antunes P. Application of FTIR-ATR spectroscopy to the quantification of sugar in honey Food Chemistry. **(2015)**, 169: p. 218-223.

[11] Ilaslan K., Boyaci I.H., Topcu A. Rapid analysis of glucose, fructose and sucrose contents of commercial soft drinks using Raman spectroscopy. Food Control **(2015)**, 48: p. 56-61.

[12] Schulz H., Baranska M., Baranski R. Potential of NIR-FT-Raman spectroscopy in natural carotenoid analysis. Biopolymers **(2005)**, 77(4): p. 212-221.

[13] Schulz H., Baranska M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vibrational Spectroscopy **(2007)**, 43(1): p. 13-25.

[14] Killeen D.P., Sansom C.E., Lill R.E., Eason J.R., Gordon K.C., Perry N.B. Quantitative Raman Spectroscopy for the Analysis of Carrot Bioactives. Journal of Agricultural and Food Chemistry **(2013)**, 61(11): p. 2701-2708.

[15] Pierna J.A.F., Abbas O., Dardenne P., Baeten V. Discrimination of Corsican honey by FT-Raman spectroscopy and chemometrics. Biotechnologie Agronomie Societe Et Environnement **(2011)**, 15(1): p. 75-84.

[16] Gierlinger N., Schwanninger M. The potential of Raman microscopy and Raman imaging in plant research. Spectroscopy-an International Journal **(2007)**, 21(2): p. 69-89.

[17] Mazurek S., Szostak R. Quantification of aspartame in commercial sweeteners by FT-Raman spectroscopy. Food Chemistry **(2011)**, 125(3): p. 1051-1057.

[18] Yang Y., Nie P., Yang H., He Y., Yang Y., Nie P.C., Yang H.Q., He Y. Transactions of the Chinese Society of Agricultural Engineering **(2010)**, 26(3): p. 238-242.

[19] Davis A.R. Influence of floral visitation on nectar-sugar composition and nectary surface changes in Eucalyptus. Apidology **(1997)**, 28: p. 27-28.

[20] Williams S., King J., Revell M., Manley-Harris M., Balks M., Janusch F., Kiefer M., Clearwater M., Brooks P. and Dawson M. Regional, annual, and individual variations in the Dihydroxyacetone content of the nectar of Mānuka (*Leptospermum scoparium*) in New Zealand. Journal of Agricultural and Food Chemistry **(2014)**, 62(42): p. 10332-10340.

[21] Atrott J., Haberlau S., Henle T. Studies on the formation of methylglyoxal from dihydroxyacetone in Manuka (*Leptospermum scoparium*) honey. Carbohydrate Research **(2012)**, 361: p. 7-11.

[22] Subhashree V., Subashini D. Estimation of total flavonoids, phenols and antioxidant activity of local and New Zealand manuka honey. Journal of Pharmacy Research **(2011)**, 4(2): p. 464-466.

[23] Stephens J.M., Schlothauer R.C., Morris B.D., Yang D., Fearnley L., Greenwood D. R., Loomes K.M. Phenolic compounds and methylglyoxal in some New Zealand manuka and kanuka honeys. Food Chemistry **(2010)**, 120: p. 78-86.

[24] Donarski J.A, Roberts D.P.T, Charlton A.J. Phenolic compounds and methylglyoxal in some New Zealand manuka and kanuka honeys. Analytical Methods **(2010)**, 2(10): p. 1479-1483.

[25] Mavric E., Wittmann S., Barth G., Henle T. Identification and quantification of methylglyoxal as the dominant antibacterial constituent of Manuka (*Leptospermum scoparium*) honeys from New Zealand. Molecular Nutrition & Food Research **(2008)**, 52(4): p. 483-489.

[26] Stephens J.M.C., Molan P.C., Clarkson B.D. A review of *Leptospermum scoparium* (Myrtaceae) in New ZealandNew Zealand Journal of Botany **(2005)**, 43(2): p. 431-449.

[27] Esbensen K.H., Geladi P. Principles of Proper Validation: use and abuse of re-sampling for validation. Journal of Chemometrics **(2010)**, 24(3-4): p. 168-187.

[28] Liu Y., Lyu Q., He S., Yi S., Liu X., Xie R., Zheng Y., Deng L. Prediction of nitrogen and phosphorus contents in citrus leaves based on hyperspectral imaging. International Journal of Agricultural and Biological Engineering **(2015)**, 8(2): p. 80-88.

Chapter 9: Discussion and Future Work

9.1 Introduction

The work described in this thesis was done as part of an ongoing comprehensive research programme that is attempting to research the technical and economic feasibility of combining improved plant genetics with plantation husbandry techniques along with the development of predictive tools to increase and manage Mānuka honey supply.

This chapter summarises findings from research into phenotypic differences between a range of *Leptospermum scoparium* cultivars, the plant and soil interactions of *Leptospermum scoparium* (Mānuka) and various New Zealand soils, along with the novel application of spectroscopy techniques to screen *L. scoparium* plants for dihydroxyacetone (DHA) potential. The figure below (Figure 9.1) gives an overall illustration of the research outcomes displaying the interplay of environmental versus genetic influences on the main parameters of interest to this research program. The figure represents the key influences on nectar and plant growth parameters with the outer boxes representing the influences and the inner circle showing the important parameters of relevance to increasing UMF honey yields. (G representing genetic influences and E representing environmental influences).

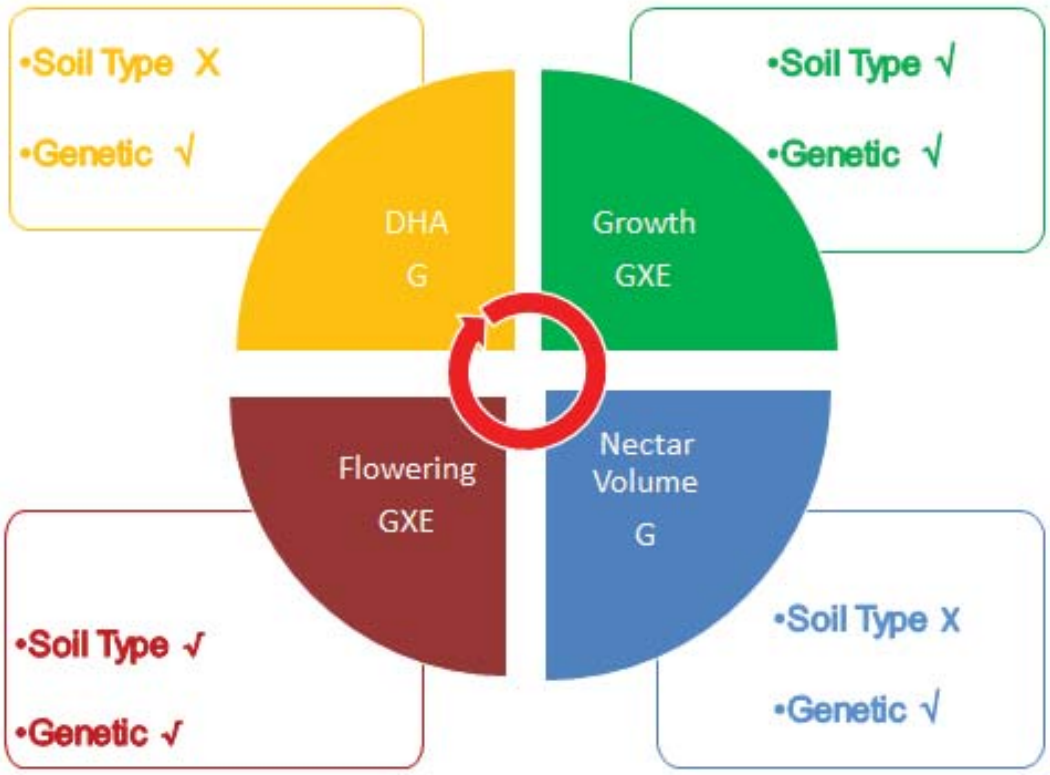

Figure 9.1. Illustration representing the interplay of influences on plant response factors relevant to the UMF honey yield.

9.2 Research Objectives

The general objectives of this research were to:

1) Assess phenotypic variability in ten different cultivars of *Leptospermum scoparium* with regard to plant growth, nectar composition and nectar yield.
2) Assess a range of environmental influences on growth and nectar production and composition in *Leptospermum scoparium* with particular reference to the influence of soil type.

3) Assess the capability of analytical spectroscopy techniques to chemotype leaf metabolites in *Leptospermum scoparium* allowing facile identification of relevant cultivars.

9.3 Discussion

The outcomes of the general research objectives as described in section 9.2 can discussed within four key themes that detail the main findings of the experimental work investigating influences on nectar composition and yield in *Leptospermum scoparium.*

9.3.1. Research of available literature to establish standard methodologies for nectar collection and decide on best practises for experimental design for assessing nectar composition, plant growth, flowering and overall nectar yield in *L. scoparium* (Chapters 3 and 5)

Standard methods of nectar collection were established and adopted for this thesis. Experimental design is an important concept not to be understated. Good experimental design allows accurate information to be collected and data to have statistically relevant conclusions drawn. Glasshouse layouts and replicate plant sample numbers were established to account for any known and unknown variables in the experiments. The nectar collection method was standardised and collection times were also set within a fixed time frame per day. A standard flower development stage (Appendix A) was chosen for nectar collection from literature references reporting flower development in the Myrtaceae family to which the species *L. scoparium* belongs.

All standardisation of methods were established from research on the available literature together with observations made during preliminary studies of the plants in the glasshouse. Additional information from field site data from other research undertaken within the same PGP program was also taken into consideration. A benefit from working within a larger scale research group such as the primary growth partnership scheme is that collaborative information, experience and results from the group's collective work can be discussed with colleagues and methodologies agreed upon and adapted for future work.

9.3.2 Identification of the variance of specific nectar and growth parameters within several cultivars of *L. scoparium* to establish the potential of breeding programmes for increasing UMF honey yield from *L. scoparium*. Model the effects of a range of climatic conditions on nectar composition and yield in several cultivars of *L. scoparium* to allow identification of localities where climatic factors may be beneficial or detrimental to overall UMF honey yield (Chapter 5 and 6)

Genetics is shown to be the dominant determining factor of nectar composition and yield in *Leptospermum scoparium* cultivars and therefore breeding programmes have the most potential to increase overall UMF honey yield. Indications from this research are that in addition to using high yield cultivars, higher soil nutrients, sunshine hours and radiation levels will further improve UMF yield to the honey industry.

The significant difference observed in flowering phenology between cultivars has a critical impact on UMF yield but also a follow on impact in terms of food resources available to bees. Flowering phenology i.e. when plants flower and how long they flower for is of crucial importance for beehive management. Plants need to flower when bees are able to take advantage of nectar resources. Bees are dormant over

winter and active in warmer spring temperatures through summer into mid-autumn, depending on daytime temperatures, therefore cultivars that flower during these conditions and in sufficient numbers are essential to support the industry. Longer flowering period cultivars in addition to greater overall nectar yield will have the added benefit of reduced hive transport costs to the beehive industry, short term flowering period cultivars will mean increased transport of hives to various available resources. For these reasons, breeding plants that flower at the right temperatures for bee activity along with long flowering periods in addition to higher nectar and DHA yield is an important priority for the future of the UMF honey industry.

The importance of assessing not just DHA concentration in deducing cultivar value, but overall nectar yield and plant parameters such as plant growth and flowering phenology cannot be underestimated. The findings of this thesis suggest that these key features must be explored when assessing *L. scoparium* plants within breeding programs, prior to cultivar selection for large-scale field production of high UMF Mānuka honey.

9.3.3 Identification the effect of the influence of the environment in terms of soil composition on nectar, plant growth and flowering parameters to allow decisions to be made regarding best soil sites to establish *L. scoparium* plantations. (Chapters 5 and 6)

Environmental factors also need to be considered when assessing plants for high UMF potential. An innate ability to adapt to the environment is a requirement for plants to be successful in establishment. Environmental conditions such as soil type and climatic factors have been shown to have a significant influence on some cultivars in terms of plant growth, flowering and overall nectar yield. Modelling of the environmental

influences on UMF yield for any specific cultivar being considered for plantation planting would be productive. There is little published work in this field with relevance to *Leptospermum* species therefore this research adds value to the understanding the influence of soil composition in this species, knowledge of the influence of environment on cultivars has potentially positive outcomes for managing the future of the developing Mānuka honey industry.

Data from the soil experiments investigating the effect of soil type on three different cultivars of *Leptospermum scoparium* showed that variable flowering phenology, nectar yield and plant growth occurred in Mānuka on a number of soils under conditions of constant and controlled environment. The experimental design of the soils experiment removed variable climatic influences such as are present in field locations and also present as microclimates within glasshouses, with the design allowing the investigation of soil composition factors without possible confounding climatic influences. This afforded a more detailed analysis of soil composition factors influencing plant growth, flowering and nectar composition in *L. scoparium.* Experimental data showed that, contrary to published literature that indicate that *L. scoparium* prefers poorer soils, *L. scoparium* actually has improved growth and flowering on soils with higher nutritional value with a significant effect on flowering with respect to floral density and flowering period in some cultivars. This result supports the theory that soil factors influence plant development and flowering when climatic conditions are kept constant and that soil type in land areas being considered for Mānuka plantations is an essential aspect to consider when growing high yield Mānuka for the UMF industry in New Zealand. Results indicated that the effect of soil type does depend on the cultivar for some parameters and so consideration needs to

be given to choosing relevant cultivars for planting. Importantly Dihydroxyacetone concentrations in nectar were not affected by soil type. Soil components; phosphorus, sulphate, ferric and chloride were shown to be the common nutrients influencing the plant parameters measured, further work elucidating fertilisation regimes that best promote growth and flowering in *L. scoparium* would be very productive to the Mānuka UMF industry. The effect of soil type on flowering phenology in *L. scoparium* has not been previously reported, along with using the plant soil interaction data to develop modelling techniques advantageous in predicting Mānuka honey supply.

Data from the experimental work showed that some cultivars of *L. scoparium* in this study were robust against environmental influences and some were not. This result indicates that in establishing plantations of Mānuka the ideal environmental growing conditions for specific cultivars must be assessed. Cultivars being considered for plantation establishment should be chosen on the grounds of whether the climatic conditions of the region and aspect of the land and soil nutrition favours higher UMF potential for any specific cultivar. It should also be considered that cultivars of *L. scoparium* can also be chosen on the grounds of robustness against influencing factors in regions considered for plantation planting as robustness against environmental conditions has the advantage to ensure a steady predictable yield to the UMF honey industry.

9.3.4 Research into the analytical development of FT-Raman spectroscopy to chemotype cultivars of *L. scoparium* and investigate the potential for linking metabolites from chemotype profiles produced by FT-Raman data to UMF potential in *L. scoparium* and test the capability of FT-Raman and ATR-FTIR spectroscopy along with chemometric modelling techniques to quantify nectar components of *L. scoparium* (Chapters 4, 7 and 8)

Adapting HPLC methods used for honey analysis proved a successful method for accurate quantification of nectar components. Spectroscopic techniques FT-Raman, Attenuated Total Reflectance (ATR)-FTIR techniques were compared in terms of their capability to quantitatively assay for DHA. FT-RAMAN, ATR-FTIR alongside chemometrics, principal component analysis (PCA) and partial least square regression analysis (PLS) were shown to be quantitative analytical techniques with the capability to quantify the nectar composition of *L. scoparium*. There are multiple benefits to be gained for the honey industry by taking advantage of techniques such as analytical spectroscopy in combination with chemometrics to screen for DHA potential in the nectar of *L. scoparium* cultivars. Along with its application to quantify nectar components of interest rapidly and non-destructively, facile analytical methods are usually cost effective due to ease of use and the lower cost of consumables.

The analytical spectroscopy experiments on leaf grinds illustrated that FT-Raman spectroscopy is a practical tool for chemotyping *Leptospermum scoparium* cultivars. In depth analysis of the PCA loadings from the principal component analysis demonstrates that the main chemical components differentiating between the cultivars are carotenoids, flavonoids, anthocyanins and terpenes. PLS regression techniques demonstrated that FT-Raman analysis of leaf material also correlated the

chemical groups; carotenoids, flavonoids and terpenes with DHA potential. The correlation of metabolites, carotenoids, flavonoids and terpenes with dihydroxyacetone (DHA) is not necessarily a cause and effect relationship but could be an overall metabolic correlation. Results showed that increases in sunshine hours led to an increase in nectar yield in some cultivars. Flavonoids are a group of metabolites associated with light interactions in plants, signalling metabolic reactions such as stress responses and have been shown to be connected to metabolic signal transduction in some plant processes. Possibly there is some metabolic signal inherent in some of the cultivars that are influenced by climate factors such as sunshine and radiation and lead to an increase metabolic processes in the whole plant therefore also increasing nectar metabolism possibly through the glycolytic pathway in nectaries although this connection is yet to be proven. To elucidate these processes, metabolomic and genetic studies would need to be undertaken. Future work could include a full analysis of leaf pigment composition coupled with isolation of individual compounds to elucidate exactly which chemical constituents are correlating with the nectar composition in *L. scoparium,* this may give insights into which metabolic pathways could be connected with nectar metabolism.

Development of spectroscopy methods in this thesis have allowed the potential to develop a screening model for plant metabolites that correlate with high DHA potential. The PLS regression shows that Ft-Raman methods can be used to screen for DHA concentrations in the range 3300-7600 mg/kg plus or minus a 20% standard error, therefore using regression FT-Raman models low, medium and high DHA synthesis plants can be distinguished within the DHA range studied.

Taking advantage of these techniques along with further work to develop a more inclusive model that can be applied across a larger range of DHA potential in the *L. scoparium* species will allow breeding programmes to be fast tracked. Screening of leaf material negates the need to wait for plants to grow and flower substantially enough for nectar collection. Screening of leaf material, as opposed to growing plants from seeds or cuttings to the stage where they have enough biomass for sufficient flowers to be available for nectar collection and analysis, would save up to 2 years each breeding round. Application of spectral methods will allow new plants from wild plantations to be screened for DHA potential for faster inclusion into breeding programmes. These factors are of benefit to developing the industry and shorter breeding programmes would be very beneficial to fast-track the development of high yield DHA plants for the New Zealand industry.

In addition screening for leaf metabolites could be beneficial to other commercial products made from Mānuka such as tea tree oils, the compounds of use in the essential oil industry could easily be screened using FT-Raman in preference to complex methods such as LCMS by applying chemotype fingerprints of plants with the highest potential for leaf oil extraction.

9.4 Recommendations for Future Work

Utilisation of the out comes from this work have the potential to have a key influence on the UMF honey industry in New Zealand and contribute significantly towards meeting the requirements of the PGP contract to research the technical and economic feasibility of combining improved genetics with Mānuka husbandry techniques to

achieve productivity gains. Four key recommendations that will help the UMF industry in New Zealand to reach its targets are suggested below:

1) Use the methods developed in this thesis to assess plants from breeding programmes for their overall potential UMF yield based on floral density and nectar volumes along with dihydroxyacetone potential and suitability to the climate in various regions considered for plantation planting.
2) Adopt facile analytical methods to develop models for screening for dihydroxyacetone potential in *Leptospermum scoparium* to allow fast tracking of breeding programmes.
3) Adopt facile methods such as FTIR and FT-Raman, which are regularly used in the food industry, to assay for sugar content and DHA in nectar to reduce analytical costs and time for nectar analysis.
4) Investigate the effects of fertilisation regimes on growth and flowering of *Leptospermum scoparium* in field conditions.

Application of the scientific knowledge gained from this body of research will contribute substantially towards the goal of establishing high UMF producing Mānuka plantations.

Appendix A

A.1 Flower stage development for nectar collection.

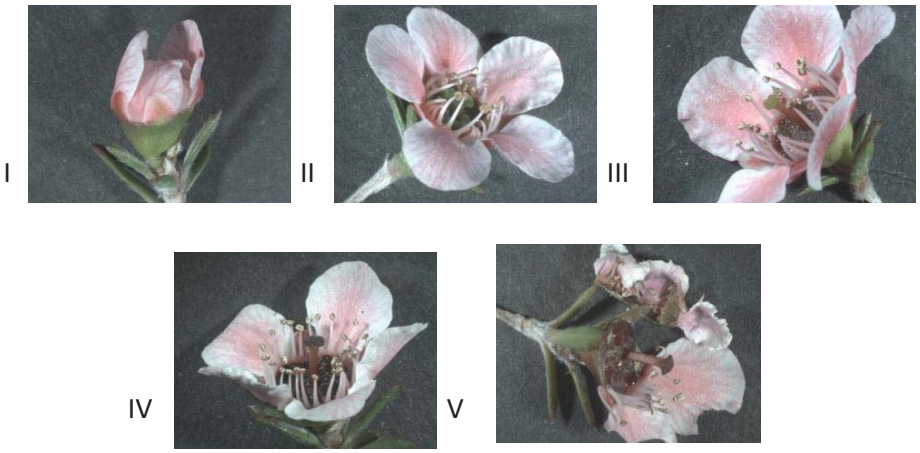

Figure A.1. Illustrating the five stages of flower development in *Leptospermum scoparium,* nectar was collected at stage IV.

Figure A.2. Photograph of a cross section through the middle of a *Leptospermum scoparium* flower. Dark red pigmented area is the hypanthium surface containing the nectaries. The moisture visible on the hypanthium surface is nectar.